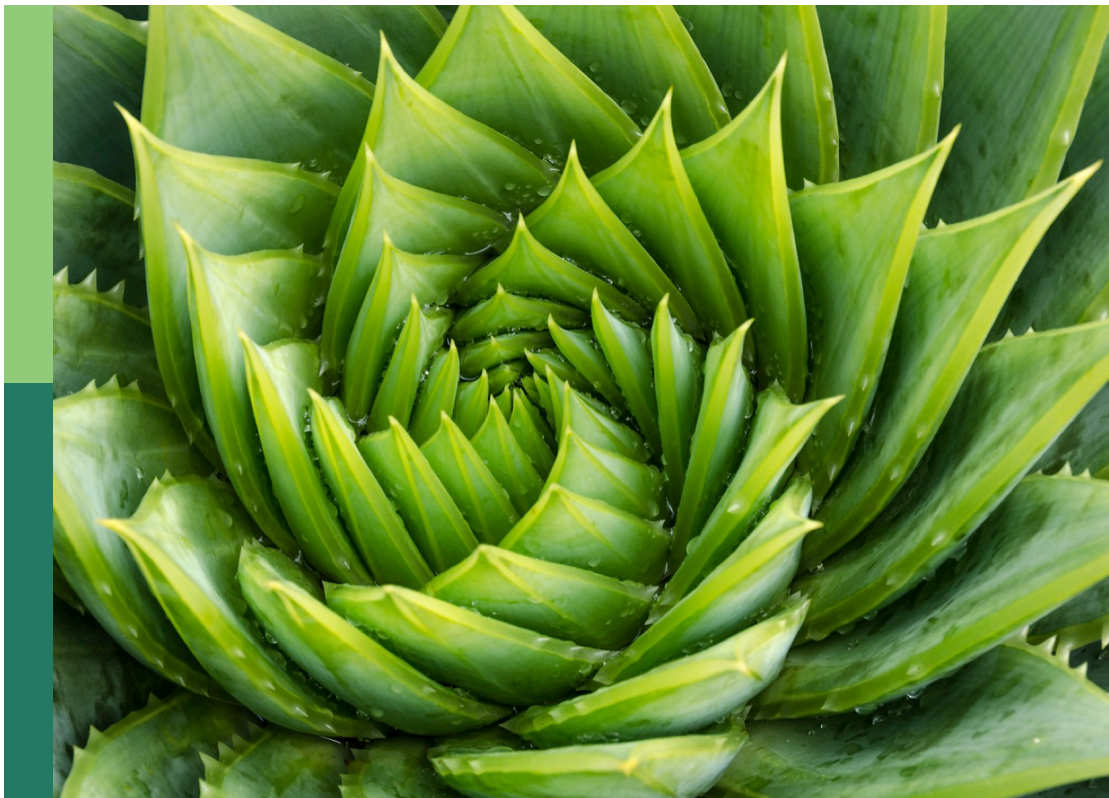
frontiers | Research Topics

Genetic improvement of *Triticeae* crops based on high-throughput phenotyping: Molecular design for yield, resistance and tolerance

Edited by
Xiaoli Fan, Yunfeng Xu, Fa Cui, Pengtao Ma and Shuyu Liu

Published in
Frontiers in Plant Science



ISSN 1664-8714
ISBN 978-2-83251-992-9
DOI 10.3389/978-2-83251-992-9

About Frontiers

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

Frontiers journal series

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

Dedication to quality

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

What are Frontiers Research Topics?

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

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

Genetic improvement of *Triticeae* crops based on high-throughput phenotyping: Molecular design for yield, resistance and tolerance

Topic editors

Xiaoli Fan — Chengdu Institute of Biology, Chinese Academy of Sciences (CAS), China

Yunfeng Xu — Kansas State University, United States

Fa Cui — Ludong University, China

Pengtao Ma — Yantai University, China

Shuyu Liu — Texas A&M University System, United States

Citation

Fan, X., Xu, Y., Cui, F., Ma, P., Liu, S., eds. (2023). *Genetic improvement of* Triticeae *crops based on high-throughput phenotyping: Molecular design for yield, resistance and tolerance*. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-83251-992-9

Table of contents

frontiers | Frontiers in Plant Science

TYPE Editorial
PUBLISHED 09 March 2023
DOI 10.3389/fpls.2023.1165461

OPEN ACCESS

EDITED AND REVIEWED BY
Nunzio D'Agostino,
University of Naples Federico II, Italy

*CORRESPONDENCE
Pengtao Ma
ptma@ytu.edu.cn

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

RECEIVED 14 February 2023
ACCEPTED 03 March 2023
PUBLISHED 09 March 2023

CITATION
Fan X, Ma P, Cui F, Xu Y and Liu S (2023) Editorial: Genetic improvement of *Triticeae* crops based on high-throughput phenotyping: Molecular design for yield, resistance and tolerance.
Front. Plant Sci. 14:1165461.
doi: 10.3389/fpls.2023.1165461

Editorial: Genetic improvement of *Triticeae* crops based on high-throughput phenotyping: Molecular design for yield, resistance and tolerance

Xiaoli Fan[1], Pengtao Ma[2]*, Fa Cui[3], Yunfeng Xu[4] and Shuyu Liu[5]

[1]Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu, China, [2]School of Life Sciences, Yantai University, Yantai, Shandong, China, [3]School of Agriculture, Ludong University/Key Laboratory of Molecular Module-Based Breeding of High Yield and Abiotic Resistant Plants in Universities of Shandong, Yantai, China, [4]Agricultural Genomics Institute at Shenzhen (AGIS), Chinese Academy of Agricultural Sciences (CAAS), Shenzhen, China, [5]Texas A&M AgriLife Research, Texas A&M University System, Amarillo, TX, United States

KEYWORDS

molecular design, high-throughput phenotyping, stress tolerance, yield, *Triticeae* crops

Editorial on the Research Topic
Genetic improvement of *Triticeae* crops based on high-throughput phenotyping: Molecular design for yield, resistance and tolerance

Introduction

Triticeae are the foremost staple foods worldwide and also the most widely cultivated crops. Their resistance to biotic and abiotic stresses is always important during genetic improvement and thus greatly influences their adaptability and cultivation ranges. However, still-evolving pathogens and pests, harsh weather, reduced resources, etc., all threaten sustainable crop production, making resistance and tolerance improvement the most important concerns in current *Triticeae* breeding programs. Meanwhile, with the global population increasing, crop production is facing great growth challenges to meet food security demands, further emphasizing the importance of overcoming these adverse conditions that jeopardize crop production. Fortunately, in recent years, with the accelerated progress of omics research, resistance/tolerance-related genes have been rapidly discovered, and associated diagnostic single nucleotide polymorphisms have been effectively developed, making the application of germplasm resources for molecular design breeding more efficiently, and providing sustainable sources of resistance and yield improvement for *Triticeae* crops to ensure future food security.

In relation to this topic, a total of 140 authors contributed their recent advances in disease resistance, stress tolerance and yield performance studies on *Triticeae* crops, which are presented in 13 research articles.

Biotic stress response in *Triticeae* crops

Powdery mildew, a globally epidemic disease in *Triticeae* crops caused by the biotrophic fungus *Blumeria graminis* f. sp. *tritici* (*Bgt*), can severely affect yield and quality. Even though more than 100 powdery mildew (*Pm*) genes/alleles have been found in *Triticeae* crops, only a few of them have been applied in wheat improvement (Han et al., 2022). These crops are facing strong selection pressure, and continual identification and utilization of *Pm* genes/alleles from various germplasm resources are needed for rational deployment. Li et al. reported a novel *Pm* gene, *PmSN0293*, on wheat chromosome 6A, putatively from *Thinopyrum ponticum*, and developed two wheat–*Th. ponticum* introgression lines carrying this new *PmSN0293* and the previously reported *Pm2* and *Pm52*, which exhibited excellent application potential in wheat breeding programs. Jin et al. identified a new splicing variant of Pm4, PmYAV, in synthetic hexaploid wheat, which was developed by hybridization of diploid Aegilops and tetraploid wheat, and four molecular markers available for marker-assisted selection (MAS) were screened. Liu et al. identified a major quantitative trait locus (QTL) *QPm.cas-7D* for adult plant resistance to wheat powdery mildew in a well-known wheat–*Agropyron cristatum* introgression line, PuBing3228. It was subsequently deduced that QPm.cas-7D was Pm38 and that a 3-bp InDel between resistant and susceptible haplotypes was responsible for the resistance. Qiu et al. identified a resistant haplotype carrying the Pm locus, PmH962, which was confirmed to be the reported Pm5e and had no negative effect on wheat agronomic performance.

In addition to powdery mildew, resistance to other destructive diseases, such as leaf rust and stem rust, were also presented. Zhang et al. applied 112 wheat accessions introduced from the U.S. National Plant Germplasm System to evaluate resistance to the Chinese predominant races causing leaf rust and their resistance genes in order to explore more effective resistance resources for overcoming wheat leaf rust. Kataria and Kaundal performed a comparative analysis of protein-protein interactions (PPIs) between two pathogen races, *Pgt* 21-0 and *Pgt* Ug99, both causing wheat stem rust, and elucidated the functional differences between these two races, thus providing the strain-specific information for the development of durable, disease-resistant crop lines.

In addition to the genetic basis for pathogen stress, Luo et al. also reported the fine mapping study of a *Hairy glume* gene, responsible for trichomes on wheat glumes, which is largely involved in resistance to various biotic and abiotic stresses, as well as in defense to against insect pests.

Abiotic stress response in *Triticeae* crops

Natural cultivation conditions, such as light, nutrients, water, and salinity, determine crop adaptability and affect their yield and quality. Yang et al. discovered that low-light stressed wheat could alter its pollination type to enable outcrossing with heterologous pollen by increasing lemma and glume angles, which finally compensated for the 2.1–18.0% loss in grain number. Liu et al. analyzed the transcriptional mechanism of the response to low-nitrogen stress of a previously located major stable QTL for wheat root growth, QMrl-7B, and found genes encoding NO^{3-} transporters, etc, composing the complex regulatory network for root determination. Li et al. also performed the transcriptome profiling of the well-known transcription factor gene family, WRKYs, in Tritiprum and the response of TtWRKY256 to salt stress. Moreover, considering that the chlorophyll content could directly impact photosynthesis and affect crop health, Wang et al. proposed an unmanned aerial vehicle (UAV)-based approach to rapidly and efficiently predict chlorophyll content under irrigation and drought stress to provide insights into the capacity of UAV-based remote sensing for phenotyping to improve crop breeding.

Yield-related trait response to stresses in *Triticeae* crops

Maintaining yield performance under various unfavorable conditions is the main concern for breeders during genetic improvement of resistance or tolerance to both biotic and abiotic stresses (Deng et al., 2017). Wang et al. studied a set of 18 yield-related and agronomic traits in wheat under different irrigation regions by QTL analysis to provide insight into the genomic regions contributing to high yield in water-limited conditions and reported a novel QTL stably controlling wheat kernel length under all tested environments, facilitating the future MAS for pyramiding the favorable loci for high-yield improvement. Zhang et al. identified three QTLs for wheat spikelet nodes per spike and investigated incorporation of their favorable alleles across different wheat agroecological production zones of China, such as Chengdu of the Southwestern Wheat Zone and Yuncheng and Linfen of the Huai River Valleys Facultative Wheat Zone, where the natural environmental factors were significantly different, advancing our understanding of the genetic basis of natural variation in spikelet development and its adaptation to environmental change.

In summary, innovative and attractive advances have been made in understanding both biotic and abiotic stress responses of *Triticeae* crops and their contribution to yield potential, which will facilitate wheat genetic improvement and food safety.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

This work is supported by the Major Program of National Agricultural Science and Technology of China (NK20220607) and the Major Science and Technology Achievement Transformation of Central Universities and Institutes in Sichuan Projects (2022ZHCG0131).

Acknowledgments

We are greatly appreciated for the contributions from all the authors and reviewers as well as the support of the editorial office of *Frontiers in Plant Science*.

Conflict of interest

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

Publisher's note

References

Deng, Y. W., Zhai, K. R., Xie, Z., Yang, D. Y., Zhu, X. D., Liu, J. Z., et al. (2017). Epigenetic regulation of antagonistic receptors confers rice blast resistance with yield balance. *Science* 355, eaai8898. doi: 10.1126/science.aai8898

Han, G. H., Yan, H. W., Wang, J., Cao, L. J., Liu, S. Y., Li, X. Q., et al. (2022). Molecular cytogenetic identification of a new wheat-rye 6R addition line and physical localization of its powdery mildew resistance gene. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.889494

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 03 May 2022
doi: 10.3389/fpls.2022.892642

Genetic Incorporation of the Favorable Alleles for Three Genes Associated With Spikelet Development in Wheat

Xiaojun Zhang[1†], Linyi Qiao[1,2†], Xin Li[1], Zujun Yang[3], Cheng Liu[4], Huijuan Guo[1], Jun Zheng[5], Shuwei Zhang[1], Lifang Chang[1], Fang Chen[1], Juqing Jia[1], Liuling Yan[2] and Zhijian Chang[1]**

[1] State Key Laboratory of Sustainable Dryland Agriculture (in Preparation), College of Agronomy, Shanxi Agricultural University, Taiyuan, China, [2] Department of Plant and Soil Sciences, Oklahoma State University, Stillwater, OK, United States, [3] School of Life Sciences and Technology, University of Electronic Science and Technology of China, Chengdu, China, [4] Crop Research Institute, Shandong Academy of Agricultural Sciences, Jinan, China, [5] Institute of Wheat Research, Shanxi Agricultural University, Linfen, China

OPEN ACCESS

Edited by:
Pengtao Ma,
Yantai University, China

Reviewed by:
Hongjie Li,
Institute of Crop Sciences (CAAS), China
Jian Ma,
Sichuan Agricultural University, China

***Correspondence:**
Liuling Yan
liuling.yan@okstate.edu
Zhijian Chang
wrczj@126.com

[†] These authors have contributed equally to this work

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

Received: 09 March 2022
Accepted: 14 April 2022
Published: 03 May 2022

Citation:
Zhang X, Qiao L, Li X, Yang Z, Liu C, Guo H, Zheng J, Zhang S, Chang L, Chen F, Jia J, Yan L and Chang Z (2022) Genetic Incorporation of the Favorable Alleles for Three Genes Associated With Spikelet Development in Wheat. Front. Plant Sci. 13:892642. doi: 10.3389/fpls.2022.892642

The number of spikelets per spike is an important trait that directly affects grain yield in wheat. Three quantitative trait loci (QTLs) associated with spikelet nodes per spike (SNS) were mapped in a population of recombinant inbred lines generated from a cross between two advanced breeding lines of winter wheat based on the phenotypic variation evaluated over six locations/years. Two of the three QTLs are *QSns.sxau-2A* at the *WHEATFRIZZY PANICLE* (*WFZP*) loci and *QSns.sxau-7A* at the *WHEAT ORTHOLOG OF APO1* (*WAPO1*) loci. The *WFZP-A1b* allele with a 14-bp deletion at *QSns.sxau-2A* was associated with increased spikelets per spike. *WAPO-A1e*, as a novel allele at *WAPO1*, were regulated at the transcript level that was associated with the SNS trait. The third SNS QTL, *QSns.sxau-7D* on chromosome 7D, was not associated with homoeologous *WAPO-D1* or any other genes known to regulate SNS. The favorable alleles for each of *WZFP-A1*, *WAPO-A1,* and *QSns.sxau-7D* are identified and incorporated to increase up to 3.4 spikelets per spike in the RIL lines. Molecular markers for the alleles were developed. This study has advanced our understanding of the genetic basis of natural variation in spikelet development in wheat.

Keywords: spikelet node number per spike, spike development, *WFZP* gene, *WAPO* gene, wheat

INTRODUCTION

Wheat (*Triticum aestivum* L., AABBDD, $2n = 6x = 42$) is the most widely grown crop in the world, and its grain is a staple food of humans. Wheat grain yield can be divided into three yield components: spikes per unit of area, grain number per spike, and grain weight (Zhang et al., 2018). Grain number per spike can be subdivided into spikelet number per spike and grain number per spikelet, and the number of spikelets per spike is thus an important subcomponent that can be directly used to improve grain yield (Guo et al., 2015).

Common wheat has a normal spike head that looks a square or "*Q*" (Muramatsu, 1963; Faris et al., 2003). Compared the square spike, a spike that is shorter or compact is called club spike

in *T. aestivum* ssp. *Compactum*, whereas a spike that is longer or loose is called spelt spike in *T. aestivum* ssp. *spelta* (Faris et al., 2003; Simons et al., 2006; Johnson et al., 2008; Cui et al., 2012; Faris et al., 2014). The gene controlling the square shape is "*Q*" on chromosomes 5A that was cloned in previous studies (Faris et al., 2003; Simons et al., 2006). The gene controlling the club spike is "*C*," and the "*C*" is not known yet but was mapped on chromosomes 2D (Johnson et al., 2008; Fan et al., 2019).

A normal wheat spike has one single spikelet on one node of the spike rachis (McMaster, 2005). However, a spike may have a supernumerary spikelet on the same node in a vertical or horizontal pattern, forming twin spikelets and even triple spikelets (Wang et al., 2016). The supernumerary spikelets are generated on the basal and central rachis nodes but not on the upper rachis nodes on the same spike (Sears, 1954). A spike of a normal wheat cultivar may generate 15–25 spikelets, each of which develops as a terminal meristem on the spike (Rawson, 1970). The developed spikelet may advance to differentiate into the floral meristem consisting of 5–10 florets encompassed by two small bract leaves to form fertile spikelets containing grains or may abort to form sterile spikelets without grains (Pennell and Halloran, 1984; Gonzalez et al., 2011; Guo and Schnurbusch, 2015; Sakuma et al., 2019).

Numerous efforts have been made to identify quantitative trait loci (QTLs) for spikelets per spike due to their relatively higher heritability than other grain yield components/subcomponents (Zhang et al., 2018; Kuzay et al., 2019). Many QTLs have been identified in association analyses of wheat panels with various genetic backgrounds (Boeven et al., 2016; Würschum et al., 2018; Voss-Fels et al., 2019; Ward et al., 2019) and biparental populations in both hexaploid and tetraploid wheat (Quarrie et al., 2006; Faris et al., 2014; Xu et al., 2014; Luo et al., 2016; Zhai et al., 2016). In two recent independent studies (Kuzay et al., 2019; Voss-Fels et al., 2019), the *TraesCS7A01G481600* gene on the long arm of chromosome 7A, which is orthologous to *ABERRANT PANICLE ORGANIZATION 1* (*APO1*), known to affect panicle development and spikelet number in rice (Ikeda et al., 2013), was reported to be the candidate gene regulating spikelets per spike in wheat. The *TraesCS7A01G481600* locus was found by fine mapping of segregation populations and association analyses of numerous cultivars.

Supernumerary spikelets are genetically controlled by a single recessive gene (Patil, 1958; Martinek and Bednar, 2001; Aliyeva and Aminov, 2011) or two recessive genes (Sharman, 1967; Millet, 1986; Peng et al., 1998; Yang et al., 2005). A major recessive gene on chromosome 2A and numerous minor genes, including one on chromosome 2B, are reported to be involved in the development of supernumerary spikelets (Klindworth et al., 1990; Zhang et al., 2013). Strong inhibitors of supernumerary spikelets may be located on chromosomes 2DS and 2AL in Chinese Spring (Sears, 1954; Zhang et al., 2013). All of these previous studies have pointed to the homoeologous group 2 chromosomes in the control of supernumerary spikelets in wheat. The wheat *FRIZZY PANICLE* (*WFZP*) gene, which encodes an APETALA2/ethylene response transcription factor, may be involved in supernumerary spikelets in common wheat (Dobrovolskaya et al., 2015). Non-sense mutations in *WFZP-A1* on chromosome arm 2AS and *WFZP-D1* on chromosome arm 2DS cause supernumerary spikelets or a so-called multirow spike (Dobrovolskaya et al., 2017). In addition, the *TB1* gene, which encodes a transcription factor in a class II TCP (TEOSINTE BRANCHED1/CYCLOIDEA/PCF1), may control paired spikelet development in a dosage-dependent manner (Dobrovolskaya et al., 2015; Poursarebani et al., 2015; Dixon et al., 2018).

Previous studies identified the functional genes for spikelet development, laying a foundation for allelic variation in the candidate genes for QTLs for the same traits that are mapped in different populations. In this study, we mapped three major QTLs for spikelet nodes per spike (SNS) in a biparental population in common wheat: the first is associated with *WFZP-A1* on chromosome arm 2AS, the second is associated with *WHEAT ORTHOLOG OF APO1* on chromosome arm 7AL (*WAPO-A1*), and the third is in a previously uncharacterized region on chromosome 7D. These findings have facilitated pyramiding the favorable alleles for the genes/QTLs into a single genotype using molecular markers to increase spikelets per spike in breeding new wheat varieties.

MATERIALS AND METHODS

Plant Materials

"SY95-71" (with a pedigree of Fan6//Fan6/Eronga83) was developed by the Triticeae Research Institute of Sichuan Agricultural University in the 1990s, and "CH7034" (with a pedigree of Jing411/Xiaoyan7430//Zhong8601) was developed by the College of Agronomy of Shanxi Agricultural University. The two advanced breeding lines have obvious differences in spike morphology. A population of 184 $F_{2:10}$ recombinant inbred lines (RILs) derived from the cross SY95-71 × CH7034 was developed to map genes for important traits.

Phenotypic Identification

The 184 RILs of the SY95-71 × CH7034 population were tested in six locations/years, including Chengdu (104°13′E, 30°47′N) in 2014 and 2015, Yuncheng (110°53′E, 35°00′N) in 2016, and Linfen (111°31′E, 36°54′N) in 2015 and 2016. The field experiments were arranged in a randomized complete block design with three replications in all of the trials. Each line was grown in a single 2-m row with 0.25 m between rows. Twenty seeds were planted in each row with plant spacing of 0.1 m. Nitrogen and superphosphate fertilizers in the six environmental trials were applied at sowing at rates of 120 and 130 kg ha^{-1}, respectively. Field management was carried out with standard agronomic practices according to wheat production throughout the growing season to obtain even crop stands (Ma et al., 2019). The number of spikelet nodes per spike was characterized in the main spikes of 15 randomly selected plants from each line at physiological maturity and was calculated as the mean. In addition, three plants from each RIL were planted in the greenhouse of Oklahoma State University in 2019 to determine whether the supernumerary spikelet phenotype was associated with the QTLs for the SNS trait. To avoid any confusion, the number of spikelet nodes per spike in a regular pattern is referred

to as SNS, and the total number of spikelets per spike in a pattern with supernumerary spikelets is referred to as SSS.

Development of Single-Nucleotide Polymorphism Markers

The 184 RILs of the SY95-71/CH7034 population and their parents were genotyped using 17K SNP chips with 17,526 markers and 35K DArT SNP chips containing 36,420 markers at Diversity Arrays Technology Pty Ltd. (Canberra, Australia[1]). The single-nucleotide polymorphism (SNP) marker sequences were used to BLAST the Chinese Spring sequence [International Wheat Genome Consortium (IWGSC) RefSeq v1.1] to determine the physical positions of each marker. The SNP markers that had less than 20% missing values and contained different recombinant information were selected in the final mapping.

Linkage Group Construction and Quantitative Trait Loci Analysis

The SNP markers were used to construct linkage groups using JoinMap 4.0 (Van-Ooijen, 2006). The genetic distance between markers was estimated using the Kosambi mapping function, and a QTL for SNSs was claimed when the value of the logarithm of the odds (LOD) threshold for significance was above 2.5 in the interval mapping program in MapQTL 6.0 (Van-Ooijen, 2009). A marker with the maximum LOD value was used to determine the chromosomal position of the QTL peak.

Identification of Allelic Variation in Candidate Genes Associated With Quantitative Trait Loci

Those genes under or close to the peak of the identified QTLs were considered candidates causing the QTLs. Specific primers were designed to isolate and sequence the candidate genes from each of the two parents to determine the allelic variation.

A diagnostic molecular marker for allelic variation in *WFZP-A1* was developed. The primers WFZP-Indel-adF2 (5′-GCCAGCCAACCTCACTTCACTTC-3′) and WFZP-Indel-aR2 (5′-TGGCCGAGGACGCGGCGT-3′) were used to amplify the 449-bp CH7034 allele and the 463-bp SY95-71 allele with the following PCR amplification conditions: 94°C for 3 min, 40 cycles of 94°C for 30 s, 64°C for 30 s, and then 72°C for 40 s, with a final extension of 72°C for 10 min. The PCR products was differentiated on a 2.5% agarose gel at 180 V for 1.5 h.

WAPO-A1 was isolated using the forward primer WAPO-A1F2 (5′-AGGAGATCAATGTTTTTAGGACAATAGGA-3′) and the reverse primer WAPO-A1R6 (5′-GAGCAGGGTGTAGCAACTAGT-3′) with the following PCR amplification conditions: 94°C for 3 min, 40 cycles of 94°C for 30 s, 58°C for 30 s, and then 72°C for 2 min, with a final extension of 72°C for 10 min. A 2-bp deletion was found to distinguish between the two alleles after the PCR products were digested with restriction enzyme HypCH4V. The primers WAPOA1-Indel-F1 (5′-CACAAACATATTGAATGAAAGTTCC-3′) and WAPOA1-Indel-R3 (5′-CATTATATCTTAAACTATATGATGTCC-3′) were used to amplify a fragment containing the 2-bp deletion with the following PCR amplification conditions: 94°C for 3 min, 38 cycles of 94°C for 30 s, 58°C for 30 s, and then 72°C for 30 s, with a final extension of 72°C for 10 min. The digested PCR products were 42 and 70 bp for the CH7034 allele but 110 bp for the SY95-71 allele. The PCR products were separated and scored on an 8% non-denaturing polyacrylamide gel.

WAPO-D1 was isolated using the forward primer WAPO-D1F1 (5′-CAGACGCAGAAAGGAAAACCCTAAGGCAACTT-3′) and the reverse primer WAPO-D1R2 (5′-TGCTAAGAACAGAACAACAAATATACTACCC-3′) with the following PCR amplification conditions: 94°C for 3 min, 40 cycles of 94°C for 30 s, 58°C for 30 s, and then 72°C for 3 min, with a final extension of 72°C for 10 min.

Gene Expression

Quantitative real-time PCR (qRT–PCR) was used to determine the transcriptional levels of *WAPO*-A1 and *WAPO-D1* in SY95-71 and CH7034 cells with SYBR Green PCR Master Mix, and actin was used as an endogenous control. Total RNA was extracted from the leaves using TRIzol® reagent (Invitrogen, Carlsbad, CA, United States). cDNA was synthesized from 1 μg of RNA treated with deoxyribonuclease I using a SuperScript^{TM} II Reverse Transcriptase kit and with an oligo $(dT)_{20}$ primer (Invitrogen). Primers for the gene transcripts were UFO-A-RT-F2 (5′-CTCACTCACTCTCACTCCACG-3′) and UFO-A-RT-R2 (5′-GGTGGTGAGGCAGTAGGTTC-3′) (Kuzay et al., 2019) for *WAPO-A1* and WAPO1-D1-RT-F (5′-CTCACTCTCCCTCCCACCACA-3′) and WAPO1-D1-RT-R (5′-GGTGGTGAGGCAGTAGGTTC-3′) for *WAPO-D1*. qRT–PCR was performed on a CFX96^{TM} Real-time PCR System (Bio–Rad Laboratories, Hercules, CA, United States) using iQTM SYBR® Green Supermix (Bio–Rad Laboratories, Hercules, CA, United States), with actin used as an endogenous control. Gene transcript levels were calculated by the $2^{-\Delta\Delta CT}$ method, where CT is the threshold cycle.

RESULTS

Spikelet Nodes per Spike in Parental Lines and Their Recombinant Inbred Lines

No supernumerary spikelet on a spikelet node was observed in either of the two parental lines SY95-71 and CH7034 when tested in different years/locations (**Figure 1A**). However, supernumerary spikelets were observed in some RILs derived from the two parental lines (**Figure 1B**), indicating that the supernumerary spikelets were generated due to recombination of genes from the two parental lines. The SNS trait in the SY95-71/CH7034 RIL population showed clear segregation, as illustrated in **Figure 1C**.

Statistical analysis showed that among the 184 RILs the SNS association under the field conditions was significant ($T > |t_{0.09}|$, **Supplementary Table 1**), indicating that major gene(s) might be

[1]https://www.diversityarrays.com

FIGURE 1 | Morphological traits of spikes in SY95-71 and CH7034. **(A)** Comparison of spike between SY95-71 and CH7034. The image shows no supernumerary spikelets in either SY95-71 or CH7034. **(B)** Dissection of supernumerary spikelets. RIL spikes were segregated with supernumerary spikelets (the left side) and without supernumerary spikelets (right). Supernumerary spikelets are shown on spikes within a red square (upper) and dissected spikes (lower). **(C)** Spikelet nodes on spikes. RIL were segregated to have less SNS (left) and more SNS (right). Bars, 1 cm.

responsible for SNS segregation in this population, regardless of the difference in changes in climate or environments in the field.

Genetic Mapping of Single-Nucleotide Polymorphism Markers in the SY95-71/CH7034 RIL Population

A total of 17,512 SNP markers was eventually generated from the 17K SNP chips, and a total of 36,420 SNP markers was generated from the 35K DArT SNP chips for the 184 RILs of the SY95-71/CH7034 population. According to their sequences, these SNP markers were assembled into 21 chromosomes, forming genetic maps for the RIL population. The total length of the 21 chromosomes containing the 2,347 SNP markers was 4,192 cM, with a marker density of 1.78 cM per marker. Detailed information on the sequence and genetic distances of the SNP markers on the three SNS QTLs mapped in this study is provided in **Supplementary Table 2**. The SX plus 8-digit SNP codes served as the reference number for each SNP from 17K SNP chips, and the SX-D plus 8-digit SNP codes served as the reference number for each SNP from the 35K DArT chips.

Allelic Variation in *WFZP-A1* Was Associated With a Spikelet Nodes per Spike Quantitative Trait Locus on Chromosome 2A

A total of 1,738 SNP markers, including 797 from 17K SNP chips and 941 from 35K DArT SNP chips, were assembled into chromosome 2A, which was confirmed by the IWGSC RefSeq v1.1 sequences. On the basis of whole-chromosome QTL scanning using interval mapping (IM) analysis, a QTL, *QSns.sxau-2A*, was found on this chromosome (**Figure 2A**). The LOD value at the peak position of *QSns.sxau-2A* ranged from 4.89 to 11.13 in six experimental locations/years, and this QTL alone explained 11.65 or 24.20% of the total phenotypic variation. The 48 markers on chromosome 2A were selected as a linkage group to indicate the genetic distance and physical location of the *QSns.sxau-2A* locus (**Figure 2B**).

The *WFZP-A1* gene at position 66,848 kb was close to marker *SX-D1098510*, which was under the peak of *QSns.sxau-2A* (**Figure 2C**); hence, *WFZP-A1* was considered a candidate gene causing *QSns.sxau-2A*. Sequencing results revealed that the SY95-71 allele showed the same sequence as the *WFZP-A1* gene of Chinese Spring, which is referred to as the *WFZP-A1a* allele, whereas the CH7034 allele had a 14-bp deletion in its coding region, which is referred to as the *WFZP-A1b* allele (**Figure 2D**). The 14-bp insertion in the *WFZP-A1a* allele could cause it to gain its function at the protein level, or the 14-bp deletion in the *WFZP-A1b* allele could cause it to lose its function at the protein level (**Figure 2E**). No association was found between the *WFZP-A1* marker and the SSS trait in this study.

On average, over the six experimental locations/years, the CH7034 *WFZP-A1b* allele (22.6 $\pm$ 0.1, n = 593) set 1.7 more spikelet nodes per spike than the SY95-71 *WFZP-A1a* allele (20.9 $\pm$ 0.1, n = 450) at the *QSns.sxau-2A* locus (p = 8.68E-27) (**Figure 2F**). Based on the 14-bp indel, a diagnostic marker for *WFZP-A1* was developed (**Figure 2G**).

Allelic Variation in *WAPO-A1* Was Associated With a Spikelet Nodes per Spike Quantitative Trait Locus on Chromosome 7A

A QTL analysis revealed that SNS was associated with 65 markers on chromosome 7A (*QSns.sxau-7A*). The 1,730 SNP markers including 844 from 17K SNP chips and 886 from 35K DArT SNP chips were assembled into chromosome 7A, but only 43 markers on chromosome arm 7AL were selected as a linkage group to indicate the genetic distance and physical location of the *QSns.sxau-7A* locus (**Figure 3A**). The LOD values at the peak position of *QSns.sxau-7A* ranged from 3.20 to 11.00 in five experimental locations/years, and this QTL alone explained 9.17–28.18% of the total phenotypic variation.

The *WAPO-A1* gene (*TraesCS7A02G481600*) was considered a candidate gene causing *QSns.sxau-7A*, as it was close to marker *SX-D1102483* that was under the peak of *QSns.sxau-7A* (**Figures 3B,C**). Primers that were reported to specifically amplify *WAPO-A1* (Kuzay et al., 2019) were used to isolate this gene from two parental lines. No difference was found between the CH7034 and *WAPO-A1b* alleles in the sequenced 2,439-bp region, including the 938-bp promoter region, 1,457 bp from the start codon to the stop codon, and 42 bp at the 3′ end of the gene. Compared with CH7034, SY95-71 had a 2-bp deletion at positions –891 to –892 bp, which was only the difference between the two parental lines (**Figure 3D**). Compared with the four reported alleles (*WAPO-A1a* through *-A1d*), the *WAPO-A1* allele in SY95-71 was novel and was thus named *WAPO-A1e* (**Figure 3D**).

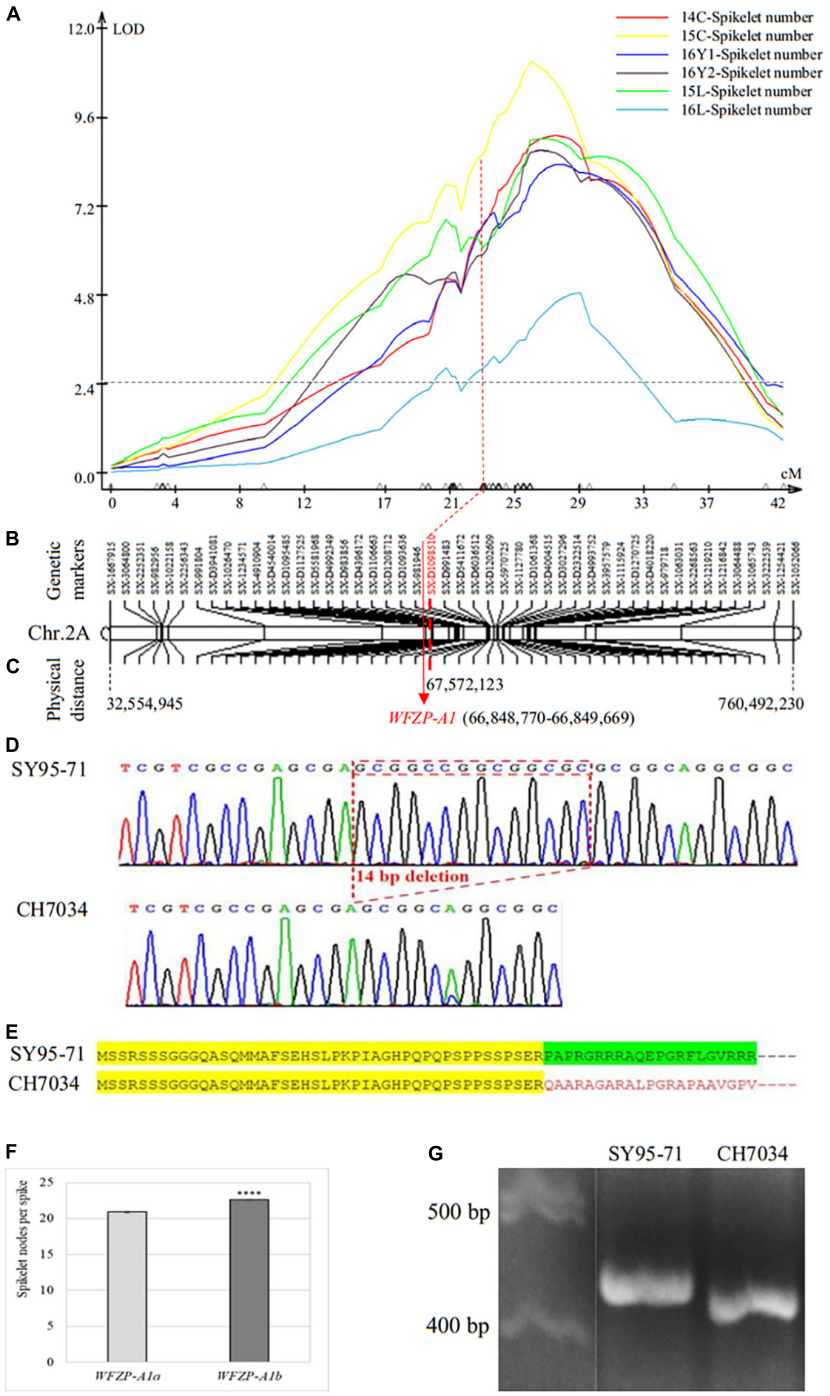


FIGURE 2 | Genetic and physical maps of SNS QTL *QSns.sxau-2A* and allelic variation in the candidate gene *WFZP-A1*. **(A)** Mapping of the *QSns.sxau-2A* locus. The phenotypic data was collected in Chengdu in 2014 (14C) and 2015 (15C), two experimental locations in Yuncheng in 2016 (16Y1 and 16Y2), and Linfen in 2015 (15L) and 2016 (16L). LOD value is indicated on Y axis, and genetic distance (cM) is indicated on X axis. **(B)** Genetic map of *QSns.sxau-2A*. The SX plus 8-digit SNP codes are markers from SNP from 17K SNP chips, and the SX-D plus 8-digit SNP codes are markers from the 35K DArT chips. **(C)** Physical map of *QSns.sxau-2A*. The physical locations of markers and candidate gene *WFZP-A1* are shown based on the Chinese Spring genome sequence of RefSeq v1.1. **(D)** Allelic variation in *WFZP-A1*. The 14-bp sequence that is placed in a red square is present in SY95-71 but absent in CH7034. **(E)** Deduced WFZP-A1 proteins in SY95-71 and CH7034. Conserved amino acids before the deletion in both SY95-71 and CH7034 are highlighted in yellow, functional amino acids after the deletion in SY95-71 are highlighted in green, and altered amino acids after the deletion in CH7034 are indicated in red. Dots indicate more amino acids toward to C terminus. **(F)** Comparison of SNS between the SY95-71 allele and the CH7034 allele. A *t*-test was performed to show significant level in SNS between the two alleles. The bars indicate standard error. "****" indicates significant difference ($p < 0.0001$). **(G)** A diagnostic PCR marker for *WFZP-A1* between the SY95-71 and CH7034. The marker was developed to distinguish between the SY95-71 allele without the 14-bp deletion and the CH7034 allele with the 14-bp deletion.

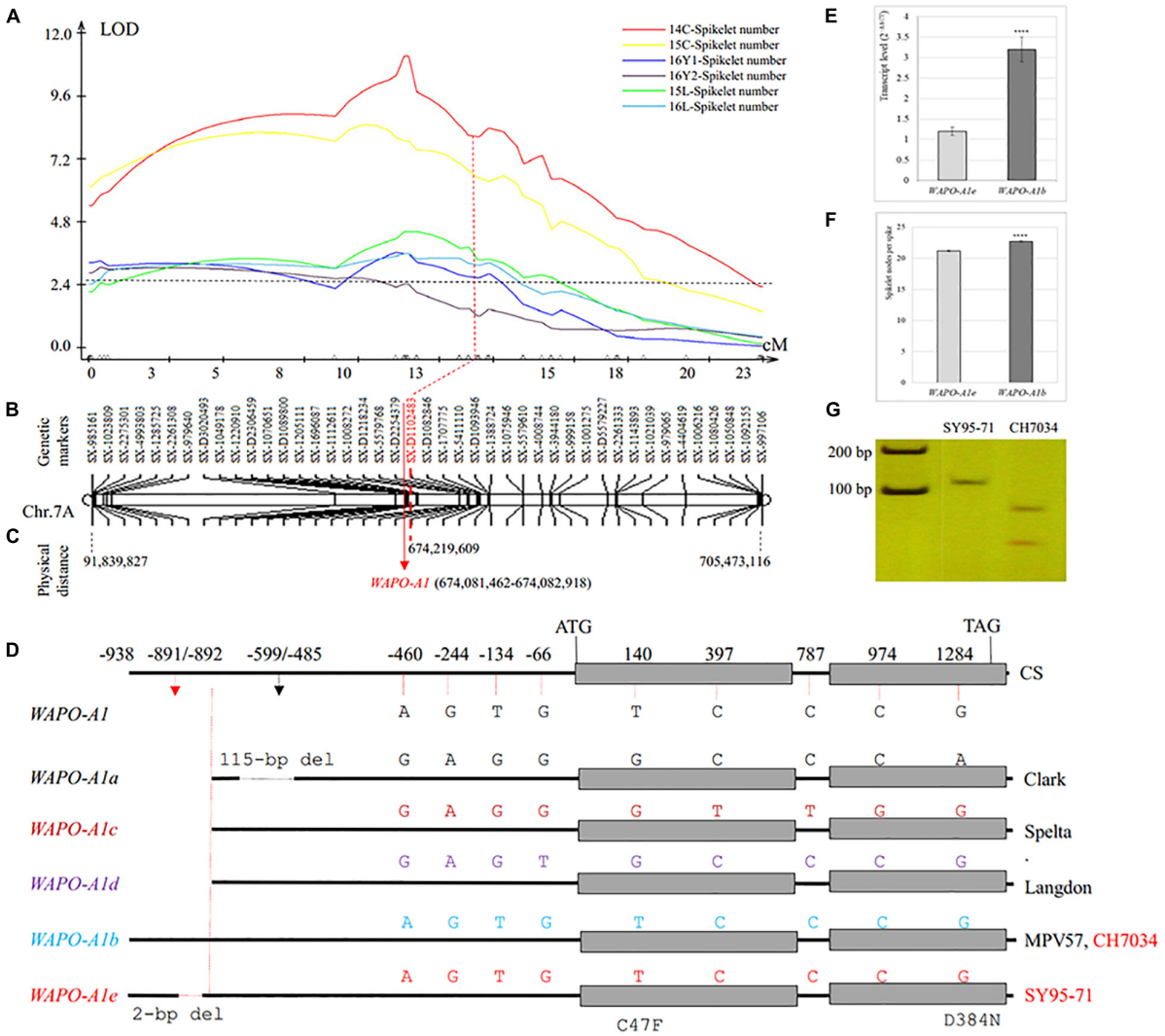


FIGURE 3 | Genetic and physical maps of SNS QTL *QSns.sxau-7A* and allelic variation in the candidate gene *WAPO-A1*. **(A)** Mapping of the *QSns.sxau-7A* locus. The phenotypic data was collected in Chengdu in 2014 (14C) and 2015 (15C), two experimental locations in Yuncheng in 2016 (16Y1 and 16Y2), and Linfen in 2015 (15L) and 2016 (16L). LOD value is indicated on Y axis, and genetic distance (cM) is indicated on X axis. **(B)** Genetic map of *QSns.sxau-7A*. The SX plus 8-digit SNP codes are markers from SNP from 17K SNP chips, and the SX-D plus 8-digit SNP codes are markers from the 35K DArT chips. **(C)** Physical map of *QSns.sxau-7A*. The physical locations of markers and candidate gene *WAPO-A1* are shown based on the Chinese Spring genome sequence of RefSeq v1.1. **(D)** A diagram of differences in multiple *WAPO-A1* alleles. Allelic variation in *WAPO-A1a* through *WAPO-A1d* is cited from the previous study (Kuzay et al., 2019). CH7034 has the same *WAPO-A1b* allele as MPV57. SY95-71 has *WAPO-A1e*, which is a novel allele. The only difference between *WAPO-A1b* and *WAPO-A1e* is that *WAPO-A1e* has a 2-bp deletion in its promoter. **(E)** Comparison of transcript levels between the SY95-71 *WAPO-A1e* allele and the CH7034 *WAPO-A1b* allele. **(F)** Comparison of SNS between the SY95-71 *WAPO-A1e* allele and the CH7034 *WAPO-A1b* allele. **(E–F)** A *t*-test was performed to show differences in the transcript levels between the two alleles, with "*****" indicating a significant difference ($p < 0.0001$). The bars indicate standard error. **(G)** A diagnostic PCR marker for *WAPO-A1* between the SY95-71 and CH7034. The PCR products were digested with restriction enzyme HypCH4V to distinguish between the SY95-71 *WAPO-A1e* allele with the 2-bp deletion and the CH7034 *WAPO-A1b* allele without the 2-bp deletion.

The *WAPO-A1* transcript levels in 1-cm young spikes were compared between the *WAPO-A1b* and *WAPO-A1e* alleles. The CH7034 *WAPO-A1b* allele showed 2.6 times higher transcript levels than the *WAPO-A1e* allele ($p = 0.0001$) (**Figure 3E**). This result suggests that either the *WAPOA1b* promoter might increase its transcript levels due to the absence of the 2-bp deletion or the *WAPO-A1e* promoter reduces its transcript levels due to the presence of the 2-bp deletion.

The CH7034 *WAPO-A1b* allele (22.7 ± 0.1, $n = 554$) showed 1.5 more spikelet nodes per spike than the SY95-71 *WAPO-A1e* allele (21.2 ± 0.1, $n = 435$) at the *QSns.sxau-7A* locus ($p = 1.54E-19$), based on the averaged SNS from the six experimental locations/years (**Figure 3F**). The results showed an association of the SNS phenotype with the 2-bp deletion genotype. A diagnostic marker was developed to distinguish *WAPO-A1e* that had the

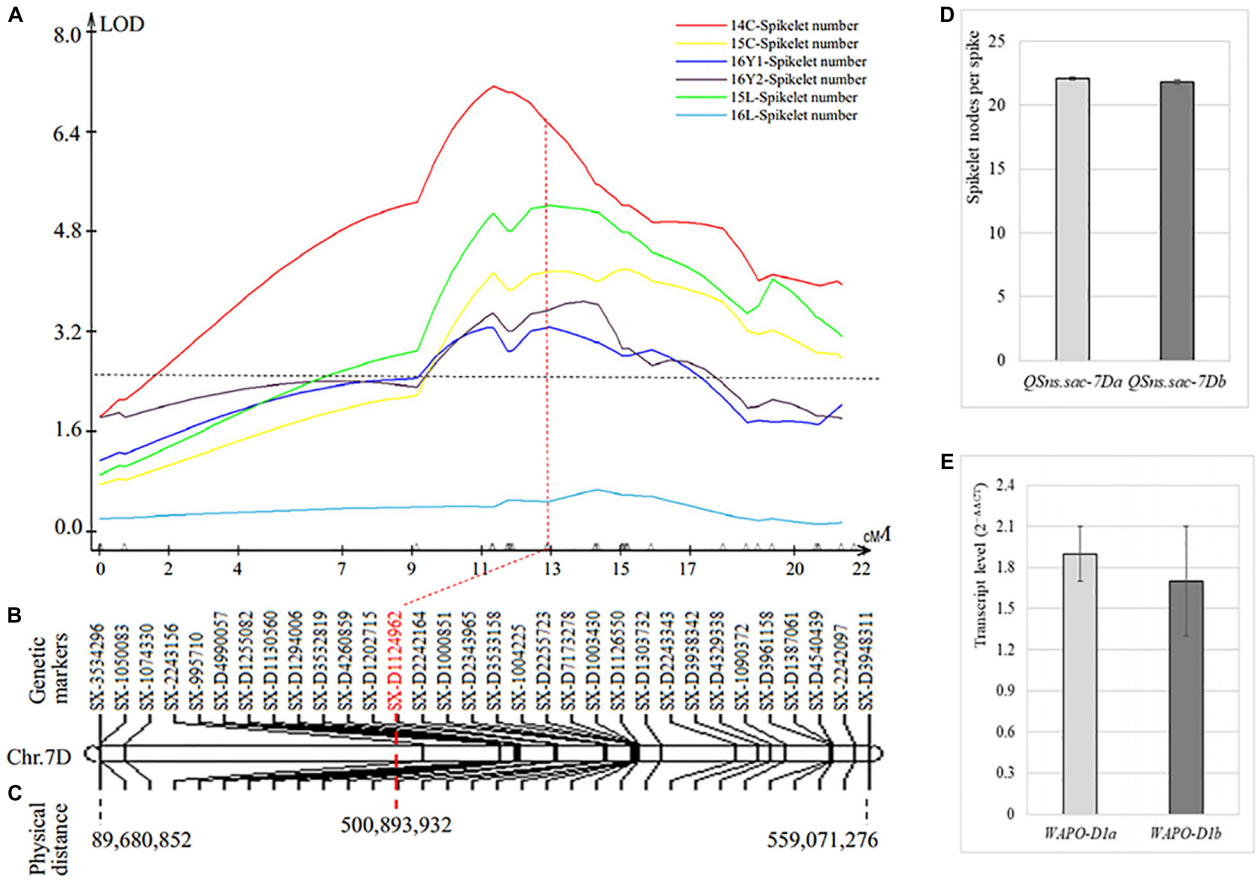


FIGURE 4 | Genetic and physical maps of SNS QTL *QSns.sxau-7D*. **(A)** Mapping of the *QSns.sxau-7D* locus. The phenotypic data was collected in Chengdu in 2014 (14C) and 2015 (15C), two experimental locations in Yuncheng in 2016 (16Y1 and 16Y2), and Linfen in 2015 (15L) and 2016 (16L). LOD value is indicated on Y axis, and genetic distance (cM) is indicated on X axis. **(B)** Genetic map of *QSns.sxau-7D*. The SX plus 8-digit SNP codes are markers from SNP from 17K SNP chips, and the SX-D plus 8-digit SNP codes are markers from the 35K DArT chips. **(C)** Physical map of *QSns.sxau-7D*. The physical locations of markers are shown based on the Chinese Spring genome sequence of RefSeq v1.1. **(D)** Comparison of SNS between the SY95-71 *QSns.sxau-7D* allele and the CH7034 *QSns.sxau-7D* allele. **(E)** Comparison of *WAPO-D1* transcript levels between the SY95-71 and the CH7034, without significant difference.

2-bp deletion from *WAPO-A1b* that did not have the 2-bp deletion (**Figure 3G**).

A New Spikelet Nodes per Spike Quantitative Trait Locus on Chromosome Arm 7DL

The third SNS QTL was located on chromosome 7D (*QSns.sxau-7D*), where only SNP markers, including 8 from 17K SNP chips and 24 from 35K DArT SNP chips, were assembled into the whole chromosome, spanning 22 cM in genetic distance (**Figure 4A**). *QSns.sxau-7D* explained up to 17.92% of the total phenotypic variation in five of the six experimental locations/years, and the effect was not significant in one experimental location/year. The marker *SX-D1124962* closest to the peak of *QSns.sxau-7D* was associated with the SNS trait, the SY95-71 allele (22.1 ± 0.1, $n = 561$) versus the CH7034 allele (21.8 ± 0.1, $n = 536$) ($p = 0.091$) (**Figures 4B–D**).

The complete gene of homoeologous *WAPO-D1* on chromosome 7D (*TraesCS7D02G468700*) was isolated and sequenced, and the sequenced region, including 564 bp upstream from the start codon, 1,488 bp from the start codon to the stop codon, and 196 bp at the 3′ end of the gene, showed no difference between the SY95-71 and CH7034 alleles. Both SY95-71 and CH7034 had the same *WAPO-D1* sequence as Chinese Spring. *WAPO-D1* showed no significant difference in its transcript level between the two alleles (**Figure 4E**). The SNP at *SX-D1124962* may be the closest marker to the gene causing *QSns.sxau-7D*.

Allele Combinations of the Three Spikelet Nodes per Spike Quantitative Trait Loci

Further analyses were performed on each of the three SNS QTLs as independent genetic factors. Eight genotypes generated from these three genes/QTLs, each with two alleles, were analyzed for their SNS phenotypes.

These eight genotypes were classified for the three QTLs in the order of *WFZP-A1*, *WAPO-A1*, and *QSns.sxau-7D*. The SY95-71 allele was designated A, and the CH7034 allele was designated B. The contribution of each genotype to the phenotype

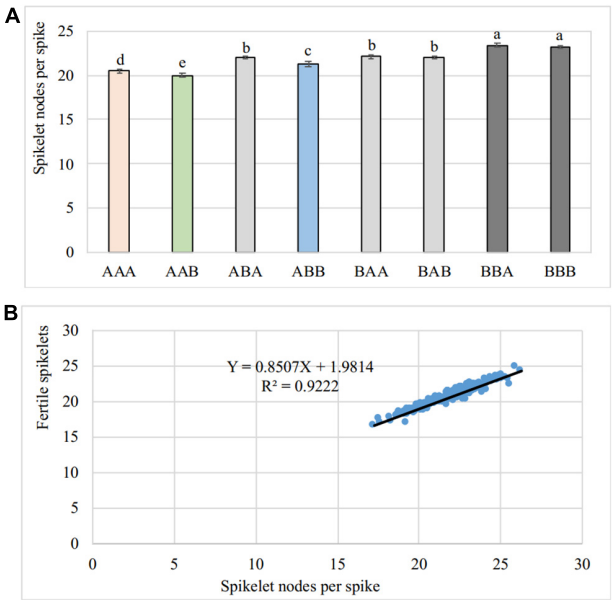


FIGURE 5 | Overall analyses of SNS and fertile spikelets. **(A)** Comparison of SNS among eight genotypes. These eight genotypes are in the order of *WFZP-A1*, *WAPO-A1*, and *QSns.sxau-7D*, with A for the SY95-71 allele and B for the CH7034 allele. The averaged SNS over the six experimental locations were comparisons. A *t*-test was performed to show significant levels in difference of SNS between the two alleles, with *a* though *e* for a significant difference ($P < 0.05$). **(B)** Correlation of SNS and fertile spikelets. The averaged SNS and fertile spikelets over the six experimental locations were analyzed their correlations.

was assessed by comparing the SNS mean, given the lack of epistatic interactions. The genotype expected to generate the highest SNS should be the recombinant (B_B_A), containing all the *WFZP-A1b* and *WAPO-A1b* alleles from CH7034 and *QSns.sxau-7Da* from SY95-71, whereas the genotype expected to generate the lowest SNS should be the recombinant (A_A_B), containing all the *WFZP-A1a* and *WAPO-A1e* alleles from SY95-71 and *QSns.sxau-7Da* from CH7034. The incorporated B_B_A genotype generated 23.4 ± 0.2 (n = 149), which was 3.4 more spikelets than the incorporated A_A_B genotype 20 ± 0.2 (n = 124). The remaining six genotypes showed intermediate SNS scores (**Figure 5A**).

The recently released genome sequences of 10 wheat cultivars provided haplotype information for the frequencies of favorable alleles, *WFZP-A1* and *WAPO-A1*, from CH7034. Only cultivar Norin61 had the same *WFZP-A1b* allele with the 14-bp deletion as CH7034. All other cultivars showed 100% identity to SY95-71 without the 14-bp deletion, except Spelt, which had a SNP in the *WFZP-A* gene region.

CH7034 showed 100% identity in the complete 2,439-bp sequence of the *WAPO-A1b* allele as three wheat cultivars: Mace, Norin61, and Stanley. None of these ten cultivars showed 100% identity to the complete 2,437 bp sequence of the SY95-71 *WAPO1-A1e* allele, but the 2-bp deletion in *WAPO1-A1e* was found to exist in six wheat cultivars: Arina, Jaggar, Julius, Lancer, Landmark, and SY_Mattis. These cultivars were found to have the *WAPO1-A1a* allele, but the sequence of the *WAPO1-A1a* allele did not include the 2-bp deletion region identified in a previous study (**Figure 3D**; Kuzay et al., 2019; Voss-Fels et al., 2019). The last cultivar, Spelt, had no 2-bp deletion, but had a *WAPO1-A1* that was different from *WAPO1-A1a* through *WAPO1-A1e*.

Relationships Between the Total Spikelets, Fertile Spikelets, and Sterile Spikelets

Correlation analysis examined the correlation of the total SNS with fertile spikelets per spike and sterile spikelets per spike over the six experimental locations/years. The total SNS was positively correlated with fertile spikelets per spike (R^2 = 0.922, n = 184) (**Figure 5B**). The results indicated that fertile spikelets were increased through spikelet development in wheat.

DISCUSSION

Allelic variations in the two advancing lines were identified being associated with three QTLs for SNS. Among the three SNS QTLs identified in this study, *QSns.sxau-2A* was most stably and consistently observed under six different experimental locations/years, but *QSns.sxau-7D* was mostly regulated by the environments. The study provides genetic information for grains per spike that can be increased through regular spikelets per spike, in addition to the spikelet fertility rate and grain number per spikelet in wheat (Guo et al., 2015; Zhang et al., 2018).

Previous studies have suggested that mutants in *WFZP* genes are important resources for increasing supernumerary spikelets to increase grains per spike in bread wheat (Dobrovolskaya et al., 2016; Dobrovolskaya et al., 2017; Zhang et al., 2017; Du et al., 2021; Li et al., 2021). While *WFZP-D1* lesions were combined with a frame shift mutation in *WFZP-A1* due to the 14-bp deletion, supernumerary spikelets were detectable. In this study, the CH7034 *WFZP-A1b* allele was not associated with the supernumerary spikelet trait that was observed in the segregation population. No QTL for supernumerary spikelets was observed in the orthologous *WFZP-D* region or any other genomic region. These results indicated that only the CH7034 *WFZP-A1b* allele was unable to generate supernumerary spikelets and that the supernumerary spikelet trait may be regulated by other unknown genetic factors in common wheat. Surprisingly, the CH7034 *WFZP-A1b* allele with the 14-bp deletion was found to positively regulate the SNS number. This observation suggested that wild-type *WFZP-A1a* might be a repressor of spikelet development, although the possibility that mutated *WFZP-A1b* may gain new function due to the presence of deduced new amino acids cannot be excluded. Gene editing technology can be used to edit three homoeologous *WFZP* genes to increase grain number to improve bread wheat yield in future studies. *WAPO1* on wheat chromosome arm 7AL was identified as the best candidate gene regulating spikelet number per spike based on a fine genetic map, a genome-wide association analysis, and the known functions of the orthologous *APO1* gene in rice (Kuzay et al., 2019; Voss-Fels et al., 2019). Since then, more robust evidence

from numerous materials with various genetic backgrounds that were tested under different environments and from mutants and transgenic wheat (Ma et al., 2019; Ding et al., 2021; Jones et al., 2021; Kuzay et al., 2022) has demonstrated that *WAPO-A1* is the causal gene for the previously reported QTLs. *WAPO-A1b* in haplotypes such as Berkut had more than threefold higher transcript levels than *WAPO-A1a* with the 115-bp deletion and three SNPs in the promoter in haplotypes such as RAC875, suggesting that the reduced transcript level of *WAPO-A1a* was due to the 115-bp deletion and three SNPs (Kuzay et al., 2019; Voss-Fels et al., 2019). A previous study also suggested that the reduced SNS by *WAPO-A1a* may be due to the effects of reduced transcript levels or two point mutations in its protein (C47F and D384N) that cannot be separated in the experiment. In this study, *WAPO-A1b* and *WAPO-A1e* encoded the same protein sequence but showed differences in their effects on the SNS trait, excluding the possibility that the SNS trait was regulated at the protein level.

WAPO-A1b and *WAPO-A1e* showed differences at the transcript level, indicating that the SNS trait was regulated at the transcript level. *WAPO-A1e* in SY95-71 has the same DNA sequence as *WAPO-A1b* that does not have the previously reported 115-bp deletion or three SNP regions in CH7034. However, further sequencing showed that *WAPO-A1e* had a 2-bp deletion in upstream of the promoter, facilitating mapping of *WAPO-A1* in this study. It is noteworthy that *WAPO-A1* in another SY95-71 showed the 115-bp deletion (Ma et al., 2019). It is likely that the two SY95-71 breeding lines might have the same source but different genetic backgrounds that were segregated in progeny plants. *WAPO-A1b* in CH7034 showed higher transcript levels than *WAPO-A1e* with the 2-bp deletion in the promoter in SY95-71, suggesting that the transcript level of *WAPO-A1e* was reduced, as observed in *WAPO-A1a*. It is likely that the promoter is involved not only in the previously reported 115-bp deletion and three SNP regions ranging from positions −599 to −485 but also in the 2-bp deletion at positions −891 and −892. Further sequencing results at positions up to −2,565 bp revealed seven SNPs further upstream, including six SNPs at positions 2,210, 2,217, 2,243, 2,255, 2,272, and 2,341 bp between the SY95-71 and CH7034 alleles at *WAPO-A1*. Further work is needed to identify actual binding sites in the *WAPO1* promoter.

The tight association between the two SNS QTLs and two genes, *WFZP-A1* and *WAPO-A1* involved in spikelet development, suggests that each of the spikelet development genes was considered as a candidate gene for one SNS QTL. However, the candidate genes need to be validated in large populations. In the present study the limited size of the RIL population was used for mapping only. Specific RILs can be selected to develop backcross populations, in which only one locus of the three QTLs is heterozygous while the other two are fixed for the same allele; therefore, each of these backcross populations can be used to clone each of the three QTLs in future projects. Diagnostic markers for the *WFZP-A1b* allele, which should be selected, and the *WAPO1e* allele, which should be selected against, will provide molecular tools to support the modification of wheat spikes in wheat breeding.

CONCLUSION

In summary, this study identified allelic variation in three QTLs associated with SNS in the same mapping population, two of which were associated with the genes known to involve in spikelet development in wheat. The findings that *WFZP-A1b* with the 14-bp deletion increased spikelets per spike and the *WFZP-A1e* as a novel allele was regulated at the transcript level advanced understanding of these known genes. The molecular markers for these allelic variations can be used to incorporate and pyramid the favorable alleles into single lines for breeding of new varieties. Improved understanding of the genetic basis of natural variation in the spikelet development will be particularly important to wheat production worldwide.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

XZ and LQ performed most of the phenotyping and molecular experiments. XL, CL, HG, SZ, LC, FC, and JJ undertook part of the field and laboratory work. ZY participated in the construction of the RILs. JZ contributed the RILs' phenotypic data in Linfen. ZC designed and directed the field experiments. LY directed molecular research work. LY, ZC, and LQ wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was financially supported by the Research Program Sponsored by State Key Laboratory of Sustainable Dryland Agriculture (in Preparation) of Shanxi Agricultural University (202002-3), Shanxi Scholarship Council of China (2021-070), the Biological Breeding Engineering of Shanxi Agricultural University (YZGC093), the Service Project from Ministry of Agriculture and Rural Affairs of China (19211089), and the Key Research and Development Program of Shanxi Province (201903D221073).

ACKNOWLEDGMENTS

We thank Ennian Yang in Sichuan Academy of Agricultural Sciences for providing the experimental field in Sichuan.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.892642/full#supplementary-material

REFERENCES

Aliyeva, A. J., and Aminov, N. K. (2011). Inheritance of the branching in hybrid populations among tetraploid wheat species and the new branched spike line 166-schakheli. *Genet. Resour. Crop Evol.* 58, 621–628. doi: 10.1007/s10722-011-9702-9

Boeven, P. H. G., Longin, C. F. H., Leiser, W. L., Kollers, S., Ebmeyer, E., and Würschum, T. (2016). Genetic architecture of male floral traits required for hybrid wheat breeding. *Theor. Appl. Genet.* 129, 2343–2357. doi: 10.1007/s00122-016-2771-6

Cui, F., Ding, A., Li, J., Zhao, C., Wang, L., Wang, X., et al. (2012). QTL detection of seven spike-related traits and their genetic correlations in wheat using two related RIL populations. *Euphytica* 186, 177–192. doi: 10.1007/s10681-011-0550-7

Ding, P. Y., Zhou, J. G., Zhao, C. H., Tang, H. P., Mu, Y., and Tang, L. W. (2021). Dissection of haplotypes, geographical distribution and breeding utilization of WAPO1 associated with spike development in wheat. *Acta. Agron. Sin.* 11, 1–16.

Dixon, L. E., Greenwood, J. R., Bencivenga, S., Zhang, P., Cockram, J., Mellers, G., et al. (2018). TEOSINTE BRANCHED1 regulates inforescence architecture and development in bread wheat (*Triticum aestivum*). *Plant Cell* 30, 563–581. doi: 10.1105/tpc.17.00961

Dobrovolskaya, O. B., Amagai, Y., Popova, K. I., Dresvyannikova, A. E., Martinek, P., Krasnikov, A. A., et al. (2017). Genes wheat frizzy panicle and sham ramification 2 independently regulate differentiation of floral meristems in wheat. *BMC Plant Biol.* 17:252. doi: 10.1186/s12870-017-1191-3

Dobrovolskaya, O. B., Pont, C., Orlov, Y. L., and Salse, J. (2016). Development of new SSR markers for homoeologous WFZP gene loci based on the study of the structure and location of microsatellites in gene rich regions of chromosomes 2AS, 2BS, and 2DS in bread wheat. *Russ. J. Genet.* 19, 330–337. doi: 10.1134/S2079059716030023

Dobrovolskaya, O. B., Pont, C., Sibout, R., Martinek, P., Badaeva, E., Murat, F., et al. (2015). FRIZZY PANICLE drives supernumerary spikelets in bread wheat. *Plant Physiol.* 167, 189–199. doi: 10.1104/pp.114.250043

Du, D. J., Zhang, D. X., Yuan, J., Feng, M., Li, Z. J., Wang, Z. H., et al. (2021). FRIZZY PANICLE defines a regulatory hub for simultaneously controlling spikelet formation and awn elongation in bread wheat. *New Phytol.* 231, 814–833. doi: 10.1111/nph.17388

Fan, X. L., Cui, F., Ji, J., Zhang, W., Zhao, X. Q., Liu, J. J., et al. (2019). Dissection of pleiotropic QTL regions controlling wheat spike characteristics under different nitrogen treatments using traditional and conditional QTL mapping. *Front. Plant Sci.* 10:187. doi: 10.3389/fpls.2019.00187

Faris, J. D., Fellers, J. P., Brooks, S. A., and Gill, B. S. (2003). A bacterial artificial chromosome contig spanning the major domestication locus Q in wheat and identification of a candidate gene. *Genetics* 164, 311–321. doi: 10.1093/genetics/164.1.311

Faris, J. D., Zhang, Q. J., Chao, S. M., Zhang, Z. C., and Xu, S. S. (2014). Analysis of agronomic and domestication traits in a durum × cultivated emmer wheat population using a high-density single nucleotide polymorphism-based linkage map. *Theor. Appl. Genet.* 127, 2333–2348. doi: 10.1007/s00122-014-2380-1

Gonzalez, F. G., Miralles, D. J., and Slafer, G. A. (2011). Wheat floret survival as related to pre-anthesis spike growth. *J. Exp. Bot.* 62, 4889–4901. doi: 10.1093/jxb/err182

Guo, J., Zhang, Y., Shi, W. P., Zhang, B. Q., Zhang, J. J., Xu, Y. H., et al. (2015). Association analysis of grain-setting rate at the apical and basal spikelets in bread wheat (*Triticum aestivum* L.). *Front. Plant Sci.* 6:1029. doi: 10.3389/fpls.2015.01029

Guo, Z. F., and Schnurbusch, T. (2015). Variation of floret fertility in hexaploid wheat revealed by tiller removal. *J. Exp. Bot.* 19, 5945–5958. doi: 10.1093/jxb/erv303

Ikeda, M., Miura, K., Aya, K., Kitano, H., and Matsuoka, M. (2013). Genes offering the potential for designing yield-related traits in rice. *Curr. Opin. Plant Biol.* 16, 213–220. doi: 10.1016/j.pbi.2013.02.002

Johnson, E. B., Nalam, V. J., Zemetra, R. S., and Riera-Lizarazu, O. (2008). Mapping the compactum locus in wheat (*Triticum aestivum* L.) and its relationship to other spike morphology genes of the triticeae. *Euphytica* 163, 193–201. doi: 10.1007/s10681-007-9628-7

Jones, B. H., Blake, N. K., Heo, H. Y., Martin, J. M., Torrion, J. A., and Talbert, L. E. (2021). Allelic response of yield component traits to resource availability in spring wheat. *Theor. Appl. Genet.* 134, 603–620. doi: 10.1007/s00122-020-03717-7

Klindworth, D. L., Williams, N. D., and Joppa, L. R. (1990). Chromosomal location of genes for supernumerary spikelet in tetraploid wheat. *Genome* 33, 515–520. doi: 10.1139/g90-076

Kuzay, S., Lin, H., Li, C., Chen, S., Woods, D. P., Zhang, J., et al. (2022). WAPO-A1 is the causal gene of the 7AL QTL for spikelet number per spike in wheat. *PLoS Genet.* 18:e1009747. doi: 10.1371/journal.pgen.1009747

Kuzay, S., Xu, Y. F., Zhang, J. L., Katz, A., Pearce, S., Su, Z. Q., et al. (2019). Identification of a candidate gene for a QTL for spikelet number per spike on wheat chromosome arm 7AL by high-resolution genetic mapping. *Theor. Appl. Genet.* 132, 2689–2705. doi: 10.1007/s00122-019-03382-5

Li, Y. P., Li, L., Zhao, M. C., Guo, L., Guo, X. X., Zhao, D., et al. (2021). Wheat frizzy panicle activates vernalization1-A and HOMEOBOX4-A to regulate spike development in wheat. *Plant Biotechnol. J.* 19, 1141–1154. doi: 10.1111/pbi.13535

Luo, W., Ma, J., Zhou, X. H., Sun, M., Kong, X. C., Wei, Y. M., et al. (2016). Identification of quantitative trait loci controlling agronomic traits indicates breeding potential of tibetan semiwild wheat (*Triticum aestivum* ssp. tibetanum). *Crop Sci.* 56, 2410–2420. doi: 10.2135/cropsci2015.11.0700

Ma, J., Ding, P. Y., Liu, J. J., Li, T., Zou, Y. Y., Habib, A., et al. (2019). Identification and validation of a major and stably expressed QTL for spikelet number per spike in bread wheat. *Theor. Appl. Genet.* 132, 3155–3167. doi: 10.1007/s00122-019-03415-z

Martinek, P., and Bednar, J. (2001). "Changes of spike morphology (multirow spike-MRS, long glumes-LG) in wheat (*Triticum aestivum*L.) and their importance for breeding," in *Proceedings of the International Conference "Genetic Collections, Isogenic and Alloplasmic Lines"*, Novosibirsk, 192–194.

McMaster, G. S. (2005). Phytomers, phyllochrons, phenology and temperate cereal development. *J. Agr. Sci.* 143, 137–150. doi: 10.1017/S0021859605005083

Millet, E. (1986). Genetic control of heading date and spikelet number in common wheat (*Triticum aestivum* L.) line 'Noa'. *Theor. Appl. Genet.* 72, 105–107. doi: 10.1007/BF00261463

Muramatsu, M. (1963). Dosage effect of the spelta gene q of hexaploid wheat. *Genetics* 48:469–482. doi: 10.1093/genetics/48.4.469

Patil, J. A. (1958). Inheritance study in wheat. *Curr. Sci.* 27, 404–405.

Peng, Z. S., Liu, D. C., Yen, C., and Yang, J. L. (1998). Genetic control of supernumerary spikelet in common wheat line LYB. *Wheat Inform. Ser.* 86, 6–12.

Pennell, A. L., and Halloran, G. M. (1984). Influence of time of sowing, photoperiod and temperature on supernumerary spikelet expression in wheat (*Triticum*). *Can. J. Bot.* 62, 1687–1692. doi: 10.1139/b84-228

Poursarebani, N., Seidensticker, T., Koppolu, R., Trautewig, C., Gawronski, P., Bini, F., et al. (2015). The genetic basis of composite spike form in barley and 'miracle-wheat'. *Genetics* 201, 155–165. doi: 10.1534/genetics.115.176628

Quarrie, S. A., Quarrie, S. P., 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, 2627–2637. doi: 10.1007/s00284-007-9035-2

Rawson, H. M. (1970). Spikelet number, its control and relation to yield per ear in wheat. *Aust. J. Biol. Sci.* 23, 1–15. doi: 10.1071/bi9700001

Sakuma, S., Golan, G., Guo, Z. F., Ogawa, T., Tagiri, A., Sugimoto, K., et al. (2019). Unleashing floret fertility in wheat through the mutation of a homeobox gene. *Proc. Natl. Acad. Sci. U.S.A.* 116, 5182–5187. doi: 10.1073/pnas.1815465116

Sears, E. R. (1954). *The Aneuploids of Common Wheat*. Columbia, IN: University of Missouri, 7–11.

Sharman, B. C. (1967). Interpretation of the morphology of various naturally occurring abnormalities of the inflorescence ofwheat (*Triticum*). *Can. J. Bot.* 45, 2073–2080. doi: 10.1139/b67-224

Simons, K. J., Fellers, J. P., Trick, H. N., Zhang, Z., Tai, Y. S., Gill, B. S., et al. (2006). Molecular characterization of the major wheat domestication gene Q. *Genetics* 172, 547–555. doi: 10.1534/genetics.105.044727

Van-Ooijen, J. W. (2006). *JoinMap® 4.0: Software for the Calculation of Genetic Linkage Maps In Experimental Populations*. Wageningen: Kyazma BV.

Van-Ooijen, J. W. (2009). *MapQTL® 6.0: Software for the Mapping Fo Quantitative Trait Loci In Experimental Populations of Diploid Species*. Wageningen: Kyazma BV.

Voss-Fels, K. P., Keeble-Gagnère, G., Hickey, L. T., Tibbits, J., Nagornyy, S., Hayden, M. J., et al. (2019). High-resolution mapping of rachis nodes per rachis, a critical determinant of grain yield components in wheat. *Theor. Appl. Genet.* 132, 2707–2719. doi: 10.1007/s00122-019-03383-4

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

Ward, B. P., Brown-Guedira, G., Kolb, F. L., Van Sanford, D. A., Tyagi, P., Sneller, C. H., et al. (2019). Genome-wide association studies for yield-related traits in soft red winter wheat grown in Virginia. *PLoS One* 14:e0208217. doi: 10.1371/journal.pone.0208217

Würschum, T., Leiser, W. L., Langer, S. M., Tucker, M. R., and Longin, C. F. H. (2018). Phenotypic and genetic analysis of spike and kernel characteristics in wheat reveals long-term genetic trends of grain yield components. *Theor. Appl. Genet.* 131, 2071–2084. doi: 10.1007/s00122-018-3133-3

Xu, Y. F., Wang, R. F., Tong, Y. P., Zhao, H. T., Xie, Q. G., Liu, D. C., et al. (2014). Mapping QTLs for yield and nitrogen-related traits in wheat: influence of nitrogen and phosphorus fertilization on QTL expression. *Theor. Appl. Genet.* 127, 59–72. doi: 10.1007/s00122-013-2201-y

Yang, W. Y., Lu, B. R., Hu, X. R., Yu, Y., and Zhang, Y. (2005). Inheritance of the triple-spikelet character in a Tibetan landrace of common. *Genet. Resour. Crop Evol.* 52, 847–851. doi: 10.1007/s10722-003-6089-2

Zhai, H. J., Feng, Z. Y., Li, J., Liu, X. Y., Xiao, S. H., Ni, Z. F., et al. (2016). QTL analysis of spike morphological traits and plant height in winter wheat (*Triticum aestivum* L.) using a high-density SNP and SSR-based linkage map. *Front. Plant Sci.* 7:1617. doi: 10.3389/fpls.2016.01617

Zhang, J. L., Gizaw, S. A., Bossolini, E., Hegarty, J., Howell, T., Carter, A. H., et al. (2018). Identification and validation of QTL for grain yield and plant water status under contrasting water treatments in fall-sown spring wheats. *Theor. Appl. Genet.* 131, 1741–1759. doi: 10.1007/s00122-018-3111-9

Zhang, R. Q., Hou, F., Chen, J., Chen, S. L., Xing, L. P., Feng, Y. G., et al. (2017). Agronomic characterization and genetic analysis of the supernumerary spikelet in tetraploid wheat (*Triticum turgidum* L.). *J. Integ. Agric.* 16, 1304–1311. doi: 10.1016/S2095-3119(16)61469-7

Zhang, R. Q., Wang, X. E., and Chen, P. D. (2013). Inheritance and mapping of gene controlling four-rowed spike in tetraploid wheat (*Triticum turgidum* L.). *Acta. Agron. Sin.* 39, 29–33. doi: 10.3724/SP.J.1006.2013.00029

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 declared a past co-authorship with one of the authors CL.

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

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 27 May 2022
doi: 10.3389/fpls.2022.896408

Prediction of Chlorophyll Content in Multi-Temporal Winter Wheat Based on Multispectral and Machine Learning

Wei Wang[1,2], *Yukun Cheng*[1], *Yi Ren*[1], *Zhihui Zhang*[1] *and Hongwei Geng*[1]*

[1] *High-Quality Special Wheat Crop Engineering Technology Research Center, College of Agronomy, Xinjiang Agricultural University, Ürümqi, China,* [2] *Department of Computer Science and Information Engineering, Anyang Institute of Technology, Anyang, China*

To obtain the canopy chlorophyll content of winter wheat in a rapid and non-destructive high-throughput manner, the study was conducted on winter wheat in Xinjiang Manas Experimental Base in 2021, and the multispectral images of two water treatments' normal irrigation (NI) and drought stress (DS) in three key fertility stages (heading, flowering, and filling) of winter wheat were obtained by DJI P4M unmanned aerial vehicle (UAV). The flag leaf chlorophyll content (CC) data of different genotypes in the field were obtained by SPAD-502 Plus chlorophyll meter. Firstly, the CC distribution of different genotypes was studied, then, 13 vegetation indices, combined with the Random Forest algorithm and correlation evaluation of CC, and 14 vegetation indices were used for vegetation index preference. Finally, preferential vegetation indices and nine machine learning algorithms, Ridge regression with cross-validation (RidgeCV), Ridge, Adaboost Regression, Bagging_Regressor, K_Neighbor, Gradient_Boosting_Regressor, Random Forest, Support Vector Machine (SVM), and Least absolute shrinkage and selection operator (Lasso), were preferentially selected to construct the CC estimation models under two water treatments at three different fertility stages, which were evaluated by correlation coefficient (r), root means square error (RMSE) and the normalized root mean square error (NRMSE) to select the optimal estimation model. The results showed that the CC values under normal irrigation were higher than those underwater limitation treatment at different fertility stages; several vegetation indices and CC values showed a highly significant correlation, with the highest correlation reaching.51; in the prediction model construction of CC values, different models under normal irrigation and water limitation treatment had high estimation accuracy, among which the model with the highest prediction accuracy under normal irrigation was at the heading stage. The highest precision of the model prediction under normal irrigation was in the RidgeCV model (r = 0.63, RMSE = 3.28, NRMSE = 16.2%) and the highest precision of the model prediction under water limitation treatment was in the SVM model (r = 0.63, RMSE = 3.47, NRMSE = 19.2%).

Keywords: winter wheat, UAV, multispectral, machine learning, vegetation index

OPEN ACCESS

Edited by:
Fa Cui,
Ludong University, China

Reviewed by:
Guangchen Liu,
Ludong University, China
Yonggui Xiao,
Chinese Academy of Agricultural Sciences (CAAS), China

***Correspondence:**
Hongwei Geng
hw-geng@163.com

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

Received: *15 March 2022*
Accepted: *19 April 2022*
Published: *27 May 2022*

Citation:
Wang W, Cheng Y, Ren Y, Zhang Z and Geng H (2022) Prediction of Chlorophyll Content in Multi-Temporal Winter Wheat Based on Multispectral and Machine Learning. Front. Plant Sci. 13:896408. doi: 10.3389/fpls.2022.896408

INTRODUCTION

Soil plant analysis development (SPAD) can directly reflect the relative chlorophyll content in leaves (Netto et al., 2005), which was the most important pigment in photosynthesis, and its content was intimately related to the photosynthesis of plants (Zhang et al., 2019a) and the changes in its concentration directly affected the health of crops (Gitelson, 2005; Shestakova et al., 2020). Winter wheat was one of the world's major food crops. High yield and quality were the goals pursued by many breeders (Lesk et al., 2016; Sun et al., 2019). In recent years, extreme weather occurs frequently in the world and drought had directly affected wheat yields. It was estimated that drought and hot weather worldwide would reduce annual yields by 9–10% (Mondal et al., 2013). Crop growth can be predicted by constructing characteristics of nutrients and canopy spectra. Chlorophyll content played a large role in guiding the drought resistance and yield of wheat. Therefore, the study of chlorophyll content in wheat provided a basis for judging the growth of crops. Currently, remote sensing technology provided new ideas for the estimation of the chlorophyll content of crops, and the research mainly focused on spectral information indices and spectral information obtained by different sensors combined with data from the ground to predict the chlorophyll content.

Remote-sensing technology was currently showing strong competitiveness for precision agriculture in different experimental environments, especially that the convenient application of multispectral imaging technology on UAVs has accelerated the development of the technology (Kaivosoja et al., 2013; Yang et al., 2019). The acquisition of characteristic data on chlorophyll content at the ground level was usually destructive (Telmo, 2017). In addition, the ground acquired data was selected from a few limited points, and it was difficult to use these points to represent the characteristics of the whole area, so the acquisition of traditional ground phenology data was limited in scope. Remote-sensing data can be acquired at high throughput and large scale, but the influence of spatial image resolution made it difficult to grasp some local features, so UAV-based remote-sensing technology made up for this deficiency.

In recent years, with the compactness and convenience of UAVs and the ability to customize their missions, they had played an extremely important role in information-based agriculture (Sampson et al., 2003; Sun et al., 2021). The UAV was used as a spaceflight vehicle to carry various sensors, such as hyperspectral sensors, multispectral sensors, RGB cameras, and thermal infrared sensors (Zhang et al., 2021). Especially, hyperspectral and multispectral sensors were more common for the prediction of nutrient elements in crops. Currently, hyperspectral characterization data had been used to some extent for agricultural traits, but the popularity of hyperspectral use was far from multispectral due to some economic and complex reasons (Taghvaeian et al., 2012). Multispectral sensors carrying different wavelengths (Blue, Green, Red, Red_edge, and Nir) had a wide range of applications in many areas of crops. For instance, Bendig et al. (2015). predicted the biomass of crops by drones and obtained better results. The LAI, planting density, and photosynthetic characteristics of canola, barley, and wheat were well predicted by the UAV. In addition, the multispectral images from the UAV were a good reference for determining the emergence rate and rising potential of spring wheat. Moreover, there are also multispectral images from UAVs that serve as a good reference for the determination of seedling emergence, as well as the rise of spring wheat. Recently, some scholars have judged the maturity of wheat, as well as sorghum under drought conditions by UAV-based multispectral indices (Guillen-Climent et al., 2012; Verger et al., 2014; Jin et al., 2017). Hunt et al. (2013) constructed the Green Normalized Difference Vegetation Index (GNDVI) from multispectral images obtained by UAV and inversed the leaf area index of wheat through the vegetation index.

Combining ground phenotype and UAV multispectral image data for chlorophyll content inversion of crops was an innovative application of UAV multispectral sensors. Machine learning was the science of how to use computers to simulate or implement human learning activities, and it was the most intelligent feature of artificial intelligence, which can be used to integrate the data that had been generated for learning, and then, go on to make scientific predictions for the unknown world. The regression models within machine learning had shown strong data prediction capabilities for both linear and nonlinear, structured and unstructured data. For example, the least-squares algorithm, Random Forest algorithm, Support Vector Machine algorithm, decision tree algorithm, and Naïve Bayesian algorithm have been used to varying degrees in agricultural remote sensing(Garg et al., 2016; Grinberg et al., 2020; Shafiee et al., 2021).

Machine learning can not only perform predictive analysis on traditional data but also embodied great advantages in handling noise and anomalies data (Witten et al., 2016). For chlorophyll content prediction, multispectral images from UAV remote-sensing combined with machine learning algorithms provided excellent thoughts. Currently, machine learning combined with different vegetation indices had shown powerful advantages in agricultural remote-sensing, but studies combining a large number of machine learning algorithm models with preferred vegetation indices had rarely been reported. Next, classic machine learning algorithms for regression were employed to analyze the test data to find the best algorithm. Therefore, this article combined high-throughput UAV remote sensing images with preferential vegetation index and CC data to predict different fertility stages of winter wheat under different water treatments to achieve an intelligent level of wheat detection. This article focused on the following issues:(1) how to prefer vegetation indices for inverse model construction; (2) the effect of chlorophyll content on water response under different irrigation conditions; and(3) the response of different machine learning algorithms for different water treatments at different fertility stages to the prediction model of chlorophyll content in winter wheat.

MATERIALS AND METHODS

Study Area and Experimental Design

The winter wheat field experiment was located in Manas, Xinjiang, China(86°12'52.2"N, 44°18'15.77"E), which had a mid-temperate continental arid semi-arid climate with severely

TABLE 1 | Unmanned aerial vehicle (UAV) flight parameters.

Parameters	Parameter values
Flight altitude	12 m
Flight Speed	5.4 km/h
Course overlap ratio	75%
Lateral overlap rate	75%
Spectral type	Blue, Green, Red, Red_edge, Nir

TABLE 2 | Manas UAV multispectral images and chlorophyll content (CC) data acquisition program.

Date	Stages	Data
2021.5.8	Heading	CC and Multispectral Images
2021.5.21	Flowering	CC and Multispectral Images
2021.5.28	Filling	CC and Multispectral Images

cold winters, hot summers, dryness and low rainfall, sufficient sunshine, high evaporation, and low precipitation. In this study, there were 2 irrigation water treatments: DI and DS. A total of186 plots were selected for each irrigation treatment, 62 wheat varieties were selected for the experiment, each plot contained 1 variety, randomized group design, 3 repeats, 1-row zone(each of size 1.5 m × 0.3 m), and water restriction treatment was not watering during the wheat heading, flowering, filling, and maturity stages. The field management was by the local conventional cultivation management mode, and the wheat field grows well. Fertilizer application, drip irrigation, insect control, and weed control were also the same as the local field management.

UAV Platform and Flight Configuration

In this study, the UAV platform was DJI Phantom 4 (Shenzhen Dajiang Technology Co. Ltd., Shenzhen, China), which carried a multispectral camera to collect multispectral images of the winter wheat canopy. The multispectral camera carried a camera with one RGB and multispectral channels with five wavelengths, centered at 450 nm (Blue), 560 nm (Green), 650 nm (Red), 730 nm (Red_edge), and 840 nm (Nir). Besides, the UAV is equipped with a light intensity sensor and a gray plate for radiation correction.

In the process of multispectral image acquisition, clear and windless weather at noon was selected, the UAV flew autonomously and recorded images according to the set route, and the multispectral camera lens was vertically downward, and the flight parameters were shown in **Table 1**.

Data Collection Plan

The data collection included CC of winter wheat canopy and UAV multispectral images. These data were collected at three different stages of wheat fertility: heading, flowering, and filling. Also, all the data in the acquisition plan contain two different water treatments. The specific acquisition plan for CC and UAV multispectral images were shown in **Table 2**.

Measurement of Chlorophyll Content

The correlation coefficients between SPAD values and chlorophyll content of wheat leaves had a significant level and can reflect the high and low levels of chlorophyll content of the crop (Netto et al., 2005). The measurement periods were three different fertility stages of wheat: heading stage, flowering stage, and filling stage. The relative chlorophyll content of different genotypes was measured simultaneously on the same day of the UAV flight by taking five uniformly growing wheat plants from each variety and using the SPAD-502 Plus chlorophyll meter, an instrument manufactured by Minolta Camera, Japan, which had been used by many scholars to obtain wheat CC data (Wang et al., 2012; Zhang et al., 2019b). The CC values were measured and recorded at the top, middle, and bottom of the flag leaf of selected winter wheat plants in the experiment field, and the average of the chlorophyll content of the three parts was taken as the CC values of the plant winter wheat, and then, the average of the CC values of five winter wheat plants was calculated as the CC values of this variety of winter wheat.

Image Processing

In this study, Pix4Dmapper software (Version 1.4, Pix4d, Lausanne, Switzerland) (https://pix4d.com/) was used to stitch the acquired multispectral images of the UAV in 5 bands. The multispectral images were firstly corrected with the corresponding ground control point data to generate Digital Orthophoto Map (DOM); then, the reflectance correction of the multispectral images was performed with the gray plate to obtain the test site reflectance images, which were stored in TIF format; finally, ARCGIS software (Version 10.3.1, Esri, USA) (http://www.esri.com/arcgis/about-arcgis) was used to extract the vector surface of the cell and obtain the spectral reflectance images in 5 bands, and the average reflectance of this study area was extracted as the spectral reflectance of the sample in this band. The specific image processing flow was shown in **Figure 1**.

Selection of Vegetation Index (VI)

The combination of changes in reflectance of different bands constitutes vegetation indices, which can reduce the degree of influence of factors, such as background soil on vegetation spectra, to a certain extent and improve the accuracy of estimating chlorophyll content. In this article, the importance of vegetation indices was evaluated by 13 vegetation indices in **Table 3** except Green Band Optimized Soil Conditioning Vegetation Index (GOSAVI) using the Random Forest algorithm (Breiman, 2001). First, finding the top five vegetation indices in terms of contribution under different treatments, then, the vegetation indices were preferentially selected by combining the correlation between 14 vegetation indices and CC values, and, finally, the model inversion and prediction of CC values are carried out by using the preferential vegetation indices. In total, we selected 14 vegetation indices. The vegetation index calculation formula is shown in **Table 3**.

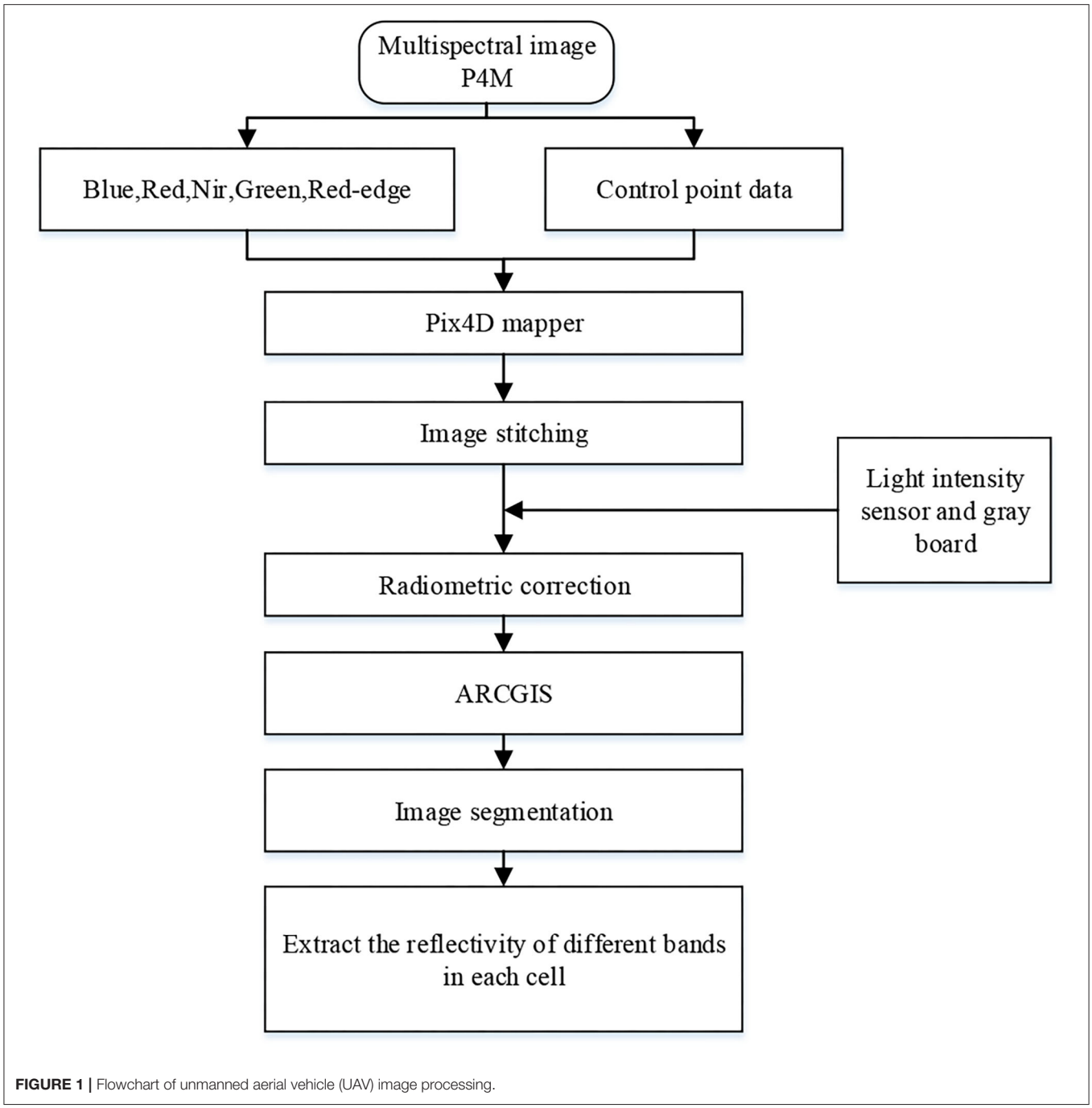


FIGURE 1 | Flowchart of unmanned aerial vehicle (UAV) image processing.

Modeling Methods

Using remote-sensing images that predicted the chlorophyll content of ground crops by modeling, the analysis from mathematical models was the process of predicting chlorophyll by observing specific variables. Since the 1980s, machine learning has attracted wide interest in the artificial intelligence community as a way to achieve artificial intelligence, especially in the last decade or so, a rapid development of research work in the field of machine learning, and it has become one of the important topics of artificial intelligence. The algorithms of traditional machine learning had all shown strong advantages in data prediction regression. To predict the chlorophyll content of winter wheat, this study investigated the most classical machine learning algorithms of RidgeCV (Pelckmans et al., 2005), Ridge (Houwelingen, 1992), Adaboost Regression (Freund and Schapire, 1997), Bagging_Regressor (Hall and Turlach, 2007), K_Neighbor (Cover and Hart, 1967), Gradient_Boosting_Regressor (Friedman, 2001), Random Forest (Breiman, 2001), and SVM, Lasso (Tibshirani, 2011), and then, model inversion was performed with the studied data, and

TABLE 3 | Vegetation index and its calculation formula.

Vegetation index	Formula to calculate	Reference
Normalized vegetation index (NDVI)	$NDVI = (R_{Nir} - R_{Red})/(R_{Nir} + R_{Red})$	Schnell, 1974
Green Normalized Vegetation Index (GNDVI)	$GNDVI = (R_{Nir} - R_{Green})/(R_{Nir} + R_{Green})$	Wagner, 1996
Normalized Green and Blue Difference Index (NGBDI)	$NGBDI = (R_{Green} - R_{Blue})/(R_{Green} + R_{Blue})$	Hunt et al., 2005
Green Band Optimized Soil Conditioning Vegetation Index (GOSAVI)	$GOSAVI = 1.16 * [(R_{Nir} - R_{Green})/(R_{Nir} + R_{Green} + 0.16)]$	Gilabert et al., 2002
Red edge optimized soil Regulating Vegetation Index (REOSAVI)	$REOSAVI = 1.16 * [(R_{Nir} - R_{Red})/(R_{Nir} + R_{Red} + 0.16)]$	Kim et al., 1994
Optimization of Soil Conditioning Vegetation Index (OSAVI)	$OSAVI = (R_{Nir} - R_{Red})/(R_{Nir} + R_{Red} + 0.16)$	Rondeaux et al., 1996
Over Green Index (EXG)	$EXG = 2R_{Green} - R_{Red} - R_{Blue}$	Torres-Sánchez et al., 2014
Green band atmospheric impedance vegetation index (VARIgreen)	$VARIgreen = (R_{Green} - R_{Red})/(R_{Green} + R_{Red} - R_{Blue})$	Gitelson et al., 2002
Red-band atmospheric impedance vegetation index (VARIred)	$VARIred = (R_{Red\_edge} - 1.7 * R_{Red} + 0.7 * R_{Blue})/(R_{Red\_edge} + 2.3 * R_{Red} - 1.3 * R_{Blue})$	Gitelson et al., 2002
Modified simple ratio(MSR)	$MSR = \frac{R_{Nir}/R_{Red}-1}{\sqrt{R_{Nir}/R_{Red}}+1}$	Chen, 1996
Simple Ratio(SR)	$SR = R_{Nir}/R_{Red}$	Wagner, 1996
Green Chlorophyll Index(GCI)	$GCI = R_{Nir}/R_{Green} - 1$	Gitelson, 2005
Normalized Difference Red-edge Index(NDREI)	$NDREI = (R_{Red\_edge} - R_{Green})/(R_{Red\_edge} + R_{Green})$	Muhammad et al., 2018
Normalized Green-Red Variance Index(NDRGI)	$NDRGI = (R_{Green} - R_{Red})/(R_{Green} + R_{Red})$	Schnell, 1974

R_{Blue}, R_{Green}, R_{Red}, R_{Red_edge} and R_{Nir} represent the reflectance of Blue wave band, Green band, Red band, Red_edge band, and Nir band, respectively.

through an experimental cross-reference of the data, an attempt was made to find out the optimal regression learning algorithm that was most suitable for this study and provided model support for subsequent data prediction. The specific flow chart of the program implementation was shown in **Figure 2**.

To discuss the prediction model of CC in winter wheat at different fertility stages with different water treatments, data sets of different genotypes were used to construct the crop chlorophyll inversion model. In the process of model construction, the datasets were partitioned separately, and the datasets were distributed in a 7:3 ratio according to the random selection method of validation set and testers.

Accuracy Evaluation

Pearson correlation coefficient (r), root mean square error (RMSE), and the normalized root mean square error (NRMSE) were used as evaluation indexes for different models, where the closer r was to 1, the lower the RMSE value indicates that the predicted and measured values of the model agreed. Also, the smaller the NRMSE value, the higher the accuracy of its estimation model and the better the effect. When the NRMSE is less than 10%, the model accuracy is very high, and the accuracy of the model is relatively high when the NRMSE is between 10 and 20%. The accuracy is at a normal level, when the NRMSE is between 20 and 30%, and when the NRMSE is more than 30%, the accuracy is poor. All data statistics experiments were implemented in Spyder by using Python 3.8.8 on a workstation with an Intel i7-6800 K 3.40 GHz CPU, 16 GB memory, and an Nvidia GeForce GTX 2080Ti graphics, running the Win10 operating system. Pandas 1.3.2, Matplotlib 3.4.2, and Scikit-Learn 0.24.2 were chosen for statistical analysis, correlation analysis, and significance of differences test for wheat CC.

RESULTS

Reliability Verification of UAV Multispectral Imaging Data

The images of the UAV were extracted for the reflectance of different brands of spectra according to the plots, as shown in **Figure 3**, the spectral reflectance of different fertility periods were all in the range of 450–550-nm band, and the spectral reflectance curves of different regions showed an increasing trend, and the phenomenon of green light wave peak appeared around 550 nm, and this result was more consistent with the results of literature (Fu et al., 2019). The position of the green wave peak appeared differently in different fertility stages, and the wavelengths were from large to small in the filling stage, flowering stage, and heading stage. A red trough appeared between 630 and 670 nm, and the pattern of the red trough was the same as that of the green peak. In the range of 466–830 nm, the reflectance of multispectral data has high accuracy, and this result was more consistent with the results of literature (Aasen et al., 2015), and the five multispectral bands selected in this article were all in this range, which can estimate the canopy chlorophyll of winter wheat.

Distribution of CC in the Winter Wheat

The CC values of winter wheat, under different water treatments of normal irrigation and water-limited treatment, were obtained

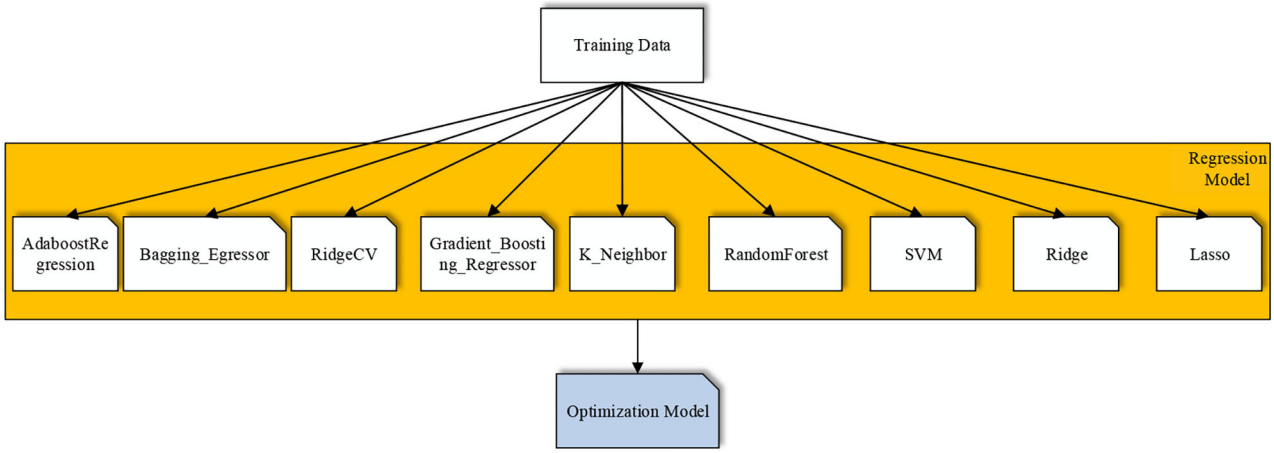


FIGURE 2 | Flow chart of the inverse model of multiple machine learning algorithms.

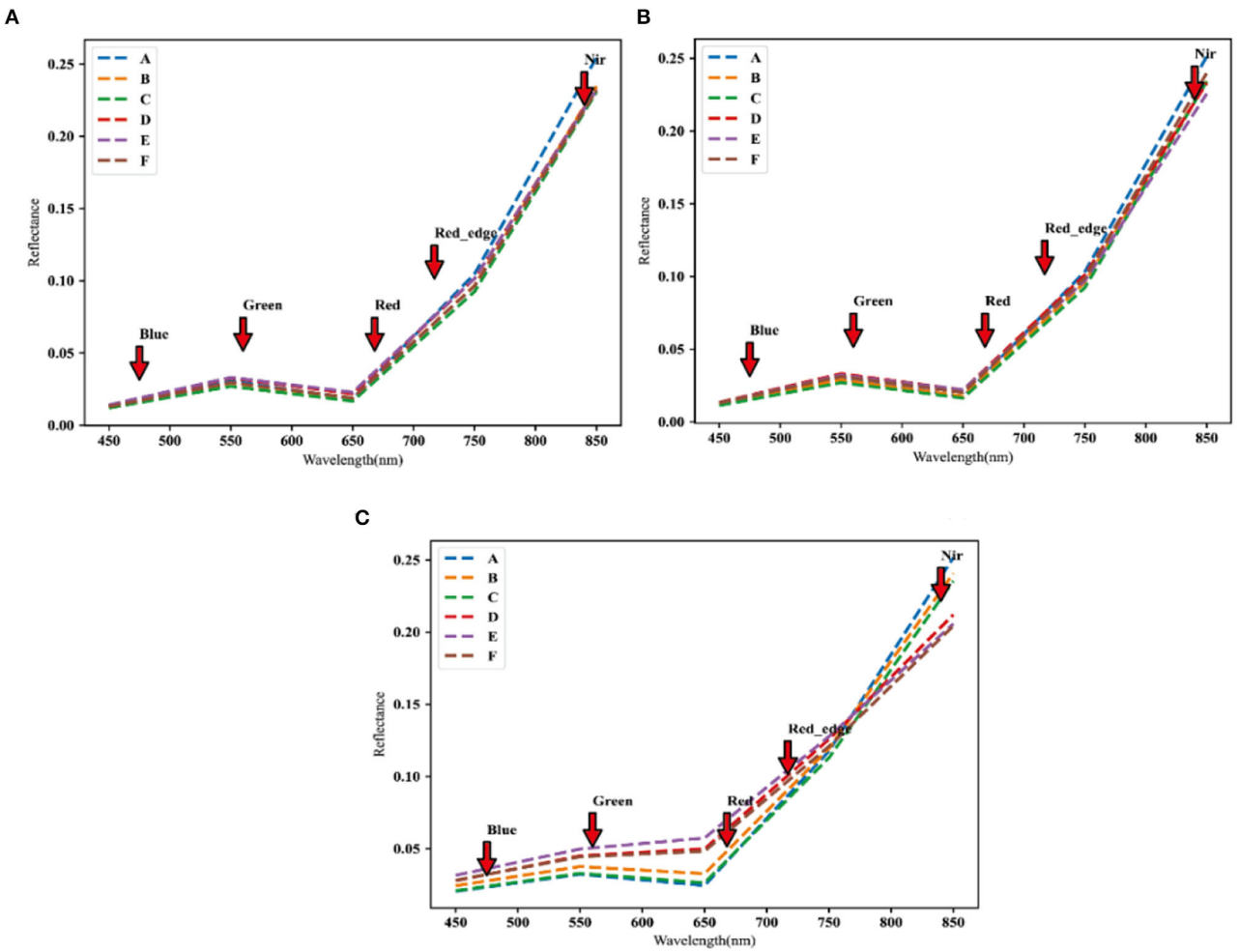


FIGURE 3 | Different spectral reflectance maps of five bands of Manas winter wheat drones at different growing stages: **(A)** Manas winter wheat heading stage; **(B)** Manas winter wheat flowering stage; and **(C)** Manas winter wheat filling stage. A–C denote three replicates under normal irrigation; D–F denote three replicates under water-limited treatment.

at the heading, flowering, and filling stages, respectively. It was evaluated by four dimensions: mean value was expressed by μ, median by a median, coefficient of variation by cv, and standard deviation by σ. From **Figure 4A**, it can be seen that the mean values of CC in winter wheat at the heading stage were distributed between 54.39-56.38, the median ranged from 56.26-57.02, the σ ranged from 3.71-4.66, and the cv ranged 6.2%-8.4%. The cv of the CC under normal irrigation ranged from 6.6 to 7.6% as seen in Graph A, B, and C in **Figure 4A**. While the cv of the population under water limitation treatment ranged from 6.2 to 8.4%, as seen in Graphs D, E, and F. From **Figure 4B**, it can be shown that the mean values of CC in winter wheat at the flowering stage ranged from 55.59 to 56.39, the median ranged from 56.06 to 57.22, the σ ranged from 3.88 to 4. 22, and the cv ranged from 6.9 to 7.9%. The coefficient of variation of the population under normal irrigation ranged from 6.9%-7.3% as seen in plots A, B, and C in **Figure 4B**. While in plots D, E, and F, the cv of the population under water limitation treatment ranged from 6.9 to 7.9%. From **Figure 4C**, it can be observed that the mean values of CC in winter wheat at the filling stage ranged from 56.41 to 59.88, the median ranged from 56.7 to 60.36, the σ ranged from 3.56 to 4.66, and the cv ranged from 5.9 to 8%. The cv of the population under normal irrigation ranged from 5.9 to 6.2% as seen in plots A, B, and C in **Figure 4C**, while in plots D, E, and F, the cv of the population under water limitation treatment ranged from 6.4 to 8%.

In summary, the overall dispersion and variation of the data were large, indicating that the population showed great variation in CC at the heading, flowering, and filling stages, and indicating that the population was rich in genetic variation. In terms of water and drought treatments, the range of variation was 1.2% for the water treatment over normal irrigation at the heading stage, 2.6% for the water treatment over normal irrigation at the flowering stage, and 1.3% for the water treatment over normal irrigation at filling stage. The range of variation became larger from the heading stage to the flowering stage, and then gradually decreased with the extension of the reproductive period.

Preferred Vegetation Index (VI)

Many selections of vegetation indices were made based on empirical values, and the visualization of the selection process was rarely given. In this study, experiments on the contribution of vegetation indices relative to CC were conducted in combination with the random forest algorithm at the heading, flowering, and filling stages of wheat under normal irrigation and water-limited treatment, respectively. It can be seen from **Figure 5** that the contribution of vegetation index to CC was different under different water treatments in different periods. From **Figure 5A**, it can be seen that the magnitude of contribution under normal irrigation at the heading stage was as follows: REOSAVI>VARIgreen>NDREI>NDRGI>NDVI>VARIred> NGBDI>OSAVI>MSR>SR>EXG>GCI>GNDVI; from **Figure 5B**, we can see that the magnitude of contribution under water irrigation at the heading stage was: VARIgreen> OSAVI> NDREI> NDRGI> EXG> REOSAVI> VARIred> NGBDI> NDVI> MSR> GCI> SR> GNDVI. As seen in **Figure 5C**, the magnitude of contribution under normal irrigation during flowering was in the following order: NDVI>GCI>VARIgreen>NDREI>MSR>SR>NDRGI> OSAVI>VARIred>EXG>REOSAVI>.

NGBDI; as seen in **Figure 5D**, the magnitude of contribution under water irrigation during flowering was in the following order: REOSAVI > NDRGI > NDREI > VARIgreen> VARIred > GCI> GNDVI> NGBDI> EXG> NDVI> OSAVI> MSR> SR. From **Figure 5E**, it can be observed that the magnitude of contribution under normal irrigation during the filling period is in the following order: VARIgreen > NDRGI > NDVI > VARIred > GNDVI > GCI > NGBDI > REOSAVI > OSAVI > SR > MSR > EXG > NDREI; from **Figure 5F**, it can be seen that the magnitude of contribution under water irrigation during the filling period is in the following order: VARIgreen > NDRGI > VARIred > REOSAVI > NDREI > OSAVI > EXG > GCI > GNDVI.

In general, the contribution of vegetation indices under two different water treatments can be found in winter wheat at the heading stage, and the overall contribution was ranked in the top 5 with three vegetation indices, VARIgreen, NDREI, and NDRGI, which were used as the priority vegetation indices when the model was constructed at the heading stage. In the flowering stage of winter wheat, the contribution of vegetation indices under two different water treatments was found to be different, and the overall contribution was ranked in the top 5 with two vegetation indices, VARIgreen and NDRGI, which were used as the priority vegetation indices in the model construction of the flowering stage. In winter wheat, the contribution rates of vegetation indices under two different water treatments were found to be different in the filling stage, and the overall contribution rates were ranked in the top 5 with three vegetation indices, VARIgreen, NDRGI, and VARIred, which were used as the priority vegetation indices in the model construction of the filling stage.

Correlation Analysis Between SPAD and Vegetation Index of Winter Wheat

The spectral parameters of the three fertility stages of heading, flowering, and filling were correlated with winter wheat CC and the results were shown in **Figures 6–8**. From **Figure 6A**, it can be seen that most of the spectral vegetation indices selected under normal irrigation at the heading stage reached highly significant levels ($p < 0.0001$). Among them, Normalized Green and Blue Difference Index (NGBDI) showed no significant correlation during the heading period, while the rest of the parameters showed correlation, among which the highest positive correlations were NDVI, MSR, and SR with correlation coefficients r reaching.5, followed by NDREI with correlation coefficient r reaching.48. It can be seen from **Figure 6B** that most of the selected spectral vegetation indices under water limitation treatment reached highly significant levels ($p < 0.0001$). Among them, Over Green Index (EXG) and NGBDI showed no significant correlation at the heading stage, while the rest of the parameters showed correlation, with the highest negative correlation being NDVI and VARIred, with a correlation

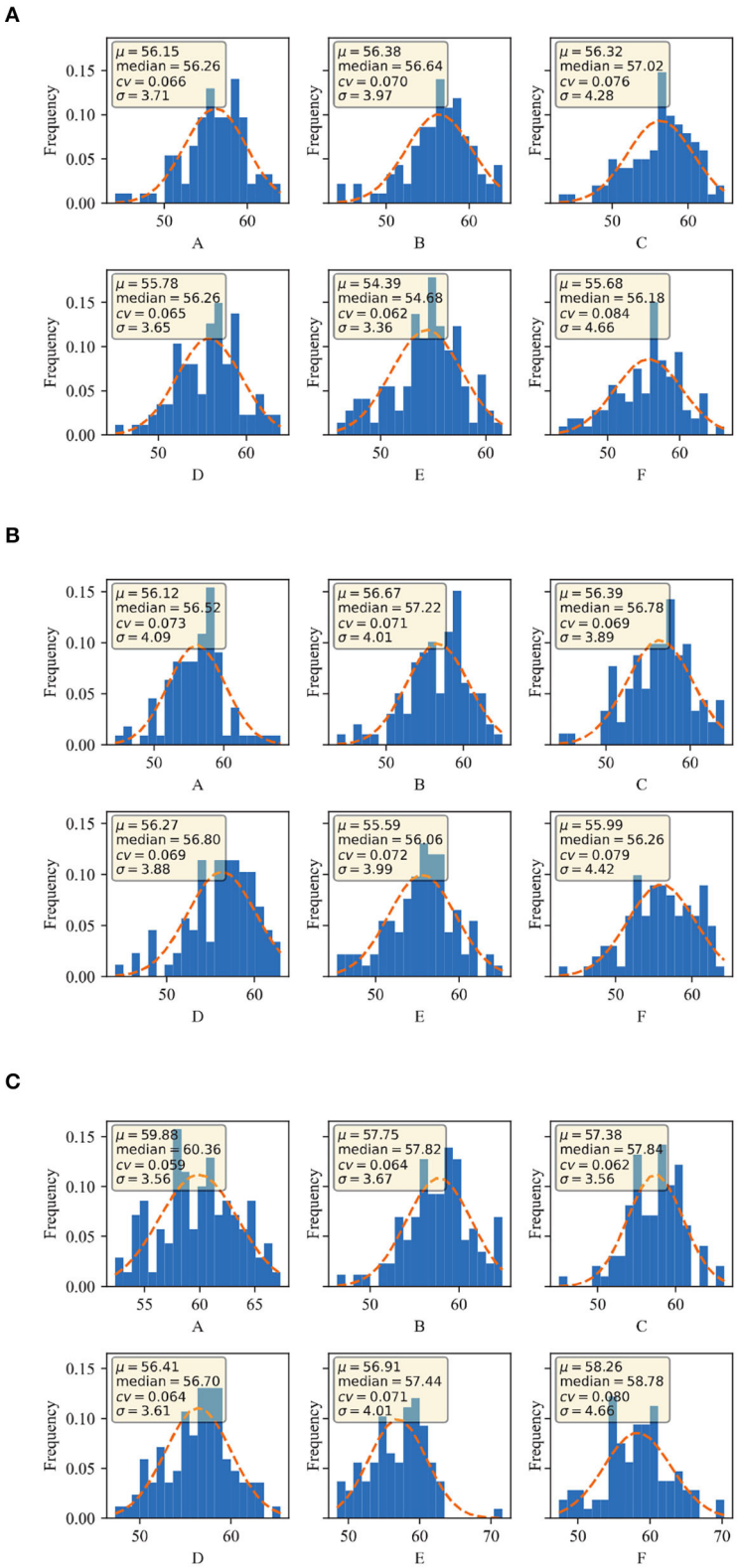


FIGURE 4 | CC distribution of different growing stages of Manas winter wheat: **(A)** heading stage of Manas winter wheat; **(B)** flowering stage of Manas winter wheat; and **(C)** filling stage of Manas winter wheat. A–C denote three replicates under normal irrigation; D–F denote three replicates under water-limited treatment.

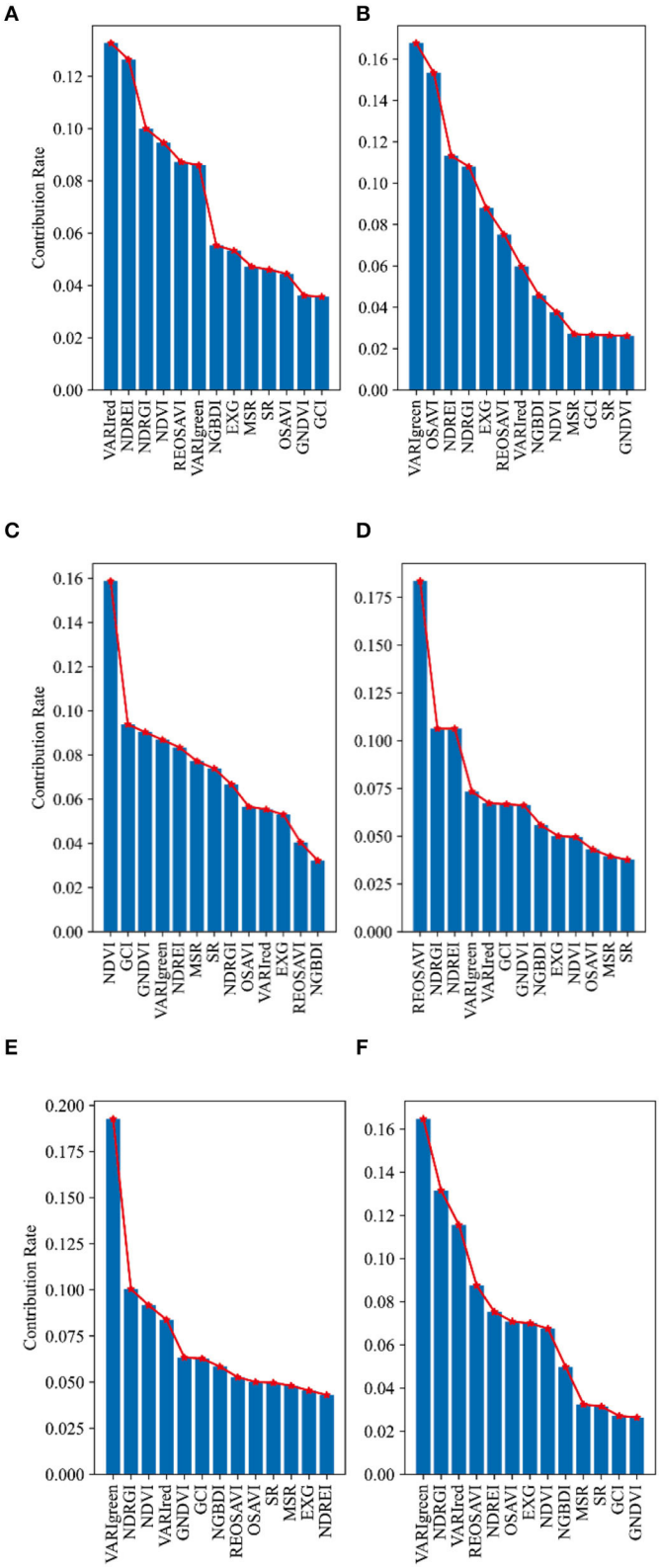


FIGURE 5 | Contribution rate distribution of vegetation indices relative to CC at different growing stages of winter wheat: **(A)** normal irrigation at heading stage of Manas winter wheat; **(B)** water limited treatment at heading stage of Manas winter wheat; **(C)** normal irrigation at flowering stage of Manas winter wheat; **(D)** water limited treatment at flowering stage of Manas winter wheat; **(E)** normal irrigation at filling stage of Manas winter wheat; and **(F)** water limited treatment at filling stage of Manas winter wheat. All vegetation indices that appear in the figure are explained in **Table 3**.

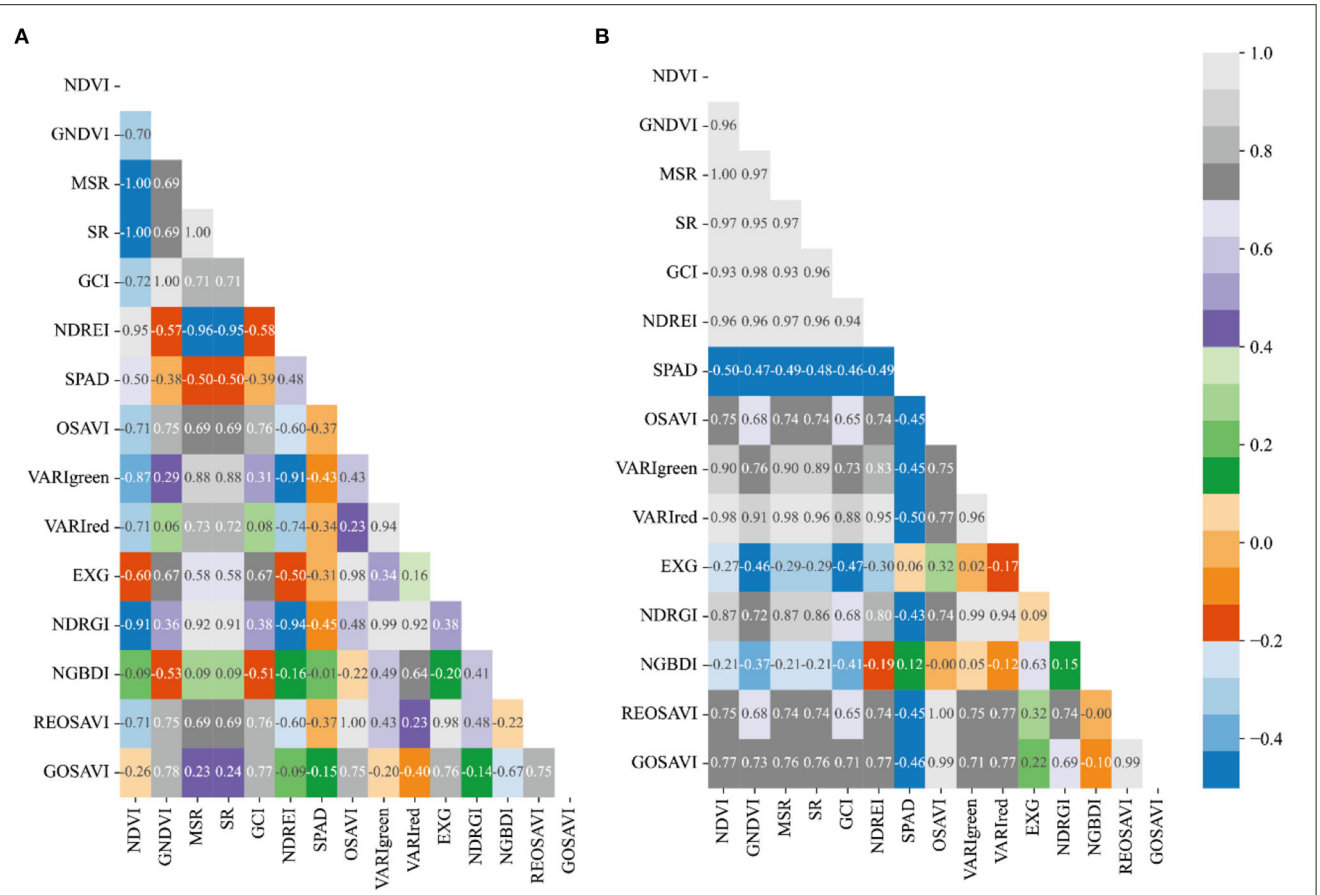


FIGURE 6 | Plot of correlation between different vegetation indices and CC of winter wheat during heading stage: **(A)** normal irrigation of Manas winter wheat during heading stage, and **(B)** water limited treatment of Manas winter wheat during heading stage. All vegetation indices that appear in the figure are explained in **Table 3**.

coefficient *r* reaching −0.5, followed by MSR and NDREI, with a correlation coefficient *r* reaching −0.49.

While the correlation coefficient *r* of VARIgreen in section Preferred Vegetation Index (VI), using the preferred vegetation index of the random forest, was −0.43 under normal irrigation and −0.45 under water-limited treatment; the correlation coefficient *r* of NDREI was 0.48 under normal irrigation and −0.49 under water-limited treatment; the correlation coefficient *r* of NDRGI was −0.45 under normal irrigation and −0.43 under water limitation treatment; all three vegetation indices reached a highly significant level ($p < 0.0001$).

Most of the spectral vegetation indices selected under normal irrigation during flowering reached a highly significant level ($p < 0.0001$) as can be seen in **Figure 7A**. Of these, EXG and NGBDI showed no significant correlation, while the rest of the parameters showed correlation, with the highest positive correlation being NDVI and MSR, with SR correlation cefficient *r* reaching 0.5, followed by VARIred, with correlation coefficient *r* reaching 0.46. From **Figure 7B** it can be seen that most of the spectral vegetation indices selected under water limitation treatment reached highly significant levels ($p < 0.0001$). Among them, EXG showed no significant correlation, while the rest of the parameters showed correlation, with the highest positive correlation being NDVI and MSR, with a correlation coefficient *r* reaching 0.5, followed by GNDVI, SR, and VARIred, with a correlation coefficient *r* reaching −0.49.

In contrast, in the preferred vegetation index using the random forest in Section Preferred Vegetation Index, the correlation coefficient *r* for VARIgreen was 0.45 and 0.33, respectively, under normal irrigation and water-limited treatment; the correlation coefficient *r* for Red Edge Index (NDREI) was 0.46 under normal irrigation and 0.48 under water-limited treatment, and both 2 vegetation indices reached a highly significant level ($p < 0.0001$).

It is evident from **Figure 8A** that most of the spectral vegetation indices selected under normal irrigation during the filling period reached a highly significant level ($p < 0.0001$). The positive correlation was highest for NDVI with a correlation coefficient *r* of 0.51, followed by VARIgreen with a correlation coefficient *r* of 0.42, and the negative correlation was highest for GNDVI with a correlation coefficient *r* of −0.50. It can be seen from **Figure 8B** that the spectral vegetation indices selected under water limitation treatments all reached highly significant

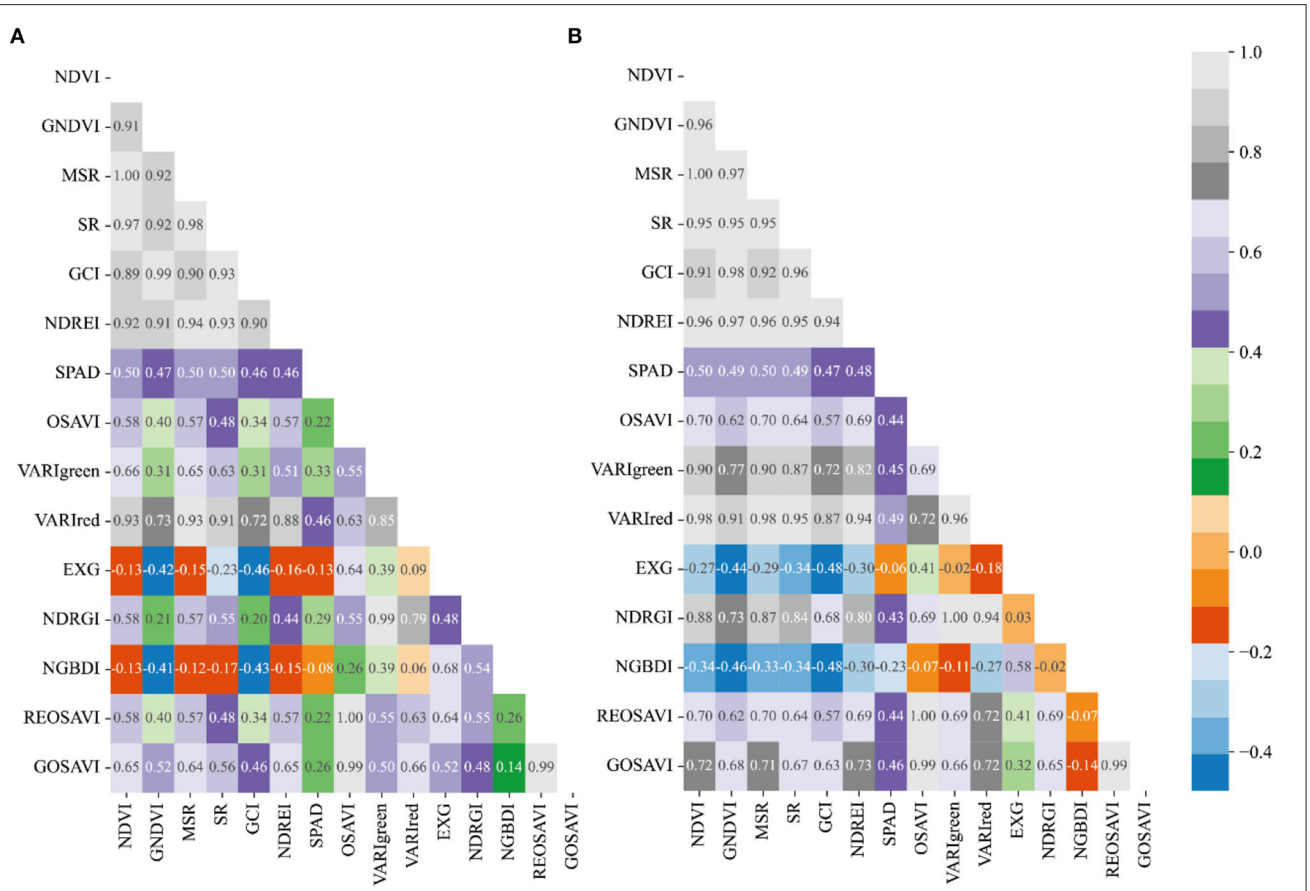


FIGURE 7 | Plot of correlation between different vegetation indices and CC of winter wheat during flowering stage: **(A)** Normal treatment of Manas winter wheat during flowering stage, and **(B)** water limited treatment of Manas winter wheat during flowering stage. All vegetation indices that appear in the figure are explained in **Table 3**.

levels ($p < 0.0001$). All the vegetation indices selected under the water limitation treatment reached a highly significant level ($p < 0.0001$). The highest positive correlations were OSAVI and REOSAVI with a correlation coefficient r of 0.51, followed by MSR and GOSAVI with a correlation coefficient r of 0.50. The highest negative correlation was NDVI, and the correlation coefficient r reached −0.50.

Meanwhile, the correlation coefficient r of VARIgreen under normal irrigation and −0.44 under water-limited treatment in the preferred vegetation index using the random forest in section Preferred Vegetation Index was 0.42; the correlation coefficient r of NDRGI was 0.42 under normal irrigation and −0.44 under water-limited treatment, and the correlation coefficient r of VARIred under normal irrigation was −0.41 under normal irrigation and.46 underwater limitation treatment, all three vegetation indices reached significant levels ($p < 0.0001$).

Therefore, the correlation analysis of vegetation indices and CC of winter wheat for the three fertility periods of winter wheat was combined, and the correlations of the vegetation indices preferred in the previous section all reached significant levels, and in addition, the model estimation was carried out by combining the vegetation indices with the highest correlation in that fertility period.

Algorithm Development for CC Estimation

The model inversions were conducted using nine machine learning algorithms, Adaboost Regression, Bagging_Regressor, Gradient_Boosting_Regressor, K_Neighbor, Random Forest, SVM, Lasso, RidgeCV, and Ridge, for the SPAD values of winter wheat at the heading, flowering, and filling stages under two water treatments. The results in **Table 4** show that the correlation coefficients between predicted and true values under normal irrigation at the heading stage ranged from 0.36 to 0.63 for r, 3.28–3.67 for RMSE, and 16.2–18.1% for NRMSE. The highest correlation was the RidgeCV model with correlation coefficient $r = 0.63$, which had RMSE = 3.28 and NRMSE = 16.2% for both Random Forest and RidgeCV in terms of model accuracy. Overall, it shows that the RidgeCV model has the best prediction accuracy and prediction effect under normal water treatment at the heading stage. In contrast, the correlation coefficients r between the predicted and true values under the water-limited treatment at the heading stage ranged from 0.41 to 0.63, RMSE from 3.44 to 3.95, and NRMSE from 18.8 to 21.9%. The highest prediction correlation is the SVM model with a correlation coefficient of $r = 0.63$, RMSE = 3.47 and NRMSE = 19.2%, which is still very good in terms of prediction accuracy. In terms of the prediction accuracy of the model, Adaboost Regression has

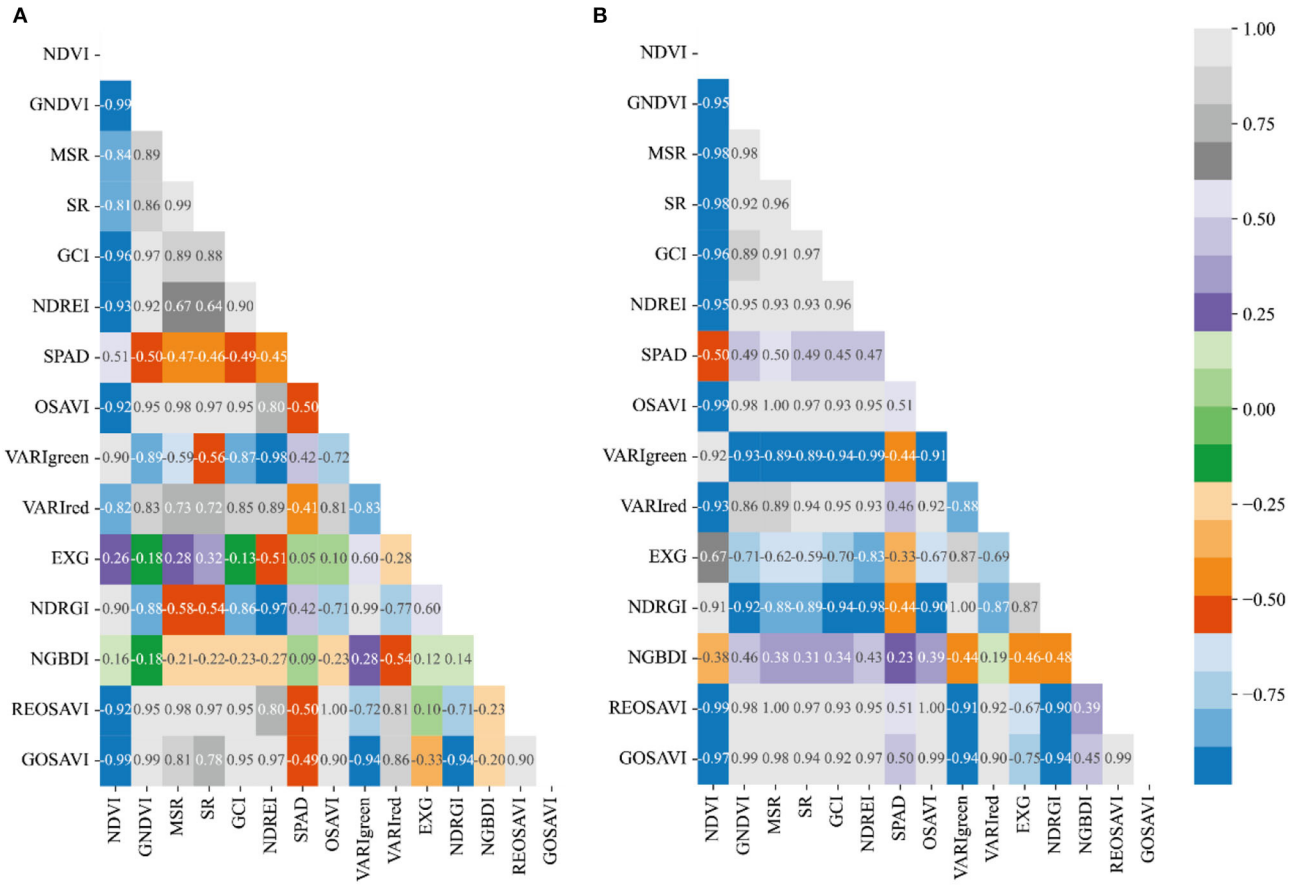


FIGURE 8 | Plot of correlation between different vegetation indices and CC of winter wheat during the filling period: **(A)** Normal treatment of Manas winter wheat during the filling period, and **(B)** Water treatment of Manas winter wheat during the filling period. All vegetation indices that appear in the figure are explained in **Table 3**.

TABLE 4 | Model analysis of CC for vegetation index prediction at the heading stage.

Models	DI			DS		
	r	RMSE	NRMSE (%)	*r*	RMSE	NRMSE (%)
Adaboost Regression	0.44	3.5	17.3	0.60	3.39	18.8
Bagging_Regressor	0.49	3.39	16.8	0.48	3.77	20.9
Gradient_Boosting_ Regressor	0.36	3.67	18.1	0.41	3.95	21.9
K_Neighbor	0.50	3.38	16.7	0.58	3.44	19.1
Random Forest	0.55	3.28	16.2	0.56	3.5	19.4
SVM	0.62	3.58	17.3	0.63	3.47	19.2
Lasso	0.61	3.37	16.7	0.60	3.48	19.3
RidgeCV	0.63	3.28	16.2	0.61	3.44	19.1
Ridge	0.61	3.47	17.1	0.60	3.47	19.2

DI, normal irrigation; DS, limited water treatment.

the smallest NRMSE of 18.8%, while the prediction correlation is $r = 0.60$.

The results of the model analysis of the predicted CC of vegetation index under normal irrigation and water limitation treatments during flowering are shown in **Table 5**. The correlation coefficients *r* between predicted and true values under normal irrigation ranged from 0.27 to 0.50, RMSE from 2.79 to 3.20, and NRMSE from 17 to 19.5%. The highest model prediction correlation is the SVM model, which has a correlation coefficient of $r = 0.50$, RMSE = 2.79 and NRMSE = 17%, and the SVM also has the lowest NRMSE in terms of model accuracy. The correlation coefficients R^2 between predicted and true values under water-limiting treatment ranged from 0.42 to 0.50, RMSE from 2.90 to 3.03, and NRMSE from 19.1 to 20.3%. The highest correlation predicted by the models was the Bagging_Regressor model, which had a correlation coefficient of $r = 0.50$, RMSE = 2.90, and NRMSE = 19.1%, with Bagging_Regressor having the lowest NRMSE as far as the accuracy of the model is concerned. Overall, it shows that the Bagging_Regressor model has the best prediction accuracy and prediction under flowering duration water treatment.

Analysis of the model for predicting CC of vegetation index under normal irrigation and water limitation treatments during the irrigation period is shown in **Table 6**. The correlation coefficients *r* between the predicted and true values under normal irrigation ranged from 0.21 to 0.43, RMSE from 2.91 to 3.49, and NRMSE from 21.6 to 25.8%. The model with the highest model prediction correlation is the SVM, which has a correlation

TABLE 5 | Model analysis of CC for predicting vegetation index during flowering.

Models	DI			DS		
	r	RMSE	NRMSE (%)	*r*	RMSE	NRMSE (%)
Adaboost Regression	0.34	3.09	18.8	0.44	3.03	19.9
Bagging_Regressor	0.40	2.96	18.0	0.50	2.90	19.1
Gradient_Boosting_ Regressor	0.49	2.81	17.1	0.43	3.08	20.3
K_Neighbor	0.27	3.20	19.5	0.44	3.02	19.9
Random Forest	0.46	2.83	17.2	0.45	3.01	19.8
SVM	0.50	2.79	17.0	0.46	2.98	19.6
Lasso	0.47	2.81	17.1	0.43	3.02	19.8
RidgeCV	0.47	2.81	17.1	0.42	3.03	20.0
Ridge	0.47	2.81	17.1	0.44	3.00	19.7

DI, normal irrigation; DS, limited water treatment.

TABLE 6 | Determination coefficient (*r*), root mean square error (RMSE), and relative error (RE) of the algorithms for modeling estimation of chlorophyll content (CC) of wheat in different filling period and normal irrigation (NI) and drought stress (DS) conditions.

Models	DI			DS		
	r	RMSE	NRMSE (%)	*r*	RMSE	NRMSE (%)
Adaboost Regression	0.34	3.08	22.8	0.35	4.08	18.6
Bagging_Regressor	0.21	3.49	25.8	0.37	4.04	18.4
Gradient_Boosting_ Regressor	0.27	3.36	24.9	0.32	4.17	19.0
K_Neighbor	0.41	2.97	22.0	0.43	3.78	17.2
Random Forest	0.26	3.28	24.3	0.40	3.91	17.8
SVM	0.43	2.91	21.6	0.51	3.57	16.3
Lasso	0.41	2.96	21.9	0.48	3.69	16.8
RidgeCV	0.41	2.94	21.8	0.48	3.69	16.8
Ridge	0.40	2.96	22.0	0.49	3.68	16.8

DI, normal irrigation; DS, limited water treatment.

coefficient of $r = 0.43$, RMSE = 2.91, and NRMSE = 21.6%, and the SVM also has the lowest NRMSE in terms of the accuracy of the model. The correlation coefficients *r* between predicted and true values under water-limiting treatment ranged from 0.32 to 0.51, RMSE from 3.57 to 4.17, and NRMSE from 16.3 to 19%. The highest correlation predicted by the models was the SVM model, which had a correlation coefficient of $r = 0.51$, RMSE = 3.57, and NRMSE = 16.3%, and the SVM had the lowest NRMSE in terms of the accuracy of the model. Overall it shows that the SVM model has the best prediction accuracy and prediction under normal irrigation and water limitation treatment during the irrigation period.

In general, among the prediction models of CC using nine machine learning algorithms for three different fertility stages, namely, heading, flowering, and filling, the correlation coefficients of the RidgeCV model under normal irrigation and the SVM model underwater restriction treatment were the highest in the heading stage; the correlation coefficients of the SVM model under normal irrigation and the SVM model underwater restriction treatment were the highest in the flowering stage; the correlation coefficients of the Bagging_Regressor model were the highest in both normal irrigation and water restriction treatments. The correlation coefficient of the SVM model was the highest under normal irrigation and the SVM model was the highest underwater restriction. In terms of prediction accuracy, Random Forest and RidgeCV models had the highest prediction accuracy under normal irrigation in the heading stage, and the Adaboost Regression model had the highest prediction accuracy under water restriction treatment; the SVM model had the highest prediction accuracy under normal irrigation in the flowering stage, and Bagging_Regressor model had the highest prediction accuracy under water restriction treatment. The highest prediction accuracy was achieved by the SVM model under normal irrigation and water restriction treatment at the flowering stage, and the highest prediction accuracy was achieved by the SVM model underwater restriction treatment.

DISCUSSION

Effect of Water and Drought Treatment on Chlorophyll

In drought environments, plants themselves evolve a series of mechanisms for self-protection and adaptation and resistance to unfavorable environmental stresses, and their phenotypic characteristics are significantly altered to minimize the impact of the adverse environment on their growth and development. At the same time, drought stress causes complex effects on the population structure and physiology of crops in various ways (Roessner, 2012). Concerning the reproductive stages of wheat, the whole reproductive period is divided into four stages: early-stage (sowing-pulling), developmental stage (pulling-heading), middle stage (pulling-potting), and late-stage (potting-harvest). Previous studies on the effects of drought stress on different fertility stages of wheat have suggested that drought stress affects the internal physiological and biochemical phenotypes of wheat to different degrees, and these changes are manifested in changes in chlorophyll content (Cao, 2010), indicating that chlorophyll content is closely related to drought resistance and yield traits in wheat.

In this study, for the analysis of canopy chlorophyll content of winter wheat, the results in **Figure 9** showed that normal irrigation conditions increased the canopy chlorophyll content of wheat from the heading stage to the filling stage very significantly, compared with drought stress. In addition, chlorophyll is the most important pigment for photosynthesis, which affects the physiological and biochemical processes in the crop under drought stress. Drought stress causes reactive oxygen species produced by the plant body to disrupt cell membranes, which hinders chlorophyll synthesis and accelerates degradation, thus reducing chlorophyll content.

Meanwhile, previous studies had found that drought stress leads to increased accumulation of malondialdehyde (MDA) and peroxide dismutase (POD) in plants (Shao et al., 2006). MDA was a class of highly reactive lipid peroxides that cross-link and

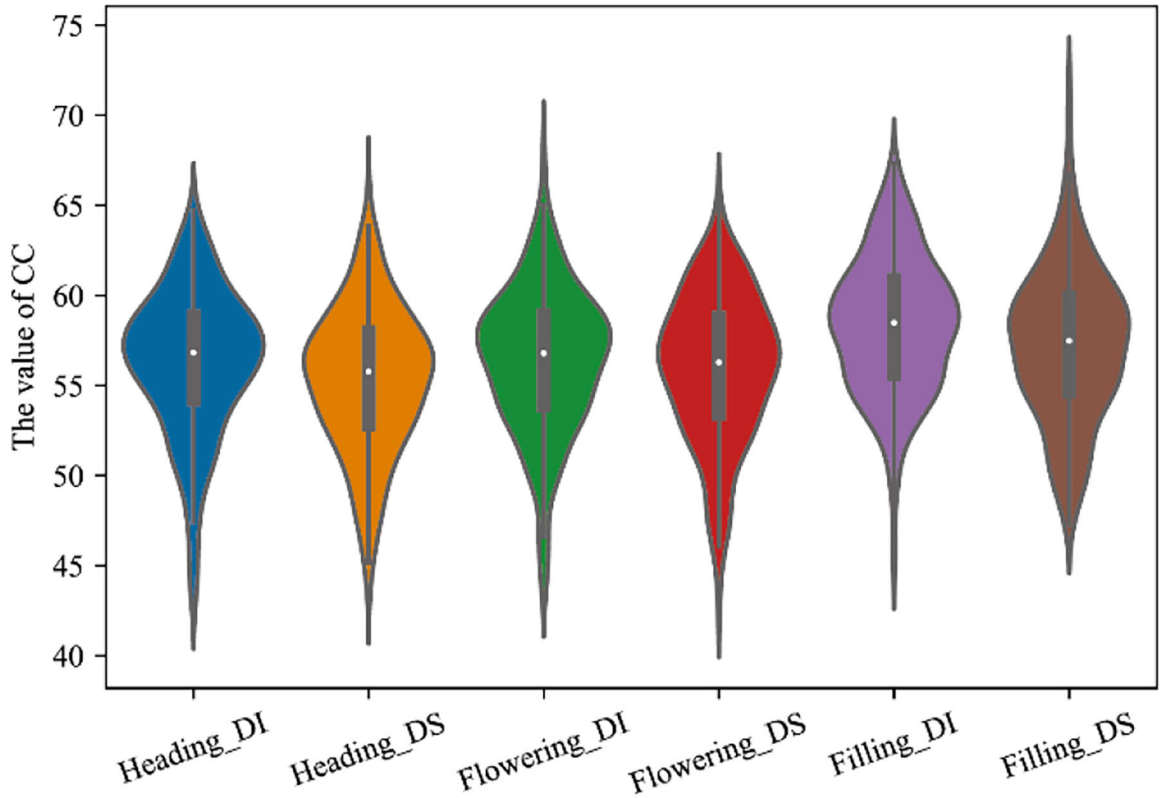


FIGURE 9 | CC distribution of winter wheat at different fertility stages under two water treatments. Heading_DI, Heading_DS, Flowering_DI, Flowering_DS, Filling_DI, and Filling_DS refer to normal treatment at heading stage, water treatment at heading stage, normal treatment at flowering stage, water treatment at flowering stage, normal treatment at filling stage, and water treatment at filling stage, respectively.

polymerize lipid nucleic acids, proteins, etc., and affected the components of cytoplasmic membranes, including chloroplast lamellae. POD can generate reactive oxygen species and trigger lipid membrane peroxidation under longer drought stress. Both substances can lead to changes in membrane structure and affect water metabolism by causing water loss in the chloroplasts, and thus the rate of chlorophyll synthesis was reduced. It has been suggested that drought stress can lead to reduced chlorophyll content in wheat, accelerated leaf senescence, and reduced green leaf area, resulting in reduced wheat yield (Verma et al., 2004). Also, previous studies had shown that the CC values of wheat flag leaves under drought stress tend to decrease and that wheat varieties with higher CC values under stress have higher dry matter quality and better drought resistance (Sun et al., 2019). These findings were consistent with the results of this article that the CC values of wheat flag leaves under drought stress showed a decreasing trend.

Generality of CC Inversion Model

Current UAV multispectral with high spectral resolution and flexible mobility played an important role in crop high-throughput phenotyping studies. In this article, we use UAVs with multispectral cameras for ground image data acquisition and estimation of ground CC content. The reflectance extraction of the image data revealed that the spectral reflectance curves of different fertility stages studied in this article and the phenomenon of a green light wave peak at a wavelength of about 550 nm can be seen, and this result was more consistent with the results of the literature (Aasen et al., 2015). The positions of the green light peaks differed among the different fertility stages, with the wavelengths of the peaks appearing at the filling, flowering, and heading stages ranging from large to small. A red trough appeared between 630 and 670 nm, and the pattern of the red trough was consistent with that of the green peak. In the range of 466–830 nm, the reflectance of the multispectral data has high accuracy, and this result is more consistent with the results of the literature (Aasen et al., 2015).

A single vegetation index does not adequately reflect the crop growth, but too many vegetation indices as input parameters of the model will lead to an increase in the complexity of the model. Therefore, the optimal vegetation indices for different fertility periods were obtained by a random forest algorithm before model construction, and the vegetation indices for different treatments involved in model construction at different periods were determined by combining the correlation of CC and vegetation indices later. The vegetation indices involved in model construction under normal treatment at the heading stage were REOSAVI, VARIgreen, NDREI, NDVI, MSR, and SR, and vegetation indices under water limitation treatment were VARIgreen, OSAVI, NDREI, NDVI, and VARIred; the vegetation

indices involved in model construction under normal treatment at the flowering stage were NDVI, GCI, VARIgreen, NDVI, and VARIred. GCI, VARIgreen, NDREI, MSR, and SR, and the vegetation indices under the water-limited treatment were REOSAVI, NDRGI, NDREI, NDVI, and MSR; the vegetation indices involved in the model construction under the normal treatment at flowering were VARIgreen, NDRGI, NDVI, EXG, and NGBDI, and the vegetation indices under the water-limited treatment were. The preferred vegetation indices under different water treatments in three different periods were different, but NDVI was not selected only in the water treatment during the filling period, while all other models were involved, which also indicates the prevalence of NDVI vegetation indices in crop model construction and the importance of NDVI vegetation indices. This is also consistent with many current types of research using NDVI vegetation indices in modeling studies.

In terms of CC prediction models, this article investigates CC prediction models by nine machine learning algorithmic models for three different fertility stages, namely, heading, flowering, and filling, under normal irrigation and water limitation treatments, respectively. The prediction models were found to be different for different water treatments at different fertility stages, but the model with the higher correlation between both predicted and true values under different treatments at different fertility stages was the SVM model, which embodied a strong fit and accuracy among all the models. The modeling of SVM under the normal treatment at the heading stage ($r = 0.62$, RMSE = 3.58, NRMSE = 17.3%) ranked second in correlation, and the modeling of SVM under the water limitation treatment ($r = 0.63$, RMSE = 3.47, NRMSE = 19.2%) ranked first in correlation; the modeling of SVM under the normal treatment at the flowering stage ($r = 0.50$, RMSE = 2.79, NRMSE = 17%), ranked second in correlation, modeling of SVM under water limitation treatment ($r = 0.46$, RMSE = 2.98, NRMSE = 18.6%), ranked second in correlation, and modeling of SVM under normal treatment at filling stage ($r = 0.43$, RMSE = 2.91, NRMSE = 21.6%), ranked first in correlation, and modeling of SVM under water-limited treatment ($r = 0.51$, RMSE = 3.57, NRMSE = 16.3%), ranked first in correlation. From the overall point of view, SVM showed the most advantage in the filling stage, this is related to the adaptability of the model under different water treatments at different fertility stages, that is, the prediction effect of different models applying different water treatments at different fertility stages is different. From the distribution of SPAD, the distribution of SPAD in the heading stage, flowering stage, and filling stage were significantly different, the distribution of the filling stage was more stable, and the time nodes of the population in the filling stage were more consistent with the fertility stage, and the prediction model of SVM had higher accuracy.

CONCLUSION

The multispectral images acquired by the UAV were used to extract the reflectance of five spectra of different genotypes (Blue, Green, Red, Red_edge, and Nir) and calculate different vegetation indices, combined with the ground canopy data collected by the handheld CC instrument underwater and dry treatments at the heading, flowering, and filling stages. Then the relevant research analysis was carried out, and the analysis of the reflectance curve of the spectrum showed a phenomenon that the green wave peak appeared at around 550 nm, and the position of the red wave valley appeared once between 630 and 670 nm, obviously, and the occurrence law of the red wave trough is consistent with that of the green wave peak. It shows that the reflectance data obtained in this study are of high quality and have good accuracy. In the study of CC phenotype distribution, it can be found that the range of variation of this population under different water treatments at different fertility periods is large, the genetic variation of the population is rich and the CC content under normal irrigation is significantly higher than that of water-limited treatment. The vegetation indices under different water treatments at different fertility periods were selected by combining the preferred vegetation indices and the correlation evaluation of CC with vegetation indices. This study examined a series of machine learning algorithms, including Adaboost Regression, Bagging_Regressor, Gradient_Boosting_Regressor, K_Neighbor, Random Forest, SVM, Lasso, RidgeCV, and Ridge in the high-throughput phenotyping context. The results showed that the highest predicted correlation under normal irrigation at the heading stage was the RidgeCV model with correlation coefficient $r = 0.63$, which had RMSE = 3.28 and NRMSE = 16.2%, and the one with highest correlation under water limitation treatment was the SVM model with correlation coefficient $r = 0.63$, RMSE = 3.47 and NRMSE = 19.2%; under normal irrigation at the flowering stage, the highest correlation was from the SVM model, which had a correlation coefficient of $r = 0.50$, RMSE = 2.79, and NRMSE = 17%, and the model with the highest correlation was Bagging_Regressor under water restriction treatment, which had a correlation coefficient of $r = 0.50$, RMSE = 2.90, and NRMSE = 19.1%; and under normal irrigation at the filling stage, the highest correlation came from the SVM model, which had a correlation coefficient of $r = 0.43$, RMSE = 2.91, NRMSE = 21.6%, and also the SVM model has the highest correlation under water limitation treatment with coefficient of $r = 0.51$, RMSE = 3.57, NRMSE = 16.3%. The results of this study showed that the prediction model constructed using the SVM model under different water treatments at different fertility stages could better invert the chlorophyll content of winter wheat canopies with different growth differences.

Many machine learning and empirical models can be selected to correlate hyperspectral reflectance with CC; therefore. it was worth investigating which models worked better and whether the combination of individual regression techniques can provide better predictive performance. The study of CC model inversion by a large number of machine learning algorithms also provided a reference for machine learning in model prediction applications. The cumulative data obtained through field trials were still empirical models obtained through statistics, which have some limitations in the spatial and temporal domain. Altogether, our results provide insights into the capacity of UAV-based remote sensing for switchgrass high-throughput phenotyping in the field, which will be useful for breeding and cultivar development. Moreover, the UAV-based approaches proposed in this study,

including the wheat's SPAD-phenotyping and predicting model, facilitated high-throughput, and precise phenotype mapping, which should have an impact on wheat breeding as well as practical use in the field. In the future, we will try to add environmental factors while improving the accuracy of UAV remote sensing images to reduce the limitation of environment on the model and give full play to the advantages of UAV high-throughput phenotype acquisition.

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/s.

AUTHOR CONTRIBUTIONS

HG contributed to conceptualization, project administration, resources, supervision, writing the original draft, and writing–reviewing and editing the manuscript. WW contributed to formal analysis, writing the original draft, and writing–reviewing and editing the manuscript. YC, YR, and ZZ contributed to investigation, methodology, and data curation. All authors have read and agreed to the published version of the manuscript.

FUNDING

The present study was funded by the Xinjiang Major Science and Technology Special Project (2021A02001-1).

REFERENCES

Aasen H, Burkart A, Bolten A, Baretha G. (2015). Generating 3D hyperspectral information with lightweight UAV snapshot cameras for vegetation monitoring: from camera calibration to quality assura*nce. ISPRS J. Photogramm. Rem. Sens*. 108, 245–259. doi: 10.1016/j.isprsjprs.2015.08.002

Bendig, J., Aasen, Y. u. K., Bolten, H., Bennertz, A., and Broscheit, S. J, et al. (2015). Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *Int. J. Appl. Earth Observ. Geoinf.* 39, 79–87. doi: 10.1016/j.jag.2015.02.012

Breiman, L. (2001). Random forests. *Mach. Learn*. 45, 5–32. doi: 10.1023/A:1010933404324

Cao, D. W. (2010). Effect of shading on morphology, physiology and grain yield of winter wheat. *Eur. J. Agron.* 33, 267–275. doi: 10.1016/j.eja.2010.07.002

Chen, J. M. (1996). Evaluation of vegetation indices and a modified simple ratio for boreal applications. *Can. J. Rem. Sens*. 22:229–42. doi: 10.1080/07038992.1996.10855178

Cover, T., and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* 13, 21–27. doi: 10.1109/TIT.1967.1053964

Freund, Y., and Schapire, R. E. (1997). A decision of on-line learning and an application to boosting. *J. Comput. Syst. Sci*. 55, 119–139. doi: 10.1006/jcss.1997.1504

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Ann. Stat.* 29, 1189–1232. doi: 10.1214/aos/1013203451

Fu, P., Meacham, K., Guan, K., and Bernacchi, C. J. (2019). Hyperspectral leaf reflectance as proxy for photosynthetic capacities: an ensemble approach based on multiple machine learning algorithms. *Front. Plant Sci*. 10:730. doi: 10.3389/fpls.2019.00730

Garg, B., Kirar, N., Menon, S., and Sah, T. (2016). A performance comparison of different back propagation neural networks methods for forecasting wheat production. *Csi Transac. Ict*. 4, 305–311. doi: 10.1007/s40012-016-0096-x

Gilabert, M. A., González-Piqueras, J., Garcaharo, F. J., and Meliá, J. (2002). A generalized soil-adjusted vegetation in*dex. Rem. Sens. Environ.* 82, 303–310. doi: 10.1016/S0034-4257(02)00048-2

Gitelson, A. A. (2005). Remote estimation of canopy chlorophyll content in crops. *Geophys. Res. Lett*. doi: 10.1029/2005GL022688

Gitelson, A. A., Kaufman, Y. J., Stark, R., and Rundquist, D. (2002). Novel algorithms for remote estimation of vegetation fraction. *Remote Sensing of Environment*, 80, 76–87. doi: 10.1016/S0034-4257(01)00289-9

Grinberg, N. F., Orhobor, O. I., and King, R. D. A. (2020). evaluation of machine-learning for predicting phenotype: studies in yeast, rice, and wheat. *Mach. Learn.* 109, 251–277. doi: 10.1007/s10994-019-05848-5

Guillen-Climent, M. L., Zarco-Tejada, P. J., Berni, J., North, P., and Villalobos, F. J. (2012). Mapping radiation interception in row-structured orchards using 3D simulation and high-resolution airborne imagery acquired from a UAV. *Precis. Agricult.* 13, 473–500. doi: 10.1007/s11119-012-9263-8

Hall, P., and Turlach, B. A. (2007). *Bagging in the presence of outliers.*

Houwelingen, S. L. C. C. V. (1992). Ridge Estimators in Logistic Regression. *J. Royal Stat. Soc. Ser. C*, 41, 191–201. doi: 10.2307/2347628

Hunt, E. R., Cavigelli, M., Daughtry, C., Mcmurtrey, J. E., and Walthall, C. L. (2005). Evaluation of digital photography from model aircraft for remote sensing of crop biomass and nitrogen status. *Precis. Agric.* 6, 359–378. doi: 10.1007/s11119-005-2324-5

Hunt, E. R., Doraiswamy, P. C., Mcmurtrey, J. E., Daughtry, C. S. T., Perry, E. M., and Akhmedov, B. (2013). A visible band index for remote sensing leaf chlorophyll content at the canopy *scale. Int. J. Appl. Earth Observ. Geoinform*. 21, 103–112. doi: 10.1016/j.jag.2012.07.020

Jin, X., Liu, S., Baret, F., Hemerlé, M., and Comar, A. (2017). Estimates of plant density of wheat crops at emergence from very low altitude UAV imagery. *Rem. Sens. Environ.* 198, 105–114. doi: 10.1016/j.rse.2017.06.007

Kaivosoja, J., Pesonen, L., Kleemola, J., Plnen, I., and Rajala, A. (2013). A case study of a precision fertilizer application task generation for wheat based on classified hyperspectral data from UAV combined with farm history data. *Proc. SPIE Int. Soc. Opt. Eng.* 8887. doi: 10.1117/12.2029165

Kim, M. S., Daughtry, C. S. T., Chappelle, E. W., Mcmurtrey, J. E., and Walthall, C. L. T. (1994). use of high spectral resolution bands for estimating absorbed photosynthetically active radiation. In: *ISPRS Sixth International Colloquium on Physical Measurements and Signatures in Remote Sensing, Val d'Isère, France*. p. 17–21.

Lesk, C., Rowhani, P., and Ramankutty, N. (2016). Influence of extreme weather disasters on global crop production. *Nature*. 529, 84–87. doi: 10.1038/nature16467

Mondal, S., Singh, R. P., Crossa, J., Huerta-Espino, J., and Joshi, A. K. (2013). Earliness in wheat: A key to adaptation under terminal and continual high temperature stress in south Asia. *Field Crops Res*. 151, 19–26. doi: 10.1016/j.fcr.2013.06.015

Muhammad, H., Yang, M., Awais, R., Jin, X., Xia, X., and Xiao, Y. (2018). Time-series multispectral indices from unmanned aerial vehicle imagery reveal senescence rate in bread wheat. *Rem. Sens.* 10, 809. doi: 10.3390/rs10060809

Netto, A. T., Campostrini, E., Oliveira, J. G. D., and Bressan-Smith, R. E. P. (2005). pigments, nitrogen, chlorophyll a fluorescence and SPAD-502 readings in coffee leaves. *entia Horticult.* 104, 199–209. doi: 10.1016/j.scienta.2004.08.013

Pelckmans, K., Suykens, J., and De, B. (2005). *A Convex Approach to Learning the Ridge Based on CV.*

Roessner, U. (2012). Drought responses of leaf tissues from wheat cultivars of differing drought tolerance at the metabolite level. *Mol. Plant*, 5, 418–429. doi: 10.1093/mp/ssr114

Rondeaux, G., Steven, M., and Baret, F. (1996). Optimization of soil-adjusted vegetation indices. *Remote Sensing of Environment*, 55, 95–107. doi: 10.1016/0034-4257(95)00186-7

Sampson, P. H., Zarco-Tejada, P. J., Mohammed, G. H., Miller, J. R., and Noland, T. L. (2003). Hyperspectral remote sensing of forest condition: estimation of chlorophyll content in tolerant

hardwoods. *For. Sci.* 49, 381–391. doi: 10.1046/j.1439-0329.2003.00323.x

Schnell, J. A. (1974). *Monitoring the Vernal Advancement and Retrogradation (Greenwave Effect) of Natural Vegetation*. Nasa/gsfct Type Final Report.

Shafiee, S., Lied, L. M., Burud, I., Dieseth, J. A., and Lillemo, M. (2021). Sequential forward selection and support vector regression in comparison to LASSO regression for spring wheat yield prediction based on UAV imagery. *Comput. Electron. Agricul.* 183, 106036. doi: 10.1016/j.compag.2021.106036

Shao, H. B., Liang, Z. S., and Shao, M. A. O. (2006). regulation of 10 wheat (*Triticum aestivum L.*) genotypes at soil water deficits. *Coll. Surf B Biointerf.* 47, 132–139. doi: 10.1016/j.colsurfb.2005.11.028

Shestakova, E., Eroshenko, F., Storchak, I., Oganyan, L., and Chernova, I. (2020). Influence of various elements of cultivation technology on the chlorophyll content in winter wheat plants and its yield. *Agrarian Bull.* 196, 27–37. doi: 10.32417/1997-4868-2020-196-5-27-37

Sun, S., Yang, X., Lin, X., Zhao, J., and Liu, Z, Zhang T et al. (2019). Seasonal variability in potential and actual yields of winter wheat in china. *Field Crops Res.* 240, 1–11. doi: 10.1016/j.fcr.2019.05.016

Sun, B., Wang, C., Yang, C., Xu, B., and Zhang, J. (2021). Retrieval of rapeseed leaf area index using the PROSAIL model with canopy coverage derived from UAV images as a correction parameter. *Int. J. Appl. Earth Observ. Geoinf.* 102, 102373. doi: 10.1016/j.jag.2021.102373

Taghvaeian, S., Chávez, J. L., and Hansen, N. C. (2012). Infrared thermometry to estimate crop water stress index and water use of irrigated maize in Northeastern Colorado. *Rem. Sens.* 4, 3619–3637. doi: 10.3390/rs4113619

Telmo, A. (2017). Hruka Joná, Pádua Luís, Bessa José, Emanuel P, Raul M. Hy|UAV-based sensors, data processing and applications for agriculture and forestry. *Rem. Sens.* 9, 1110. doi: 10.3390/rs9111110

Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: a retrospective. *J. Royal Stat. Soc. Ser. B (Stat. Methodol.)*, 73, 267–288. doi: 10.1111/j.1467-9868.2011.00771.x

Torres-Sánchez, J., Peña, J. M., De-Castro, A. I., and López-Granados, F. (2014). Multi-temporal mapping of the vegetation fraction in early-season wheat fields using images from UAV. *Comput. Electron. Agricult.*103, 104–113. doi: 10.1016/j.compag.2014.02.009

Verger, A., Vigneau, N., Chéron, C., Gilliot, J. M., and Baret, F. (2014). area index from an unmanned aerial system over wheat and rapeseed crops. *Rem. Sens. Environ.* 152, 654–664. doi: 10.1016/j.rse.2014.06.006

Verma, V., Foulkes, M. J., Sylvester-Bradley, R., Caligari, P. D. S., and Snape, J. W. (2004). Mapping quantitative trait loci for flag leaf senescence as a yield determinant in winter wheat under optimal and drought-stressed environments. *Euphytica.* 135, 255–263. doi: 10.1023/B:EUPH.0000013255.31618.14

Wagner, H. D. T. (1996). Residual stress in composites with anisotropic interphases. *Phys. Rev. B Condens. Matter.* 53, 5055. doi: 10.1103/PhysRevB.53.5055

Wang, C., Shen, S., Wang, R., and Zhao, H. (2012). Impact of soil moisture on the winter wheat SPAD readings in semi-humid climate zone. *J Arid Land Resour Environ*. 26, 137–141. doi: 10.13448/j.cnki.jalre.2012.12.010

Witten, I. H., Frank, E., Hall, M. A., and Pal, C. J. (2016). *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann: Burlington, MA, USA.

Yang, Q., Shi, L., Han, J., Zha, Y., and Zhu, P. (2019). Deep convolutional neural networks for rice grain yield estimation at the ripening stage using UAV-based remotely sensed images. *Field Crops Res.* 235, 142–153. doi: 10.1016/j.fcr.2019.02.022

Zhang, J., Cheng, T., Guo, W., Xu, X., and Ma, X. (2021). Leaf area index estimation model for UAV image hyperspectral data based on wavelength variable selection and machine learning methods. *Plant Meth.* 17:49. doi: 10.1186/s13007-021-00750-5

Zhang, J., Fengler, K., Van Hemert, J. L., Gupta, R., Mongar, N., and Sun, J. (2019a). Identification and characterization of a novel stay-green QTL that increases yield in maize. *Plant Biotechnol. J.* 17, 2272–2285. doi: 10.1111/pbi.13139

Zhang, S., Zhao, G., Lang, K., Su, B., Zhang, H., and Integrated, S. (2019b). Unmanned Aerial Vehicle (UAV) and ground inversion of the SPAD of winter wheat in the reviving stage. *Sensors*, 19, 1485. doi: 10.3390/s19071485

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

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

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 08 June 2022
doi: 10.3389/fpls.2022.918559

The *Pm5e* Gene Has No Negative Effect on Wheat Agronomic Performance: Evidence From Newly Established Near-Isogenic Lines

Dan Qiu[1†], Jiang Huang[2†], Guanghao Guo[3,4], Jinghuang Hu[1], Yahui Li[1,5], Hongjun Zhang[1], Hongwei Liu[1], Li Yang[1], Yang Zhou[1], Benzhou Yang[6], Yudan Zhang[6], Zhiyong Liu[3,4] and Hongjie Li[1]**

[1] The National Engineering Laboratory of Crop Molecular Breeding, Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, China, [2] School of Intelligent Medicine and Biotechnology, Guilin Medical University, Guilin, China, [3] State Key Laboratory of Plant Cell and Chromosome Engineering, Institute of Genetics and Developmental Biology, The Innovative Academy of Seed Design, Chinese Academy of Sciences, Beijing, China, [4] College of Advanced Agricultural Science, University of Chinese Academy of Sciences, Beijing, China, [5] College of Life and Environmental Sciences, Minzu University of China, Beijing, China, [6] Institute of Wheat Sciences, Jining Academy of Agricultural Sciences, Jining, China

Wheat genotypes resistant to powdery mildew (*Blumeria graminis* f. sp. *tritici*, *Bgt*) provide a sustainable means for disease control. We developed a pair of near-isogenic lines H962R and H962S with contrasting reactions to powdery mildew from a residue heterozygous line. H962R was resistant to 127 out of the 136 *Bgt* isolates collected from the major wheat-producing regions of China and showed a similar virulence/avirulence pattern as Fuzhuang 30, Xiaobaidong, and Hongquanmang carrying resistance allele of *Pm5e*, but H962S was resistant to none of them. A dominant gene was responsible for the powdery mildew resistance of H962R as revealed by the genetic analysis using segregating populations derived from a cross between H962R and H962S. Molecular marker analysis detected a resistance locus, designated *PmH962*, on a genetic interval of the chromosome arm 7BL where *Pm5e* resides. This locus was co-segregated with the functional marker of *Pm5e*. The PCR-based sequence alignment of *Pm5e* demonstrated that H962R had an identical sequence as Fuzhuang 30 (haplotype HapGA), and H962S possessed the same sequence as the powdery mildew susceptible cultivar Kenong 199. The genomic compositions of lines H962R and H962S were highly comparable as evidenced by only a small percentage of SNP variations detected by the 16K Genotyping by Target Sequencing (GBTS) SNP array and the 90K Illumina iSelect Wheat SNP array. The two lines performed similarly in the yield-related and plant growth traits investigated, except for greater kernel weight in H962R than in H962S. This indicates that *Pm5e* has no deleterious effect and can be served as an excellent disease resistance gene in wheat breeding.

Keywords: ***Triticum aestivum*, *Blumeria graminis* f. sp. *tritici*, *Pm5e*, resistance gene, agronomic traits**

OPEN ACCESS

Edited by:
Pengtao Ma,
Yantai University, China

Reviewed by:
Guohao Han,
Center for Agricultural Resources Research, Institute of Genetics and Developmental Biology (CAS), China
Cheng Liu,
Shandong Academy of Agricultural Sciences, China

***Correspondence:**
Hongjie Li
lihongjie@caas.cn
Zhiyong Liu
zyliu@genetics.ac.cn

† These authors have contributed equally to this work

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

Received: *12 April 2022*
Accepted: *26 April 2022*
Published: *08 June 2022*

Citation:
Qiu D, Huang J, Guo G, Hu J, Li Y, Zhang H, Liu H, Yang L, Zhou Y, Yang B, Zhang Y, Liu Z and Li H (2022) The Pm5e Gene Has No Negative Effect on Wheat Agronomic Performance: Evidence From Newly Established Near-Isogenic Lines. Front. Plant Sci. 13:918559. doi: 10.3389/fpls.2022.918559

INTRODUCTION

Wheat (*Triticum aestivum* L.) is exposed to various pathogen infections, among which *Blumeria graminis* f. sp. *tritici* (*Bgt*), the causal agent of wheat powdery mildew, is of global importance causing an estimated yield loss of below 10% in average and could be up to 30–40% when occurred severely (Singh et al., 2016; Savary et al., 2019; Bhavani et al., 2021). This fungal disease has been epidemic in China since the 1970s and currently occurs in an area of ca. 6–8 million hectares across both winter and spring wheat regions. In recent years, climate change enhances the spread of powdery mildew throughout the wheat fields across the country (Tang et al., 2017; Zhang et al., 2017).

The economic importance of powdery mildew promotes extensive research toward its sustainable management. The obligate biotrophic nature of the pathogen makes it survive only in alive up-ground tissues of plants such as leaves, stems, and spikes. So, it is feasible to restrict the multiplication and spread of the fungus with host resistance and, in turn, control epidemics of the disease. After several decades of efforts on gene mining research, more than 100 powdery mildew resistance genes or alleles with designation symbols of *Pm1* to *Pm68* and many others with temporal designations have been documented. Most of these genetic loci confer race-specific or all-stage resistance. Also, many quantitative trait loci (QTL), which are non-race specific and effective at the adult plant stage only, have been characterized (Kang et al., 2020).

Multiple alleles that occur in different wheat genotypes have been detected in several powdery mildew resistance loci. For example, locus *Pm5* has five alleles designated *Pm5a* through *Pm5e*, all of which are located on the long arm of chromosome 7B with a recessive or semi-dominant mode of inheritance. *Pm5a* was introgressed from Yaroslav Emmer (*Triticum dicoccum* Schübl) into an American wheat cultivar Hope (Lebsock and Briggle, 1974). Hsam et al. (2001) reported the characterization of *Pm5b*, *Pm5c*, and *Pm5d*. Also, they found the presence of *Pm5a* in several other wheat cultivars from different countries across Europe, North America, and Asia in addition to Hope. *Pm5b* in a series of German wheat cultivars Ibis, Aquila, Flanders, and Rothwell Perdix may originate from wheat line DHE516 from Hindu Kush Mountains (Bennett, 1984; Hsam et al., 2001). Kolandi, an accession of *Triticum sphaerococcum* var. *rotundatum* Perc., carries allele *Pm5c* (Hsam et al., 2001). *Pm5d* was transferred from a Chinese wheat landrace CI 10904 (PI 83395) into Sweden spring wheat cultivars Prins and Starke resulting in the production of lines IGV 1-455 and IGV 2-556 (Leijerstam, 1972; Hsam et al., 2001). *Pm5e* was originally detected in Fuzhuang 30 (Huang et al., 2003). Several alleles of *Pm5e* were identified in different Chinese wheat landraces. *Pm5e* encodes a nucleotide-binding domain leucine-rich-repeat (NLR) containing protein, and the functional markers were developed to facilitate its application in wheat breeding (Xie et al., 2020).

Despite powdery mildew resistance genes identified with clear positions on chromosomes and traced with molecular markers, not all of them are widely used in breeding. Breeders prefer to use the resistance genes with promising agronomic traits and no yield penalty. Any disease resistance genes must be accompanied by proper yield, quality, and agronomic traits, and even other disease and stress resistance, to ensure its successful application in breeding programs. Linkage and competitive drags are the major factors that determine the deployment of powdery mildew resistance genes in wheat breeding and production (Summers and Brown, 2013). Usually, pre-breeding steps are necessary to eliminate the deleterious genes that may link to the disease resistance genes of interest.

We developed a pair of near-isogenic lines with contrasting responses to powdery mildew from residue heterozygous lines. Line H962R was resistant to most *Bgt* isolates collected from China, in contrast to the susceptible line H962S. These lines provide a useful source of resistance in breeding, disclosure of disease resistance mechanism, and evaluation of the yield penalty of the disease resistance gene. This study was initiated to characterize their phenotypic reactions, genetic compositions, and genetic control of the disease resistance for these lines.

MATERIALS AND METHODS

Plant Materials

H962 is an F_8 line selected from a cross between wheat-*Thinopyrum ponticum* partial amphiploid Xiaoyan 693 ($2n = 56$) [pedigree: Xiaoyan 2/(Lin 7/*Th. ponticum*) F_1] and common wheat line Ji87050 (pedigree: 72180/Jimai 30) (Li et al., 2003). This line was found to be heterozygous in the response to powdery mildew, and its residue heterozygous lines H962R (Resistant) and H962S (Susceptible) with distinct responses to powdery mildew were crossed to produce F_1, F_2 (461 plants), $F_{2:3}$ (299 families), and $F_{2:9}$ recombinant inbred line (RIL) (187 lines) populations for genetic analysis and molecular mapping of the powdery mildew resistance gene. Wheat genotypes with known powdery mildew resistance genes in the *Pm5* locus on the chromosome arm 7BL, including Hope (*Pm5a*), CI 14125 (*Pm5*c), Fuzhuang 30 (*Pm5e*), Hongquanmang (*PmH*), and Xiaobaidong (*mlxbd*), as well as another 17 wheat differential genotypes, were used to compare the reactions to different *Bgt* isolates from various wheat fields in China. Susceptible cultivars Zhongzuo 9504 and Chancellor were used to increase *Bgt* isolates and served as the susceptible control in all the powdery mildew tests. Selected Chinese Spring (CS) homoeologous group 7 nullisomic-tetrasomic lines (CSN7A-T7D, CSN7B-T7A, CSN7B-T7D, CSN7D-T7A, and CSN7D-T7), ditelosomic lines (CSDt7BS and CSDt7BL), and deletion lines (7BS-1, 7BL-6, 7BL-7, and 7BL-10) were used to determine the chromosomes on which the molecular markers linked to the powdery mildew resistance gene in H962R reside.

Agronomic Trait Evaluation

A field trial was conducted in a farm of the Institute of Crop Sciences, Chinese Academy of Agricultural Sciences (CAAS), Beijing (116°249′ E, 40°179′ N) during the cropping seasons of 2010–2011. Lines H962R and H962S were arranged in a randomized complete block design with three replicates. Each plot consisted of three rows of 2 m in length and 0.3 m between rows. Dates from sowing to heading, anthesis, and maturity

were recorded on a plot basis. Ten plants from each plot were randomly harvested to investigate plant height, spike, and kernel traits, such as spike length, number of spikelets per spike, number of kernels per spike, thousand-kernel weight, glume, and kernel color (**Supplementary Table 1**).

Powdery Mildew Assessments

The *Bgt* isolates were originally collected from wheat fields of Shandong, Hebei, Henan, Beijing, and other provinces (**Supplementary Table 2**). They were subjected to at least three rounds of single-pustule culture prior to the preparation of inocula for disease assessments. A total of 136 isolates were applied in the seedling tests. Two-leaf stage seedlings grown in plastic pots (5 × 5 × 5 cm in dimension) were dusted with freshly increased conidiospores, incubated in dew plastic bags for 24 h in the dark, and allowed for disease symptom development for 2 weeks in a greenhouse set at 20 ± 2°C, 80% relative humidity, and a photoperiod of 14 h light/10 h dark. The phenotypic categorization of wheat plants was carried out based on the infection types (ITs) of the primary leaves rated on a 0–4 scale. Wheat plants were regarded to have a resistant phenotype when their ITs were 0 (immune), 0; (hypersensitive reaction), 1 (highly resistant), and 2 (moderately resistant), or to have a susceptible phenotype when their ITs were 3 (moderately susceptible) and 4 (highly susceptible) (Liu et al., 1999). At least two parallel tests for each wheat genotype were tested.

Analysis of Genomic Compositions for Lines H962R and H962S

Total genomic DNA of H962R and H962S was extracted with a Hi-DNA secure Plant Kit (DP350, Tiangen Biotech Co., Beijing, China). After examination of quality and concentration using an agarose gel electrophoresis and NanoDrop One (Thermo Fisher Scientific, Waltham, MA, United States), DNA samples of the two lines were genotyped with the 16K Genotyping by the Target Sequencing (GBTS) SNP array and the 90K Illumina iSelect Wheat SNP array (MolBreeding Biotechnology Co., Ltd., Shijiazhuang, China). Raw reads generated by the 16K GBTS SNP array were filtered using fastp software (version 0.20.0, parameter: -n 10 -q 20 -u 40). Clean reads were aligned to the Chinese Spring genome sequence using Burrows-Wheeler Aligner (BWA) software (Li and Durbin, 2009). The UnifiedGenotyper and VariantFiltration modules in GATK software were used for mutation detection and data filtering. Software ANNOVAR was used for functional annotation of the gene mutations detected. For the 90K SNP array (Wang et al., 2014), SNP calling was carried out in the Illumina iScan reader with the Genome Studio software (version 2011.1; Illumina Inc., San Diego, CA, United States). The genotypic clusters for each SNP were determined based on the data for the two lines using the manual option of Genome Studio software. The diagrams of chromosomal compositions for the two lines were drawn by the self-written program in the R package (MolBreeding Biotechnology Co., Ltd., Shijiazhuang, China).

As Xiaoyan 693 and 72180 have *Th. ponticum* in their pedigree (Zhang et al., 1989; Li et al., 2003), the presence of *Th. ponticum* chromatin in the H962 lines, if any, was detected using genomic *in situ* hybridization (GISH) and the *Thinopyrum* genome-specific marker analyses as described previously (Li et al., 2004). The mitotic chromosomes were prepared from root tip cells for GISH analysis. The total genomic DNA from *Th. ponticum* was labeled with biotin-14-dATP as a probe, and the total genomic DNA from Chinese Spring wheat was sheared as a blocker. The hybridization signals were visualized under a fluorescent microscope (Carl Zeiss, Oberkochem, Germany) after staining with fluorescein isothiocyanate (FITC)-avidin DN and biotinylated goat anti-avidin D (Vector Laboratories, Inc., Burlingame, CA, United States). Propidium iodide was used to counterstain the wheat chromosomes. A pair of primers, 2P1 and 2P2, which was derived from a repetitive sequence and can specifically detect chromatin of genus *Thonopyrum* (Wang and Wei, 1995), was used to detect *Th. ponticum* genome-specific DNA. Since the parental cultivar Jimai 30 is a T1BL·1RS translocation cultivar (Wei et al., 2013), markers ω-sec-P1 and ω-sec-P2 specific for the ω-seclin gene on the chromosome arm 1RS of rye (*Secale cereal* L.) were used to detect this translocation (Froidmont, 1998). The amplification products in a Biometra T3000 Thermocycler (Applied Biosystems, New York, NY, United States) were visualized on a 1.5% agarose gel.

Molecular Marker Analysis

Polymorphisms of the WMC, MVG, GWM, CFD, and BARC series of SSR markers and the EST-based markers on chromosome arm 7BL were determined using H962R and H962S. To develop more markers, RNA sequencing was performed on an Illumina Hi2500 platform (BGI, Shenzhen, China) with leaves of H962R and H962S at 0, 4, 8, 16, 24, 48, and 72 h post-inoculation (hpi) with *Bgt* isolate E09. Raw reads were purified by trimming adapters and removing low-quality sequences. Adaptor-trimmed reads comprising 40,349 unigene sequences with a total of 31,671,110 bases were mapped to *T. aestivum* genome sequence (International Wheat Genome Sequencing Consortium [IWGSC], 2018). The number of clean reads in each library was normalized to the Transcripts Per kilobase of exon model per Million (TPM) mapped reads to obtain the normalized gene expression level. The differentially expressed genes (DEGs) between H962R and H962S were determined with the parameters of false discovery rate (FDR) $\leq$ 0.001 and threshold absolute $\log_2$-fold change $\geq$ 1 for the sequence counts across the libraries using DESeq2.[1] Sequences of the DEGs were used to design primer pairs using online website Primer3Plus.[2]

Ten RILs from cross H962R × H962S with contrasting responses to each powdery mildew were separately pooled to construct the resistant and susceptible DNA bulks for bulked segregant analysis (BSA). Each reaction mixture (10 μl) consisted of 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 250 μM each of the dNTP, forward and reverse primer 0.2 μM

[1]https://bioconductor.org/packages/DESeq2/

[2]https://www.primer3plus.com/

each, 0.5 U *Taq* DNA polymerase, and 50 ng DNA template. Amplification of DNA was programmed at 94°C for 5 min, 38 cycles of 94°C for 35 s, 50–62°C for 35 s, and 72°C for 1 min, and a final extension at 72°C. Products amplified were separated on 8% non-denatured polyacrylamide gels with a 29:1 of acrylamide and bisacrylamide and visualized by silver staining. Polymorphic markers were used to genotype the RIL population to determine the linkage with the target gene. The KASP marker Pm5e-KASP, which was designed based on the sequence of *Pm5e* (Xie et al., 2020), was also used to genotype the RIL population to examine the linkage between *Pm5e* and the powdery mildew resistance in H962R.

Alignment of *Pm5e* Sequences

Total genomic DNA was extracted using a Hi-DNA secure Plant Kit (DP350, Tiangen Biotech Co., China) for amplifying the *Pm5e* sequences with primers as listed in **Supplementary Table 3**. The resulting sequences from H962R and H962S were aligned to the *Pm5* allelic sequences from Fuzhuang 30 (*Pm5e*, HapGA), Baimangduomai (*PmBMDM*, HapGG), Chinese Spring (*Pm5-CS*, HapCG), Non-gda 015 (*Pm5-ND*, HapCG), Hope (*Pm5a*), and Mission (*Pm5b*), as well as the wheat 10 + genomes (Walkowiak et al., 2020), Kenong 199 (*Pm5-KN199*), and Fielder (Sato et al., 2021). Multiple sequence alignment was performed using Snapgene (version 4.1.9),[3] and the TriticeaeGeneTrebe (TGT) website was used for homologous gene search, structure alignment, and clustering.[4]

Statistical Analysis and Genetic Map Construction

All statistical analyses were performed using the SAS package (Version 9.2, SAS Institute, Cary, NC, United States). ANOVA for the agronomic traits was carried out to test the significance of the differences between H962R and H962S using the least significant difference (LSD) at $P < 0.05$. The IT values were recorded as a data matrix composed of 1 (resistant) and 2 (susceptible). A dendrogram was drawn using the online website HIPLOT,[5] to compare the relationship between the target gene and the known powdery mildew resistance genes. The chi-squared test (χ^2) was performed to determine the goodness of fit for the observed data from the expected segregation ratios of the segregating populations. The linkage relationship between the polymorphic markers identified by BSA and the target resistance gene was determined with Mapmaker Version 3.0 software (Lincoln et al., 1993). The genetic distance between markers and the target gene was calculated using the Kosambi function (threshold LOD $\geq$ 3.0), and the maximum linkage distance was set at 50.0 (Kosambi, 1943). The genetic linkage map was constructed by Mapdraw (Version 2.1) (Liu and Meng, 2003).

[3]https://www.snapgene.com/

[4]http://wheat.cau.edu.cn/TGT/

[5]https://hiplot.com.cn

RESULTS

Development of Isogenic Lines H962R and H962S

Wheat-*Th. ponticum* partial amphiploid Xiaoyan 603 ($2n = 56$) was crossed with wheat line Ji87050 in 1989 and self-pollinated repeatedly to obtain line H962 from the F_8 progenies. This line was segregated for powdery mildew reactions when inoculated with *Bgt* isolate E03, and the resistant and susceptible plants were subjected to nine generations of self-pollination. Two near-isogenic lines, designated H962R (Resistant) and H962S (Susceptible), were finally obtained for characterization of their genomic composition and genetic control of disease resistance.

Characterization of Genomic Compositions for H962R and H962S

Two versions of the SNP array were applied to characterize the genomic compositions of lines H962R and H962S. The majority of SNP loci (14,315 out of 14,387, 99.5%) were common between the two lines generated by the 16K GBTS SNP array. Seventy-two (0.5%) polymorphic SNP loci were scattered on different wheat chromosomes with similar numbers in the A (22), B (26), and D (24) genomes (**Figure 1A**). Similarly, the 90K Illumina iSelect Wheat SNP array identified 136 (0.2%) polymorphic loci among 79,955 SNP loci detected (**Figure 1B**). These results demonstrate the high level of genomic similarity between H962R and H962S.

The GISH with the *Th. ponticum* genomic DNA as a probe and Chinese Spring genomic DNA as a blocker detects a complete set of 42 wheat genome chromosomes without any *Th. ponticum* chromosomes or chromosome segments in H962R and H962S (data not shown). Meanwhile, *Thinopyrum* genome-specific primers 2P1 and 2P2 did not amplify any *Thinopyrum* chromatin in the H962 siblings (**Supplementary Figure 1**). However, the wheat-rye translocated chromosome T1BL·1RS was detected in both H962R and H962S *via* molecular markers specific for the rye chromosome arm 1RS (**Supplementary Figure 2**).

Seedling Reaction Patterns of H962R and H962S to *Bgt* Isolates

An array of 103 *Bgt* isolates collected from different provinces during 2011–2016 were used to compare the seedling reaction patterns between the two H962 lines. These isolates showed different virulence patterns against 22 wheat accessions carrying different known powdery mildew resistance genes (**Supplementary Table 2**). H962R was resistant to 94 *Bgt* isolates, while H962S was as susceptible to all of them as that of the susceptible controls Zhongzuo 9504 and Chancellor. The virulence/avirulence patterns of H962R to these *Bgt* isolates were most similar as those of Fuzhuang 30 (*Pm5e*), Xiaobaidong (*mlxbd*), and Hongquanmang (*PmH*), which carry the *Pm5e* locus (**Figure 2** and **Supplementary Table 2**). The reactions of H962R differed from these *Pm5e*-carrying genotypes in only a few isolates. Another 33 isolates recently collected during 2018–2020 were also tested, and all of them were avirulent on H962R but virulent on H962S (**Supplementary Table 4**).

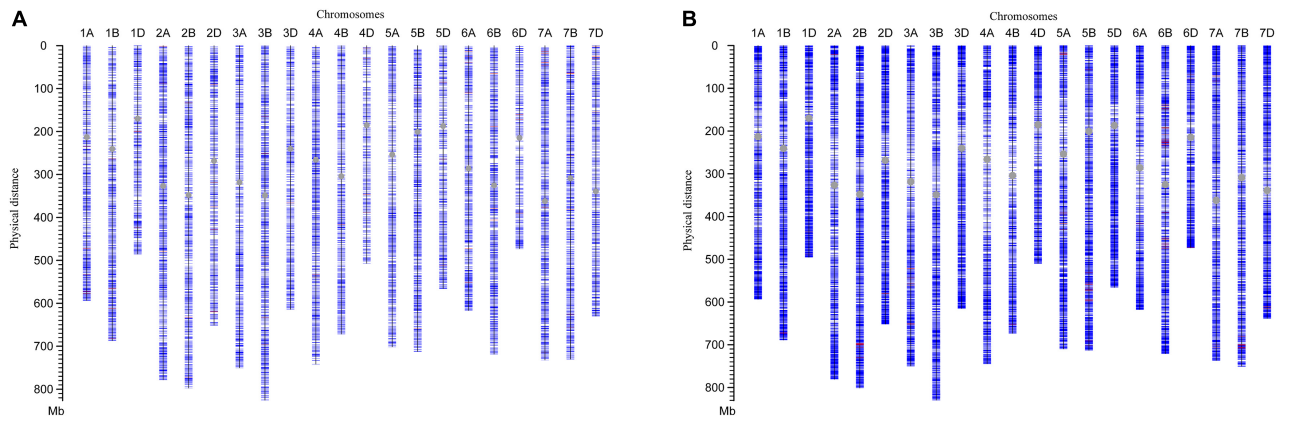


FIGURE 1 | Diagram of chromosomal compositions for H962R and H962S by 16K GBTS SNP array **(A)** and 90K Illumina iSelect Wheat SNP array **(B)**. The blue and red lines represent the same and different SNPs between H962R and H962S, respectively. The centromeres are indicated by the gray dots.

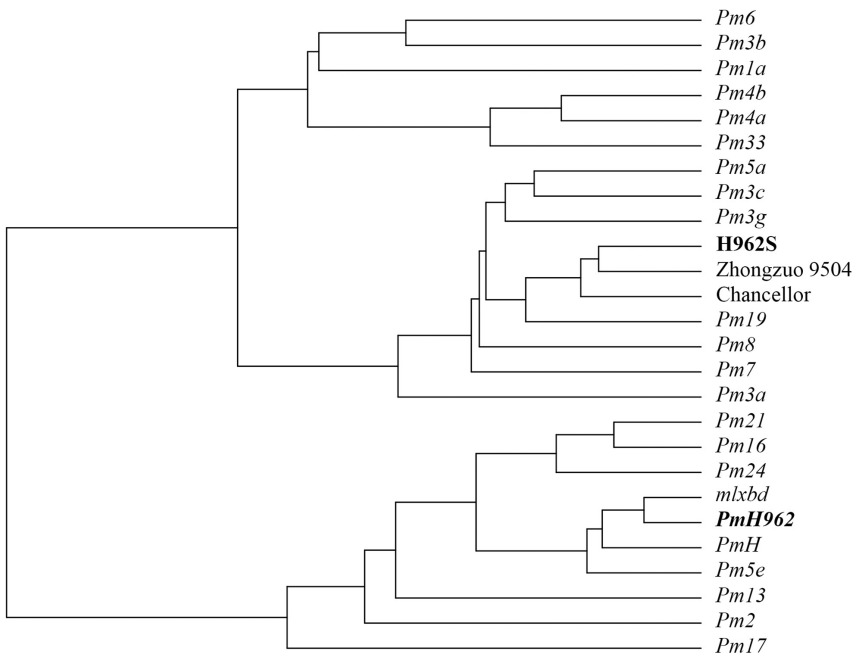


FIGURE 2 | A dendrogram based on reactions of *PmH962* and differential wheat cultivars/lines possessing known powdery mildew resistance genes against 103 *Blumeria graminis* f. sp. *tritici* isolates.

Inheritance of the Powdery Mildew Resistance in H962R

When challenged with *Bgt* isolate E09 at the seedling stage, H962R (IT 0) was distinct from H962S (IT 4) in the response to powdery mildew (**Figure 3**). So, this isolate was used to determine the inheritance of powdery mildew resistance using the populations developed from the cross between H962R and H962S. The phenotypic segregation patterns for the resistant and susceptible plants (or lines) in the F_1 (1:0), F_2 (3:1), $F_{2:3}$ (1:2:1), and $F_{2:9}$ (1:1) populations clearly demonstrated that a single dominant gene, designated *PmH962*, governs the resistance to *Bgt* isolate E09 in H962R (**Table 1**).

Identification of SSR and EST-SSR Markers Linked to *PmH962*

Due to high similarity in the genomic compositions between the two lines, only about 6% of 1,380 SSR and EST markers screened were polymorphic between H962R and H962S. Two markers *Xgwm783* and *XTC17* (**Supplementary Table 3**) anchored on chromosome 7BL were further polymorphic between the resistant and susceptible DNA bulks. After genotyping the $F_{2:3}$ population, they were linked to *PmH962* with genetic distances of 8.3 and 6.4 cM, respectively. The examination of additional 180 EST-derived markers on chromosome 7BL identified another *PmH962*-linked marker *Xics-X27* (**Supplementary Table 3**).

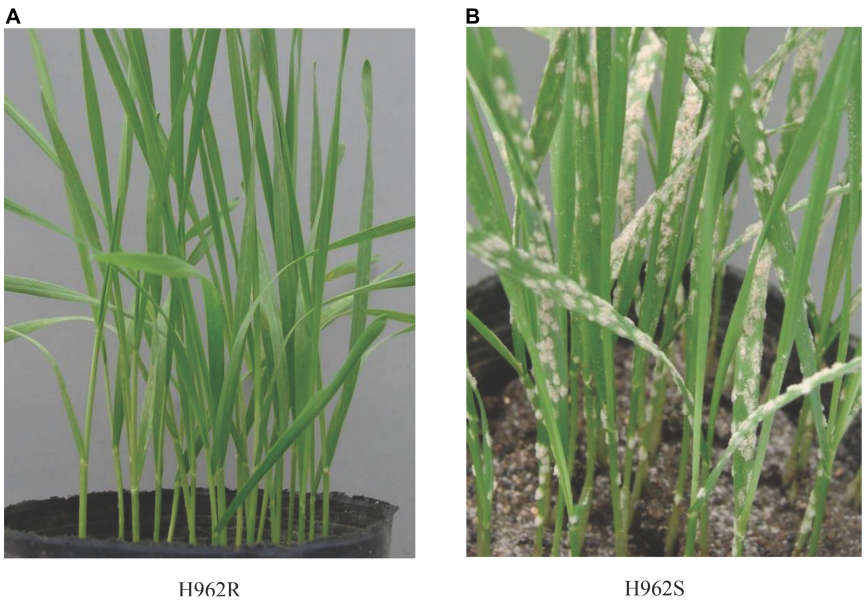


FIGURE 3 | Seedling phenotypes of H962R **(A)** and H962S **(B)** to *Blumeria graminis* f. sp. *tritici* isolate A49.

TABLE 1 | Inheritance analysis of lines H962R and H962S and their F_1, F_2, $F_{2:3}$, and $F_{2:9}$ progenies for resistance to *Blumeria graminis* f. sp. *tritici* isolate E09.

Parents or cross	Generation[a]	Total number of plants/families	Observed ratio		Expected ratio	χ^2	*P*-value
			R	S			
H962R	P_R	28	28				
H962S	P_S	30		30			
H962R × H962S	F_1	20	20				
H962R × H962S	F_2	461	346	116	3:1	0.003	0.957
H962R × H962S	$F_{2:3}$	117	A:H:B = 32:57:28[b]		1:2:1	0.350	0.839
H962R × H962S	$F_{2:9}$	187	A:B = 97:90		1:1	0.262	0.609

[a] *P_R, resistant parent; P_S, susceptible parent.*
[b] *A, homozygous resistant; H, segregating (heterozygous resistant); B, homozygous susceptible.*

Development of Molecular Markers From Differentially Expressed Genes Generated by RNA-Seq Analysis

Due to the low polymorphism for the available SSR markers between lines H962R and H962S, RNA-Seq was applied to isolate E09-inoculated leaves of the two lines to identify the DEGs for developing more polymorphic molecular markers. The statistical results of the RNA-Seq are shown in **Supplementary Table 5**. The 155 top DEGs were selected and 12 of them were anchored on chromosome 7BL. Sequences of these DEGs were used to design 188 gene-specific primer pairs. Markers *Xics-NL2* and *Xics-X5*, specific for genes *TraesCS7B01G441600.1* and *TraesCS7B01G441700.1* (**Supplementary Table 3**), respectively, were polymorphic and linked to *PmH962* after examining the contrasting DNA bulks constructed with the RIL population. Physical mapping of markers *Xics-NL2* and *XTC17* with the Chinese Spring nulli-tetrasomic lines, ditelosomic lines, and deletion lines for the homoeologous group 7 chromosomes indicated that they were located on bin 7BL10-0.78-1.00 at the end of chromosome 7BL (**Figure 4**). Although the wheat-rye translocated chromosome T1BL·1RS carrying *Pm8* was present in both H962R and H962S, this translocation was not associated with the resistance of H962R to powdery mildew as evidenced by 1RS-specific marker genotyping the RIL population of cross H962R × H962S (data not shown).

Sequence Characteristics of *Pm5e* in H962R and H962S

Since marker *Xics-X5* was derived from gene *TraesCS7B01G441700*, which was annotated as the *NLR* gene, *Pm5e*, the functional marker of this cloned gene, Pm5e-KASP, was used to genotype the RIL population of cross H962R × H962S. The result confirmed the co-segregation of Pm5e-KASP with the powdery mildew resistance in this population (**Figure 5**).

The haplotypes of H962R and H962S at the *Pm5* locus were compared based on their sequences. The amplified sequence of *Pm5e* from H962R (3,948 bp) was identical

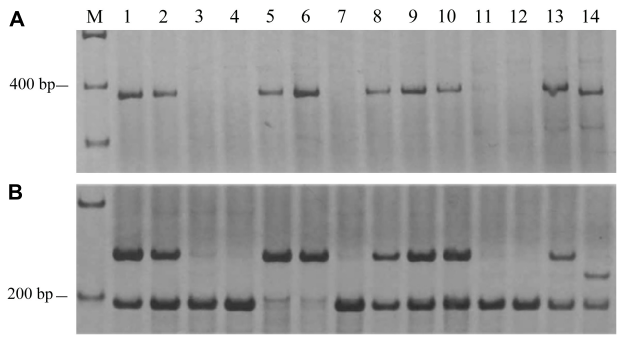


FIGURE 4 | The banding patterns of *PmH962*-linked markers *NL2* **(A)** and *XTC17* **(B)** in the Chinese Spring (CS) homoeologous group 7 nulli-tetrasomics, ditelosomics, and deletion lines. 1, CS; 2, CSN7A-T7D; 3, CSN7B-T7A; 4, CSN7B-T7D; 5, CSN7D-T7A; 6, CSN7D-T7B; 7, CSDt7BS; 8, CSDt7BL; 9, 7BS-1; 10, 7BL-6; 11, 7BL-7; 12, 7BL-10; 13, H962R; 14, H962S.

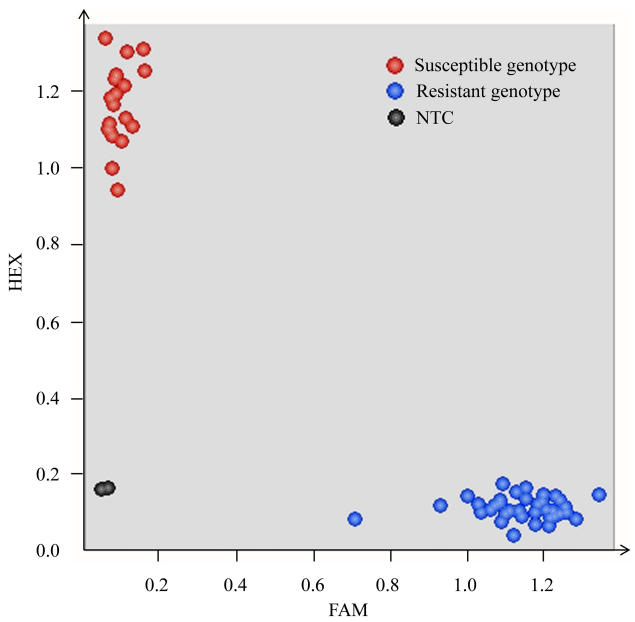


FIGURE 5 | Scatter plots of the RIL population from cross H962R × H962S for the Kompetitive allele-specific PCR (KASP) assay using the *Pm5e* functional marker Pm5e-KASP. The blue and red dots represent the lines with the disease-resistant and susceptible genotypes, respectively. The non-template control is indicated by the black dots.

with that from Fuzhuang 30, belonging to the resistant haplotype HapGA (**Figure 6**). It differed from the susceptible haplotypes HapCG from Chinese Spring, Hope (*Pm5a*), Julius, ArinaLrFor, Landmark, Norin61, SY Mattis, Mattis, Mace, Kariega, LongReach Lance, Claire, Robigus, and Weebill, as well as haplotype HapGG, represented by Mission (*Pm5b*). The *Pm5e* allelic sequence from H962S (3,980 bp) shared the same sequence as Kenong 199 and had 91% sequence identity compared with H962R with many insertions, deletions, and SNP variations.

Impacts of *Pm5e* on Agronomic Performance

Line H962R (51.30 ± 0.39 g) produced greater thousand-kernel weight than H962S (45.84 ± 0.62 g), with a significant difference of 5.46 ± 0.94 ($P < 0.05$). Examination of the RIL population of H962R × H962S revealed that the lines carrying *Pm5e* had a mean thousand-kernel weight of 35.46 ± 6.34 g with a range of 22.23–48.84 g, which was significantly greater than that of the lines without *Pm5e* (32.12 ± 6.34 g, with a range of 19.20–47.04 g) (**Figure 7**). A significant difference in kernel width but not in kernel length was observed between the two groups of RILs. Lines H962R and H962S did not differ in several other spike and kernel traits, e.g., number of spikelets and kernels per spike, as well as glume and kernel colors (**Supplementary Table 1**). The two near-isogenic lines were highly comparable in the plant traits, such as dates from sowing to heading, flowering, maturity, and plant height.

DISCUSSION

We developed a pair of near-isogenic lines, H962R and H962S, from the residue heterozygous line of H962, a progeny originating from a cross between wheat-*Th. ponticum* partial amphiploid Xiaoyan 693 and common wheat line Ji87050. H962R and H962S possess nearly identical genomic constitutions, as evidenced by the small proportions of SNP variations detected by the 16K GBTS and 90K iSelect SNP arrays. However, the two lines were distinct in their reactions to powdery mildew. Line H962R was resistant to most *Bgt* isolates collected in a wide range of wheat fields in China, while H962S was resistant to none of them. Genetic and molecular mapping identified a dominant gene *PmH962* on the chromosome arm 7BL that is responsible for its disease resistance. This gene was mapped to the genomic region where *Pm5e* resides and was co-segregated with the *Pm5e* functional marker Pm5e-KASP. H962R shared the same sequence of *Pm5e* and performed similar virulence/avirulence patterns to an array of *Bgt* isolates as Fuzhuang 30, the *Pm5e* carrier. H962S had an identical sequence as the susceptible cultivar Kenong 199 at the *Pm5* locus. Collectively, the powdery mildew resistance of H962R is conferred by gene *Pm5e*.

Pm5e was initially detected in Fuzhuang 30, a cultivar selected from Jinghui 30 that was developed from a cross between two Chinese wheat landraces Liquanheshangtou and Huaxianqisifeng in Shaanxi Province in the 1940s (Huang et al., 2003). Subsequent studies identified several genes allelic to *Pm5e* in wheat landraces, e.g., Hongquanmang (*PmH*) (Zhou et al., 2005), Hongyoumai (*pmHYM*) (Fu et al., 2017), Baiyouyantiao (*PmBYYT*) (Xu et al., 2018a), Shangeda (*PmSGD*) (Xu et al., 2018b), Dahongtou (*pmDHT*) (Qie et al., 2019), Youbailan (*pmYBL*) (Xu et al., 2020), Xiaobaidong (*mlxbd*) (Xue et al., 2009), and Duanganmang (*PmDGM*) (Wu et al., 2021). Most of these landraces were ever grown in the Guanzhong Plain, Shaanxi Province, or nearby regions (Xie et al., 2020). The parents of H962, 72180 and Xiaoyan 693, were also developed in Shaanxi Province (Zhang et al., 1989). Although both of them have *Th. ponticum* in their pedigrees, GISH and genome-specific PCR analyzes did

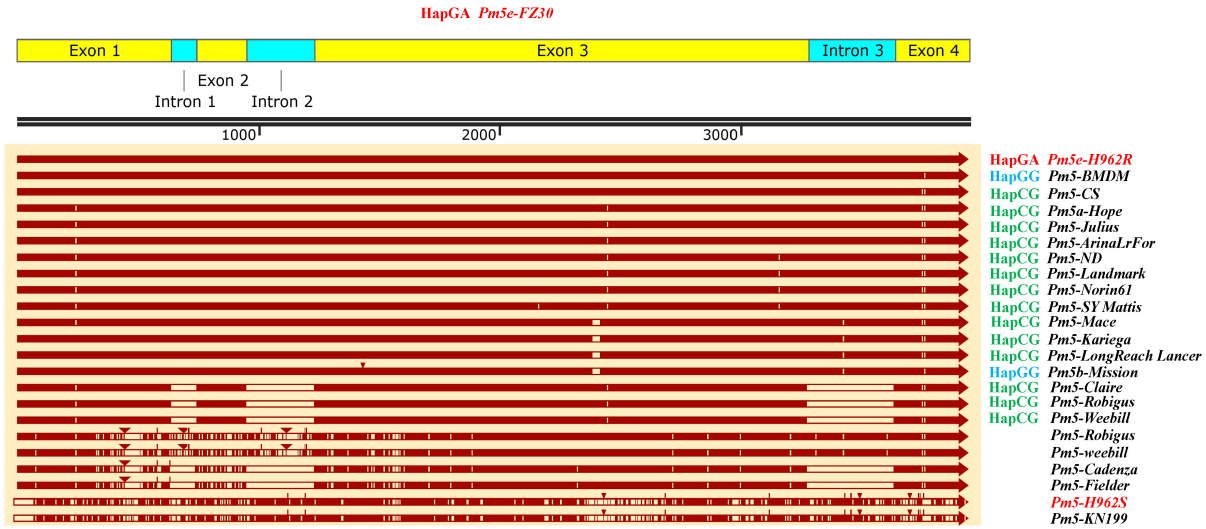


FIGURE 6 | Sequence alignment of *Pm5e* from different genotypes.

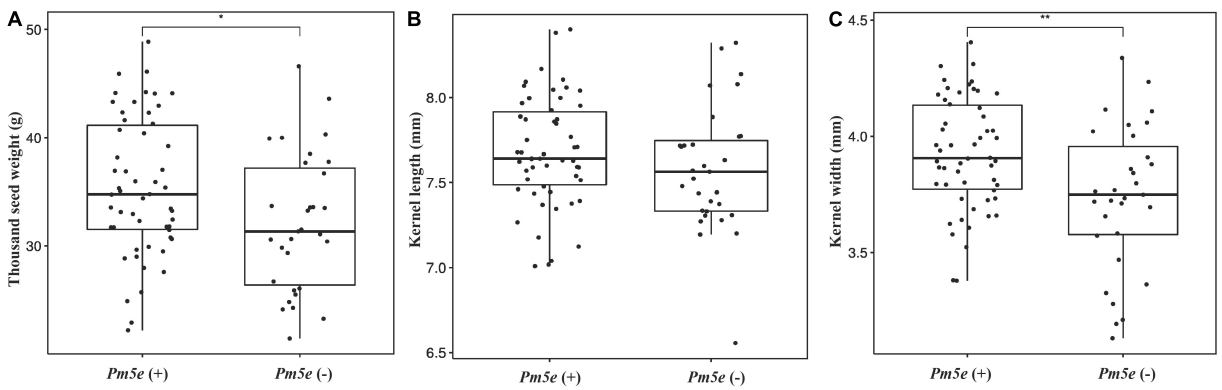


FIGURE 7 | Box plots of thousand-kernel weight **(A)**, kernel length **(B)**, and kernel width **(C)** for the lines with or without *Pm5e* from the cross H962R × H962S.

not detect any evidence that the H962 lines possessed any *Thinopyrum* chromatin. Therefore, it is reasonable to speculate that the origin of the *Pm5* allele in H962R is derived from common wheat grown in the region where the other *Pm5e*-carrying landraces originated.

Most alleles of the *Pm5* locus perform a recessive mode of inheritance, as was observed in the majority of wheat landraces known to carry the *Pm5e* alleles. Exceptions are cultivar Tangmai 4 and landrace Duanganmang with the dominant alleles. Sequence analysis confirms that both of them carry *Pm5e* (Xie et al., 2017, 2020; Wu et al., 2021). Genetic analysis of the powdery mildew resistance also revealed a dominant inheritance in line H962R, which also possesses *Pm5e*. The genomic region that flanks *Pm5e* contains several NLR genes (Xie et al., 2020). This indicates that there might exist uncharacterized genetic factors that coordinate the expression of *Pm5e* in different genetic backgrounds. Although the T1BL·1RS chromosome translocation was present in both H962R and H962S, *Pm8* on this chromosome appears to be unlikely to affect the phenotypes of powdery mildew resistance. More studies are needed to elucidate the different expression patterns of *Pm5e* in the H962 sib-lines and other *Pm5e* carriers.

A broad spectrum of disease resistance has been observed in most *Pm5e* carriers including H962R. These *Pm5e*-carrying wheat landraces had once been used in agriculture and did not have deleterious traits. For example, the original *Pm5e* carrier, Fuzhuang 30, and its parental cultivar Jinghui 30 were widely grown in Shaanxi Province (Xie et al., 2020). To date, these landraces are not grown due to their poorer productivity. Because the functional marker of *Pm5e* is available, it is easy to incorporate these *Pm5e* alleles into modern wheat breeding programs with the aid of molecular marker-assisted selection.

Success of powdery mildew resistance in breeding for resistant wheat cultivars depends on not only its disease resistance but also the backgrounds in which it resides. For example, *Pm16* failed to be deployed in commercial wheat cultivars due to a significant yield penalty (Summers and Brown, 2013). Gene *Pm5e* in line H962R did not link to deleterious genes for poor

agronomic traits, as revealed by the similar performances in most traits associated with growth and yield-related parameters observed. Moreover, the RILs of the H962R × H962S that carry *Pm5e* had greater kernel weight than those without *Pm5e* in the field without powdery mildew pressure, although we cannot determine whether *Pm5e* is linked to gene or QTL for kernel weight yet. This provides further evidence that *Pm5e* does not exert negative effects on wheat genotypes and has great potential in wheat breeding.

DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the National Genomics Data Center repository (https://bigd.big.ac.cn/gsa/browse/CRA006702), accession number: CRA006702.

AUTHOR CONTRIBUTIONS

HJL and ZL conceived and designed the research. DQ, JAH, GG, JHH, YL, HZ, HWL, LY, YZ, BY, and YDZ conducted the experiments and analyzed the data. HJL, DQ, and ZL wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The financial supports provided by the National Natural Science Foundation of China (Grant Nos. 32172001 and 31871621) and the Agricultural Science and Technology Innovation Program (Grant No. CAAS-ZDRW202002) are gratefully appreciated.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.918559/full#supplementary-material

Supplementary Figure 1 | Detection of *Th. ponticum* gene using primer 2P1 and 2P2. M: 100 bp DNA ladder; 1: Xiaoyan 693; 2: H962R; 3: H962S.

Supplementary Figure 2 | Detection of *Th. ponticum* gene using primer ω-sec-P1 and ω-sec-P2. M: 100 bp DNA ladder; 1: H962R; 2: H962S; 3 and 4: *Secale cereale*; 5: 1R-2; 6: 1RL.

REFERENCES

Bennett, F. G. A. (1984). Resistance to powdery mildew in wheat: a review of its use in agriculture and breeding programmes. *Plant Pathol.* 33, 279–300. doi: 10.1111/j.1365-3059.1984.tb01324.x

Bhavani, S., Singh, P. K., Qureshi, N., He, X., Biswal, A. K., Juliana, P., et al. (2021). "Globally important wheat diseases: status, challenges, breeding and genomic tools to enhance resistance durability," in *Genomic Designing for Biotic Stress Resistant Cereal Crops*, ed. C. Kole (Cham: Springer Nature), 59–128. doi: 10.1007/978-3-030-75879-0_2

Froidmont, D. (1998). A co-dominant marker for the 1BL/1RS wheat-rye translocation via mutiplex PCR. *J. Cereal Sci.* 27, 229–232. doi: 10.1006/jcrs.1998.0194

Fu, B. S., Zhang, Z. L., Zhang, Q. F., Wu, X. Y., Wu, J. Z., and Cai, S. B. (2017). Identification and mapping of a new powdery mildew resistance allele in the Chinese wheat landrace Hongyoumai. *Mol. Breed.* 37:133. doi: 10.1007/s11032-017-0728-3

Hsam, S. L. K., Huang, X. Q., and Zeller, F. J. (2001). Chromosomal location of genes for resistance to powdery mildew in common wheat (*Triticum aestivum* L. em Thell.) 6. Alleles at the Pm5 locus. *Theor. Appl. Genet.* 102, 127–133. doi: 10.1007/s001220050848

Huang, X. Q., Wang, L. X., Xu, M. X., and Röder, M. S. (2003). Microsatellite mapping of the powdery mildew resistance gene Pm5e in common wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 106, 858–865. doi: 10.1007/s00122-002-1146-3

International Wheat Genome Sequencing Consortium [IWGSC] (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361:eaar7191. doi: 10.1126/science.aar7191

Kang, Y. C., Zhou, M. X., Merry, A., and Barry, K. (2020). Mechanisms of powdery mildew resistance of wheat – a review of molecular breeding. *Plant Pathol.* 69, 601–617. doi: 10.1111/ppa.13166

Kosambi, D. D. (1943). The estimation of map distance from recombination values. *Ann. Eugen.* 12, 172–175. doi: 10.1111/j.1469-1809.1943.tb02321.x

Lebsock, K. L., and Briggle, L. W. (1974). Gene *Pm5* for resistance to *Erysiphe gramini* f. sp. *tritici* in Hope wheat. *Crop Sci.* 14, 561–563. doi: 10.2135/cropsci1974.0011183X001400040020x

Leijerstam, B. (1972). Studies in powdery mildew on wheat in Sweden. IV. Inventory of sources of resistance to the Scandinavian race population complex. *St. Växtsk. Meddn.* 15, 251–269.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with burrows–wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, H. J., Arterburn, M., Jones, S. S., and Murray, T. D. (2004). A new source of resistance to *Tapesia yallundae* associated with a homoeologous group 4 chromosome in *Thinopyrum ponticum*. *Phytopathology* 94, 932–937. doi: 10.1094/PHYTO.2004.94.9.932

Li, H. J., Chen, Q., Conner, R. L., Guo, B. H., Zhang, Y. M., Graf, R. J., et al. (2003). Molecular characterization of a wheat-*Thinopyrum ponticum* partial amphiploid and its derivatives for resistance to leaf rust. *Genome* 46, 906–913. doi: 10.1139/g03-053

Lincoln, S. E., Daly, M. J., and Lander, E. S. (1993). *Constructing Linkage Maps with MAPMAKER/Exp Version 3.0. A Tutorial Reference Manual*, 3rd Edn. Cambridge, MA: Whitehead Institute for Medical Research.

Liu, R. H., and Meng, J. L. (2003). MapDraw: A Microsoft excel macro for drawing genetic linkage maps based on given genetic linkage data. *Yi Chuan Hered.* 25, 317–321.

Liu, Z. Y., Sun, Q. X., Ni, Z. F., Yang, T. M., and McIntosh, R. A. (1999). Development of SCAR markers linked to the *Pm21* gene conferring resistance to powdery mildew in common wheat. *Plant Breed* 118, 215–219. doi: 10.1046/j.1439-0523.1999.118003215.x

Qie, Y. M., Sheng, Y., Xu, H. X., Jin, Y. L., Ma, F. F., Li, L. H., et al. (2019). Identification of a new powdery mildew resistance gene *pmDHT* at or closely linked to the *Pm5* locus in the Chinese wheat landrace Dahongtou. *Plant Dis.* 103, 2645–2651. doi: 10.1094/PDIS-02-19-0401-RE

Sato, K., Abe, F., Mascher, M., Haberer, G., Gundlach, H., Spannagl, M., et al. (2021). Chromosome-scale genome assembly of the transformation-amenable common wheat cultivar 'Fielder'. *DNA Res.* 28:dsab008. doi: 10.1093/dnares/dsab008

Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., and Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nat. Ecol. Evol.* 3, 430–439. doi: 10.1038/s41559-018-0793-y

Singh, R. P., Singh, P. K., Rutkoski, J., Hodson, D. P., He, X. Y., Jørgensen, L. N., et al. (2016). Disease impact on wheat yield potential and prospects of genetic control. *Annu. Rev. Phytopathol.* 54, 303–322. doi: 10.1146/annurev-phyto-080615-095835

Summers, R. W., and Brown, J. K. M. (2013). Constraints on breeding for disease resistance in commercially competitive wheat cultivars. *Plant Pathol.* 62, 115–121. doi: 10.1111/ppa.12165

Tang, X. L., Cao, X. R., Xu, X. M., Jiang, Y. Y., Luo, Y., Ma, Z. H., et al. (2017). Effects of climate change on epidemics of powdery mildew in winter wheat in China. *Plant Dis.* 101, 1753–1760. doi: 10.1094/PDIS-02-17-0168-RE

Walkowiak, S., Gao, L., Monat, C., Haberer, G., Kassa, M. T., Brinton, J., et al. (2020). Multiple wheat genomes reveal global variation in modern breeding. *Nature* 588, 277–283. doi: 10.1038/s41586-020-2961-x

Wang, R. R.-C., and Wei, J. Z. (1995). Variations of two repetitive DNA sequences in several Triticeae genomes revealed by polymerase chain reaction and sequencing. *Genome* 38, 1221–1229. doi: 10.1139/G95-160

Wang, S. C., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., et al. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnol. J.* 12, 787–796. doi: 10.1111/pbi.12183

Wei, X. J., Zhang, N., Hu, Y. Y., Wang, F., Peng, Q. H., and Yang, W. X. (2013). Detection of 1BL/1RS by molecular markers in 53 wheat varieties. *Mol. Plant Breed* 11, 503–508. doi: 10.3969/mpb.011.000503

Wu, Y. N., Yu, X. T., Zhang, X., Yan, L. J., Gao, L., Hao, Y. Q., et al. (2021). Characterization of *PmDGM* conferring powdery mildew resistance in Chinese wheat landrace Duanganmang. *Plant Dis.* 105, 3127–3133. doi: 10.1094/PDIS-12-20-2719-RE

Xie, J. Z., Guo, G. H., Wang, Y., Hu, T. Z., Wang, L., Li, J. T., et al. (2020). A rare single nucleotide variant in *Pm5e* confers powdery mildew resistance in common wheat. *New Phytol.* 228, 1011–1026. doi: 10.1111/nph.16762

Xie, J. Z., Wang, L. L., Wang, Y., Zhang, H. Z., Zhou, S. H., Wu, Q. H., et al. (2017). Fine mapping of powdery mildew resistance gene *PmTm4* in wheat using comparative genomics. *J. Integr. Agric.* 16, 540–550. doi: 10.1016/S2095-3119(16)61377-1

Xu, X. D., Feng, J., Fan, J. R., Liu, Z. Y., Li, Q., Zhou, Y. L., et al. (2018a). Identification of the resistance gene to powdery mildew in Chinese wheat landrace Baiyouyantiao. *J. Integr. Agric.* 17, 37–45. doi: 10.1016/S2095-3119(16)61610-6

Xu, X. D., Li, Q., Ma, Z. H., Fan, J. R., and Zhou, Y. L. (2018b). Molecular mapping of powdery mildew resistance gene *PmSGD* in Chinese wheat landrace Shangeda using RNA-seq with bulk segregant analysis. *Mol. Breed* 38, 23. doi: 10.1007/s11032-018-0783-4

Xu, X. D., Liu, W., Liu, Z. Y., Fan, J. R., and Zhou, Y. J. (2020). Mapping powdery mildew resistance gene *pmYBL* on chromosome 7B of Chinese wheat (*Triticum aestivum* L.) landrace Youbailan. *Plant Dis.* 104, 2411–2417. doi: 10.1094/PDIS-01-20-0118-RE

Xue, F., Zhai, W. W., Duan, X. Y., Zhou, Y. L., and Ji, W. Q. (2009). Microsatellite mapping of powdery mildew resistance gene in wheat land-race Xiaobaidong. *Acta Agron Sin.* 35, 1806–1811. doi: 10.3724/SP.J.1006.2009.01806

Zhang, L., Yang, B. Y., Li, S., and Guo, A. H. (2017). Disease-weather relationships for wheat powdery mildew under climate change in China. *J. Agric. Sci.* 155, 1239–1252. doi: 10.1017/S0021859617000442

Zhang, X. Y., Li, Z. S., and Chen, S. Y. (1989). Studies on the nullisomic backcrossing procedures for producing alien substitution lines of common wheat. *Acta Genet. Sin.* 16, 420–429.

Zhou, R. H., Zhu, Z. D., Kong, X. Y., Huo, N. X., Tian, Q. Z., Li, P., et al. (2005). Development of wheat near-isogenic lines for powdery mildew resistance. *Theor. Appl. Genet.* 110, 640–648. doi: 10.1007/s00122-004-1889-0

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

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

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 09 June 2022
doi: 10.3389/fpls.2022.924565

Wheat Escapes Low Light Stress by Altering Pollination Types

Hong Yang[1], Yongpeng Li[1], Dongxiao Li[2], Liantao Liu[2], Yunzhou Qiao[1], Hongyong Sun[1], Wenwen Liu[1,3], Wenjun Qiao[1,3], Yuzhao Ma[1], Mengyu Liu[1], Cundong Li[2] and Baodi Dong[1]**

[1] Key Laboratory of Agricultural Water Resources, Hebei Laboratory of Agricultural Water-Saving, State Key Laboratory of North China Crop Improvement and Regulation, Center for Agricultural Resources Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Shijiazhuang, China, [2] State Key Laboratory of North China Crop Improvement and Regulation, College of Agronomy, Hebei Agricultural University, Baoding, China, [3] College of Advanced Agricultural Sciences, University of Chinese Academy of Sciences, Beijing, China

Although low light stress seriously affects florets fertility and grain number during the reproductive period, crops can be fertilized by heterologous pollen to alleviate the reduction of grain number. However, wheat is strongly autogamous, how to change to outcross after low light remains unclear. To understand the mechanisms of this change process, an approach combined morphological, physiological, and transcriptomic analyses was performed under low light stress imposed at the young microspore stage the booting stage from tetrad to uni-nucleate microspores stage. The results showed that low light stress caused pollen abortion, and the unfertilized ovary is fertilized by heterologous pollen after floret opening. Compared to control, the opening angle of lemma and glume were increased by 11.6–48.6 and 48.4–78.5%, respectively. The outcross of stressed wheat compensated for the 2.1–18.0% of grain number loss. During this process, phytohormones played an important role. Jasmonic acid (JA) and methyl jasmonate (MeJA) levels in spikelets were increased. Meanwhile, lignin and cellulose content decreased, and genes associated with cell wall related GO terms were enriched. Among the differentially expressed genes (DEGs), were identified 88-710 transcription factors genes, of which some homologs in Arabidopsis are proposed to function in lignin and cellulose, influencing the glume and lemma opening. Our finding can provide new insight into a survival mechanism to set seeds through pollination way alteration in the absence of self-fertilization after the stress of adversity.

Keywords: wheat, floret, grain number, low light stress, pollination way

OPEN ACCESS

Edited by:
Pengtao Ma,
Yantai University, China

Reviewed by:
Lijiang Hou,
Anyang Institute of Technology, China
Jiajia Liu,
Northwest A&F University, China

****Correspondence:***
Cundong Li
nxylcd@hebau.edu.cn
Baodi Dong
dongbaodi@126.com

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

Received: *20 April 2022*
Accepted: *16 May 2022*
Published: *09 June 2022*

Citation:
Yang H, Li Y, Li D, Liu L, Qiao Y, Sun H, Liu W, Qiao W, Ma Y, Liu M, Li C and Dong B (2022) Wheat Escapes Low Light Stress by Altering Pollination Types. Front. Plant Sci. 13:924565. doi: 10.3389/fpls.2022.924565

INTRODUCTION

Wheat (*Triticum aestivum* L.) is the third largest food crop for approximately 60% of the world's population (Food and Agriculture Organization [FAO], 2013). By 2050, wheat production will need to be doubled to meet the demands of the world's growing population (Foulkes et al., 2011). Generally, crop production depends on a moderate natural environment for successful reproductive development (Wang et al., 2015). Unfortunately, extreme climate events, such as global dimming (reduction of global radiation), cold or drought, have occurred frequently in the past decades (Cardona et al., 2012), which limited crop growth and caused severe losses in grain yield (Ji et al., 2011). For instance, from 1960 to 2000, solar radiation at the global

surface decreased by 1.3% per decade (Solomon et al., 2007). In China, solar radiation decreased by 5–30% over these years in some major wheat producing areas (Chameides et al., 1999), resulting in a 5.2% reduction in annual wheat production potential (Chen et al., 2013).

Grain number is one of the main components of yield. Stresses that happened during reproductive stages have the greatest effect on the formation of grain number, since it could dramatically induce floret sterility (Fischer, 2011). In self-fertilizing cereals, the successful development of pollen and ovary is essential for grain production, and abiotic stresses which interfere with the early development of pollen formation and ovaries can result in substantial grain loss (Jain et al., 2007; Endo et al., 2009). It has been shown that although low light stress leads to failure of florets (Dong et al., 2019), wheat produces seeds by capturing heterologous pollen. To overcome yield loss due to pollen failure caused by stresses, capturing heterologous mature pollen is one of the few strategies for yield stability under an adverse environment. Therefore, it is important to understand the mechanism of transition from self-fertilization to heterologous pollen fertilization of wheat as it links directly to ensure food security in extreme climates.

Wheat is a strict self-pollinating crop, with less than 1% probability of cross-pollination (Griffin, 1987; Martin, 1990; Hucl, 1996), which is due to its cleistogamic (closed) floral morphologies (Nguyen et al., 2015). Wheat florets are hermaphrodites. The small anthers and ovary are located at the base of the floret, surrounded by the palea and lemma, and several florets grew on one spikelet with two stiff and large glumes outside. This structure is suited for selfing. Normally at anthesis, the lemma and palea are pushed apart temporally (lasting for 8–30 min) by lodicules swelling, and the pollen dehiscences from anthers and falls on the ovary to accomplish fertilization (de Vries, 1971). However, some studies have found that wheat glumes and lemmas could open to receive heterologous pollen when the ovary is not pollinated within a few days post-anthesis (Kirby et al., 1983; Pickett, 1993). Moreover, wheat's outcrossing rate is related to the degree of glume and lemma opening (Hucl, 1996), and the exsertion of stigmas (Xie et al., 2003).

It is known that the glume opening (first opening) in cereals is controlled by the swelling and subsequent withering of lodicules (Wang et al., 1988; Heslop-Harrison and Heslop-Harrison, 1996). This was mainly related to the weakening of the cell wall (Wang and Gu, 2012; Christiaens et al., 2016). Meanwhile, jasmonic acid (JA) and methyl jasmonate (MeJA) have been reported to influence the opening of glume (Yang J. et al., 2020). Okada et al. (2018) suggested that floret opening after anthesis (second opening) was related to the ovary size but not to the lodicules. The unfertilized ovary increases in radial dimension and produces lateral thrust of the rigid lemma and palea. Some studies are shown that the floret opening was mainly related to the uneven thickness of the cell wall in many cases of flowering plants (Ostergaard, 2009; Llorens et al., 2016; Zajczkowska et al., 2021). The cell wall composition and glume toughness was related to cellulose and lignin biosynthesis (Zou et al., 2015). Fu et al. (2017) found that carbohydrate metabolism, including the degradation of polysaccharides (such as cellulose and lignin) in cell walls, may play an important role in regulation of glume-unclosing after anthesis. However, it is still not clear whether the lemma and glume undergo deformations which could increase the efficiency of outcrossing during the "second opening". Overall, the physiological and biological mechanisms of wheat glume opening have so far been poorly studied.

Extreme weather has become an increasing threat to wheat production. Thus, an intensive understanding of the mechanism of how wheat avoids or resists these stresses is important to reduce crop yield losses under adverse conditions. In this study, we used an integrated strategy combining transcriptome sequencing and morphophysiological analyses of the wheat spike at two stages (under low light stress period and stress removal until early grain filling stage) with following aims to evaluate the compensation effects of cross-pollination after low light stress; and to determine the mechanisms of how wheat switched the pollinating ways to avoid the low light stress. Our results provide insights into the mechanism of glume opening for cross-pollination after low light stress in this important crop.

MATERIALS AND METHODS

Wheat Cultivars, Growing Conditions, and Stress Treatments

Two cultivars of winter wheat, namely 'Henong825' ('HN825') and 'Kenong9204' ('KN9204') were used, both of them were widely planted in North China Plain and were identified with different degrees of shade tolerance by our previous study (Dong et al., 2019).

The field experiments were conducted at the Luancheng Agro-Ecosystem Experimental Station of Chinese Academy of Sciences (37°53′N, 114°41′E, altitude at 50 m), Hebei Province, China during the 2017/2018 and 2018/2019 winter wheat growing seasons. The soil was loam containing 21.41 g kg^{-1} organic matter, 109.55 mg kg^{-1} alkaline nitrogen, 1.44 g kg^{-1} total N, 15.58 mg kg^{-1} available phosphorus, and 220 mg kg^{-1} rapidly available potassium in the topsoil (0–40 cm) of the experimental plots. The study area is in a monsoon climate zone. In a previous manuscript (Yang H. et al., 2020), we have shown the mean temperature, total precipitation, and solar radiation in the winter wheat-growing season. The experiment was a randomized block design with three plot replicates. Each plot size was 6 m long and 2 m wide, with 40 rows. The experiment contained a low light treatment and control (CK). The low light treatment was performed as follows: 90% shading stress treatment was applied for 5 days at the YM stage (young microspore). YM stage was determined using the auricle distance (AD, the distance between the auricle of the flag leaf and the auricle of the penultimate leaf) of the main stem. The cultivar 'HN825' reached the YM stage at 1–2 cm, whereas 'KN9204' reached the YM stage at -1-0 cm (Dong et al., 2019). Shading stress treatment was applied by placing black polyethylene screens horizontally at the height of 2 m aboveground, reducing

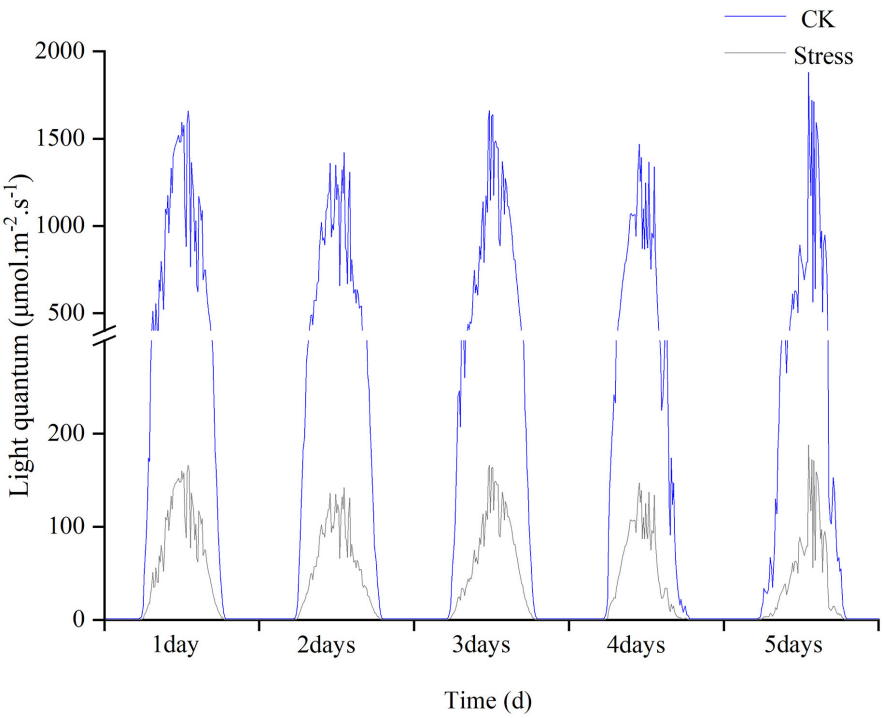


FIGURE 1 | The light intensity under control and stress conditions was monitored during the experimental period in 2017–2018.

incident radiation (measured by a portable weather station, ECA-YW0501; Beijing, China) by 90 ± 5% during treatment imposition (**Figure 1**).

Testing the Cross-Pollination Rate After Low Light Stress

After low light stress, randomly selected 60 main tillers were labeled, and 30 spikes from these main tillers were covered with white paper bags. At maturity, these spikes were harvested and threshed to determine the grain number in uncased and bagged spikes after low light stress. In addition, the compensation rate of cross-pollination after low light stress was the percentage of the yield difference between uncased and bagged spike to the yield difference between control and bagged spike:

$$\text{Compensation rate of cross} - \text{pollination} = \frac{(\text{Grain number of uncased spike} - \text{Grain number of bagged spike})}{(\text{Grain number of control spike} - \text{Grain number of control of bagged spike})}$$

Testing Anthers and Ovaries Fertility

To test the effect of low light stress on anthers and ovary fertility, stressed wheat and control wheat was used in cross-pollination experiments. Selected 20 plants were labeled from each of the stressed and control spikes in two cultivars. Ears of two wheat cultivars were covered with paper bags at heading stage to prevent uncontrolled cross-pollination. Crosses were conducted using control pollen a few days later when control ears were fully exerted and most of the florets were open. At maturity, grain number was determined and analyzed using *t*-tests.

In situ Starch Localization and the Activity of Pollen Grains Determining

Fresh florets from control and stressed wheat plants were dissected from the ear. Anthers and ovaries staining were performed using I_2-KI solution (0.20% w/w I_2 and 0.50% w/w KI). The staining solution was allowed to penetrate the cells for 1 min at room temperature. After washing with deionized water, stained florets were viewed with a stereomicroscope.

For pollen staining, four florets in the middle of spike were harvested from each plant, and pollen was released by gentle pressure in 50 μL NanoPure water. I_2-KI staining was added to the pollen, which was then incubated for 1 min at room temperature. Five individual plants from each treatment group were analyzed and viewed using a Leica DM6 microscope (Leica Microsystems, Wetzlar, Germany). Moreover, the viable pollen cells rate was measured using a pollen vitality analyzer (Ampha Z32, Switzerland).

To study young microspore development under low light stress, 20 anthers from each treatment group were harvested and then fixed using FAA (Formaldehyde-acetic acid-ethanol) stationary liquid. After 3 days, fixed anther samples were washed with 70% ethanol and stored at 4°C. The anthers were stained with 0.5 μg mL^{-1} DAPI [2-(4-Amidinophenyl)-6-indolecarbamidine dihydrochloride] to detect their nuclear

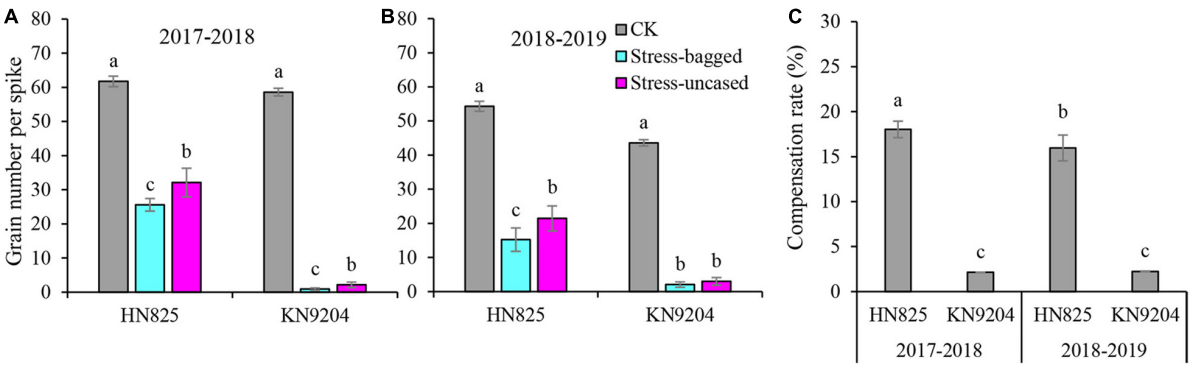


FIGURE 2 | Testing the seed setting rate of heterologous pollen pollination of spike after low light stress in 2017–2018 **(A)** and in 2018–2019 **(B)**. **(C)** Compensation rate of heterologous pollen pollination of the spike to grain number after low light stress. CK: grain number per spike of control plants were grown under normal light conditions; Stress-bagged: grain number per spike with white bags after removing low light stress; Stress-uncased: grain number per spike without white bags after removing low light stress; Different letters in each cultivar indicate the statistically significant differences between treatments ($P < 0.05$) **(A,B)**.

development and pollen fertility, and photographs were taken using a Leica DM6 microscope.

Measurements of Glume and Lemma Opening Angle

In the early grain filling stage (about 20 days after removing low light, which is the period of floret opening), ten spikelets from five labeled plants in each treatment were collected and taken images for glume and lemma opening angle measurement. The opening angle was analyzed by ImageJ software. Glume angle was measured between the tips of both glume and the connection point to the spikelet pedicel. Similarly, the lemma angle was measured between the tips of both lemma and the connection point to the spikelet pedicel.

Measurements of Plant Endogenous Hormones

On the fifth day of low light stress and early grain filling stage (20 days after removing low light), spike tissues in each treatment were harvested and immediately frozen in liquid nitrogen and then stored at −80°C for subsequent measurement of hormones content. Jasmonic acid (JA) and methyl jasmonate (MeJA) was extracted at 4°C and three technical replicates were performed for each biological replicate. Measurement of JA and MeJA levels was carried out at Servicebio (Wuhan, China) using the LC-MS system. Three biological replicates in each treatment were measured.

Lignin, Cellulose, and Hemicellulose Content Assay

Measurements of lignin content in the glume and lemma were performed with the acetylation method using a Plant Lignin assay kit (Suzhou Grace Biotech Co. Ltd., Suzhou, China) according to the manufacturer's protocol. For each treatment group, three replicates were harvested on the fifth day of low light stress and early grain filling stage. Measurements of cellulose in the glume were performed with the reaction method of anthrone reagent with furfural compounds using a Plant Cellulose assay kit (Suzhou Grace Biotech Co. Ltd., Suzhou, China) according to the manufacturer's protocol. Measurements of hemicellulose in the glume were performed using a Plant Hemicellulose assay kit (Suzhou Grace Biotech Co. Ltd., Suzhou, China) according to the manufacturer's protocol.

RNA Isolation, cDNA Construction, and Sequencing

Spikes of two cultivars were harvested on the fifth day of low light stress and early grain filling stage, respectively. Two biological replicate samples were collected for RNA-seq. Total RNA was extracted using Plant RNA Purification Reagent (Invitrogen, Carlsbad, CA, United States). The cDNA libraries were constructed using a Truseq^{TM} RNA sample prep Kit (Illumina, San Diego, CA, United States) according to the manufacturer's instructions. RNA sequencing was performed by Shanghai Majorbio Bio-pharm Technology Company using an Illumina NovaSeq 6000.

After sequencing, the clean reads were mapped to the *Triticum aestivum* genome (IWGSC, Ensembl Plants) by HISAT2. The FPKM (Fragments Per Kilobase of transcript per Million mapped reads) was calculated by RSEM, and DEGs (deferentially expressed genes) were identified using DESeq2. Genes with at least a two-fold change in expression level with a *P*-adjust value less than 0.05 were considered DEGs. Gene Ontology (GO) enrichment analysis was performed using ClusterProfiler3.0.

The transcription factor family analysis was performed as described previously (Li et al., 2018). The homologous genes in arabidopsis of DEGs were gained from Ensembl Biomart.

The scatter plots of GO enrichment results were drawn using ggplot2 in R packages, the heatmap of genes was visualized using Cluster3.0 and TreeView. Co-expression interactions of hub genes and the co-expressed genes were visualized using Cytoscape.

Statistical Analysis

All data were analyzed using the SPSS 22.0 package. One-way ANOVA was used to determine the statistical significance between control treatment and shading stress for glume and

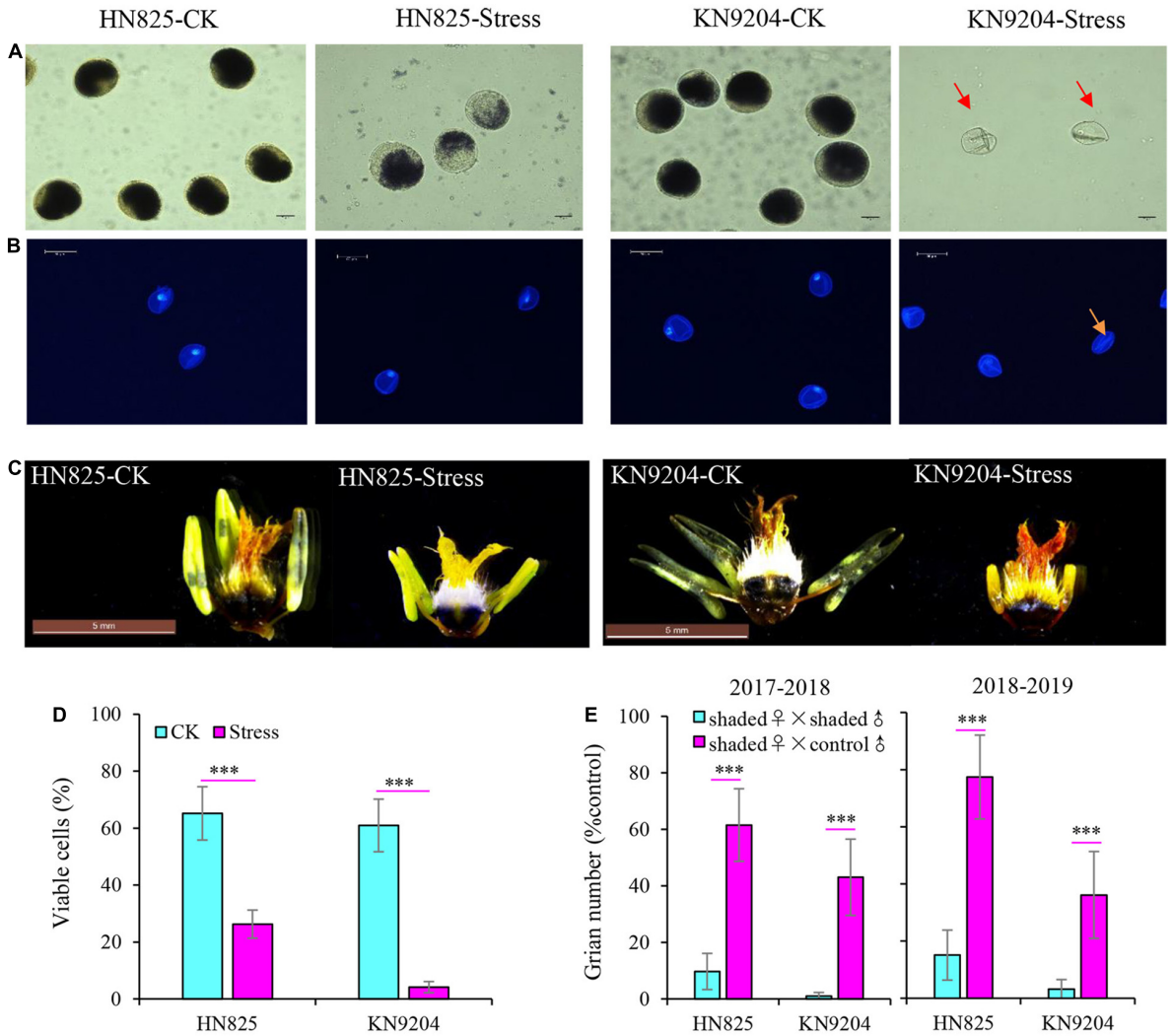


FIGURE 3 | Effects of low light stress on the anther and ovary fertility in 'HN825' and 'KN9204'. Pollen of the two cultivars was stained with I_2-KI solution **(A)** and DAPI solution **(B)** to determine fertility and development. Starch staining was carried out on anthers and ovaries of control (CK) and stressed plants **(C)**. The rate of viable pollen cells in control and stress treatments were measured at anthesis **(D)**. Stressed wheat and control wheat was used in cross-pollination experiments to test the effect of low light stress on anthers and ovary fertility **(E)**. The scale bar indicates 20 μm **(A)**, 50 μm **(B)**, and 5 mm **(C)**, respectively. Shaded ♀ × shaded ♂ indicates that stressed ovary was pollinated with stressed pollens; Shaded ♀ × control ♂ indicates that stressed ovary was pollinated with control pollens **(D)**. Significant differences between treatments are indicated by asterisks (***$P < 0.001$, Student's t-test).

lemma angle, and compensation rate, at $P < 0.05$ using the LSD value. The statistical significance of grain number, hormone content, lignin, cellulose, and hemicellulose content measurements was analyzed with a t-test. Difference was tested at *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

RESULTS

Low Light Changed the Fertilization Way of Wheat

In previous study, we found that low light stress can dramatically cause floret abortion, resulting in the reduction of grain number (Dong et al., 2019). The fact that pollen is more sensitive to adverse environments and that aborted pollen can induce ovary to receive heterogenous pollens (Chen et al., 2020). This reminded us that the switch of pollination way might exist in wheat under low light stress and could use to compensate for the yield loss. To verify it, we systematically investigated the grain number production of bagged spikes (performed after removing low light stress) and uncased spikes (**Figures 2A,B**). Compared with the bagged spike, the grain number of the uncased spike was significantly higher in both cultivars after low light stress treatment. The grain number of an uncased spike was 25.4–137.0% higher than that of bagged spike. The results suggested that low light stress can increase cross-pollination which would alleviate wheat grain number loss induced by low light. We then evaluated the compensation effect of cross-pollination on grain

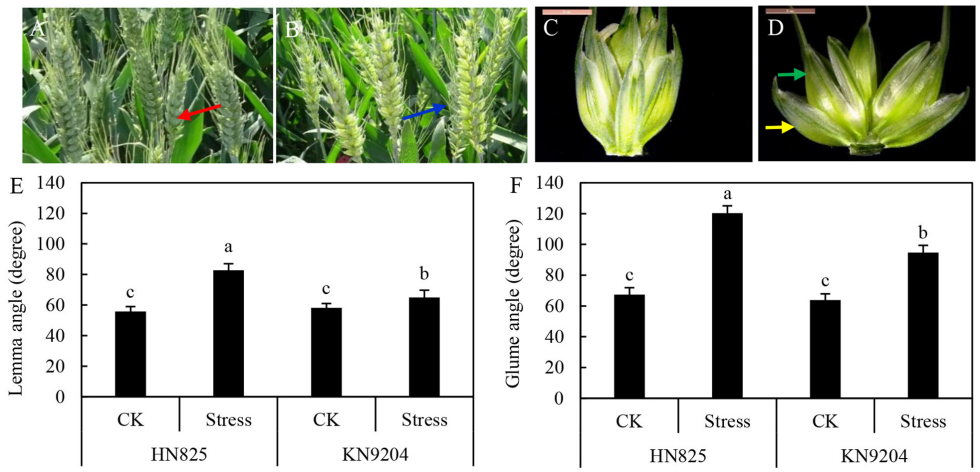


FIGURE 4 | Measurement of glume and lemma angle in 'HN825' and 'KN9204'. Spike **(A,B)** and spikelet **(C,D)** images of control **(A,C)** and low light stress **(B,D)** at early grain filling stage (20 days after low light treatment). **(E)** The lemma angle. **(F)** The glume angle. The red arrows indicate the spike in control **(A)**, the blue arrows indicate the spike in stressed spike **(B)**, the yellow arrows indicate the glume, and the green arrows indicate the lemma **(D)**. Different letters indicate the statistically significant differences between treatments ($P < 0.05$). Scale bars are 5 mm.

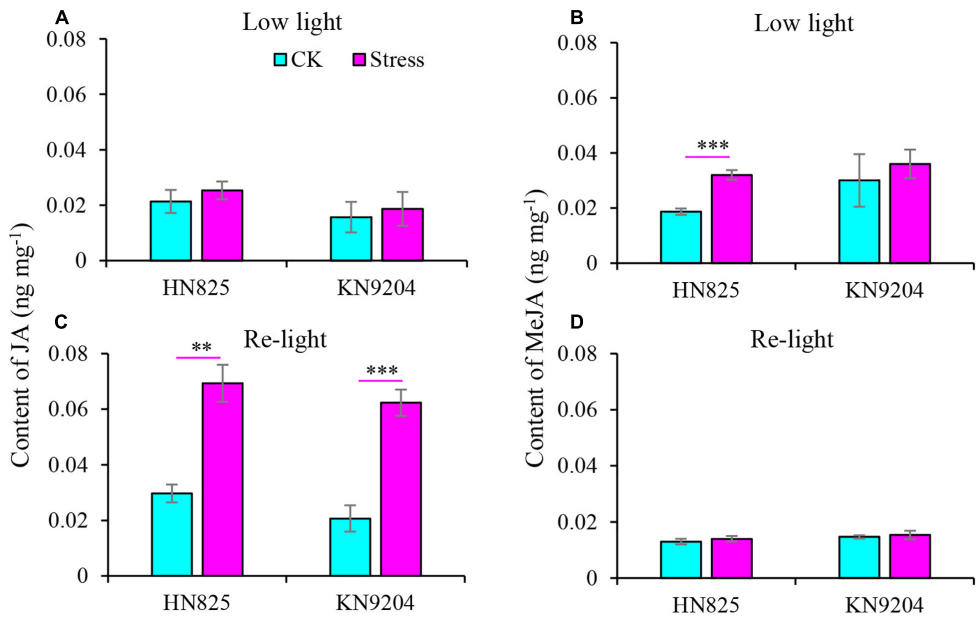


FIGURE 5 | Effect of low light on the contents of jasmonic acid (JA) **(A,C)** and methyl-jasmonate (MeJA) **(B,D)** in the spikelet of two wheat cultivars ('HN825' and 'KN9204'). Low light, under low light stress; Re-light, early grain filling stage (20 days after low light treatment). Significant differences between the CK and stress treatments are indicated by asterisks (**$P < 0.01$; ***$P < 0.001$, Student's *t*-test).

number per spike. The cross-pollination compensated 2.1–18.0% reduction of grain number per spike caused by self-pollination failure after low light stress (**Figure 2C**).

Anthers Were More Sensitive to Low Light Stress Than Ovaries

The change of grain number is closely related to florets fertility. To determine floret fertility, anthers and ovaries were stained with I_2-KI staining solution, and pollen grain development was determined by DAPI staining (**Figure 3**). For anthers, the fertility of pollen was significantly decreased under low light stress in both cultivars, and that of 'KN9204' was even completely sterile (**Figure 3A**). Moreover, the development of pollen in 'KN9204' was suppressed with the pollen nuclei disappearing under low light stress (**Figure 3B**). In addition, compared with control, the anther morphology of both cultivars was seriously abnormal, and the anther of 'KN9204' was sterile and shriveled (**Figure 3C**). Moreover, viable pollen cells rate was only 26.3 and 4.1% in 'HN825' and 'KN9204' after low light stress (**Figure 3D**). To

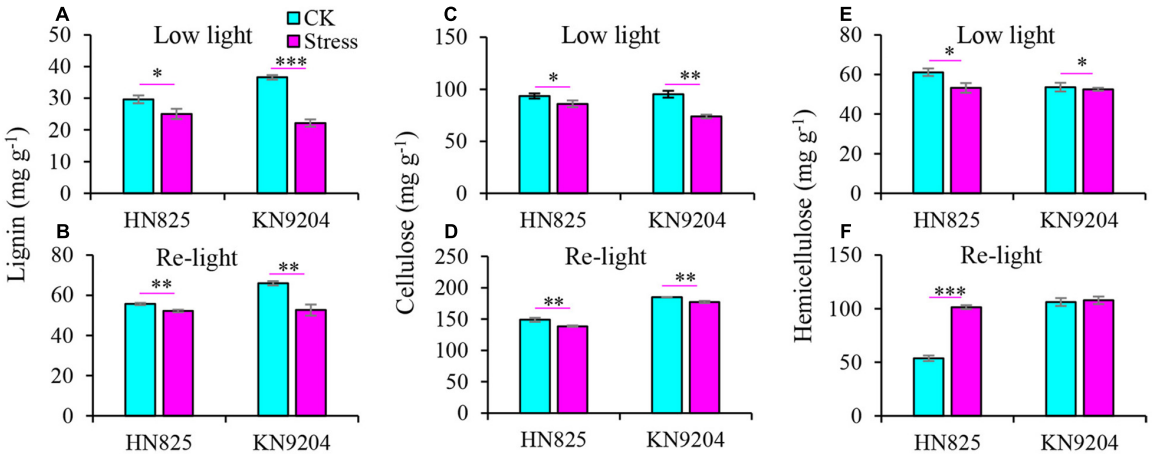


FIGURE 6 | Effect of low light on the contents of lignin **(A,B)**, cellulose **(C,D)**, and hemicellulose **(E,F)** in the spikelet of two wheat cultivars ('HN825' and 'KN9204'). Low light, under low light stress; Re-light, early grain filling stage (20 days after low light treatment). Significant differences between the CK and stress treatments are indicated by asterisks (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$, Student's *t*-test).

further determine the effect of low light stress during the YM stage on wheat anther and ovary fertility, the cross-pollination experiments between control male fertile and stress female were performed (**Figure 3E**). The grain number of the self-pollination (the stressed ovary was pollinated with the fresh pollens from the stressed plants) is significantly lower than that of the cross-pollinated plants (the stressed ovary was pollinated with the fresh pollens from the control plants) in both cultivars. The grain number of self-pollination plant was only 1.0–15.1% of that of control, while the grain number of the cross-pollinated plant was 36.1–77.4%. This result suggested that the harmful effect of low light stress on anther was greater than that of the ovary, and the ovary was still fertile after 5 days of low light stress.

Glumes and Lemma Opening After Low Light Stress

To explore the causes why wheat cross-pollination occurred, the morphology and color of stressed spikelets and control spikelets were investigated. And we found the color of stressed spike was lighter than that of the control spike at the early grain filling stage (**Figures 4A,B**). Moreover, glume and lemma in the stressed spikelet were open, while the control spikelet was tightly closed (**Figures 4C,D**). Compared to control, the opening degree of lemma and glume were increased by 11.6–48.6, 48.4–78.5%, respectively (**Figures 4E,F**). Moreover, the opening degree of the glume was larger than that of the lemma. The glume opening in 'HN825' was larger than that in 'KN9204', which is consistent with the compensation rate. These results suggested that increasing the glume opening might increase the ovary cross-pollination rate.

Content of Jasmonic Acid, Methyl Jasmonate, Lignin, and Cellulose in Spikelet

A previous study reported that JA and MeJA content increased upon the spikelet opening (Chen et al., 2020). Therefore, the JA and MeJA content was measured in a spikelet (**Figure 5**). Compared to control, the content of JA and MeJA in spike increased under low light stress (**Figures 5A,B**). Moreover, the content of JA was constantly higher in stressed spikelets compared to control from the point when the stress was removed to the early grain filling stage (**Figure 5C**). To further investigate the glume and lemma opening, we dissected glume and lemma under low light and at the early grain filling stage (20 days after low light treatment) to measure the content of lignin, cellulose, and hemicellulose. Results showed that the lignin and cellulose of glume from the low light treatment group significantly decreased as compared with CK at the two-time point (**Figures 6A–D**). Statistically, the lignin and cellulose were reduced by 15.5–39.4 and 8.0–22.5% in the low light treatment group, by 6.3–20.2 and 4.2–7.2% after low light treatment. In addition, hemicellulose of glume only significantly decreased in the low light treatment group, while it was not hampered after low light treatment (**Figures 6E,F**). This suggested that low light has a negative influence on the lignin and cellulose of glume, increasing the opening of the glume.

RNA-Seq Finding Key Terms Responsible for Floret Opening

To investigate the molecular mechanisms underlying spike fertility and the opening of the glume, RNA-seq analysis was performed using the spike tissues of the 'HN825' and 'KN9204' cultivars collected on the fifth day of low light stress and at early grain filling stage (20 days after removing low light). The DEGs between CK and stress treatment at the two-time point were identified based on their normalized expression level. The number of the DEGs in each comparison ranged from 14433 to 39205 (**Figure 7A** and **Supplementary Dataset 1**). Four clusters of genes were selected for further analysis. Cluster 1 and cluster 2 represented overlapped up-regulated or down-regulated genes under stress on the fifth day of low light stress. Cluster 3 and

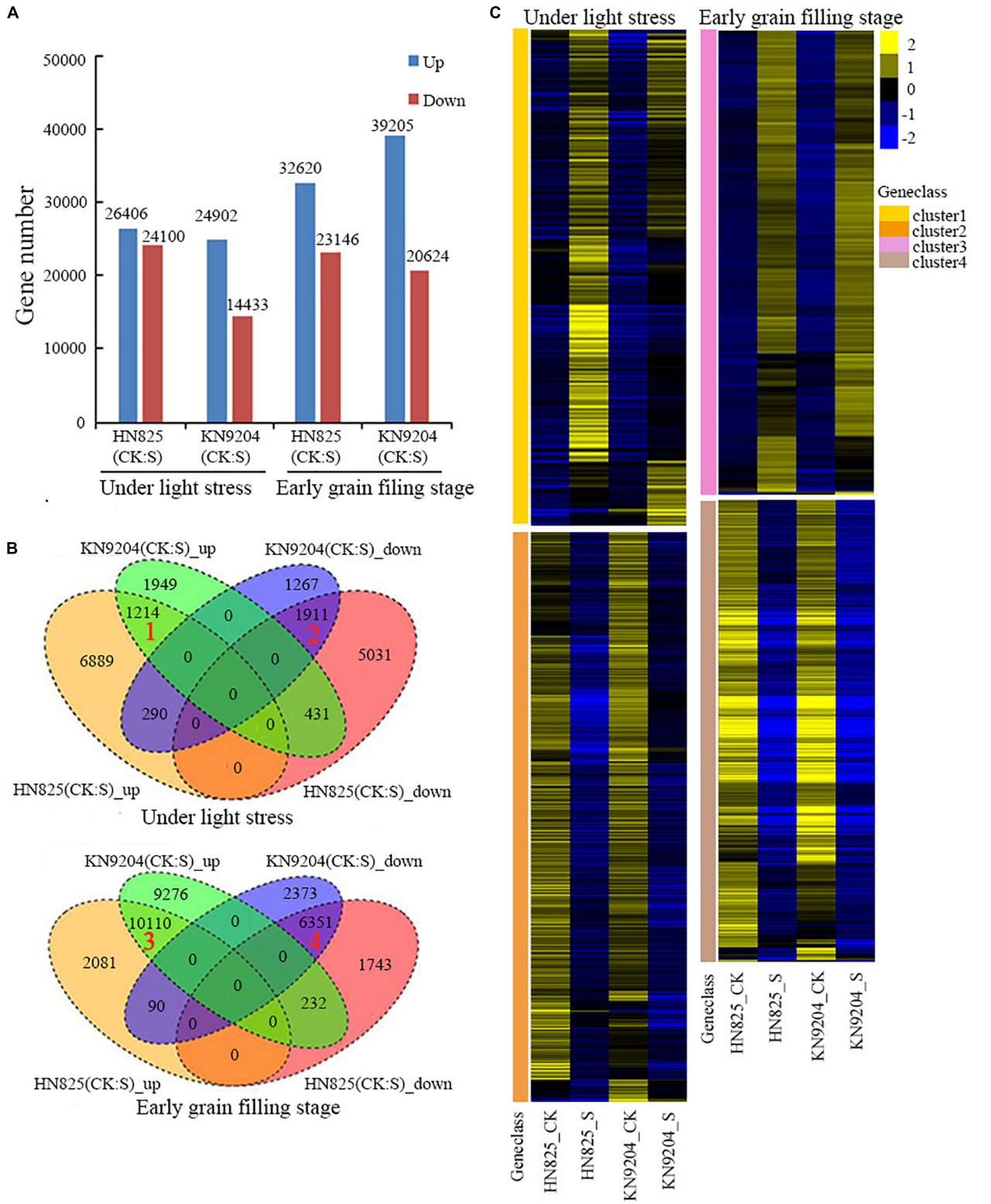


FIGURE 7 | Analysis of RNA-sequencing data of the spike in low light stress (during young microspore) and early grain filling stage (20 days after low light treatment) under low light stress (S) and control (CK) conditions. **(A)** The numbers of the differentially expressed genes (DEGs). **(B)** Venn diagram analysis of all up-regulated and down-regulated DEGs in two wheat cultivars ('HN825' and 'KN9204'). **(C)** Heatmap clustering of the up-regulated (cluster 1 and 3) and down-regulated (cluster 2, 4) in both cultivars under low light stress and early grain filling stage.

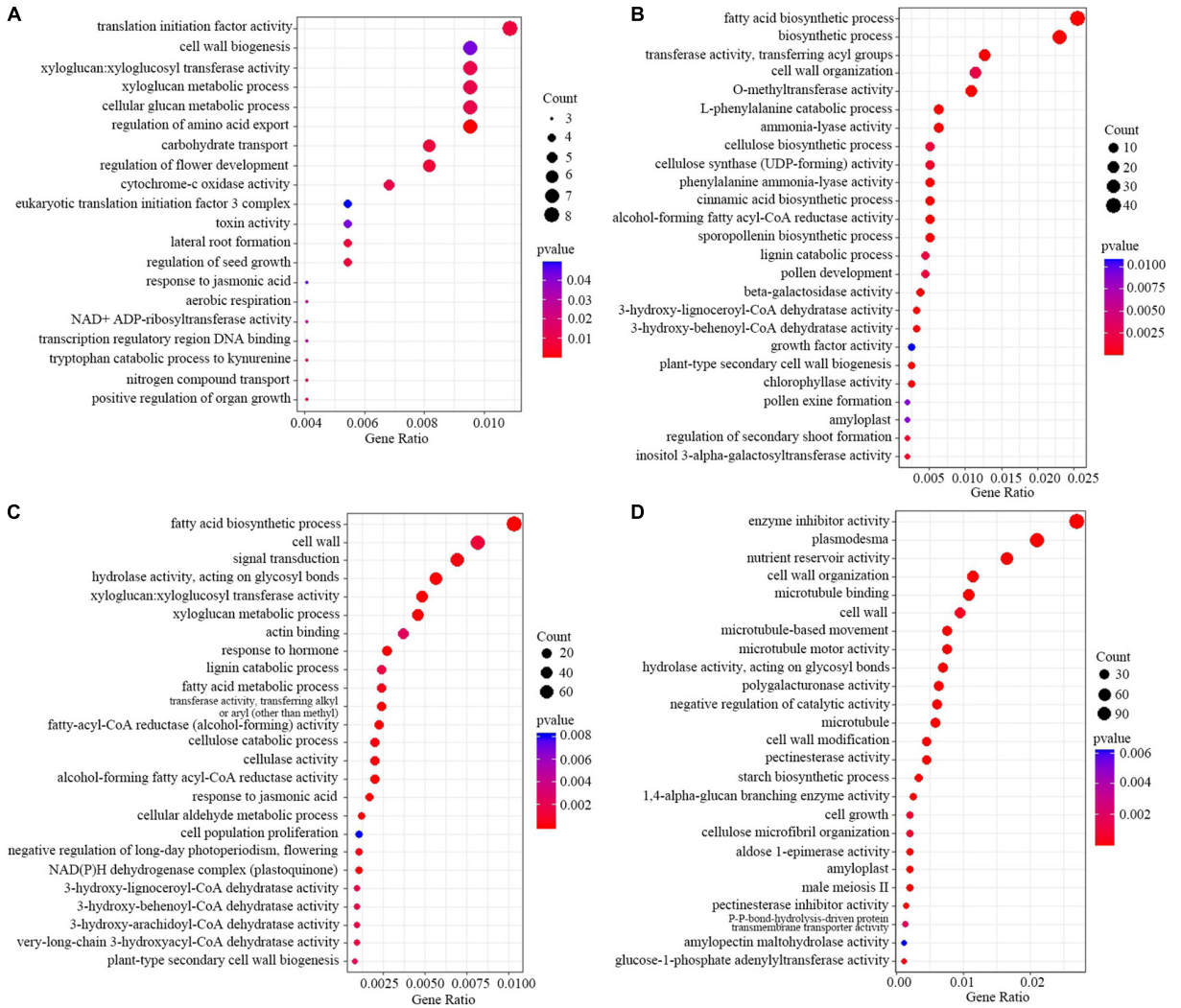


FIGURE 8 | Scatterplots showing the GO enrichment results of cluster 1 **(A)**, 2 **(B)**, 3 **(C)**, and 4 **(D)**.

cluster 4 represented overlapped up-regulated or down-regulated genes under stress on the 20th day after low light treatment (**Figures 7B,C**).

Gene ontology enrichment analysis found that several nutrients transport related GO terms such as "carbohydrate transport", "nitrogen compound transport", and "cytochrome-c oxidase activity", "regulation of flower development", and "response to jasmonic acid" enriched for up-regulated DEGs under low light group (**Figure 8A** and **Supplementary Dataset 2**). For the down-regulated genes in cluster 2, genes related to fatty acid biosynthesis metabolism pathways, and genes involved in the biosynthesis and inactivation process of the cell wall and lignin and cellulose were also enriched (**Figure 8B**). In addition, lignin and cellulose catabolic related terms, response to jasmonic acid related terms were enriched in the up-regulated genes in the early grain filling stage (**Figure 8C** and **Supplementary Dataset 3**). For the down-regulated genes in the early grain filling stage, GO terms, such as "nutrient reservoir activity", "starch biosynthetic process", "pectinesterase activity", and "pectinesterase inhibitor activity" were enriched (**Figure 8D**). Moreover, phenylalanine metabolism related KEGG term was enriched in the down-regulated genes under low light group (**Supplementary Dataset 4**). The results suggested that the lignin and cellulose degradation might play an important role in glume opening under low light conditions, consistent with the content changes of lignin and cellulose in different stages in our previous data.

In addition, we checked all the transcription factors (TFs) in the DEGs of four clusters. There were 88 and 64 TFs in the up- and down-regulated DEGs in low light treatment groups (**Figures 9A,B**), 710 and 294 TFs in the up- and down-regulated DEGs after low light treatment groups, respectively (**Figures 9C,D**). Many MYB and LBD family TFs were enriched in the down-regulated DEGs, but only a few were found in the up-regulated DEGs in low light treatment groups. Homology comparison showed that most of the

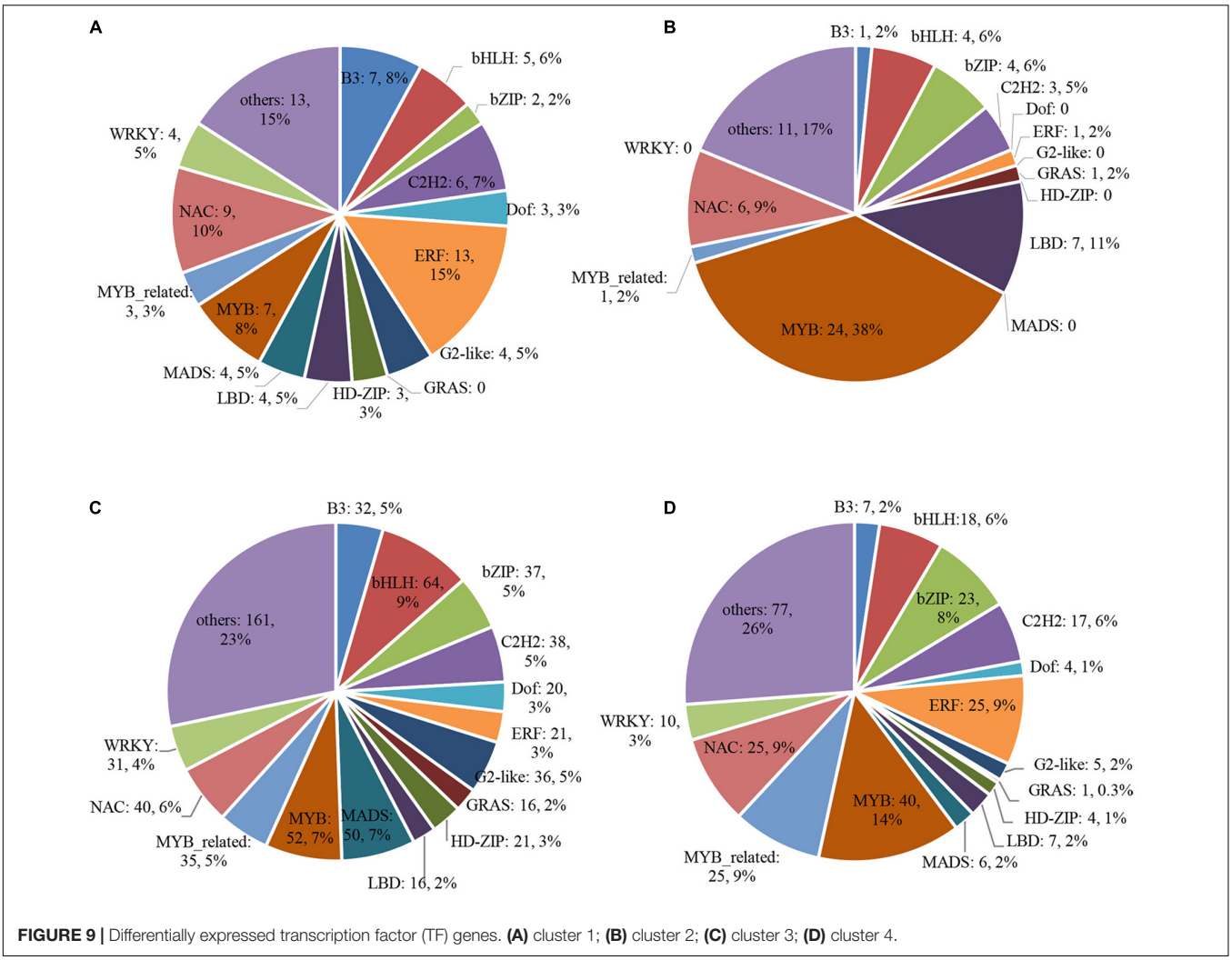


FIGURE 9 | Differentially expressed transcription factor (TF) genes. **(A)** cluster 1; **(B)** cluster 2; **(C)** cluster 3; **(D)** cluster 4.

homologous genes of down-regulated TFs in Arabidopsis could modulate xylem development or secondary cell wall biogenesis (**Supplementary Dataset 5**). This implies that they might participate in the regulation of glume opening process in wheat.

DISCUSSION

In cereals, flowering is related to the capacity of seed production, which is more sensitive to abiotic stress in booting stage (Ji et al., 2010). Due to climate change, adverse stress seriously threatened flowering and seed setting, eventually resulting in lower yields (Dong et al., 2017). However, plants exhibit various mechanisms for surviving, including morphological adaptions under abiotic stress conditions (Hasanuzzaman et al., 2013). For self-pollination wheat, when the ovary was not pollinated a few days after anthesis, the cross-pollination was carried out by changing its cleistogamic (closed) floral morphologies (Okada et al., 2018), and a similar phenomenon after low light stress was found in our study. Low light stress reduces the fertility of anther and ovaries, while the ovaries are less affected and still fertile. To receive fertile pollens for florets with normal ovaries but sterile anthers, wheat adopts the strategy of capturing heterologous pollen to finish pollination. In the present study, grain number per spike significantly decreased by 58.5–98.5% due to unsuccessful self-pollination after low light stress. However, the compensation ratio of grain number per spike due to accepting heterologous pollen grain was 2.1–18.0%. The compensatory capacity of 'KN9204' was lower than that of 'HN825', probably because the spike layer of 'KN9204' variety was neatly grown and flowered uniformly, which made it difficult for the ovary to capture heterologous mature pollen. Taken together, the change of wheat pollination ways after low light stress encountered at the YM stage could serve as a compensation strategy to alleviate the reduction of yield, which is of great significance for ensuring food security under adverse climate conditions.

For the male sterile florets in wheat, the separation angle of the glumes and lemma is important for ensuring the exsertion of stigmas (Chen et al., 2020). In present study, the glume and lemma opened greatly after low light removal till the early grain filling stage. The opening angles of lemma

and glume were increased by 11.6–48.6, 48.4–78.5% compared to control, respectively. This ensures that the ovaries can receive heterologous fertile pollens brought by the wind. Glume and lemma opening is a trait that is related not only to spikelet position within the spike, spikelet density, and ovary size but also to the glume stiffness (Okada et al., 2018). The main factors influencing glume toughness are cellulose and lignin metabolism, and cell wall metabolism (Zou et al., 2015). In this study, we found that several biological processes related to carbohydrate metabolism, response to cell wall macromolecular metabolism were involved in glume opened spikes after low light stress. Both down-regulated DEGs were significantly enriched in the cell wall organization process and up-regulated DEGs were significantly enriched in lignin metabolism, fatty acid metabolic process, and response to the jasmonic acid process. These results indicate that cell wall metabolism and the jasmonic acid process may play a key role in the opening of wheat glumes after low light stress.

Plant growth and development are modulated by complex molecular networks in which transcription factors (TFs) played a central role in regulating other gene expressions (Ramachandran et al., 1994; Riechmann et al., 2000; Mitsuda and Ohme-Takagi, 2009). We identified TFs with altered expression in our set of DEGs with common expression patterns in both cultivars. Most of the homologous genes in Arabidopsis of these TFs (such as MYB, LBD, ERF NAC, and MADS family TFs) have functions related to the cell wall biogenesis. They may be involved in regulating glume opening. Meanwhile, lignin and cellulose contents in spikelets were significantly reduced compared with the control. This further verified that lignin and cellulose might be the main factors leading to glume opening.

Phytohormones have been considered the main factor regulating the spikelet opening. JA and MeJA can promote the opening of crops glumes (Zeng et al., 1999; Liu et al., 2017; Yang J. et al., 2020). The glume opening and increased level of JA and MeJA were also found after low light stress, strongly demonstrating that JA and MeJA might play an important role in the glume opening and capturing heterogenous pollens after stresses. Thus, the application of JA or MeJA may be a useful method to guarantee grain number and grain yield when wheat encounter stresses which can induce male sterility.

CONCLUSION

Compared with female organs, anther was more sensitive to environmental stimulation, such as low light stress. When the stress was imposed during the microspore stage, the male is partially or completely sterile. Male sterility leads to failure of self-pollination, which leads to serious yield loss. However, the active ovary can compensate for the loss of yield caused by pollen abortion by receiving active pollen from different sources. In the process of glume and lemma opening, the cellulose and lignin decrease induced by down-regulated key metabolism- or signaling-related genes played an important role, which is necessary for cross-pollination. Cross-pollination of wheat provides a survival mechanism to set seeds in the absence of self-fertilization after the stress of adversity. Thus, pollination way alteration may be a method for plants to adjust to a stressful environment and can be used in agriculture production to guarantee a stable yield under extreme weather.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/sra; PRJNA835521.

AUTHOR CONTRIBUTIONS

BD, HY, and ML designed the study. HY, YQ, WL, YM, and WQ performed the experiments. HY and YL contributed data analysis. HY prepared the manuscript. BD, CL, HS, LL, and DL supervised this project and the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the National Science Foundation of China (31971848), the National Key Research and Development Program (2021YFD1901003-01), Hebei Province Key Research and Development Program (19226436D), and State Key Laboratory of North China Crop Improvement and Regulation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.924565/full#supplementary-material

REFERENCES

Cardona, O., van Aalst, M., Birkmann, J. M., Fordham, M., McGregor, G., Perez, R., et al. (2012). *Determinants of Risk: Exposure and Vulnerability, in Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC)*. Cambridge, MA: Cambridge University Press, 65–108.

Chameides, W. L., Yu, H., Liu, S. C., Bergin, M., Zhou, X., Mearns, L., et al. (1999). Case study of the effects of atmospheric aerosols and regional haze on agriculture: an opportunity to enhance crop yields in China through emission controls? *Proc. Natl. Acad. Sci. U.S.A.* 96, 13626–13633. doi: 10.1073/pnas.96.24.13626

Chen, C., Baethgen, W. E., and Robertson, A. (2013). Contributions of individual variation in temperature, solar radiation and precipitation to crop yield in the North China Plain, 1961–2003. *Clim. Change* 116, 767–788. doi: 10.1007/s10584-012-0509-2

Chen, J., Xu, Y., Fei, K., Wang, R., and Yang, J. (2020). Physiological mechanism underlying the effect of high temperature during anthesis on spikelet-opening of photo-thermo-sensitive genic male sterile rice lines. *Sci. Rep.* 10:2210. doi: 10.1038/s41598-020-59183-0

Christiaens, A., De Keyser, E., Pauwels, E., De Riek, J., Gobi, B., and Van Labek, M. C. (2016). Suboptimal light conditions influence source-sink metabolism during flowering. *Front. Plant Sci.* 7:249. doi: 10.3389/fpls.2016.00249

de Vries, A. (1971). Flowering biology of wheat, particularly in view of hybrid seed production—a review. *Euphytica* 20, 152–170. doi: 10.1007/BF00056076

Dong, B. D., Yang, H., Liu, H. P., Qiao, Y. Z., Zhang, M. M., Wang, Y. K., et al. (2019). Effects of shading stress on grain number, yield, and photosynthesis during early reproductive growth in wheat. *Crop Sci.* 59, 363–378. doi: 10.2135/cropsci2018.06.0396

Dong, B. D., Zheng, X., Liu, H. P., Able, J. A., Yang, H., Zhao, H., et al. (2017). Effects of drought stress on pollen sterility, grain yield, abscisic acid and protective enzymes in two winter wheat cultivars. *Front. Plant Sci.* 8:1008. doi: 10.3389/fpls.2017.01008

Endo, M., Tsuchiya, T., Hamada, K., Kawamura, S., Yano, K., Ohshima, M., et al. (2009). High temperatures cause male sterility in rice plants with transcriptional alterations during pollen development. *Plant Cell Physiol.* 50, 1911–1922. doi: 10.1093/pcp/pcp135

Fischer, R. A. (2011). Wheat physiology: a review of recent developments. *Crop Pasture Sci.* 62, 95–114. doi: 10.1071/CP10344

Food and Agriculture Organization [FAO] (2013). *FAO Statistical Yearbook*. Rome: FAO.

Foulkes, M. J., Slafer, G. A., Davies, W. J., Berry, P. M., Sylvester-Bradley, R., Martre, P., et al. (2011). Raising yield potential of wheat. III. Optimizing partitioning to grain while maintaining lodging resistance. *J. Exp. Bot.* 62, 469–486.

Fu, C., Wang, F., Liu, W., Liu, D., Li, J., Zhu, M., et al. (2017). Transcriptomic analysis reveals new insights into high-temperature-dependent glume-unclosing in an elite rice male sterile line. *Front. Plant Sci.* 8:112. doi: 10.3389/fpls.2017.00112

Griffin, W. B. (1987). Outcrossing in New Zealand wheats measured by occurrence of purple grain. *New Zeal. J. Agric. Res.* 30, 287–290. doi: 10.1080/00288233.1987.10421885

Hasanuzzaman, M., Nahar, K., Alam, M., Roychowdhury, R., and Fujita, M. (2013). Physiological, biochemical, and molecular mechanisms of heat stress tolerance in plants. *Int. J. Mol. Sci.* 14, 9643–9684. doi: 10.3390/ijms14059643

Heslop-Harrison, Y., and Heslop-Harrison, J. S. (1996). Lodicule function and filament extension in the grasses: potassium ion movement and tissue specialization. *Ann. Bot.* 77, 573–582. doi: 10.1093/aob/77.6.573

Hucl, P. (1996). Out-crossing rates for 10 Canadian spring wheat cultivars. *Can. J. Plant Sci.* 76, 423–427. doi: 10.4141/cjps96-075

Jain, M., Prasad, P. V., Boote, K. J. Jr., Hartwell, A. L., and Chourey, P. S. (2007). Effects of season-long high temperature growth conditions on sugar-to-starch metabolism in developing microspores of grain sorghum (*Sorghum bicolor* L. Moench). *Planta* 227, 67–79. doi: 10.1007/s00425-007-0595-y

Ji, X., Shiran, B., Wan, J., Lewis, D. C., Jenkins, C. L., Condon, A. G., et al. (2010). Importance of preanthesis anther sink strength for maintenance of grain number during reproductive stage water stress in wheat. *Plant Cell Environ.* 33, 926–942. doi: 10.1111/j.1365-3040.2010.02130.x

Ji, X. M., Dong, B. D., Shiran, B., Talbotet, M. J., Edlington, J. E., Hughes, T., et al. (2011). Control of abscisic acid catabolism and abscisic acid homeostasis is important for reproductive stage stress tolerance in cereals. *Plant Physiol.* 156, 647–662. doi: 10.1104/pp.111.176164

Kirby, E. J. M., Fellowes, G., and Appleyard, M. (eds) (1983). "Anthesis in winter barley," in *Annual Report—Plant Breeding Institute*, (Cambridge, MA: Cambridge Plant Breeding Institute), 112–113.

Li, Y., Fu, X., Zhao, M., Zhang, W., Li, B., An, D., et al. (2018). A genome-wide view of transcriptome dynamics during early spike development in bread wheat. *Sci. Rep.* 8:15338. doi: 10.1038/s41598-018-33718-y

Liu, L. I., Zou, Z., Qian, K. E., Xia, C., He, Y., Zeng, H., et al. (2017). Jasmonic acid deficiency leads to scattered floret opening time in cytoplasmic male sterile rice Zhenshan 97A. *J. Exp. Bot.* 68, 4613–4625. doi: 10.1093/jxb/erx251

Llorens, C., Argentina, M., Rojas, N., Westbrook, J., Dumais, J., and Noblin, X. (2016). The fern cavitation catapult: mechanism and design principles. *J. R. Soc. Interface* 13:20150930. doi: 10.1098/rsif.2015.0930

Martin, T. J. (1990). Outcrossing in 12 hard red winter-wheat cultivars. *Crop Sci.* 30, 59–62. doi: 10.2135/cropsci1990.0011183X003000010013x

Mitsuda, N., and Ohme-Takagi, M. (2009). Functional analysis of transcription factors in *Arabidopsis*. *Plant Cell Physiol.* 50, 1232–1248. doi: 10.1093/pcp/pcp075

Nguyen, V., Fleury, D., Timmins, A., Laga, H., Hayden, M., Mather, D., et al. (2015). Addition of rye chromosome 4R to wheat increases anther length and pollen grain number. *Theor. Appl. Genet.* 128, 953–964. doi: 10.1007/s00122-015-2482-4

Okada, T., Ridma, J. E. A., Jayasinghe, M., Nansamba, M., Baes, M., Warner, P., et al. (2018). Unfertilized ovary pushes wheat flower open for cross-pollination. *J. Exp. Bot.* 69, 399–412. doi: 10.1093/jxb/erx410

Ostergaard, L. (2009). *Fruit Development and Seed Dispersal (L. Ostergaard, Ed.)*. Oxford: Wiley.

Pickett, A. (1993). *Hybrid Wheat Results and Problems*. Berlin: Paul Parey Scientific Publishers.

Ramachandran, S., Hiratsuka, K., and Chua, N. H. (1994). Transcription factors in plant growth and development. *Curr. Opin. Genet. Dev.* 4, 642–646. doi: 10.1016/0959-437X(94)90129-Q

Riechmann, J. L., Heard, J., Martin, G., Reuber, L., Jiang, C., Keddie, J., et al. (2000). *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. *Science* 290, 2105–2110. doi: 10.1126/science.290.5499.2105

Solomon, S. D., Qin, D., Manning, M., Chen, Z., and Miller, H. L. (2007). *Climate Change 2007: The Physical Science Basis. Working Group I Contribution to the Fourth Assessment Report of the IPCC*. Cambridge, MA: Cambridge University Press.

Wang, L., Deng, F., and Ren, W. J. (2015). Shading tolerance in rice is related to better light harvesting and use efficiency and grain filling rate during grain filling period. *Field Crops Res.* 180, 54–62.

Wang, Z., and Gu, Y. J. (2012). The process of opening and closing of florets in rice and its influencing factors. *Sci. Pap.* [Epub ahead of print].

Wang, Z., Lu, C. M., Gu, Y. J., and Gao, Y. Z. (1988). Studies on the mechanism of the anthesis of rice I. Effect of temperature on spikelet opening and pollen vigor. *Acta Agron. Sin.* 14, 14–21.

Xie, G. S., Zeng, H. L., Lin, X. H., Tao, A. L., and Zhang, D. P. (2003). Effect of temperature on the fertility restoration of different two-line hybrid rice. *J. Genet. Genomics* 30, 142–146.

Yang, H., Dong, B. D., Wang, Y. K., Qiao, Y. Z., Shi, C. H., Jin, L. L., et al. (2020). Photosynthetic base of reduced grain yield by shading stress during the early reproductive stage of two wheat cultivars. *Sci. Rep.* 10:14353. doi: 10.1038/s41598-020-71268-4

Yang, J., Fei, K., Chen, J., Wang, Z., Zhang, W., and Zhang, J. (2020). Jasmonates alleviate spikelet-opening impairment caused by high temperature stress during anthesis of photo-thermo-sensitive genic male sterile rice lines. *Food Energy Secur.* 6:e233. doi: 10.1002/fes3.233

Zajczkowska, U., Denisow, B., Lotocka, B., Dolkin-Lewko, A., and Rakoczy-Trojanowska, M. (2021). Spikelet movements, anther extrusion and pollen production in wheat cultivars with contrasting tendencies to cleistogamy. *BMC Plant Biol.* 21:136. doi: 10.1186/s12870-021-02917-7

Zeng, X. C., Zhou, X., Zhang, W., Murofushi, N., Kitahara, T., and Kamuro, Y. (1999). Opening of rice floret in rapid response to methyl jasmonate. *J. Plant Growth Regul.* 18, 153–158. doi: 10.1007/PL00007063

Zou, H., Tzarfati, R., Hubner, S., Krugman, T., Fahima, T., Abbo, S., et al. (2015). Transcriptome profiling of wheat glumes in wild emmer, hulled landraces and modern cultivars. *BMC Genom.* 16:777. doi: 10.1186/s12864-015-1996-0

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

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

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 21 June 2022
doi: 10.3389/fpls.2022.895480

Deciphering the Crosstalk Mechanisms of Wheat-Stem Rust Pathosystem: Genome-Scale Prediction Unravels Novel Host Targets

Raghav Kataria[1] *and Rakesh Kaundal*[1,2,3]*

[1] Department of Plants, Soils, and Climate, College of Agriculture and Applied Sciences, Utah State University, Logan, UT, United States, [2] Bioinformatics Facility, Center for Integrated BioSystems, Utah State University, Logan, UT, United States, [3] Department of Computer Science, College of Science, Utah State University, Logan, UT, United States

Triticum aestivum (wheat), a major staple food grain, is affected by various biotic stresses. Among these, fungal diseases cause about 15–20% of yield loss, worldwide. In this study, we performed a comparative analysis of protein-protein interactions between two *Puccinia graminis* races (*Pgt* 21-0 and *Pgt* Ug99) that cause stem (black) rust in wheat. The available molecular techniques to study the host-pathogen interaction mechanisms are expensive and labor-intensive. We implemented two computational approaches (interolog and domain-based) for the prediction of PPIs and performed various functional analysis to determine the significant differences between the two pathogen races. The analysis revealed that *T. aestivum-Pgt* 21-0 and *T. aestivum-Pgt* Ug99 interactomes consisted of ∼90M and ∼56M putative PPIs, respectively. In the predicted PPIs, we identified 115 *Pgt* 21-0 and 34 *Pgt* Ug99 potential effectors that were highly involved in pathogen virulence and development. Functional enrichment analysis of the host proteins revealed significant GO terms and KEGG pathways such as *O*-methyltransferase activity (GO:0008171), regulation of signal transduction (GO:0009966), lignin metabolic process (GO:0009808), plastid envelope (GO:0009526), plant-pathogen interaction pathway (ko04626), and MAPK pathway (ko04016) that are actively involved in plant defense and immune signaling against the biotic stresses. Subcellular localization analysis anticipated the host plastid as a primary target for pathogen attack. The highly connected host hubs in the protein interaction network belonged to protein kinase domain including Ser/Thr protein kinase, MAPK, and cyclin-dependent kinase. We also identified 5,577 transcription factors in the interactions, associated with plant defense during biotic stress conditions. Additionally, novel host targets that are resistant to stem rust disease were also identified. The present study elucidates the functional differences between *Pgt* 21-0 and *Pgt* Ug99, thus providing the researchers with strain-specific information for further experimental validation of the interactions, and the development of durable, disease-resistant crop lines.

Keywords: wheat, stem rust, computational modeling, effectors, interolog method, domain-based approach, disease resistance

OPEN ACCESS

Edited by:
Xiaoli Fan,
Chengdu Institute of Biology (CAS), China

Reviewed by:
Li Long,
Sichuan Agricultural University, China
Xu Hongxing,
Henan University, China

***Correspondence:**
Rakesh Kaundal
rkaundal@usu.edu

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

Received: *13 March 2022*
Accepted: *31 May 2022*
Published: *21 June 2022*

Citation:
Kataria R and Kaundal R (2022) Deciphering the Crosstalk Mechanisms of Wheat-Stem Rust Pathosystem: Genome-Scale Prediction Unravels Novel Host Targets. Front. Plant Sci. 13:895480. doi: 10.3389/fpls.2022.895480

INTRODUCTION

Triticum aestivum L. (family *Poaceae*) is one of the highly cultivated staple food grains, and ranks third in terms of global production, owing to about 35% of the world's food grain produce (Obembe et al., 2021). It contributes significantly to the daily nutrient intake of the human population, thus providing plant-derived proteins, carbohydrates, calories, vitamins, and a wide range of other nutrients (Shewry and Hey, 2015; Miransari and Smith, 2019). The gradual increase in the world's human population leads to increased threat to global food security, which further demands to improve the crop yield substantially to meet the food supply of the world in the future (El Sabagh et al., 2021). Apart from the climate change, the global wheat production is also affected by various unpredictable biotic and abiotic stresses, which further leads to reduced genetic diversity of the crop (Afzal et al., 2015). Annually, the diseases caused by plant-pathogenic fungi lead to yield losses varying from 15 to 20%. Among these pathogenic fungi, the obligately biotrophic rust fungi emerge as a major threat to wheat production, leading to an economic loss of $4.3–5.0 billion dollars (Figueroa et al., 2018).

Stem (black) rust, caused by *Puccinia graminis* f. sp. *tritici* (*Pgt*), is considered as one of the highly destructive diseases of wheat. The occurrences of the disease have also been found in crops such as barley, rye, and other cereals (Dean et al., 2012). The disease can cause enormous yield losses ranging from 50 to 70% or more, depending on the environmental conditions (Saari and Prescott, 1985). Stem rust is also responsible for contraction of grain size, reduced photosynthetic area, diversion of photosynthetic assimilates, and water loss (Willocquet et al., 2021). *Pgt* consists of a wide range of strains, the most significant being the African strain "Ug99" (race TTKSK), which later evolved into variants of its own (Olivera et al., 2012; Li et al., 2019). Another *Pgt* isolate, "21-0," was found in Australia, which has been used for the comparative study of stem rust in wheat. The draft genome of *Pgt* 21-0 was built using reference-based and *de novo* assembly (Upadhyaya et al., 2015). The infection by *Pgt* occurs in a series of steps, typically initiating by the germination of urediniospores on the surface of plant stem, followed by the formation of appressorium, mitosis of nuclei, and differentiation of haustorial mother cells into haustoria, which acquires nutrients from the plant cells (Leonard and Szabo, 2005).

The tremendous losses caused by the fungal pathogens have spurred the researchers to study the in-depth infection mechanism of the pathogen. Various studies have progressed the detection and genetic mapping of genes, and QTLs that confer resistance to *Pgt* in wheat (Duplessis et al., 2011), but a frequent resistance breakdown has been observed, owing to mutations in the *Pgt* isolates (Stokstad, 2007). The fungicides are an effective way against the fungal pathogens, but these pathogens develop resistance against the fungicides/chemicals, and also the fungicides have a negative impact on human health and environment (Van de Wouw et al., 2014). The protein-protein interactions (PPIs) in plant cells perform various functions, involving immune responses against biotic or abiotic stresses. The pathogens secrete effector proteins into the plant cell, sabotage the intercellular mechanisms of the host cell, and cause infection (Garbutt et al., 2014). Thus, the understanding of pathogen infection and the subsequent plant cell defense response is highly crucial. Computational prediction of PPIs reveals relationship among the proteins on a genome-wide scale. Various computational methods exist for the prediction of host-pathogen interactions (HPIs) such as protein sequence homology-based interolog approach, domain-based approach, gene co-expression, phylogenetic profiles, and others (Matthews et al., 2001; Ng et al., 2003; Sun et al., 2007; Piya et al., 2014; Kataria et al., 2022). In the present study, we delineated the PPIs between *T. aestivum* and *Puccinia* species proteins by employing two most widely used computational approaches, i.e., interolog (homology-based) and domain-based approach. Different molecular strategies for PPI detection are available, but those are expensive, time-consuming, and labor-intensive (Chen et al., 2008). Our research is mainly focused on elucidating genome-wide scale PPIs to unravel the complex intermolecular networks of *T. aestivum-Puccinia* interactome.

MATERIALS AND METHODS

Data Source

The whole proteomes of *T. aestivum*, *Pgt* isolate 21-0, and *Pgt* isolate Ug99 were obtained from Ensembl Plants,[1] National Center for Biotechnology Information (NCBI),[2] and Ensembl Fungi,[3] respectively. All the proteomes were analyzed with CD-HIT (Fu et al., 2012) at 100% to cluster the identical proteins. The total number of proteins are described in **Table 1**. In the research analysis, the proteins with prefixes "Traes," "KAA," and "GMQ" refer to *T. aestivum*, *Pgt* 21-0, and *Pgt* Ug99 proteins, respectively.

Interactome Prediction Between *Triticum aestivum* and *Puccinia* Species

The HPIs between *T. aestivum* and *Puccinia* species were predicted using two most widely implemented computational approaches: interolog-based, and domain-based. Interolog method is based on sequence homology that determines the conserved interactions between protein pairs of two species (Nourani et al., 2015). The interolog-based approach employs seven protein-protein interaction (PPI) databases, viz., BioGRID (Chatr-Aryamontri et al., 2017), DIP (Salwinski et al., 2004), HPIDB (Kumar and Nanduri, 2010), IntAct (Kerrien et al., 2012), MINT (Licata et al., 2012), PHI-base (Urban et al., 2020), and STRING (Szklarczyk et al., 2019). The interaction data from these databases was downloaded and implemented locally using SQL. The proteomes of host and pathogen species are aligned against these PPI databases using BLAST v2.7.1, followed by filtering of the results using random BLAST parameter combinations of sequence identity (30, 40, 50, and 60%), sequence coverage (40, 50, 60, and 80%), and *e*-value (1*e*-10, 1*e*-50, 1*e*-05, 1*e*-04, 1*e*-20, 1*e*-30, and 1*e*-25). In the past, there are no substantial reports for selecting an

[1]https://plants.ensembl.org/index.html

[2]https://www.ncbi.nlm.nih.gov/

[3]https://fungi.ensembl.org/index.html

TABLE 1 | Protein datasets used in the study.

Species	Number of proteins	
	Downloaded	Non-redundant
Triticum aestivum	133,346	104,701
Puccinia graminis 21-0 (*Pgt* 21-0)	37,843	35,376
Puccinia graminis Ug99 (*Pgt* Ug99)	24,524	22,524

appropriate combination of BLAST parameters to predict PPIs https://academic.oup.com/bib/article/22/3/bbz162/5842243. A study on human and *Escherichia coli* HPIs determined the homologs using 30% sequence identity, 80% coverage, and *e*-value ≤1*e*-10 (Bose et al., 2017). In Arabidopsis-*Pseudomonas* system, the researchers identified homologs with 80% coverage, 1*e*-04 *e*-value, and 50% identity (Sahu et al., 2014). In our study, using different BLAST parameters (identity, coverage, and *e*-value), 112 combinations were generated, and an optimal combination was selected based on the maximum number of effectors.

On the other hand, in the domain-based approach, three domain-domain interaction (DDI) databases were implemented locally: 3did (Mosca et al., 2014), DOMINE (Raghavachari et al., 2008), and IDDI (Kim et al., 2012). In this method, the PPIs are predicted on the basis of Pfam domain composition. The proteins of host and pathogen were analyzed against Pfam database using "hmmscan" program in HMMER v3.3.1, which identified the significant domains. To filter the results of hmmscan, an *e*-value of 1*e*-23 and coverage 0.2 was used for host proteins, while those of pathogen proteins were filtered with *e*-value and coverage of 1*e*-17 and 0.45, respectively. The identified Pfam domains were then further used for the prediction of PPIs using local SQL queries. The details of number of sequences, and interaction pairs from each database are available in **Supplementary Material 1**, Sheet 1.

Effector and Secretory Proteins Prediction

Effector proteins, secreted by the fungi, interact with host proteins and manipulate the immune responses of host cell (Sonah et al., 2016). The secretory proteins contain a secretion signal peptide, less than 300 amino acids, that employs various cell wall degrading enzymes and phytotoxins to modulate the crucial host cell defense mechanisms (Kim et al., 2016). To identify the effector proteins, we analyzed the proteomes of *Pgt* 21-0 and *Pgt* Ug99 in EffectorP 2.0[4] (Sperschneider et al., 2018), while the secretory proteins were identified using SignalP-5.0[5] (Almagro Armenteros et al., 2019).

Functional Enrichment Analysis of the Proteins

The classification of the proteins into different functional categories such as molecular function, biological process, and cellular component was carried out by obtaining the functional annotation of the proteins. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were conducted using the clusterProfiler (Yu et al., 2012) package in R. GO databases for *T. aestivum* and *Puccinia* species was created locally using *makeOrgPackage* function of the R package "AnnotationForge." GO enrichment was then performed by implying Benjamini and Hochberg test correction method (Benjamini et al., 1995), followed by filtering the enriched terms on adjusted *p*-value cutoff of ≤0.05. Similarly, KEGG enrichment was also conducted at a *p*-value cutoff of 0.05.

Subcellular Localization

According to the studies, a high correlation is observed between the protein function and its subcellular localization, which provide more insights into the protein function (Chi, 2010). The pathogens secrete effector proteins into the host cell, which then translocate to various cellular compartments, and suppress the immune system of the host (Sperschneider et al., 2017). Thus, the prediction of subcellular localization of the host and pathogen proteins is an essential component of the plant-pathogen interaction studies. The subcellular localization of *T. aestivum* proteins was performed using standalone version of Support Vector Machine (SVM)-based tool, Plant-mSubP (Sahu et al., 2021). While for the subcellular localization of *Puccinia* proteins, we employed a deep learning-based tool, DeepLoc 1.0 (Almagro Armenteros et al., 2017).

Comparison Between Host-Pathogen Interactions of *Pgt* 21-0 and *Pgt* Ug99

Different races of *P. graminis* cause stem rust infection in wheat. We were interested in comparing the PPIs between two major strains (*Pgt* 21-0 and *Pgt* Ug99) to uncover the differences between the two fungal species. With regard to this, we identified the orthologs between both the fungal species using OrthoFinder, which implements phylogenetic-based prediction of the orthologs (Emms and Kelly, 2019). The interactions from the orthologs were referred to as common subnetwork. Further, we also focused on the strain-specific functional analysis of the *Puccinia* species. For this, we analyzed the "unique proteins," i.e., the *Puccinia* species proteins that were not the orthologs of each other. This provided us more insights into the functionality of an individual strain.

Network Visualization and Analysis

The protein-protein interaction network is an extensively employed tool to study the functioning of cellular machinery by highlighting the crucial protein complexes and the relationship between the proteins, based on various network parameters such as node degree, centrality, etc. (Agapito et al., 2013). We analyzed the protein networks using the most widely used tool, Cytoscape v3.8.2 (Shannon et al., 1971). Various in-built layout algorithms and styles

[4]http://effectorp.csiro.au/

[5]https://services.healthtech.dtu.dk/service.php?SignalP-5.0

were used to analyze and enhance the visualization of the network.

RESULTS AND DISCUSSION

To predict the protein-protein interactions, the proteomes of *T. aestivum* and *Puccinia* species were randomly paired, followed by the estimation of the interaction probability of an individual pair using interolog, and domain-based approaches. The interactome was predicted using the BLAST parameter combination of 30% sequence identity, 40% sequence coverage, and *e*-value of 1*e*-04. Using both the computational approaches, the predicted interactome for *T.* aestivum-*Pgt* 21-0 consisted of 90,493,282 interactions, whereby 84,125 host proteins interact with 9,022 pathogen proteins, of which 115 proteins were effectors (**Table 2**). While the *T. aestivum-Pgt* Ug99 interactome accounted for 56,755,414 interactions, involving 84,069 host and 5,863 pathogen proteins, consisting of 34 effectors (**Table 3**). The randomly employed (112) BLAST parameter combinations, and the resulting interactions from each combination for *Pgt* 21-0 and *Pgt* Ug99 have been described in **Supplementary Material 1**, Sheets 2, 3, respectively. For clarification, the term "effectors" has been used to represent the pathogen proteins that serve both as effector and secretory proteins.

TABLE 2 | *Triticum aestivum-Pgt* 21-0 interactome.

Interaction database	Number of interactions	Number of host proteins	Number of pathogen proteins
Interolog-based			
BioGRID	22,129,912	53,400	7,214
DIP	2,484,487	27,287	4,839
HPIDB	62,505	6,844	418
IntAct	8,870,790	49,486	6,641
MINT	2,497,106	23,672	5,251
PHI-base	154	7	22
STRING	60,582,077	83,058	5,369
Total (Interolog) (I)	73,877,190	83,821	7,758
Domain-based			
3DID	2,336,648	27,053	4,891
DOMINE	11,148,777	25,649	5,130
IDDI	22,963,441	33,862	5,982
Total (Domain) (II)	27,163,377	35,734	6,305
I and II (combined)	90,493,282	84,125	9,022
I and II (consensus)	10,547,285	31,143	4,689
Interolog (unique)	63,329,905	83,816	7,755
Domain (unique)	16,616,092	34,159	6,190

Total (Interolog) (I): The predicted HPIs from all the seven interolog databases were merged and duplicates were removed.
Total (Domain) (II): The predicted HPIs from all the three domain databases were merged and duplicates were removed.
I and II (combined): The predicted HPIs from both the methods were merged and the duplicates were removed.
I and II (consensus): From both the methods, the consensus of the predicted HPIs was taken and duplicates were removed.
Interolog (unique): The unique HPIs containing the interactions only from interolog-based method.
Domain (unique): The unique HPIs containing the interactions only from domain-based method.

TABLE 3 | *Triticum aestivum-Pgt U*g99 interactome.

Interaction database	Number of interactions	Number of host proteins	Number of pathogen proteins
Interolog-based			
BioGRID	14,058,763	53,443	4,453
DIP	1,613,309	27,005	3,021
HPIDB	41,663	6,883	266
IntAct	5,752,184	48,653	4,122
MINT	1,641,961	23,550	3,286
PHI-base	77	7	11
STRING	38,432,627	83,007	3,620
Total (Interolog) (I)	46,736,430	83,767	5,104
Domain-based			
3DID	1,510,939	27,010	2,942
DOMINE	6,770,500	25,471	3,117
IDDI	13,827,014	33,824	3,602
Total (Domain) (II)	16,528,057	35,737	3,809
I and II (combined)	56,755,414	84,069	5,863
I and II (consensus)	6,509,073	30,901	2,834
Interolog (unique)	40,227,357	83,759	5,100
Domain (unique)	10,018,984	34,272	3,741

Total (Interolog) (I): The predicted HPIs from all the seven interolog databases were merged and duplicates were removed.
Total (Domain) (II): The predicted HPIs from all the three domain databases were merged and duplicates were removed.
I and II (combined): The predicted HPIs from both the methods were merged and the duplicates were removed.
I and II (consensus): From both the methods, the consensus of the predicted HPIs was taken and duplicates were removed.
Interolog (unique): The unique HPIs containing the interactions only from interolog-based method.
Domain (unique): The unique HPIs containing the interactions only from domain-based method.

Puccinia Orthologs Interactome

The ortholog analysis resulted in 1,958 proteins that are orthologs between *Pgt* 21-0 and *Pgt* Ug99. These orthologs were found to interact with 83,340 host proteins, involved in 21,901,125 interactions (referred to as "common subnetwork" throughout the analysis). For the subsequent functional analysis, the interactions from ortholog analysis were taken into consideration.

Highly Connected Protein Hubs

The host-pathogen protein-protein interaction network represents the functional clustering of the interacting proteins, which allows in-depth understanding of a specific protein with respect to the proteins in its surrounding (Wachi et al., 2005; Jonsson and Bates, 2006). The identification of the protein function helps gaining the knowledge of the disease infection mechanism by providing information about various biological processes and molecular mechanisms (Kuzmanov and Emili, 2013). In our study, the proteins hubs were determined from common subnetwork using the metric "node degree." The

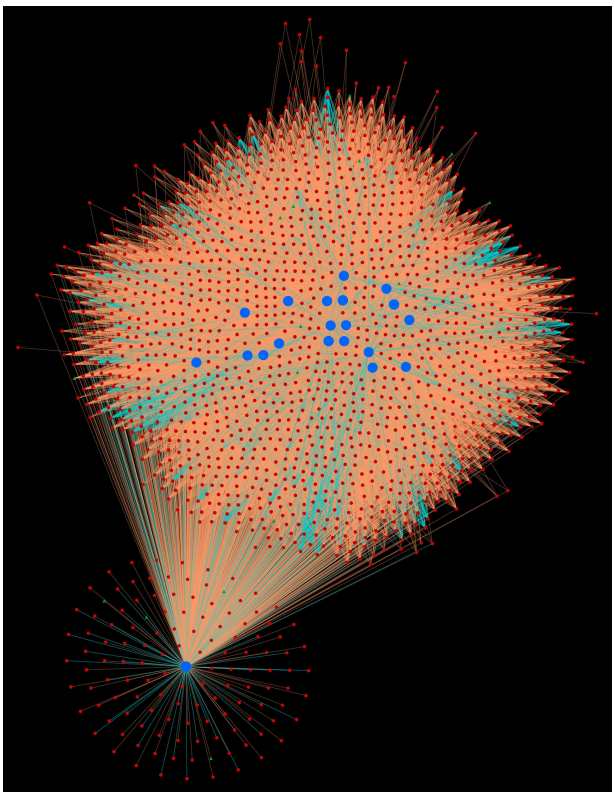


FIGURE 1 | Protein-protein interaction network for top 20 host protein hubs. Blue nodes represent host proteins, red nodes are pathogen proteins, and green nodes are effector proteins. Orange edges depict the interactions from interolog-based approach, while cyan edges belong to domain-based approach.

average degree of host and pathogen proteins was found to be 263 and 11,185, respectively (**Supplementary Material 2**, Sheets 1, 2). The pathogen proteins have higher degree in comparison to host proteins, which is in-line with the host-pathogen protein ratio obtained by the researchers in the past (Li et al., 2012; Kurubanjerdjit et al., 2013). The top 20 protein hubs for each host and pathogen have been discussed below.

Triticum aestivum Protein Hubs

The protein network analysis revealed that majority of the host proteins belonged to protein kinase domain family, of which serine/threonine (Ser/Thr) protein kinase was found to form highly interconnected hubs (TraesCS4D02G250600.1.cds1, TraesCS7A02G437400.1, TraesCS5B02G254600.1.cds1, TraesCS5A02G255500.1.cds1, TraesCS5D02G263800.1.cds1, TraesCS2D02G120200.1, TraesCS3A02G247700.1.cds1, TraesCS4B02G260700.1.cds1, TraesCS3B02G271700.1.cds1, TraesCS4B02G123100.1, and TraesCS4A02G192300.1) (**Figure 1**). The physiological importance of Ser/Thr kinases play a crucial role in the regulation of various environmental stress responses, particularly in signaling pathway (País et al., 2009). SnRK Ser/Thr kinase is divided into 3 subgroups: SnRK1, SnRK2, and SnRK3 (Mao et al., 2010). In wheat, PKABA1 (a member of SnRK2) has been reported to be induced by ABA. The levels of ABA increase during stress conditions, thus showing its role in plant defense (Johnson et al., 2002). Another most interconnected protein (TraesCS6D02G339600.1), interacting with 1,497 pathogen proteins, belonged to heat shock protein 70 (hsp70) family. The members of hsp70 are thought to play a crucial role in different cellular processes during biotic and abiotic stress conditions (Usman et al., 2017). These proteins are also involved in R protein stability, and regulation of immune signaling pathways (Van Ooijen et al., 2010; Park and Seo, 2015). Five host proteins (TraesCS5A02G295800.1, TraesCS7A02G029700.1, TraesCS4A02G336800.2, TraesCS5B02G536500.1, and TraesCS5D02G534000.2) were identified as mitogen-activated protein kinases (MAPKs), which are known to be critical in response to pathogenic infection including the generation of hypersensitive response (HR), defense hormone responses, and ROS signaling (He et al., 2020). Another major hub was formed by the proteins (TraesCS5A02G521700.1, TraesCS4B02G353600.1, and TraesCS4D02G347600.1) that function as cyclin-dependent kinases (CDKs). A study shows the regulation of Arabidopsis resistance against *Alternaria brassicicola* by CDK8 that regulates the intermediates of the secondary metabolites, hydroxycinnamic acid amides (HCAAs), that play a role in fungal resistance. Also, an increased resistance against *Botrytis cinerea* was observed in the *cdk8* mutant (Bessire et al., 2007; Zhu et al., 2014).

Puccinia Protein Hubs

The common subnetwork analysis indicated that the largest hub was formed by aurora kinase that involves the *Puccinia* species protein GMQ_15838T0/KAA1117900.1, interacting with 57,744 host proteins. Various studies reveal the role of aurora kinase in mitotic processes such as chromosome segregation and cytokinesis, thus promoting the growth of fungal pathogens (Tückmantel et al., 2011; Bavetsias and Linardopoulos, 2015). Heat shock protein 70 superfamily forms another major hub involving seven pathogen proteins GMQ_09878T0/KAA1086976.1, GMQ_23673T0/KAA1112990.1, GMQ_10843T0/KAA1076916.1, GMQ_14817T0/KAA1118905.1, GMQ_14422T0/KAA1086735.1, GMQ_13441T0/KAA1119794.1, and GMQ_05646T0/KAA1072403.1. Heat shock proteins are conserved molecular chaperones that play characteristic role in activating essential signal transducers in pathogenic fungi (Tiwari et al., 2015). In *Fusarium pseudograminearum*, 14 *FpHsp70* genes were highly expressed at the time of crown rot infection in wheat. While the knockout of a Hsp70 homolog gene (*FpLhs1*) in ER lumenal resulted in reduction of fungal growth and virulence, implying the role of HSPs in pathogenicity (Chen et al., 2019). The pathogen proteins (GMQ_03421T0/KAA1111489.1, GMQ_12489T0/KAA1117418.1, GMQ_08991T0/KAA1087502.1, GMQ_13869T0/KAA1085281.1, and GMQ_15253T0/KAA1066088.1) were found to function as calcium/calmodulin-dependent kinases (CaMKs). Based on hidden Markov model, the fungal CaMKs are classified into different families (CAMK1, CAMKL, RAD53, and CAKM-Unique), and subfamilies (Kin4,

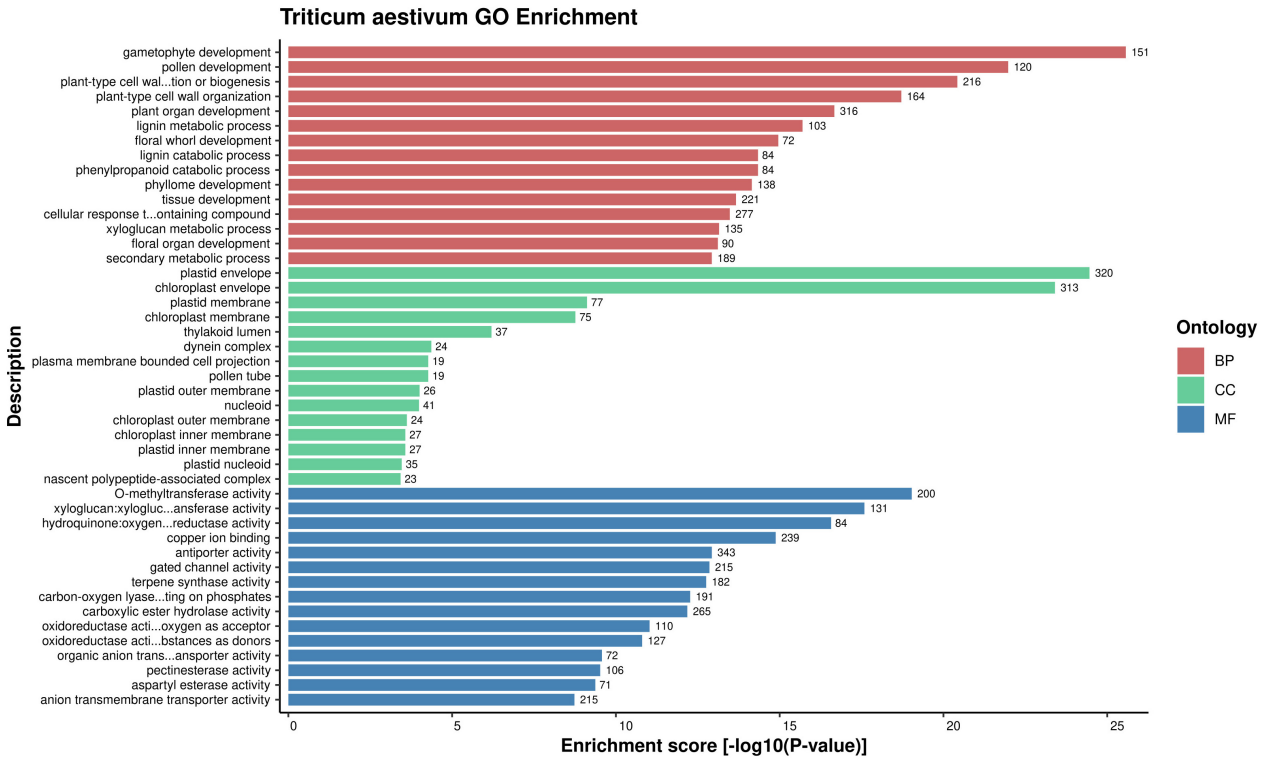


FIGURE 2 | Over-representation of top 15 GO terms from each category (molecular function, cellular component, and biological process) for the host proteins, based on enrichment score.

Kin1, GIN4, PASK, AMPK, CHK1, and MARK) (Goldberg et al., 2013; Jiao et al., 2017). *Fg*Kin1 and *Fg*Kin4 in wheat fungal pathogen, *Fusarium graminearum*, are reported to be responsible for growth and pathogenesis (Wang et al., 2011; Luo et al., 2014). Similarly, the pathogen hubs formed by two Ser/Thr protein kinases: glycogen synthase kinase (GMQ_16722T0/KAA1117762.1), and AGC kinase (GMQ_11492T0/KAA1086570.1) are critical for pathogenicity and development in fungi (Qin et al., 2015; Fabri et al., 2019). The cluster of proteins (GMQ_05648T2/KAA1078421.1, GMQ_24430T0/KAA1114232.1, and GMQ_09263T0/KAA11 11598.1) served as cyclin-dependent kinases (CDKs). Researchers in the past established that the rice blast fungus, *Magnaporthe oryzae*, requires CDK subunit Cks1 for infection-associated development (Yue et al., 2017). Two proteins, GMQ_11353T0/KAA1076537.1 and GMQ_19044T0/KAA1114913.1, were associated with RNA-binding domain/RNA recognition motif. In *Ustilago maydis*, the causal agent of smut disease in corn, RNA-binding proteins were found to be involved in the fungal growth and development during the infection process (Becht et al., 2005).

The protein hubs analysis indicated that the pathogen proteins invade and subvert the host immune machinery, while the host activates various signaling cascades and hormones in response to the pathogen attack. Additionally, the hubs significantly revealed the crucial protein domain families that are involved in the disease infection and defense mechanisms in the pathogen and host, respectively, thus suggesting the cross-talks between the host and pathogen.

Gene Ontology Analysis: Unifying the Biology of Host and Pathogen Proteins

Gene Ontology enrichment analysis is an effective approach of deciphering the underlying biological process, molecular function and cellular component of the proteins of an organism (Tomczak et al., 2018). GO enrichment of the host and pathogen proteins was carried out using enrichment score [-log10(*P*-value)]. The enrichment analysis revealed that 83,340 host proteins in the common subnetwork are involved in 3,570 GO terms, categorized into biological process (2,167), cellular component (408), and molecular function (995) (**Figure 2**). The highly enriched GO terms in different categories involve gametophyte development (GO:0048229), regulation of response to stimulus (GO:0048583), plastid envelope (GO:0009526), chloroplast envelope (GO:0009941), *O*-methyltransferase activity (GO:0008171), xyloglucan:xyloglucosyl transferase activity (GO:0016762), and other significant GO terms (**Supplementary Material 2**, Sheet 3). Various studies have reported the direct or indirect involvement of the above-mentioned significant biological processes/cellular components/molecular functions in plant defense mechanisms. Ubiquitin-conjugating enzymes (E1, E2, and E3) are associated with ubiquitination, which regulates various plant immune signals. In Arabidopsis, UBC22 (E2 subfamily) showed its involvement in female gametophyte

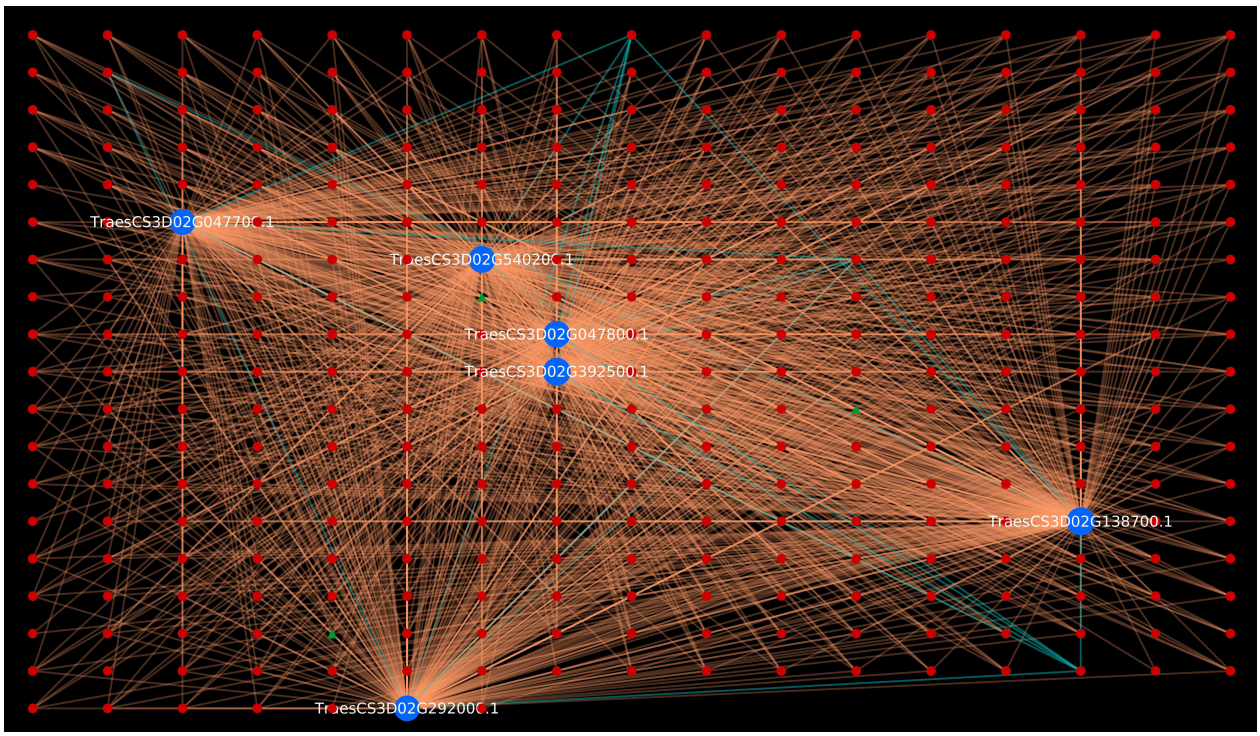


FIGURE 3 | Visualization of six *Triticum aestivum* proteins belonging to chromosome 3D in GO:0008171 (*O*-methyltransferase activity). Blue nodes represent host proteins, red nodes are pathogen proteins, and green nodes are effector proteins. Orange edges depict the interactions from interolog-based approach, while cyan edges belong to domain-based approach.

development, indicating its role in plant defense (Devoto et al., 2003; Wang et al., 2016). The enzyme hydroperoxide lyase (HPL) in plastid envelope is known to catalyze C6-aldehyde that play a role in plant defense. The attack of pathogenic fungi, *B. cinerea*, on Arabidopsis showed an upregulation of AtHPL expression, thus enhancing C6-aldehyde levels, which further inhibited the pathogen growth (Howe and Schilmiller, 2002; Kishimoto et al., 2008; Breuers et al., 2011). Caffeic acid 3-*O*-methyltransferases (COMT) are implicated in biosynthesis of lignin, which provides biotic/abiotic stress resistance to the plants (Bhuiyan et al., 2009; Kataria and Kaundal, 2021). In wheat, the COMT gene (*TaCOMT-3D*) showed significantly high expression level on infection with *Rhizoctonia cerealis*. *TaCOMT-3D* was localized in chromosome 3D (Wang et al., 2018). In our analysis, we identified six wheat proteins (TraesCS3D02G392500.1, TraesCS3D02G540200.1, TraesCS3D02G047700.1, 394TraesCS3D02G047800.1, TraesCS3D02G138700.1, and TraesCS3D02G292000.1) that belong to chromosome 3D and are associated with *O*-methyltransferase activity. These six host proteins were found interacting with 307 pathogen proteins, accounting to 1,611 interactions (**Figure 3**). This provides concrete evidence of the involvement of host proteins in plant defense against fungal attack. Furthermore, lignification also restricts the diffusion of nutrients from host to pathogen, thus suggesting the inability of haustoria to maintain the biotrophic relationship with the host, resulting in reduced pathogen infection.

On the other hand, 1958 pathogen proteins associated with the host proteins were involved in 1,362 GO terms. These include significant GO terms such as proteolysis (GO:0006508), protein peptidyl-prolyl isomerization (GO:0000413), endoplasmic reticulum (GO:0005783), GTPase activity (GO:0003924), and hydrolase activity (GO:0004553). These processes are known to be involved in pathogen virulence and development (Pogány et al., 2015; Pinter et al., 2019). The detailed GO enrichment of *Puccinia* species has been provided in **Supplementary Material 2**, Sheet 4.

Plant Defense and Immune Signaling Pathways During Biotic Stress

The in-depth knowledge of biological pathways of the proteins helps in better understanding of a PPI network. KEGG enrichment analysis of the proteins involved in PPIs was conducted. A total of 399 highly enriched KEGG pathways were obtained for the host proteins in the common subnetwork (**Supplementary Material 2**, Sheet 5). The over-represented pathways include NF-kappa B signaling pathway (ko04064), flavonoid biosynthesis (ko00941), biosynthesis of secondary metabolites (ko01110), plant-pathogen interaction (ko04626), MAPK signaling pathway (ko04016), and a few more significantly enriched pathways related to plant defense mechanism. The top 20 KEGG pathways have been represented in **Figure 4**. The nuclear factor kappa B (NF-κB) transcription factor helps in the regulation of cellular immune responses against diverse

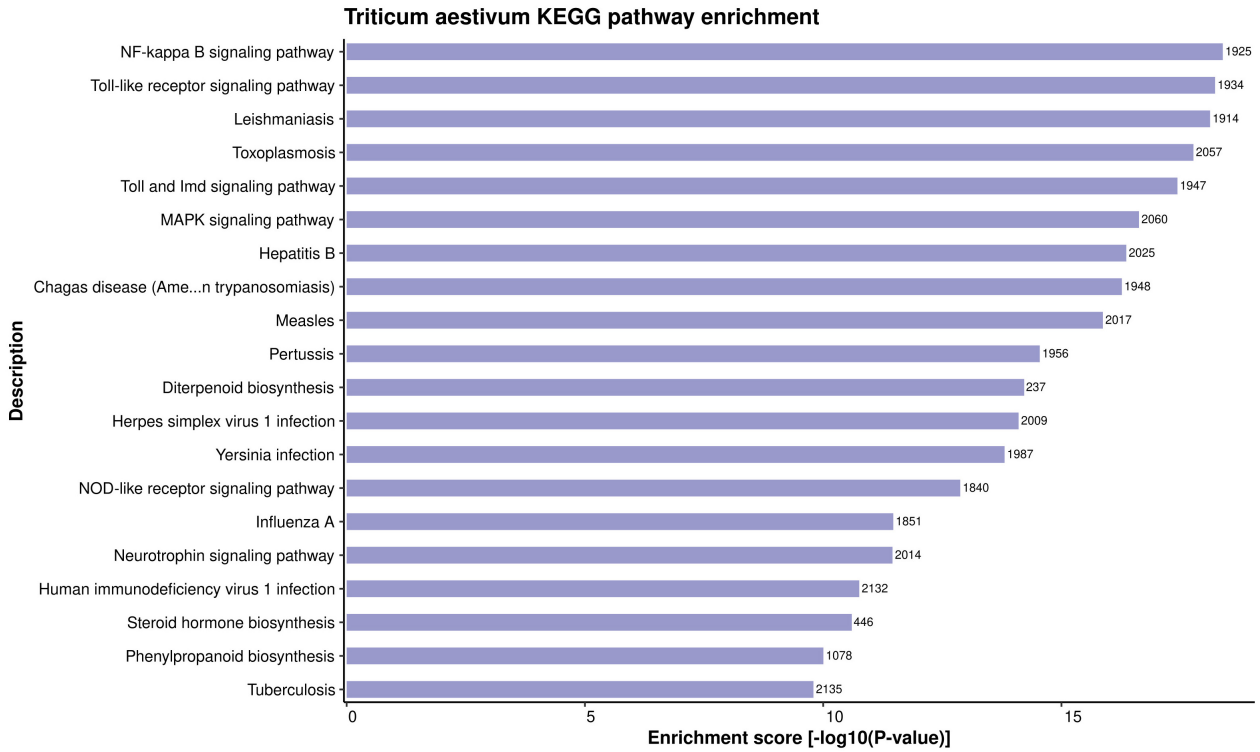


FIGURE 4 | Representation of top 20 enriched KEGG pathways for the host proteins involved in HPIs, based on enrichment score.

environmental stresses (Zhao et al., 2018). The interaction of protein NIM1 with transcription factor NF-κB has been reported to induce systemic acquired resistance (SAR) and gene-for-gene resistance against the disease in Arabidopsis (Ryals et al., 1997). The secondary metabolites are known to play a major role in plant immune responses to external stimuli. In our analysis, around 4,640 host proteins were found to be involved in biosynthesis of secondary metabolites pathway. Previous reports show the activation of secondary metabolites on recognition of the pathogen-secreted effectors by resistance proteins in the host (Ahuja et al., 2012; Piasecka et al., 2015). Flavonoids, a class of secondary metabolites, have been reported to account for plant development and defense responses against pathogens in various crops such as cotton (Mathesius, 2018). Among the aforementioned pathways, the most significant is mitogen-activated protein kinase (MAPK) signaling pathway, which is known to play a critical role in plant immune signaling during various stresses (Zhang and Klessig, 2001). A study revealed the induction of rice MAPKs, OsMKK3, and OsMPK7, during the infection process of *Xanthomonas oryzae* that causes leaf blight disease in rice. The overexpression of OsMKK3 and OsMPK7 genes during pathogenesis also suggested the probable disease resistance mechanism. Also, the silencing of overexpressed OsMPK7 lead to disease susceptible plants (Jalmi and Sinha, 2016). Also, transcription factors are known to be activated by MAPKs by the process of phosphorylation, thus regulating the immune response against the pathogens by integrating defense signals from various MAPKs (Nadal-Ribelles et al., 2019).

We also performed the KEGG pathway enrichment of the pathogen proteins, which revealed important pathways such as MAPK signaling pathway (ko04016), biosynthesis of secondary metabolites (ko01110), and others, which have a direct or indirect relationship with pathogenicity (**Supplementary Material 2**, Sheet 6). The host proteins are also involved in these pathways, which suggests the potential interaction of host and pathogen proteins during plant defense response.

Thus, the functional enrichment analysis suggested significant molecular processes and biological pathways in which the host and pathogen proteins are involved. A comprehensive analysis of the enriched processes/pathways can further enhance the study of host-pathogen interaction mechanism, and other related biological processes occurring within the host cell.

Plastid: A Primary Target for Pathogen Attack in the Host

The proteins are translocated to various subcellular compartments, where they perform specific biological functions, thus revealing the physiology of the cell. Some proteins are also distributed to multiple cellular locations, depending on the sorting signal (Briesemeister et al., 2010). The subcellular localization is also known to be statistically correlated with the protein function, and its gene expression levels (Garapati et al., 2020). A significant number of novel proteins have been identified using sequencing technologies, but their subcellular location remains unknown.

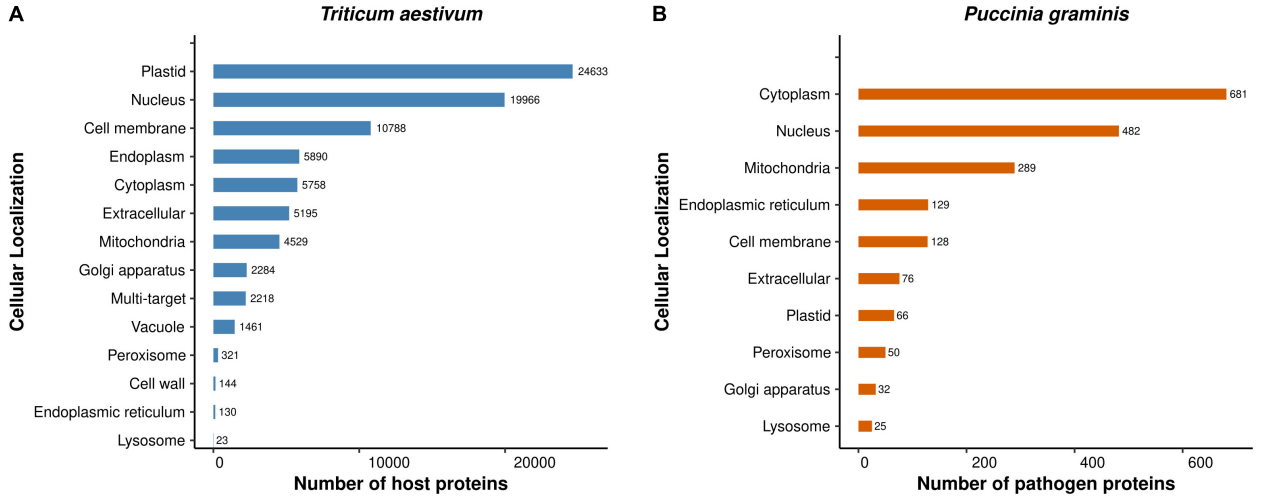


FIGURE 5 | Subcellular Localization of the **(A)** host proteins, and **(B)** pathogen proteins involved in the common subnetwork.

TABLE 4 | Stem rust resistance genes and QTLs mapped on various *T. aestivum* chromosomes.

	Gene/QTL	Chromosome number	References
Stem rust resistance genes	*Sr2*	3BS	Spielmeyer et al., 2003
	Sr12	3B	Crossa et al., 2007
	Sr19	2B	
	Sr23	2B	
	Sr26	6AL	
	Sr31	1BS	
	Sr36	2B	
	Sr40	2B	
	Sr13	6AL	Admassu et al., 2011
	Sr22	7AL	Upadhyaya et al., 2014
	Sr28	2BL	Rouse et al., 2012
	Sr30	5DL	Hiebert et al., 2010
	Sr42	6DS	Ghazvini et al., 2012
Stem rust resistance QTLs identified	*QSr-sparc-2B*	2BS	Bokore et al., 2021
	QSr-sparc-7A	7AL	
	QSr-sparc-5A	5AL	
	QSr-sparc-6A	6AS	
	QSr-sparc-7B	7BL	

We predicted the sequence-based subcellular localization of the *T. aestivum* and *Puccinia* proteins to have a better understanding of the occurrence of PPIs in a particular subcellular compartment. The subcellular localization analysis classified the *T. aestivum* proteins into 14 categories: plastid (29.56%), nucleus (23.96%), cell membrane (12.94%), endoplasm (7.07%), cytoplasm (6.91%), extracellular (6.23%), mitochondria (5.43%), golgi apparatus (2.74%), multi-target (2.66%), vacuole (1.75%), peroxisome (0.38%), cell wall (0.17%), endoplasmic reticulum (0.16%), and lysosome (0.03%) (**Figure 5A**). 2,218 proteins were found to be multi-target (moonlighting proteins), performing specific functions in various cellular organelles. 24,633 and 19,966 host proteins were localized in plastid and nucleus, respectively. Researchers in the past have reported the presence of host proteins in plastid, which plays an essential role in intracellular signaling pathways (de Dios Barajas-López et al., 2013; Caplan et al., 2015). Another study demonstrated the localization of rice OsVQ domain proteins in plastid and nucleus using rice protoplast system. OsVQ proteins are considered to be the co-regulators during immune response against biotic stress (Kim et al., 2013).

On the other hand, the *Puccinia* species proteins were localized in cytoplasm (34.78%), nucleus (24.62%), mitochondria (14.76%), endoplasmic reticulum (6.54%), cell membrane (6.54%), extracellular (3.88%), plastid (3.37%), peroxisome (2.55%), golgi apparatus (1.63%), and lysosome (1.28%) (**Figure 5B**). In *B. cinerea*, two ubiquitin-like (UBL) activating enzymes, BcAtg3 (E2) and BcAtg7 (E1), were determined to be localized in cytoplasm (Ren et al., 2018), which is coherent to our localization prediction. The detailed information of the predicted subcellular localization of the host and pathogen proteins is available in **Supplementary Material 2**, Sheets 7, 8, respectively. Furthermore, we were also interested in predicting the location of interactions between the host and pathogen proteins. The analysis indicated that the host proteins are mostly targeted by the pathogen proteins in the plastid, whereby 24,237 host proteins interact with 674 pathogen proteins, accounting to 2,630,496 interactions (**Supplementary Material 2**, Sheet 9).

Pathogenicity of Effectors During Stem Rust Infection

The rust fungal pathogens secrete effectors (using specialized structures known as haustoria) which subvert the host cell immune machinery, followed by enhancing their pathogenicity in the host cell (Ramachandran et al., 2017). Therefore, understanding the behavior of the effector proteins is a crucial step in the HPI analysis. From the common subnetwork, we

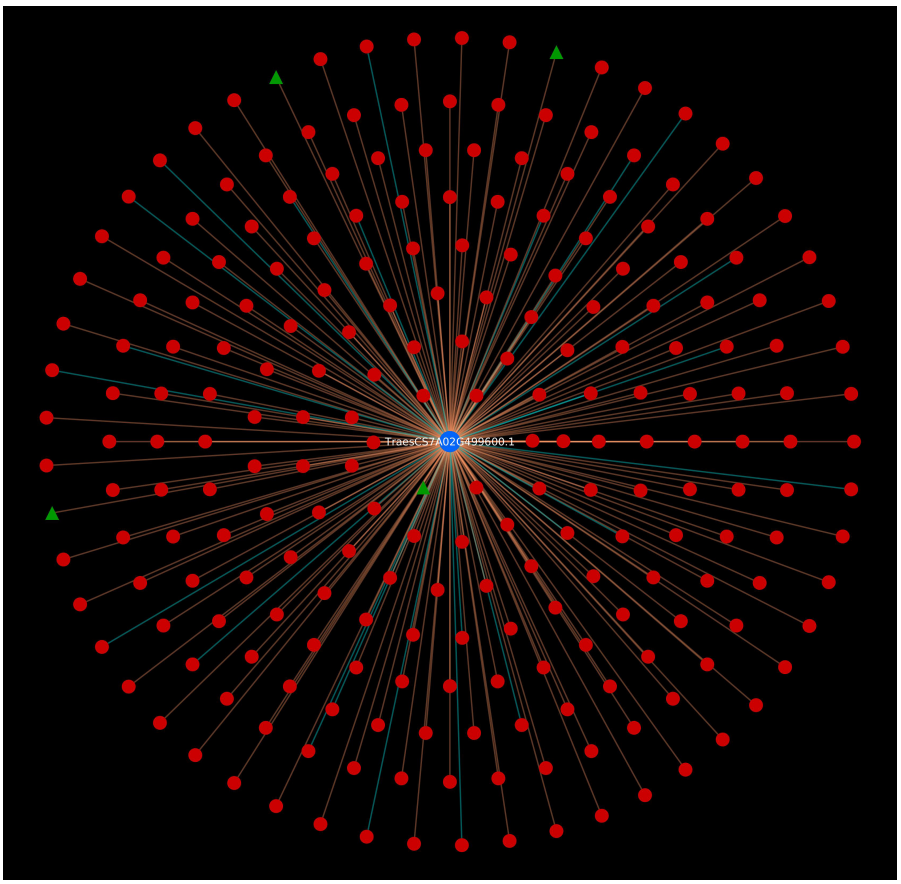


FIGURE 6 | Wheat stem resistance gene *Sr22* encoded protein. Blue nodes represent host protein, red nodes are pathogen proteins, and green nodes are effector proteins. Orange edges depict the interactions from interolog-based approach, while cyan edges belong to domain-based approach.

identified 18 effectors interacting with 43,054 host proteins, resulting in 156,529 interactions (**Supplementary Material 3**, Sheet 1). Among the 43,054 host proteins involved in interactions with effectors, 3367 were identified to be transcription factors.

The functional analysis suggested that the effector proteins in the interactions are highly enriched in superoxide metabolic process (GO:0006801), protein dephosphorylation (GO:0006470), vesicle (GO:0031982), phosphoric ester hydrolase activity (GO:0042578), and metabolic pathways (ko01100). These processes are involved in enhancing pathogenicity, development, and secretion during the host-pathogen interaction mechanism (Rodrigues et al., 2011; Tamayo et al., 2016; Rafiei et al., 2021), which helps in the survival of the pathogen under various stresses in the host cell. The host proteins associated with the effectors were found to be involved in secondary metabolic process (GO:0019748), cell wall polysaccharide metabolic process (GO:0010383), response to external biotic stimulus (GO:0043207), plant-pathogen interaction pathway (ko04626), and plant hormone signal transduction (ko04075). These processes are actively related to the plant defense and immune signaling against the biotic stresses. Thus, the predicted interactions of host proteins with the effectors can be considered of high confidence, and potential candidates for further studying the infection mechanism of stem rust in wheat.

Functional Differences Between *Pgt* 21-0 and *Pgt* Ug99

Furthermore, we were interested in deciphering the strain-specific functionalities of the *Pgt* 21-0 and *Pgt* Ug99 proteins involved in the PPIs. Since these proteins were not the orthologs of each other, hence these are referred to as unique proteins. The *T. aestivum-Pgt* 21-0 interactome predicted 68,465,557 PPIs, involving 83,962 host proteins and 7,063 unique *Pgt* 21-0 proteins, of which 100 proteins served as effectors. While the *T. aestivum-Pgt* Ug99 interactome involved 83,495 host proteins and 3,905 *Pgt* Ug99 proteins (16 effectors), accounting to 34,854,274 interactions. The low number of pathogens and predicted PPIs in *Pgt* Ug99 interactome as compared to that of *Pgt* 21-0 suggests the virulence of the pathogen, and that fewer pathogen proteins (mainly effectors) are capable of causing the infection.

Unique *Puccinia graminis* 21-0

The GO enrichment analysis of the unique pathogen proteins suggested that most of the GO terms were similar to that of

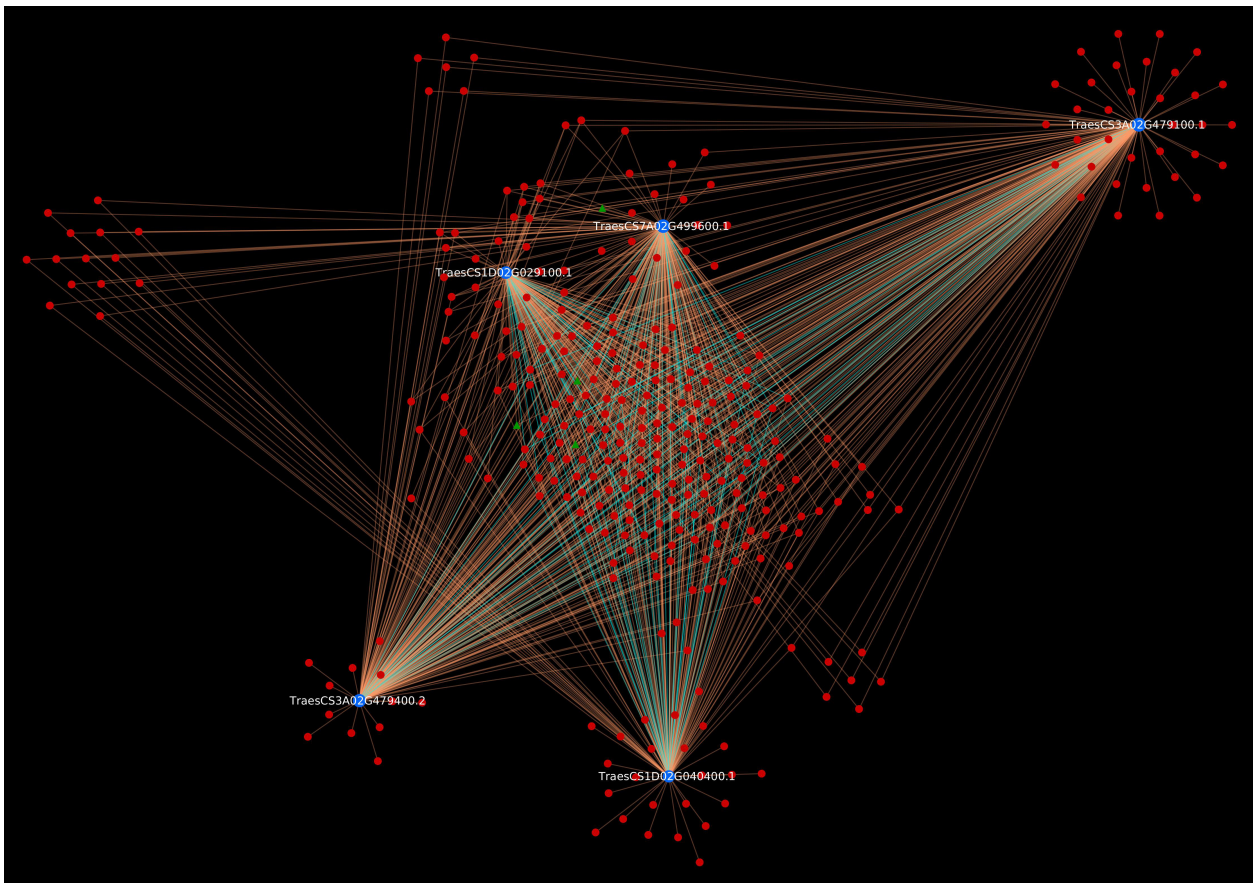


FIGURE 7 | Stem rust resistance gene complex (*Sr22*, *Sr33*, *Sr35*, and *Sr45*) causing resistance against TTKSK race. Blue nodes represent host proteins, red nodes are pathogen proteins, and green nodes are effector proteins. Orange edges depict the interactions from interolog-based approach, while cyan edges belong to domain-based approach.

the common subnetwork. But we also found 180 significant GO terms that were unique to *Pgt* 21-0 proteins. These included cell wall modification (GO:0042545), NADH dehydrogenase complex assembly (GO:0010257), cyclin-dependent protein kinase holoenzyme complex (GO:0000307), and SUMO ligase complex (GO:0106068). The unique KEGG pathways such as carotenoid biosynthesis (ko00906), and plant hormone signal transduction (ko4075) were found to be highly over-represented. Researchers in past have reported the direct or indirect role of these GO terms/KEGG pathways in pathogen virulence (Avalos et al., 2017). The detailed information of the significant GO terms and KEGG pathways for *Pgt* 21-0 is available in **Supplementary Material 3**, Sheets 2, 3, respectively. For the host proteins interacting with the unique *Pgt* 21-0 proteins, only 1 significant GO term (condensed chromosome kinetochore; GO:0000777) was identified. While no unique KEGG pathway was obtained for the host proteins.

Unique *Puccinia graminis* Ug99

In comparison with the common subnetwork, the unique *Pgt* Ug99 were highly enriched in 201 GO terms, including oxidoreductase activity (GO:0016491), regulation of response to stress (GO:0080134), snoRNA binding (GO:0030515), and GTPase complex (GO:1905360). The over-represented KEGG pathways obtained for these proteins involve monoterpenoid biosynthesis (ko00902), and polyketide sugar unit biosynthesis (ko00523), which have been shown to regulate the fungal growth and development (Chiang et al., 2009; Dallery et al., 2019; Noar et al., 2019). The significant GO terms and over-represented KEGG pathways have been detailed in **Supplementary Material 3**, Sheets 4, 5, respectively. While no significant GO terms or KEGG pathways were found for the host associated with unique *Pgt* Ug99 proteins.

Role of Transcription Factors in Plant Defense

Recent molecular studies have elucidated the role of transcription factors (TFs) in diverse cellular mechanisms such as regulating gene expression, act as transcriptional activators or repressors, and in plant defense (Seo and Choi, 2015). The plant immune signal activation is tightly controlled by the gene-specific transcription factors that bind to *cis*-elements in the promoter

region (Li et al., 2016). In line with this, we predicted the wheat proteins that serve as transcription factors using PlantTFDB v5.0 (Jin et al., 2017). This resulted in 5,577 wheat proteins that served as transcription factors in the common subnetwork, involved in 1,311,301 interactions with 1,600 *Puccinia* ortholog proteins. These transcription factors were classified into 55 TF families, of which 28 families are significantly involved in biotic and abiotic stresses. The major TF families include basic helix-loop-helix (bHLH), ethylene responsive factor (ERF), myeloblastosis related (MYB), WRKY, basic leucine zipper (bZIP), and NAM, ATAF1/2, and CUC2 (NAM). The host proteins and their respective TF family has been described in **Supplementary Material 4** (Sheet 1). Researchers in the past have extensively demonstrated the crucial role of various transcription factors in diverse biological processes, and significant immune signaling pathways in response to plant defense against pathogen attack (Asai et al., 2002; Wu et al., 2009; Zander et al., 2010; Pieterse et al., 2012; Zhao et al., 2012).

Novel Stem Rust-Resistant Host Targets

Our study on the host-pathogen interaction system focuses on understanding the disease infection mechanism, host immune response, and identifying the host targets that show resistance against stem rust disease. The resistance (*R*) genes in host are responsible for the recognition of effector proteins (secreted by pathogens), followed by the initiation of immune responses. According to gene-for-gene hypothesis, a successful resistant response requires two genes: *R* gene in the host and corresponding avirulence (*Avr*) effector gene in the pathogen, which makes resistance dependent on the specific pathogen strain (Flor, 1971). The mutations in *Avr* genes leads to the inability of the corresponding *R* genes to recognize the *Avr* genes, thus resulting in the pathogen to overcome host resistance (Ellis et al., 2014). The recent advancement in plant breeding techniques, in conjunction with increasing genomic resources, has accelerated the identification (and cloning) process of wheat resistance genes (Andersen et al., 2020). Scientists also created a wheat *R*-gene atlas to facilitate the research community with an efficient resource of resistance genes in wheat, aiming at reducing the pathogen co-evolution (Hafeez et al., 2021).

In wheat, a total of 46 *R* genes are officially designated to show resistance against stem rust, of which only 20 belong to *T. aestivum* (Leonard and Szabo, 2005). A few of the identified *R* genes in wheat include *Sr5*, *Sr13*, *Sr23*, *Sr27*, *Sr36*, *Sr40*, etc., which have varying effect on *Pgt* races. In wheat, *R* gene *Sr5* is known to limit the growth of avirulent *Puccinia* strain, while *Sr22* advances the development of *Pgt* races (Hatta et al., 2020; Wu et al., 2020). Various researchers mapped *Sr* genes on wheat chromosomes 1BS, 2B, 3B, 5DL, 6AL, 6DS, and 7AL (**Table 4**), conferring resistance against stem rust during adult-plant stage. A QTL-based study on wheat identified stable QTLs: *QSr-sparc-2B*, *QSr-sparc-7A*, *QSr-sparc-5A*, *QSr-sparc-6A*, and *QSr-sparc-7B* on chromosome 2BS, 7AL, 5AL, 6AS, and 7BL, respectively (Bokore et al., 2021). Further, we identified the PPIs associated with these chromosomes, which accounted to 9,387,396 PPIs, involving 35,809 host and 1,956 pathogen proteins. Of the 1,956 pathogen proteins in the interactions, 18 proteins were identified as effectors, involved in 66,772 PPIs (**Supplementary Material 4**, Sheet 2). The maximum number of interactions (1,305,326 PPIs) were identified on chromosome 2B, on which five *Sr* genes (*Sr19*, *Sr23*, *Sr28*, *Sr36*, and *Sr40*) have been mapped. The plant immune response against various pathogens is mediated by nucleotide-binding and leucine-rich repeat (NLR) domain proteins. Among the identified 35,809 host proteins, 2,123 proteins were found to be associated with NLR domain.

To have deeper insight into the resistance mechanism, we analyzed *Sr22* gene located on chromosome 7A and cloned using MutRenSeq (Steuernagel et al., 2016). In the predicted interactome, the protein encoded by this gene was found to be interacting with 230 pathogen proteins (230 PPIs). The host protein was actively involved in various plant defense pathways such as MAPK signaling, plant-pathogen interaction, and plant hormone signal transduction pathway. Additionally, 4 effectors were also identified (**Figure 6**). It has also been reported that the stem rust resistance genes (*Sr22*, *Sr33*, *Sr35*, and *Sr45*) form a complex that effectively confers resistance against TTKSK in wheat (Hatta et al., 2021). Further, we identified the interactions associated with this complex in the predicted interactome, which resulted in 1,051 PPIs, involving 327 pathogen proteins, of which 4 were effectors (**Figure 7**). The host proteins involved in the interactions can be considered as the novel targets for the breeders for development of disease-resistant lines. Further, the interaction of these novel host proteins with the effectors are the potential candidate pairs to understand the immune responses against the fungal pathogen attack during stem rust disease, thus giving deeper insights to the infection mechanism and host defense responses.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

RKu formulated and designed the research and contributed to writing – review and editing, supervision, project administration, and funding acquisition. RKt analyzed the data, developed prediction models, performed functional analysis of the data, validations, literature mining, etc., and contributed to writing – original draft preparation. Both authors contributed to visualization, read, and agreed to the published version of the manuscript.

FUNDING

We acknowledged the support to this study from Utah Agricultural Experiment Station (UAES), Utah State University, seed grant #A48025 to RKu. This research was also partially supported by UAES office and approved as journal paper number

9587. The funding body did not play any role in the design of this study; the collection, analysis, or interpretation of data; or in the writing of this manuscript.

ACKNOWLEDGMENTS

We appreciate the team members of KAABiL (Kaundal Artificial intelligence & Advanced Bioinformatics Lab) for their valuable time and inputs at the time of research. We profoundly thank the referees for the suggestions to improve the research article.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.895480/full#supplementary-material

REFERENCES

Admassu, B., Perovic, D., Friedt, W., and Ordon, F. (2011). Genetic mapping of the stem rust (*Puccinia graminis* f. sp. tritici Eriks. & E. Henn) resistance gene Sr13 in wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 122, 643–648. doi: 10.1007/s00122-010-1433-3

Afzal, F., Chaudhari, S. K., Gul, A., Farooq, A., Ali, H., Nisar, S., et al. (2015). "Bread wheat (*Triticum aestivum* L.) under biotic and abiotic stresses: an overview," in *Crop Production and Global Environmental Issues*, ed. K. R. Hakeem (Cham: Springer International Publishing), 293–317. doi: 10.1007/978-3-319-23162-4_13

Agapito, G., Guzzi, P. H., and Cannataro, M. (2013). Visualization of protein interaction networks: problems and solutions. *BMC Bioinformatics* 14:S1. doi: 10.1186/1471-2105-14-S1-S1

Ahuja, I., Kissen, R., and Bones, A. M. (2012). Phytoalexins in defense against pathogens. *Trends Plant Sci.* 17, 73–90. doi: 10.1016/j.tplants.2011.11.002

Almagro Armenteros, J. J., Sønderby, C. K., Sønderby, S. K., Nielsen, H., and Winther, O. (2017). DeepLoc: prediction of protein subcellular localization using deep learning. *Bioinformatics* 33, 3387–3395. doi: 10.1093/bioinformatics/btx431

Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., et al. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* 37, 420–423. doi: 10.1038/s41587-019-0036-z

Andersen, E., Lindsey, L., and Nepal, M. (2020). Genome-wide identification of disease resistance genes (R Genes) in wheat. *bioRxiv* [Preprint] doi: 10.1101/2020.07.18.210286

Asai, T., Tena, G., Plotnikova, J., Willmann, M. R., Chiu, W. L., Gomez-Gomez, L., et al. (2002). Map kinase signalling cascade in *Arabidopsis* innate immunity. *Nature* 415, 977–983. doi: 10.1038/415977a

Avalos, J., Pardo-Medina, J., Parra-Rivero, O., Ruger-Herreros, M., Rodríguez-Ortiz, R., Hornero-Méndez, D., et al. (2017). Carotenoid biosynthesis in *Fusarium*. *J. Fungi* 3, 39. doi: 10.3390/jof3030039

Bavetsias, V., and Linardopoulos, S. (2015). Aurora kinase inhibitors: current status and outlook. *Front. Oncol.* 5:278. doi: 10.3389/fonc.2015.00278

Becht, P., Vollmeister, E., and Feldbrügge, M. (2005). Role for RNA-binding proteins implicated in pathogenic development of *Ustilago maydis*. *Eukaryot. Cell* 4, 121–133. doi: 10.1128/EC.4.1.121-133.2005

Benjamini, Y., Hochberg, Y., Benjamini, R., and Yoav, H. Y. (1995). Benjamini and Y FDR.pdf. *J. R. Stat. Soc. Ser. B* 57, 289–300.

Bessire, M., Chassot, C., Jacquat, A. C., Humphry, M., Borel, S., Petétot, J. M. D. C., et al. (2007). A permeable cuticle in *Arabidopsis* leads to a strong resistance to *Botrytis cinerea*. *EMBO J.* 26, 2158–2168. doi: 10.1038/sj.emboj.7601658

Bhuiyan, N. H., Selvaraj, G., Wei, Y., and King, J. (2009). Gene expression profiling and silencing reveal that monolignol biosynthesis plays a critical role in penetration defence in wheat against powdery mildew invasion. *J. Exp. Bot.* 60, 509–521. doi: 10.1093/jxb/ern290

Bokore, F. E., Cuthbert, R. D., Hiebert, C. W., Fetch, T. G., Pozniak, C. J., N'Diaye, A., et al. (2021). Mapping stem rust resistance loci effective in Kenya in Canadian spring wheat (*Triticum aestivum* L.) lines 'AAC Prevail' and 'BW961.'. *Can. J. Plant Pathol.* 43, S263–S274. doi: 10.1080/07060661.2021.1966651

Bose, T., Venkatesh, K. V., and Mande, S. S. (2017). Computational analysis of host–pathogen protein interactions between humans and different strains of enterohemorrhagic *Escherichia coli*. *Front. Cell. Infect. Microbiol.* 7:128. doi: 10.3389/fcimb.2017.00128

Breuers, F. K. H., Bräutigam, A., and Weber, A. P. M. (2011). The plastid outer envelope - A highly dynamic interface between plastid and cytoplasm. *Front. Plant Sci.* 2:97. doi: 10.3389/fpls.2011.00097

Briesemeister, S., Rahnenführer, J., and Kohlbacher, O. (2010). Going from where to why-interpretable prediction of protein subcellular localization. *Bioinformatics* 26, 1232–1238. doi: 10.1093/bioinformatics/btq115

Caplan, J. L., Kumar, A. S., Park, E., Padmanabhan, M. S., Hoban, K., Modla, S., et al. (2015). Chloroplast stromules function during innate immunity. *Dev. Cell* 34, 45–57. doi: 10.1016/j.devcel.2015.05.011

Chatr-Aryamontri, A., Oughtred, R., Boucher, L., Rust, J., Chang, C., Kolas, N. K., et al. (2017). The BioGRID interaction database: 2017 update. *Nucleic Acids Res.* 45, D369–D379. doi: 10.1093/nar/gkw1102

Chen, L., Geng, X., Ma, Y., Zhao, J., Chen, W., Xing, X., et al. (2019). The ER lumenal Hsp70 protein FpLhs1 is important for conidiation and plant infection in *Fusarium pseudograminearum*. *Front. Microbiol.* 10:1401. doi: 10.3389/fmicb.2019.01401

Chen, P. Y., Deane, C. M., and Reinert, G. (2008). Predicting and validating protein interactions using network structure. *PLoS Comput. Biol.* 4:e1000118. doi: 10.1371/journal.pcbi.1000118

Chi, S. M. (2010). Prediction of protein subcellular localization by weighted gene ontology terms. *Biochem. Biophys. Res. Commun.* 399, 402–405. doi: 10.1016/j.bbrc.2010.07.086

Chiang, Y. M., Szewczyk, E., Davidson, A. D., Keller, N., Oakley, B. R., and Wang, C. C. C. (2009). A gene cluster containing two fungal polyketide synthases encodes the biosynthetic pathway for a polyketide, asperfuranone, in *Aspergillus nidulans*. *J. Am. Chem. Soc.* 131, 2965–2970. doi: 10.1021/ja8088185

Crossa, J., Burgueño, J., Dreisigacker, S., Vargas, M., Herrera-Foessel, S. A., Lillemo, M., et al. (2007). Association analysis of historical bread wheat germplasm using additive genetic covariance of relatives and population structure. *Genetics* 177, 1889–1913. doi: 10.1534/genetics.107.078659

Dallery, J. F., Adelin, É, Le Goff, G., Pigné, S., Auger, A., Ouazzani, J., et al. (2019). H3K4 trimethylation by CclA regulates pathogenicity and the production of three families of terpenoid secondary metabolites in *Colletotrichum higginsianum*. *Mol. Plant Pathol.* 20, 831–842. doi: 10.1111/mpp.12795

de Dios Barajas-López, J., Blanco, N. E., and Strand, Å (2013). Plastid-to-nucleus communication, signals controlling the running of the plant cell. *Biochim. Biophys. Acta Mol. Cell Res.* 1833, 425–437. doi: 10.1016/j.bbamcr.2012.06.020

Dean, R., Van Kan, J. A. L., Pretorius, Z. A., Hammond-Kosack, K. E., Di Pietro, A., Spanu, P. D., et al. (2012). The Top 10 fungal pathogens in molecular plant pathology. *Mol. Plant Pathol.* 13, 414–430. doi: 10.1111/j.1364-3703.2011.00783.x

Devoto, A., Muskett, P. R., and Shirasu, K. (2003). Role of ubiquitination in the regulation of plant defence against pathogens. *Curr. Opin. Plant Biol.* 6, 307–311. doi: 10.1016/S1369-5266(03)00060-8

Duplessis, S., Cuomo, C. A., Lin, Y. C., Aerts, A., Tisserant, E., Veneault-Fourrey, C., et al. (2011). Obligate biotrophy features unraveled by the genomic analysis of rust fungi. *Proc. Natl. Acad. Sci. U.S.A.* 108, 9166–9171. doi: 10.1073/pnas.1019315108

El Sabagh, A., Islam, M. S., Skalicky, M., Ali Raza, M., Singh, K., Anwar Hossain, M., et al. (2021). Salinity stress in wheat (*Triticum aestivum* L.) in the changing climate: adaptation and management strategies. *Front. Agron.* 3:1661932. doi: 10.3389/fagro.2021.661932

Ellis, J. G., Lagudah, E. S., Spielmeyer, W., and Dodds, P. N. (2014). The past, Present and future of breeding rust resistant wheat. *Front. Plant Sci.* 5:641. doi: 10.3389/fpls.2014.00641

Emms, D. M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20, 1–14. doi: 10.1186/s13059-019-1832-y

Fabri, J. H. T. M., Godoy, N. L., Rocha, M. C., Munshi, M., Cocio, T. A., Von Zeska Kress, M. R., et al. (2019). The AGC kinase YpkA regulates sphingolipids biosynthesis and physically interacts with SakA MAP kinase in *Aspergillus fumigatus*. *Front. Microbiol.* 10:3347. doi: 10.3389/fmicb.2018.03347

Figueroa, M., Hammond-Kosack, K. E., and Solomon, P. S. (2018). A review of wheat diseases—a field perspective. *Mol. Plant Pathol.* 19, 1523–1536. doi: 10.1111/mpp.12618

Flor, H. H. (1971). Current status of the gene-fob-gene concept. *Annu. Rev. Phytopathol.* 9, 275–296.

Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28, 3150–3152. doi: 10.1093/bioinformatics/bts565

Garapati, H. S., Male, G., and Mishra, K. (2020). Predicting subcellular localization of proteins using protein-protein interaction data. *Genomics* 112, 2361–2368. doi: 10.1016/j.ygeno.2020.01.007

Garbutt, C. C., Bangalore, P. V., Kannar, P., and Mukhtar, M. S. (2014). Getting to the edge: protein dynamical networks as a new frontier in plant-microbe interactions. *Front. Plant Sci.* 5:312. doi: 10.3389/fpls.2014.00312

Ghazvini, H., Hiebert, C. W., Zegeye, T., Liu, S., Dilawari, M., Tsilo, T., et al. (2012). Inheritance of resistance to Ug99 stem rust in wheat cultivar Norin 40 and genetic mapping of Sr42. *Theor. Appl. Genet.* 125, 817–824. doi: 10.1007/s00122-012-1874-y

Goldberg, J. M., Griggs, A. D., Smith, J. L., Haas, B. J., Wortman, J. R., and Zeng, Q. (2013). Kinannote, a computer program to identify and classify members of the eukaryotic protein kinase superfamily. *Bioinformatics* 29, 2387–2394. doi: 10.1093/bioinformatics/btt419

Hafeez, A. N., Arora, S., Ghosh, S., Gilbert, D., Bowden, R. L., and Wulff, B. B. H. (2021). Creation and judicious application of a wheat resistance gene atlas. *Mol. Plant* 14, 1053–1070. doi: 10.1016/j.molp.2021.05.014

Hatta, M. A. M., Arora, S., Ghosh, S., Matny, O., Smedley, M. A., Yu, G., et al. (2021). The wheat Sr22, Sr33, Sr35 and Sr45 genes confer resistance against stem rust in barley. *Plant Biotechnol. J.* 19, 273–284. doi: 10.1111/pbi.13460

Hatta, M. A. M., Ghosh, S., Athiyannan, N., Richardson, T., Steuernagel, B., Yu, G., et al. (2020). Extensive genetic variation at the sr22 wheat stem rust resistance gene locus in the grasses revealed through evolutionary genomics and functional analyses. *Mol. Plant Microbe Interact.* 33, 1286–1298. doi: 10.1094/MPMI-01-20-0018-R

He, X., Wang, C., Wang, H., Li, L., and Wang, C. (2020). The function of MAPK cascades in response to various stresses in horticultural plants. *Front. Plant Sci.* 11:952. doi: 10.3389/fpls.2020.00952

Hiebert, C. W., Fetch, T. G., and Zegeye, T. (2010). Genetics and mapping of stem rust resistance to Ug99 in the wheat cultivar Webster. *Theor. Appl. Genet.* 121, 65–69. doi: 10.1007/s00122-010-1291-z

Howe, G. A., and Schilmiller, A. L. (2002). Oxylipin metabolism in response to stress. *Curr. Opin. Plant Biol.* 5, 230–236. doi: 10.1016/S1369-5266(02)00250-9

Jalmi, S. K., and Sinha, A. K. (2016). Functional involvement of a mitogen activated protein kinase module, OsMKK3-OsMPK7-OsWRK30 in mediating resistance against *Xanthomonas oryzae* in rice. *Sci. Rep.* 6:37974. doi: 10.1038/srep37974

Jiao, M., Yu, D., Tan, C., Guo, J., Lan, D., Han, E., et al. (2017). Basidiomycete-specific PsCaMKL1 encoding a CaMK-like protein kinase is required for full virulence of *Puccinia striiformis* f. sp. Tritici. *Environ. Microbiol.* 19, 4177–4189. doi: 10.1111/1462-2920.13881

Jin, J., Tian, F., Yang, D. C., Meng, Y. Q., Kong, L., Luo, J., et al. (2017). PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Res.* 45, D1040–D1045. doi: 10.1093/nar/gkw982

Johnson, R. R., Wagner, R. L., Verhey, S. D., and Walker-Simmons, M. K. (2002). The abscisic acid-responsive kinase PKABA1 interacts with a seed-specific abscisic acid response element-binding factor, TaABF, and phosphorylates TaABF peptide sequences. *Plant Physiol.* 130, 837–846. doi: 10.1104/pp.001354

Jonsson, P. F., and Bates, P. A. (2006). Global topological features of cancer proteins in the human interactome. *Bioinformatics* 22, 2291–2297. doi: 10.1093/bioinformatics/btl390

Kataria, R., Duhan, N., and Kaundal, R. (2022). Computational systems biology of alfalfa – bacterial blight host-pathogen interactions: uncovering the complex molecular networks for developing durable disease resistant crop. *Front. Plant Sci.* 12:807354. doi: 10.3389/fpls.2021.807354

Kataria, R., and Kaundal, R. (2021). alfaNET: a database of alfalfa-bacterial stem blight protein–protein interactions revealing the molecular features of the disease-causing bacteria. *Int. J. Mol. Sci.* 22:8342. doi: 10.3390/ijms22158342

Kerrien, S., Aranda, B., Breuza, L., Bridge, A., Broackes-Carter, F., Chen, C., et al. (2012). The IntAct molecular interaction database in 2012. *Nucleic Acids Res.* 40, 841–846. doi: 10.1093/nar/gkr1088

Kim, D. Y., Kwon, S. I., Choi, C., Lee, H., Ahn, I., Park, S. R., et al. (2013). Expression analysis of rice VQ genes in response to biotic and abiotic stresses. *Gene* 529, 208–214. doi: 10.1016/j.gene.2013.08.023

Kim, K. T., Jeon, J., Choi, J., Cheong, K., Song, H., Choi, G., et al. (2016). Kingdom-wide analysis of fungal small secreted proteins (SSPs) reveals their potential role in host association. *Front. Plant Sci.* 7:186. doi: 10.3389/fpls.2016.00186

Kim, Y., Min, B., and Yi, G. S. (2012). IDDI: integrated domain-domain interaction and protein interaction analysis system. *Proteome Sci.* 10:S9. doi: 10.1186/1477-5956-10-S1-S9

Kishimoto, K., Matsui, K., Ozawa, R., and Takabayashi, J. (2008). Direct fungicidal activities of C6-aldehydes are important constituents for defense responses in *Arabidopsis* against *Botrytis cinerea*. *Phytochemistry* 69, 2127–2132. doi: 10.1016/j.phytochem.2008.04.023

Kumar, R., and Nanduri, B. (2010). HPIDB - a unified resource for host-pathogen interactions. *BMC Bioinformatics* 11:S16. doi: 10.1186/1471-2105-11-16

Kurubanjerdjit, N., Tsai, J. J. P., Sheu, C. Y., and Ng, K. L. (2013). The prediction of protein-protein interaction of A. thaliana and *X. campestris* pv. campestris based on protein domain and interolog approaches. *Plant Omics* 6, 388–398.

Kuzmanov, U., and Emili, A. (2013). Protein-protein interaction networks: probing disease mechanisms using model systems. *Genome Med.* 5, 1–12. doi: 10.1186/gm441

Leonard, K. J., and Szabo, L. J. (2005). Stem rust of small grains and grasses caused by *Puccinia graminis*. *Mol. Plant Pathol.* 6, 99–111. doi: 10.1111/j.1364-3703.2005.00273.x

Li, B., Meng, X., Shan, L., and He, P. (2016). Transcriptional regulation of pattern-triggered immunity in plants. *Cell Host Microbe* 19, 641–650. doi: 10.1016/j.chom.2016.04.011

Li, F., Upadhyaya, N. M., Sperschneider, J., Matny, O., Nguyen-Phuc, H., Mago, R., et al. (2019). Emergence of the Ug99 lineage of the wheat stem rust pathogen through somatic hybridisation. *Nat. Commun.* 10:5068. doi: 10.1038/s41467-019-12927-7

Li, Z. G., He, F., Zhang, Z., and Peng, Y. L. (2012). Prediction of protein-protein interactions between *Ralstonia solanacearum* and *Arabidopsis* thaliana. *Amino Acids* 42, 2363–2371. doi: 10.1007/s00726-011-0978-z

Licata, L., Briganti, L., Peluso, D., Perfetto, L., Iannuccelli, M., Galeota, E., et al. (2012). MINT, the molecular interaction database: 2012 Update. *Nucleic Acids Res.* 40, 857–861. doi: 10.1093/nar/gkr930

Luo, Y., Zhang, H., Qi, L., Zhang, S., Zhou, X., Zhang, Y., et al. (2014). FgKin1 kinase localizes to the septal pore and plays a role in hyphal growth, ascospore germination, pathogenesis, and localization of Tub1 beta-tubulins in *Fusarium graminearum*. *New Phytol.* 204, 943–954. doi: 10.1111/nph.12953

Mao, X., Zhang, H., Tian, S., Chang, X., and Jing, R. (2010). TaSnRK2.4, an SNF1-type serine/threonine protein kinase of wheat (*Triticum aestivum* L.), confers enhanced multistress tolerance in *Arabidopsis*. *J. Exp. Bot.* 61, 683–696. doi: 10.1093/jxb/erp331

Mathesius, U. (2018). Flavonoid functions in plants and their interactions with other organisms. *Plants* 7, 7–9. doi: 10.3390/plants7020030

Matthews, L. R., Vaglio, P., Reboul, J., Ge, H., Davis, B. P., Garrels, J., et al. (2001). Identification of potential interaction networks using sequence-based searches for conserved protein-protein interactions or "interologs.". *Genome Res.* 11, 2120–2126. doi: 10.1101/gr.205301

Miransari, M., and Smith, D. (2019). Sustainable wheat (*Triticum aestivum* L.) production in saline fields: a review. *Crit. Rev. Biotechnol.* 39, 999–1014. doi: 10.1080/07388551.2019.1654973

Mosca, R., Céol, A., Stein, A., Olivella, R., and Aloy, P. (2014). 3did: a catalog of domain-based interactions of known three-dimensional structure. *Nucleic Acids Res.* 42, 374–379. doi: 10.1093/nar/gkt887

Nadal-Ribelles, M., Solé, C., Martínez-Cebrián, G., Posas, F., and de Nadal, E. (2019). "Shaping the transcriptional landscape through MAPK signaling," in

Gene Expression and Control, ed. F. Uchiumi (Rijeka: IntechOpen), doi: 10.5772/intechopen.80634

Ng, S. K., Zhang, Z., and Tan, S. H. (2003). Integrative approach for computationally inferring protein domain interactions. *Bioinformatics* 19, 923–929. doi: 10.1093/bioinformatics/btg118

Noar, R. D., Thomas, E., Xie, D. Y., Carter, M. E., Ma, D., and Daub, M. E. (2019). A polyketide synthase gene cluster associated with the sexual reproductive cycle of the banana pathogen, *Pseudocercospora fijiensis*. *PLoS One* 14:e0220319. doi: 10.1371/journal.pone.0220319

Nourani, E., Khunjush, F., and Durmus, S. (2015). Computational approaches for prediction of pathogen-host protein-protein interactions. *Front. Microbiol.* 6:94. doi: 10.3389/fmicb.2015.00094

Obembe, O. S., Hendricks, N. P., and Tack, J. (2021). Decreased wheat production in the USA from climate change driven by yield losses rather than crop abandonment. *PLoS One* 16:e0252067. doi: 10.1371/journal.pone.0252067

Olivera, P. D., Badebo, A., Xu, S. S., Klindworth, D. L., and Jin, Y. (2012). Resistance to race TTKSK of *Puccinia graminis* f. sp. tritici in Emmer wheat. *Crop Sci.* 52, 2234–2242. doi: 10.2135/cropsci2011.12.0645

País, S. M., Téllez-Iñón, M. T., and Capiati, D. A. (2009). Serine/threonine protein phosphatases type 2A and their roles in stress signaling. *Plant Signal. Behav.* 4, 1013–1015. doi: 10.4161/psb.4.11.9783

Park, C. J., and Seo, Y. S. (2015). Heat shock proteins: a review of the molecular chaperones for plant immunity. *Plant Pathol. J.* 31, 323–333. doi: 10.5423/PPJ.RW.08.2015.0150

Piasecka, A., Jedrzejczak-Rey, N., and Bednarek, P. (2015). Secondary metabolites in plant innate immunity: conserved function of divergent chemicals. *New Phytol.* 206, 948–964. doi: 10.1111/nph.13325

Pieterse, C. M. J., Van Der Does, D., Zamioudis, C., Leon-Reyes, A., and Van Wees, S. C. M. (2012). Hormonal modulation of plant immunity. *Annu. Rev. Cell Dev. Biol.* 28, 489–521. doi: 10.1146/annurev-cellbio-092910-154055

Pinter, N., Andrea, C., Id, H., Id, M. H., Rekhter, D., Id, K. Z., et al. (2019). Signal peptide peptidase activity connects the unfolded protein response to plant defense suppression by *Ustilago maydis*. *PLoS Pogthog.* 15:e1007734.

Piya, S., Shrestha, S. K., Binder, B., Neal Stewart, C., and Hewezi, T. (2014). Protein-protein interaction and gene co-expression maps of ARFs and Aux/IAAs in *Arabidopsis*. *Front. Plant Sci.* 5:744. doi: 10.3389/fpls.2014.00744

Pogány, M., Dankó, T., Kámán-Tóth, E., Schwarczinger, I., and Bozsó, Z. (2015). Regulatory proteolysis in *Arabidopsis*-Pathogen interactions. *Int. J. Mol. Sci.* 16, 23177–23194. doi: 10.3390/ijms161023177

Qin, J., Wang, G., Jiang, C., Xu, J. R., and Wang, C. (2015). Fgk3 glycogen synthase kinase is important for development, pathogenesis, and stress responses in *Fusarium graminearum*. *Sci. Rep.* 5:8504. doi: 10.1038/srep08504

Rafiei, V., Vélëz, H., and Tzelepis, G. (2021). The role of glycoside hydrolases in phytopathogenic fungi and oomycetes virulence. *Int. J. Mol. Sci.* 22:9359. doi: 10.3390/ijms22179359

Raghavachari, B., Tasneem, A., Przytycka, T. M., and Jothi, R. (2008). DOMINE: a database of protein domain interactions. *Nucleic Acids Res.* 36, 656–661. doi: 10.1093/nar/gkm761

Ramachandran, S. R., Yin, C., Kud, J., Tanaka, K., Mahoney, A. K., Xiao, F., et al. (2017). Effectors from wheat rust fungi suppress multiple plant defense responses. *Phytopathology* 107, 75–83. doi: 10.1094/PHYTO-02-16-0083-R

Ren, W., Sang, C., Shi, D., Song, X., Zhou, M., and Chen, C. (2018). Ubiquitin-like activating enzymes BcAtg3 and BcAtg7 participate in development and pathogenesis of *Botrytis cinerea*. *Curr. Genet.* 64, 919–930. doi: 10.1007/s00294-018-0810-3

Rodrigues, M. L., Nosanchuk, J. D., Schrank, A., Vainstein, M. H., Casadevall, A., and Nimrichter, L. (2011). Vesicular transport systems in fungi. *Future Microbiol.* 6, 1371–1381. doi: 10.2217/fmb.11.112

Rouse, M. N., Nava, I. C., Chao, S., Anderson, J. A., and Jin, Y. (2012). Identification of markers linked to the race Ug99 effective stem rust resistance gene Sr28 in wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 125, 877–885. doi: 10.1007/s00122-012-1879-6

Ryals, J., Weymann, K., Lawton, K., Friedrich, L., Ellis, D., Steiner, H. Y., et al. (1997). The *arabidopsis* NIM1 protein shows homology to the mammalian transcription factor inhibitor IκB. *Plant Cell* 9, 425–439. doi: 10.1105/tpc.9.3.425

Saari, E. E., and Prescott, J. M. (1985). *World Distribution in Relation to Economic Losses*. Cambridge, MA: Academic Press, Inc, doi: 10.1016/b978-0-12-148402-6.50017-1

Sahu, S. S., Loaiza, C. D., and Kaundal, R. (2021). Plant-mSubP: a computational framework for the prediction of single- And multi-target protein subcellular localization using integrated machine-learning approaches. *AoB Plants* 12, 1–10. doi: 10.1093/AOBPLA/PLZ068

Sahu, S. S., Weirick, T., and Kaundal, R. (2014). Predicting genome-scale *Arabidopsis-Pseudomonas* syringae interactome using domain and interolog-based approaches. *BMC Bioinformatics* 15:S13. doi: 10.1186/1471-2105-15-S11-S13

Salwinski, L., Miller, C. S., Smith, A. J., Pettit, F. K., Bowie, J. U., and Eisenberg, D. (2004). The database of interacting proteins: 2004 update. *Nucleic Acids Res.* 32, 449–451. doi: 10.1093/nar/gkh086

Seo, E., and Choi, D. (2015). Functional studies of transcription factors involved in plant defenses in the genomics era. *Brief. Funct. Genomics* 14, 260–267. doi: 10.1093/bfgp/elv011

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (1971). Cytoscape: a software environment for integrated models. *Genome Res.* 13:426. doi: 10.1101/gr.1239303.metabolite

Shewry, P. R., and Hey, S. J. (2015). The contribution of wheat to human diet and health. *Food Energy Secur.* 4, 178–202. doi: 10.1002/FES3.64

Sonah, H., Deshmukh, R. K., and Bélanger, R. R. (2016). Computational prediction of effector proteins in fungi: opportunities and challenges. *Front. Plant Sci.* 7:126. doi: 10.3389/fpls.2016.00126

Sperschneider, J., Catanzariti, A. M., Deboer, K., Petre, B., Gardiner, D. M., Singh, K. B., et al. (2017). LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. *Sci. Rep.* 7:44598. doi: 10.1038/srep44598

Sperschneider, J., Dodds, P. N., Gardiner, D. M., Singh, K. B., and Taylor, J. M. (2018). Improved prediction of fungal effector proteins from secretomes with EffectorP 2.0. *Mol. Plant Pathol.* 19, 2094–2110. doi: 10.1111/mpp.12682

Spielmeyer, W., Sharp, P. J., and Lagudah, E. S. (2003). Identification and validation of markers linked to broad-spectrum stem rust resistance gene Sr2 in wheat (*Triticum aestivum* L.). *Crop Sci.* 43, 333–336. doi: 10.2135/cropsci2003.3330

Steuernagel, B., Periyannan, S. K., Hernández-Pinzón, I., Witek, K., Rouse, M. N., Yu, G., et al. (2016). Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. *Nat. Biotechnol.* 34, 652–655. doi: 10.1038/nbt.3543

Stokstad, E. (2007). Plant pathology. Deadly wheat fungus threatens world's breadbaskets. *Science* 315, 1786–1787. doi: 10.1126/science.315.5820.1786

Sun, J., Li, Y., and Zhao, Z. (2007). Phylogenetic profiles for the prediction of protein-protein interactions: how to select reference organisms? *Biochem. Biophys. Res. Commun.* 353, 985–991. doi: 10.1016/j.bbrc.2006.12.146

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 47, D607–D613. doi: 10.1093/nar/gky1131

Tamayo, D., Muñoz, J. F., Lopez, Á, Urán, M., Herrera, J., Borges, C. L., et al. (2016). Identification and Analysis of the role of superoxide dismutases isoforms in the *Pathogenesis* of *Paracoccidioides* spp. *PLoS Negl. Trop. Dis.* 10:e0004481. doi: 10.1371/journal.pntd.0004481

Tiwari, S., Thakur, R., and Shankar, J. (2015). Role of heat-shock proteins in cellular function and in the biology of fungi. *Biotechnol. Res. Int.* 2015, 1–11. doi: 10.1155/2015/132635

Tomczak, A., Mortensen, J. M., Winnenburg, R., Liu, C., Alessi, D. T., Swamy, V., et al. (2018). Interpretation of biological experiments changes with evolution of the Gene Ontology and its annotations. *Sci. Rep.* 8:5115. doi: 10.1038/s41598-018-23395-2

Tückmantel, S., Greul, J. N., Janning, P., Brockmeyer, A., Grütter, C., Simard, J. R., et al. (2011). Identification of *Ustilago maydis* Aurora kinase as a novel antifungal target. *ACS Chem. Biol.* 6, 926–933. doi: 10.1021/cb200112y

Upadhyaya, N. M., Garnica, D. P., Karaoglu, H., Sperschneider, J., Nemri, A., Xu, B., et al. (2015). Comparative genomics of australian isolates of the wheat stem rust pathogen *puccinia graminis* f. Sp. Tritici reveals extensive polymorphism in candidate effector genes. *Front. Plant Sci.* 5:759. doi: 10.3389/fpls.2014.00759

Upadhyaya, N. M., Mago, R., Staskawicz, B. J., Ayliffe, M. A., Ellis, J. G., and Dodds, P. N. (2014). A bacterial type III secretion assay for delivery of fungal effector proteins into wheat. *Mol. Plant Microbe Interact.* 27, 255–264. doi: 10.1094/MPMI-07-13-0187-FI

Urban, M., Cuzick, A., Seager, J., Wood, V., Rutherford, K., Venkatesh, S. Y., et al. (2020). PHI-base: the pathogen-host interactions database. *Nucleic Acids Res.* 48, D613–D620. doi: 10.1093/nar/gkz904

Usman, M. G., Rafii, M. Y., Martini, M. Y., Yusuff, O. A., Ismail, M. R., and Miah, G. (2017). Molecular analysis of Hsp70 mechanisms in plants and their function in response to stress. *Biotechnol. Genet. Eng. Rev.* 33, 26–39. doi: 10.1080/02648725.2017.1340546

Van de Wouw, A. P., Marcroft, S. J., Ware, A., Lindbeck, K., Khangura, R., and Howlett, B. J. (2014). Breakdown of resistance to the fungal disease, blackleg, is averted in commercial canola (*Brassica napus*) crops in Australia. *Field Crop. Res.* 166, 144–151. doi: 10.1016/j.fcr.2014.06.023

Van Ooijen, G., Lukasik, E., Van Den Burg, H. A., Vossen, J. H., Cornelissen, B. J. C., and Takken, F. L. W. (2010). The small heat shock protein 20 RSI2 interacts with and is required for stability and function of tomato resistance protein I-2. *Plant J.* 63, 563–572. doi: 10.1111/j.1365-313X.2010.04260.x

Wachi, S., Yoneda, K., and Wu, R. (2005). Interactome-transcriptome analysis reveals the high centrality of genes differentially expressed in lung cancer tissues. *Bioinformatics* 21, 4205–4208. doi: 10.1093/bioinformatics/bti688

Wang, C., Zhang, S., Hou, R., Zhao, Z., Zheng, Q., Xu, Q., et al. (2011). Functional analysis of the kinome of the wheat scab fungus *Fusarium graminearum*. *PLoS Pathog.* 7:e1002460. doi: 10.1371/journal.ppat.1002460

Wang, M., Zhu, X., Wang, K., Lu, C., Luo, M., Shan, T., et al. (2018). A wheat caffeic acid 3-O-methyltransferase TaCOMT-3D positively contributes to both resistance to sharp eyespot disease and stem mechanical strength. *Sci. Rep.* 8:6543. doi: 10.1038/s41598-018-24884-0

Wang, S., Cao, L., and Wang, H. (2016). *Arabidopsis* ubiquitin-conjugating enzyme UBC22 is required for female gametophyte development and likely involved in Lys11-linked ubiquitination. *J. Exp. Bot.* 67, 3277–3288. doi: 10.1093/jxb/erw142

Willocquet, L., Rossi, V., and Savary, S. (2021). Simulation modelling of yield losses caused by wheat stem rust. *Plant Pathol.* 71, 544–555. doi: 10.1111/ppa.13488

Wu, X. X., Lin, Q. J., Ni, X. Y., Sun, Q., Chen, R. Z., Xu, X. F., et al. (2020). Characterization of wheat monogenic lines with known sr genes and wheat lines with resistance to the ug99 race group for resistance to prevalent races of *puccinia graminis* f. sp. tritici in China. *Plant Dis.* 104, 1939–1943. doi: 10.1094/PDIS-12-19-2736-RE

Wu, Y., Deng, Z., Lai, J., Zhang, Y., Yang, C., Yin, B., et al. (2009). Dual function of *Arabidopsis* ATAF1 in abiotic and biotic stress responses. *Cell Res.* 19, 1279–1290. doi: 10.1038/cr.2009.108

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). ClusterProfiler: an R package for comparing biological themes among gene clusters. *Omi. A J. Integr. Biol.* 16, 284–287. doi: 10.1089/omi.2011.0118

Yue, X., Que, Y., Deng, S., Xu, L., Oses-Ruiz, M., Talbot, N. J., et al. (2017). The cyclin dependent kinase subunit Cks1 is required for infection-associated development of the rice blast fungus *Magnaporthe oryzae*. *Environ. Microbiol.* 19, 3959–3981. doi: 10.1111/1462-2920.13796

Zander, M., La Camera, S., Lamotte, O., Métraux, J. P., and Gatz, C. (2010). *Arabidopsis* thaliana class-II TGA transcription factors are essential activators of jasmonic acid/ethylene-induced defense responses. *Plant J.* 61, 200–210. doi: 10.1111/j.1365-313X.2009.04044.x

Zhang, S., and Klessig, D. F. (2001). MAPK cascades in plant defense signaling. *Trends Plant Sci.* 6, 520–527. doi: 10.1016/S1360-1385(01)02103-3

Zhao, M., Joy, J., Zhou, W., De, S., Wood, W. H., Becker, K. G., et al. (2018). Transcriptional outcomes and kinetic patterning of gene expression in response to NF-κB activation. *PLoS Biol.* 16:e2006347. doi: 10.1371/journal.pbio.2006347

Zhao, Y., Wei, T., Yin, K. Q., Chen, Z., Gu, H., Qu, L. J., et al. (2012). *Arabidopsis* RAP2.2 plays an important role in plant resistance to Botrytis cinerea and ethylene responses. *New Phytol.* 195, 450–460. doi: 10.1111/j.1469-8137.2012.04160.x

Zhu, Y., Schluttenhoffer, C. M., Wang, P., Fu, F., Thimmapuram, J., Zhu, J. K., et al. (2014). CYCLIN-DEPENDENT KINASE8 differentially regulates plant immunity to fungal pathogens through kinase-dependent and -independent functions in *Arabidopsis*. *Plant Cell* 26, 4149–4170. doi: 10.1105/tpc.114.128611

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

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

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 15 July 2022
doi: 10.3389/fpls.2022.943669

Characterization of Two Wheat-*Thinopyrum ponticum* Introgression Lines With Pyramiding Resistance to Powdery Mildew

Mingzhu Li[1,2†], *Yuanyuan Yuan*[3†], *Fei Ni*[1], *Xingfeng Li*[1], *Honggang Wang*[1] *and Yinguang Bao*[1*]

[1] State Key Laboratory of Crop Biology, Agronomy College of Shandong Agricultural University, Tai'an, China, [2] Bureau of Agriculture and Rural Affairs of Linqing, Liaocheng, China, [3] Crop Research Institute, Jinan Academy of Agricultural Sciences, Jinan, China

Powdery mildew is one of the most devastating foliar diseases in wheat production. The wild relative *Thinopyrum ponticum* ($2n = 10x = 70$) has been widely used in wheat genetic improvement due to its superior resistance to both biotic and abiotic stresses. In the present study, two wheat-*Th. ponticum* introgression lines named SN0293-2 and SN0293-7 were developed from the progenies of a cross between the octoploid *Trititrigia* SNTE20 and common wheat, including the elite cultivar Jimai 22. They had a novel powdery mildew resistance gene (temporarily named *PmSN0293*) putatively from *Th. ponticum* pyramided with *Pm2* and *Pm52*, exhibiting excellent *Pm* resistance at both the seedling and adult stages. Sequential GISH-FISH detected no signal of *Th. ponticum* in these two lines but a pair of T1BL·1RS in SN0293-2. Chromosomal structural variations were also observed obviously in SN0293-2 and SN0293-7. Through the Wheat 660K SNP array, 157 SNPs, 134 of which were on 6A, were found to be specific to *Th. ponticum*. Based on the data combined with DNA re-sequencing, seven specific markers, including one CAPS marker on 2B and six CAPS and Indel markers on 6A, were developed, confirming their wheat-*Th. ponticum* introgression nature. Furthermore, the two lines displayed positive plant height and produced more kernels and higher 1,000-grain weight. Excellent resistance with desirable agronomic traits makes them valuable in wheat breeding programs.

Keywords: wheat, powdery mildew, *Thinopyrum ponticum*, introgression, GISH, FISH

OPEN ACCESS

Edited by:
Pengtao Ma,
Yantai University, China

Reviewed by:
Huagang He,
Jiangsu University, China
Diaoguo An,
Institute of Genetics and Developmental Biology (CAS), China

****Correspondence:***
Yinguang Bao
baoyinguang@163.com

[†] These authors have contributed equally to this work

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

Received: *14 May 2022*
Accepted: *02 June 2022*
Published: *15 July 2022*

Citation:
Li M, Yuan Y, Ni F, Li X, Wang H and Bao Y (2022) Characterization of Two Wheat-Thinopyrum ponticum Introgression Lines With Pyramiding Resistance to Powdery Mildew. Front. Plant Sci. 13:943669. doi: 10.3389/fpls.2022.943669

INTRODUCTION

As one of the most important staple food crops, wheat provides more than 20% of the calories and the protein for the world's population (Braun et al., 2010). The need for the production of wheat is continuously growing because of the increasing population (Hawkesford et al., 2013). However, it is threatened by many diseases which reduce yield and decrease quality.

Powdery mildew, caused by *Blumeria graminis* f. sp. *tritici* (*Bgt*), is one of the most devastating wheat diseases across the world. It can affect wheat photosynthesis and consequently decrease plant growth and grain filling, resulting in yield reductions (Zhang et al., 2020). Although fungicides can effectively control this disease, they also cause environmental pollution and cost increases. Host resistance is considered to be the most economical and environment-friendly means to do so

(Wang et al., 2005; Liu et al., 2017). Therefore, it is a key step to explore and utilize resistance genes in wheat breeding programs. To date, more than 100 designated powdery mildew (*Pm*) resistance genes/alleles in 63 loci (*Pm*1-*Pm68*) have been documented (He et al., 2021; Gao et al., 2022). Some of them have been cloned, such as the broad-spectrum resistance gene *Pm21* from *Dasypyrum villosum* (2*n*= 2x = 14, VV) (He et al., 2018; Xing et al., 2018). However, with the rapid evolution of new *Bgt* isolates, the single resistance gene is easily overcome. For instance, the well-known *Pm8* on rye (*Secale cereale* L., 2*n* = 2x = 14, RR) chromosome arm 1RS has lost its function against new *Bgt* isolates, such as E09 (Ren et al., 2017). Hence, it is of great necessity to explore new *Pm* genes and pyramid multiple ones to broaden the resistance spectrum.

Wild relatives of common wheat carry many valuable genes that can be used for wheat improvement. The tall wheatgrass *Thinopyrum ponticum* (Podp.) Barkworth & D. R. Dewey (2*n* =10*x* =70, StStStStE^{e}E^{e}E^{b}E^{b}E^{x}E^{x} or JJJJJJJ^{S}J^{S}J^{S}J^{S}) has long been known to have superior resistance to both biotic and abiotic stresses, including powdery mildew, stem rust, leaf rust, stripe rust, eyespot and Fusarium head blight (Li and Wang, 2009). Because of its ability to readily be crossed with wheat, many genes for disease resistance have been introduced into wheat. Among them, eleven were formally documented, including *Lr19*, *Lr24*, *Lr29*, *Sr24*, *Sr25*, *S26*, *Sr43*, *Yr69*, *Cmc2*, *Fhb7,* and *Pm51* (Sarma and Knott, 1966; Hart et al., 1976; Whelan and Lukow, 1990; Procunier et al., 1995; Mago et al., 2005; Li and Wang, 2009; Niu et al., 2014; Zhan et al., 2014; Hou et al., 2016; Wang et al., 2020). *Pm51* is the only *Pm* gene designated officially from *Th. ponticum* so far. Due to the allodecaploid nature, *Th. ponticum* has a large and complex genome, and resistance to a disease might be associated with its different chromosomes (Chen et al., 1998; Li and Wang, 2009). Therefore, there might be novel *Pm* gene(s) to be explored in *Th. ponticum*.

SNTE20, a wheat-*Th. ponticum* partial amphiploid with powdery mildew resistance, was previously developed in our lab (He et al., 2013). In the present study, crosses between SNTE20 and common wheat were carried out, and consequently two *Trititrigia* introgression lines were generated, SN0293-2 and SN0293-7, both of which were resistant to powdery mildew. The two introgression lines were then characterized by combined methods of morphology, disease evaluation, sequential genomic *in situ* hybridization-fluorescence *in situ* hybridization (GISH-FISH), and molecular marker analyses.

MATERIALS AND METHODS

Plant Materials

Materials used in this study included *Th. ponticum*, rye, SNTE20, Yannong15 (YN15), Shannongfu63 (SNF63), Jimai22 (JM22), Shannong224 (SN224), SN0293-2, SN0293-7, CH7086, and Huixianhong (HXH). *Th. ponticum* (accession No. R431) was provided by Prof. Zhensheng Li, the former Northwest Institute of Botany, the Chinese Academy of Sciences, Yangling, China. Wheat-*Th. ponticum* octoploid SNTE20 was developed from the multiple cross *Th. ponticum*/YN15//SNF63. YN15, SNF63, and JM22 are wheat cultivars. SN224 is a T1BL·1RS translocation line with dwarf stems. SN0293-2 and SN0293-7 were generated from a hybrid of the cross SNTE20/YN15//SN224/3/JM22 (**Supplementary Figure 1**). CH7086 carrying *Pm51* was provided by Prof. Xiaojun Zhang, the former Crop Science Institute, Shanxi Academy of Agricultural Sciences, Taiyuan, China.

Assessment of Agronomic Traits

Fifteen plants were grown in each 1.5 m long row, with 25 cm spacing between the rows. Agronomic traits, including plant height, spike length, spikelets per spike, kernels per spike, and thousand kernels weight, were recorded at the Experimental Station of Shandong Agricultural University. Each trait was averaged on 10 plants.

Evaluation of Powdery Mildew Resistance

Seedling resistance assessment was performed in a growth chamber using the *Bgt* isolate E09, following the method described by Zhao et al. (2013). Seedlings were grown in rectangular plastic trays (5 cm × 5 cm; 10 plants per tray) and inoculated with fresh *Bgt* conidiospores obtained from the susceptible cultivar HXH at the one-leaf stage. After approximately 2 weeks, when symptoms were severe on HXH, infection types (ITs) on the plants were described using a 0–4 infection scale: 0–2 scores indicating resistance and 3–4 susceptibility.

At the adult stage, resistance to powdery mildew was evaluated after natural infection in field-grown plants at the Experimental Station of Shandong Agricultural University over three growing seasons (2018–2020). The most severe reaction type in a given year was considered to be the final resistance result. HXH was planted perpendicular and adjacent to the test rows to serve as an inoculum spreader and a susceptible control. The disease symptoms were recorded three times at weekly intervals after flowering, and the most severe infection score was used as the final response. The ITs of powdery mildew at the adult stage were scored using a 0–9 scale: 0–4 scores indicating resistance and 5–9 susceptibility (Li et al., 2011).

Identification of Sequential GISH-FISH

The chromosomes were prepared following the method described by Kato et al. (2004). The purified total genomic DNA extracted from *Th. ponticum* or rye was labeled with Texas Red-5-dCTP probes, with the sheared genomic DNA from YN15 as a blocker. GISH analysis was performed as described by Fu et al. (2012). For FISH analysis, oligonucleotide probes, including TAMRA (6-carboxytetramethylrhodamine)-labeled oligonucleotides pAs1-1, pAs1-3, pAs1-4, pAs1-6, AFA-3, and AFA-4 and FAM (6-carboxyfluorescein)-labeled oligonucleotides pSc119.2-1 and $(GAA)_{10}$, were used. All probes were synthesized by Sangon Biotech Co., Ltd. (Shanghai, China). FISH analysis was performed as described by Huang et al. (2018). The chromosomes were counterstained with 4, 6-diamidino-2-phenylindole (DAPI), and the images were captured with a fluorescence microscope (Olympus BX60) equipped with a CCD (charge-coupled device) camera.

Development of Molecular Markers and PCR Amplification

The Axiom Wheat 660K Genotyping Array was used to genotype SN0293-2 and their parents YN15, SNF63, JM22, SN224, SNTE20 and *Th. ponticum*. The SNP typing data were processed using Microsoft Excel 2016 software, regarding SNPs present in SN0293-2, SNTE20 and *Th. ponticum* and absent in the comment wheat parents as the putatively specific ones of *Th. ponticum*. Library construction and high-throughput sequencing of SN0293-2, SNF63, and SNTE20 were performed by Berry Genomics Company (Beijing, China). The raw reads were subjected to a quality check and then filtered by fastp to remove adapter sequences and low-quality bases (Chen et al., 2018). High-quality reads were mapped to the *Triticeae* repeat database mipsREdat 9.3p (PGSB Repeat Database) using Burrows-Wheeler Aligner in order to filter out repeat noises, non-mapped reads (from the genome region of non-repeat sequences) were fished by homemade scripts and then used for variant calling (Li and Durbin, 2010). GATK 3.8 (https://www.software.broadinstitute.org/gatk) was used to call out all the variations, including SNPs and InDels. Markers of cleaved amplified polymorphic sequences (CAPS) were designed on the basis of SNPs, and the corresponding restriction endonucleases were used for restriction digestion. InDel markers were designed according to the results of resequencing. Briefly, the corresponding 500-bp both upstream and downstream sequences of the InDels larger than 10 bp were fished from the Chinese Spring reference on the website WheatOmics (Ma et al., 2021). About 1-kb fragments corresponding the InDels were used to perform local BLAST to identify genome specific fragments for marker development (identity $>$ 95%, number of hits $\leq$ 3). The conserved sequences flanking the target region were then used to design primers.

DNA was isolated from young leaf tissues following a standard CTAB method. The PCR cycling condition was as follows: 95°C for 3 min, followed by 35 cycles of 95°C for 30 s, appropriate anneal temperature (50–60°C) for 40 s, 72°C for 1 min, and a final extension at 72°C for 5 min, and the amplification products were kept at 10°C using a Bio-Rad 9600 Thermal Cycler (Hercules, USA). PCR products amplified by CAPS markers were separated through 1% agarose electrophoresis and PCR products amplified by Indel markers were separated on 8% non-denatured polyacrylamide gels (39 acrylamide: 1 bisacrylamide).

Restriction endonuclease (*HaeIII, SacII, BstEII*) were selected for the CAPS marker and the digestion reaction was carried out according to the manufacturer's instructions (New England Biolabs, England). The digestion was conducted in a 10 μL reaction volume, contained 50 ng PCR products and digested at 37°C or 65°C (depending on the enzyme). The digestion reaction products were separated by 2% agarose electrophoresis.

RESULTS

Evaluation of Powdery Mildew Resistance

At the seedling stage, SN0293-2, SN0293-7, and its parents were inoculated with the *Bgt* isolate E09, and the disease reaction was assessed once the control HXH displayed thoroughly susceptible symptoms. *Th. ponticum*, SNTE20, JM22, SN0293-2, and SN0293-7 were all immune to E09 (IT = 0), while the common wheat parents YN15, SNF63, and SN224 appeared to be susceptible with an IT score of 4 (**Figure 1A**), suggesting that the resistance to powdery mildew at the seeding stage in SN0293-2 and SN0293-7 originated from either *Th. ponticum* or JM22. The seedling reactions of SN0293-2, SN0293-7, and 42 wheat genotypes carrying documented *Pm* genes and gene combinations to 25 *Bgt* isolates were simultaneously recorded (**Supplementary Table 1**). The susceptible cultivar Chancellor exhibited a high IT (4). SN0293-2 and SN0293-7 were resistant to 23 isolates, indicating their broad-spectrum resistance. Additionally, *Pm51*, the only documented powdery mildew resistance gene putatively derived from *Th. ponticum*, was nearly immune to the isolate E20 (Zhan et al., 2014), but SN0293-2 and SN0293-7 were susceptible to E20, indicating that the resistance genes in SN0293-2 and SN0293-7 were different from *Pm51*.

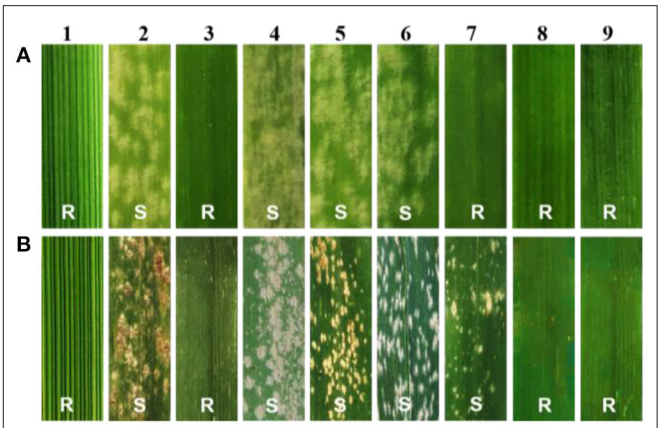


FIGURE 1 | Evaluation of powdery mildew resistance. **(A)** The seedling stage. **(B)** The adult stage. 1–9 refer to *Th. ponticum*, HXH, SNTE20, SNF63, YN15, SN224, JM22, SN0293-2, and SN0293-7, respectively.

At the adult stage, resistance to powdery mildew was tested in field over three growing seasons (2018–2020), and the most severe reaction type observed in a given year was considered to be the final result. It was found that SN0293-2, SN0293-7, *Th. ponticum,* and SNTE20 were resistant to powdery mildew, whereas YN15, SNF63, SN224, JM22, and CH7086 showed susceptibility (**Figure 1B** and **Supplementary Figure 2**). Thus, the resistance at the adult stage in SN0293-2 and SN0293-7, temporarily designated *PmSN0293*, was putatively inherited from *Th. ponticum* and different from *Pm51*. Previous studies have found that JM22 contains powdery mildew resistance genes *Pm2* and *Pm52* (Cao et al., 2010; Qu et al., 2020). To determine whether SN0293-2 and SN0293-7 had these two genes or not, we amplified the genomes of them and their parents by using the markers *Xcfd81* and *Xicssl326*, which were known to be linked with *Pm2* and *Pm52*, respectively. It was showed that both SN0293-2 and SN0293-7 contained the two genes derived from JM22 (**Figure 2**). The results above indicated that SN0293-2 and SN0293-7 might carry a new powdery mildew resistance gene *PmSN0293* from *Th. ponticum* pyramided with *Pm2* and *Pm52*.

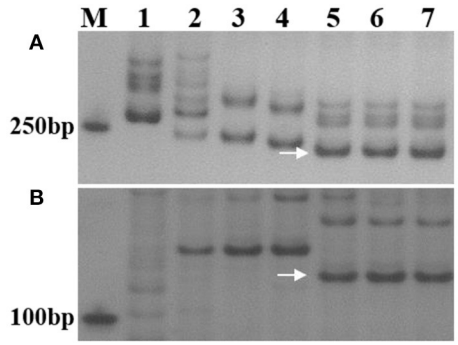


FIGURE 2 | Amplified results of specific markers *Xcfd-81* and *Xicssl326*. **(A)** *Xcfd-81*. **(B)** *Xicssl326*. Lanes: M, DL2000 marker; 1–7 refer to *Th. ponticum*, SNTE20, YN15, SN224, JM22, SN0293-2, and SN0293-7, respectively. Arrows indicate specific bands.

Cytogenetic Analyses

GISH, probed with the total genomic DNA of *Th. ponticum* and blocked with the genomic DNA of YN15, revealed that both SN0293-2 and SN0293-7 had 42 chromosomes and no alien signal was detected (**Figures 3A,B**). Because the parent SN224 is a T1BL·1RS translocation line, the rye genomic DNA was further used as a probe. Two chromosome arms with red coloration alongside the blue chromosome arms of wheat were found in SN0293-2 and no hybridization signal in SN0293-7, indicating the rye chromosome arm 1RS was present in SN0293-2 and absent in SN0293-7 (**Figures 3C,D**).

After removing the GISH signals, the same slides were subjected to FISH analysis. The 1RSs in SN0293-2 were found in the form of T1BL·1RS, inheriting from the parent SN224 (**Figure 3E**). Further, FISH patterns of SN0293-2 and SN0293-7 in **Figures 3E,F** were compared with those of their parents (**Figure 4**). Differences between SN0293-2 and its parents were detected in the terminal region of 1DS, 2DS, 2DL, 6BL, and 7AL, and the middle of 3BL. As for SN0293-7, the terminal of 1DS, 2DL, 6BS, and 7AL as well as the middle of 1BL and 3BL appeared to be different from its parents. This suggested that chromosomes underwent structural variations with the formation of SN0293-2 and SN0293-7. In addition, differences between SN0293-2 and SN0293-7 were also present apart from the chromosome T1BL·1RS. For instance, red signals existed in the terminal regions of 1DS, 6BS, and 7AS in SN0293-2, but they disappeared in the corresponding regions of SN0293-7. The green signals in the telomere of SN0293-7 differed from those of SN0293-2.

Development of Molecular Markers

A total of 157 specific SNP loci of *Th. ponticum* were obtained by using wheat 660K SNP array to analyze *Th. ponticum* and its parents, involving 9 chromosomes of three wheat subgenomes, of which 134 were located on chromosome 6A and mainly distributed in the physical interval of 60–110 Mb (**Figure 5**). DNA resequencing data showed that a total of 111 Indels larger than 10 bp were identified in the above 50 Mb region of chromosome 6A. These Indels were identical

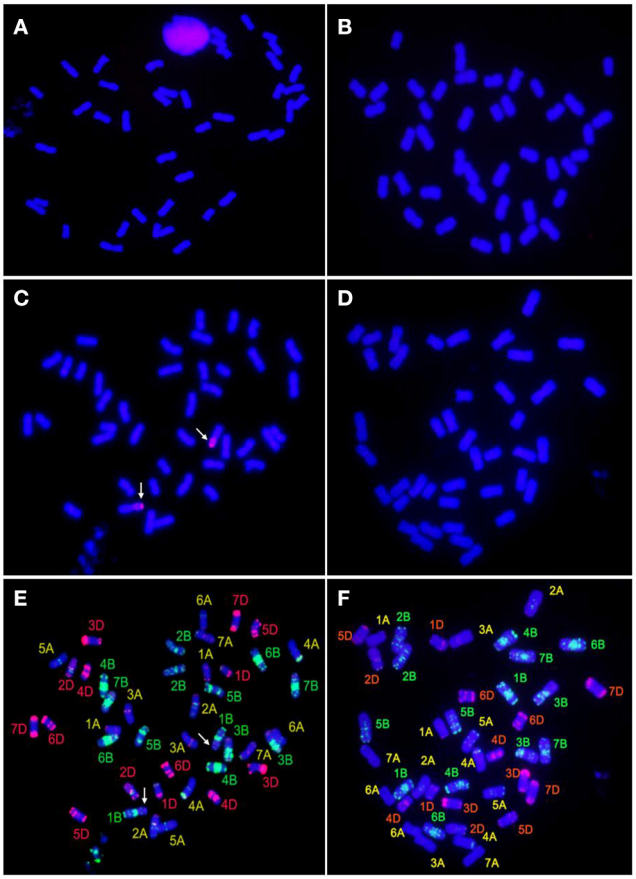


FIGURE 3 | GISH-FISH analyses of SN0293-2 and SN0293-7. **(A)** GISH patterns of SN0293-2 probed with total genomic DNA of *Thinopyrum ponticum*. **(B)** GISH patterns of SN0293-7 probed with total genomic DNA of *Th. ponticum*. **(C)** GISH patterns of SN0293-2 probed with total genomic DNA of rye. **(D)** GISH patterns of SN0293-7 probed with total genomic DNA of rye. **(E)** FISH patterns of SN0293-2. **(F)** FISH patterns of SN0293-7. Arrows indicate the translocation T1BL·1RS.

between the two lines SN0293 and SNTE20 containing *Th. ponticum* fragments, but different from the common wheat line SNF63. According to the results of Wheat 660K SNP array and resequencing data, three CAPS markers (*CAPS421-HaeIII*, *CAPS761-SacI*, *CAPS564-BstEII*) and four Indel markers (*Indel192*, *Indel101*, *Indel752*, *Indel753*) were obtained (**Table 1**). *Th. ponticum* specific bands were amplified in SN0293-2 and SN0293-7 (**Figure 6**), in which *CAPS421-HaeIII* was located on wheat chromosome 2B, *CAPS761-SacI*, *CAPS564-Bst II*, *Indel192*, *Indel101*, *Indel752,* and *Indel753* were located on chromosome 6A. The results showed that SN0293-2 and SN0293-7 inherited genetic components from *Th. ponticum*, so they were wheat-*Th. ponticum* introgression lines.

Analyses of Agronomic Traits

The agronomic traits of SN0293-2 and SN0293-7 were compared with those of their parents (**Figure 7**). SN0293-2 and SN0293-7 displayed average plant heights of 65.4 ± 3.1 cm and 69.3

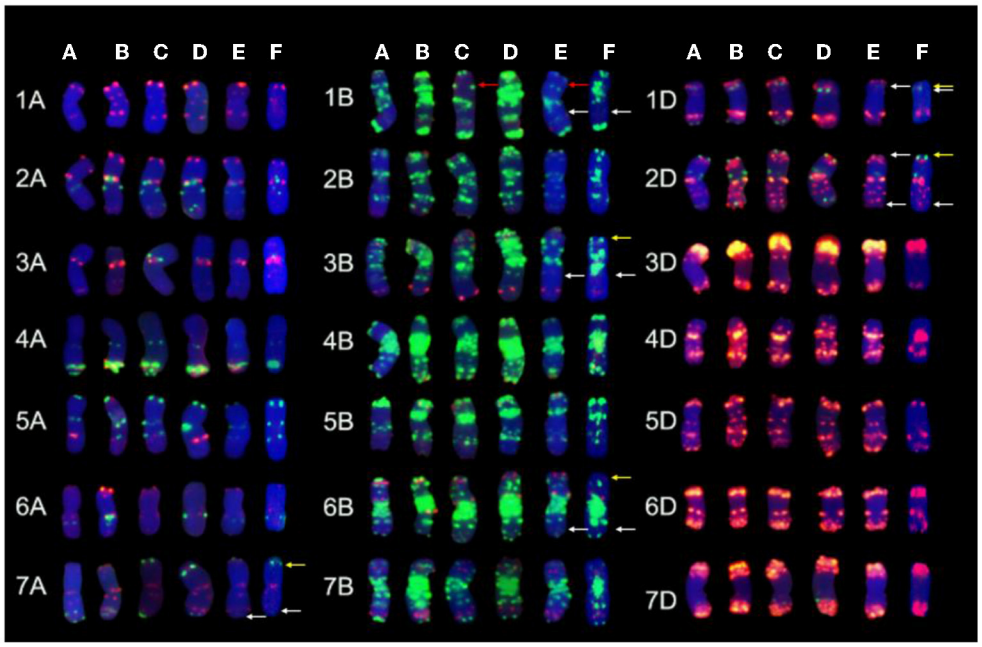


FIGURE 4 | Chromosome comparison of FISH patterns among SN0293-2, SN0293-7, and their parents. **(A)** SNTE20. **(B)** YN15. **(C)** SN224. **(D)** JM22. **(E)** SN0293-2. **(F)** SN0293-7. The red, white, and yellow arrows indicate 1RS, FISH differences of SN0293-2 and SN0293-7, respectively.

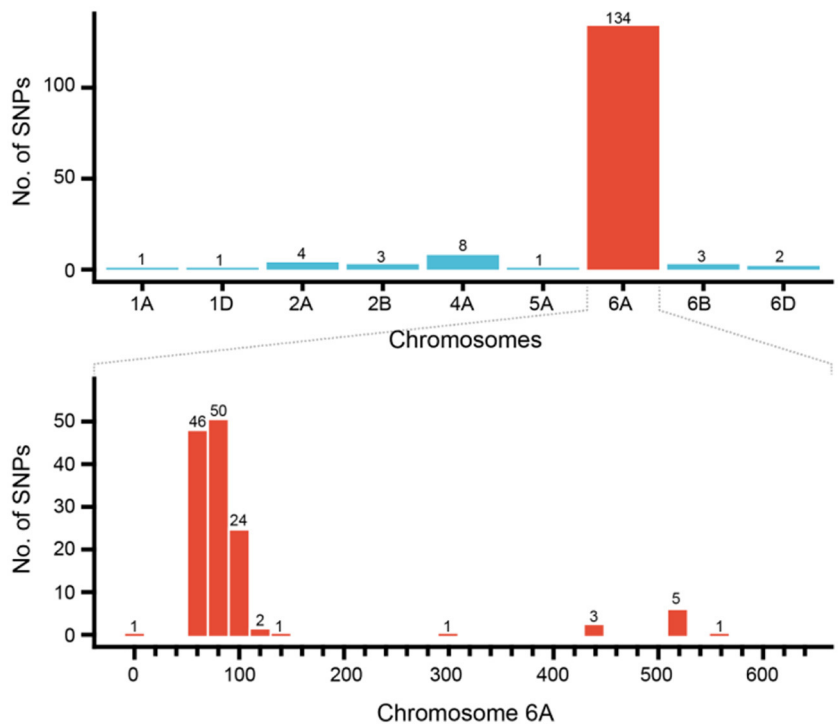


FIGURE 5 | Distribution of SNPs specific to *Thinopyrum ponticum* in SN0293-2.

± 4.2 cm, respectively, which were lower than the parents SNTE20 (105.2 ± 3.9 cm), YN15 (84.0 ± 2.4 cm) and JM22 (83.5 ± 2.7 cm) except SN224 (61.0 ± 3.6 cm). They shared about equal numbers of spike length and spikelets as the common wheat parents. However, SN0293-2 and SN0293-7 produced more kernels per spike (56 ± 5 and 67 ± 3, respectively)

TABLE 1 | Markers developed specific to *Thinopyrum ponticum*.

Markers	Homeologous groups	Sequences
CAPS421-HaeIII	2	F: 5′-GAACCCCGGATCTGAGTGTCCA-3′ R: 5′-CAACCAACTGCGCTGTCCGTC-3′
CAPS761-SacI	6	F: 5′-GAAATATCCAACCAGAACAGTGG-3′ R: 5′-CTCTGCTTGAGTGGCAGGACT-3′
CAPS564-BstEII	6	F: 5′-ATCCAAACAAGACAACCCGTCTTG-3′ R: 5′-GCTTGTCTATACCCTAGTCGCGT-3′
Indel192	6	F: 5′-ACTCCCAAGGGTGAACCTATGAT-3′ R: 5′-CGGTCAGAGGTAACTTGCTGTG-3′
Indel101	6	F: 5′-TAGACCTTCGTGGGAACCTTTG-3′ R: 5′-TAGACCTTCGTGGGAACCTTTG-3′
Indel752	6	F: 5′-TACGGCTAAAGGAGTTGACC-3′ R: 5′-TGATGCTGTGGGAACGAAA-3′
Indel753	6	F: 5′-AACGCTAAGACTGGATTGATTG-3′ R: 5′-ACCTAATGCGACAGATGGACAA-3′

and higher 1,000-grain weight (53.8 ± 1.7 and 53.0 ± 1.3, respectively). These agronomic data indicated that the two introgression lines probably carried genes that were beneficial for wheat breeding.

DISCUSSION

Wild relatives of wheat have been serving as valuable gene reservoirs due to their resistance to many biotic and abiotic stresses. In the past few decades, wide crosses have been employed to incorporate useful genes into common wheat. The tall grass *Th. ponticum* was extensively used in wheat genetic improvement. The elite cultivars, such as Xiaoyan 6, Gaoyou 503, Xiaoyan 60, and Xiaoyan 81, were widely planted in China and had played an important role in wheat production (Li et al., 2015; Luo et al., 2021; Yang et al., 2022). Meanwhile, numerous chromosome engineering materials were generated, including partial amphiploids (Zheng et al., 2015), additions (Li et al., 2016), substitutions (Wang et al., 2019; Li et al., 2021), translocations (Yang et al., 2022) and introgression lines (Zhan et al., 2014).

In the present study, *Th. ponticum*, SNTE20 and wheat-*Th. ponticum* introgression lines SN0293-2 and SN0293-7 displayed excellent resistance to powdery mildew. Up to now, 11 genes have been officially designated from *Th. ponticum* (Li and Wang, 2009; Liu et al., 2020). Among them, only one gene named *Pm51* was responsible for resistance to powdery mildew. At the seedling stage, the two introgression lines were susceptible to the *Bgt* isolate E20 (**Supplementary Table 1**), while *Pm51* was nearly immune (Zhan et al., 2014). As for the *Bgt* isolate E09, common wheat parents SNF63, YN15, and SN224 showed susceptibility but JM22 displayed immunity because JM22 carried two resistance genes *Pm2* and *Pm52* (Cao et al., 2010; Qu et al., 2020). Molecular marker analyses demonstrated that both SN0293-2 and SN0293-7 carried these two powdery mildew resistance genes as well. At the adult plant stage, JM22 and CH7086 exhibited susceptibility to powdery mildew similarly to the other common wheat parents (**Figure 1** and **Supplementary Figure 2**), suggesting that *Pm2*, *Pm51,* and *Pm52* lost resistance in Tai'an, Shandong Province, China. However, resistance of the two introgression lines remained as well as *Th. ponticum* and SNTE20. Seen from the results above, the seedling resistance of SN0293-2 and SN0293-7 originated from either *Th. ponticum* or JM22, and the adult plant resistance of them was putatively derived from *Th. ponticum* obviously different from *Pm51*. Segregating populations are being developed for mapping the adult plant resistance. In addition, 1RS, herein, had no resistance to powdery mildew since the T1BL·1RS translocation SN224 exhibited consistent susceptibility at both the seeding stage and adult plant stage.

GISH and FISH have been used as powerful tools to detect alien chromosomes or chromosome segments and chromosomal structural variations in distant hybridization. In this study, one pair of T1BL·1RS was confirmed by GISH-FISH in SN0293-2, but no cytological evidence for *Th. ponticum* was found in SN0293-2 and SN0293-7, even though some structural variations were demonstrated between these two lines and their parents. The alien segments must be too small to be detected for cytogenetic tools due to the accuracy limitation. Recently, the wheat SNP array has been employed to identify wild relatives and their derivatives with wheat. Zhou et al. (2018) constructed the genetic linkage map of *Agropyron cristatum* and characterized a number of wheat-*A. cristatum* chromosome lines. Several wheat-*Th. ponticum* translocation and introgression lines were also subjected to Wheat 660K SNP array and deletion events were detected (Li et al., 2019; Yang et al., 2021, 2022). In order to determine whether SN0293-2 and SN0293-7 inherited genetic components from *Th. ponticum* or not, the Axiom Wheat 660K Genotyping Array were also used in the present study. One hundred and fifty-seven SNPs were found to be specific to *Th. ponticum*. These SNPs referred to 9 chromosomes of wheat genomes, 134 of which were located on the wheat chromosome 6A (**Figure 5**). Then three CAPs markers were developed and showed specificity to *Th. ponticum* in SN0293-2 and SN0293-7. Similarly, four Indel markers belonging to homoeologous group 6 were obtained according to the re-sequence data (**Table 1** and **Figure 6**). It was indicated that the wheat SNP array has higher accuracy than cytological methods and could effectively work in identification of wheat-*Th. ponticum* introgression lines.

Most of the alien chromosome lines experience more or less penalties of agronomic traits because of linkage drags of the whole alien chromosomes or large alien segments they carry. For instance, many addition lines generally have longer growth duration and smaller grains due to the introduced alien chromosomes. However, introgression lines have small alien segments that GISH cannot detect, generally resulting in no or less genetic drag. Therefore, germplasms of this type with target genes usually exhibit positive agronomic traits and are considered to be the most valuable materials for wheat breeding. In order to breed disease-resistant lines with favorable agronomic traits, SN224 and JM22 were used as male parents. The former is a breeding line with dwarf stems developed in our lab. The latter is an elite cultivar with the

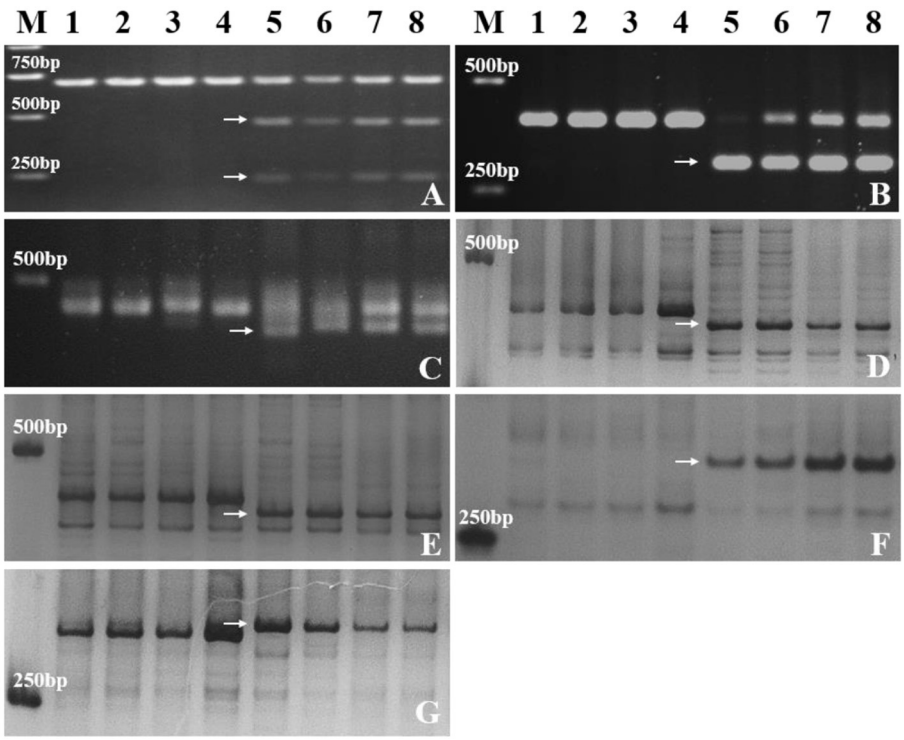


FIGURE 6 | Molecular markers analysis of SN0293-2, SN0293-7 and their parents. **(A)** *CAPS421-HaeIII*. **(B)** *CAPS761-SacI*. **(C)** *CAPS564-BstEII*. **(D)** *Indel192*. **(E)** *Indel101*. **(F)** *Indel752*. **(G)** *Indel753*. Lanes: M, DL2000 marker; 1–8 refer to SNF63, YN15, SN224, JM22, *Th. ponticum*, SNTE20, SN0293-2, and SN0293-7, respectively. Arrows indicate specific bands of *Th. ponticum*.

FIGURE 7 | Plants, spikes, and kernels of SN0293-2, SN0293-7, and their parents. **(A)** Plants. **(B)** Spikes. **(C)** Kernels. 1–6 refer to SNTE20, YN15, SN224, JM22, SN0293-2, and SN0293-7, respectively. Scale bar = 2 cm.

largest promotion area in the last decade in China (Jia et al., 2020). JM22 was also employed as the last parent to improve agronomic traits and to provide two *Pm* genes. Consequently, the wheat-*Th. ponticum* introgression lines SN0293-2 and SN0293-7 carried a new resistance gene putatively from *Th. ponticum* and pyramided *Pm2* and *Pm52* derived from JM22. They also showed positive agronomic traits, such as more kernels and higher 1,000-weight, enhancing the application potential in wheat breeding programs.

CONCLUSIONS

Two wheat-*Th. ponticum* introgression lines were developed and named SN0293-2 and SN0293-7, respectively. Besides *Pm2* and *Pm52*, they possessed a new powdery mildew resistance gene *PmSN0293* putatively from *Th. ponticum* obviously different from *Pm51* reported previously. Seven markers were obtained and confirmed to be specific to *Th. ponticum*. Superior resistance to powdery mildew at both the seedling and adult plant stages and positive agronomic traits give the two introgression lines great potential to be used in wheat breeding programs.

DATA AVAILABILITY STATEMENT

The data presented in the study are deposited in the Genome Sequence Archive in BIG Data Center repository (https://bigd.big.ac.cn/), accession number PRJCA009783

AUTHOR CONTRIBUTIONS

YB designed the research. ML and YY performed the experiments. FN and XL analyzed the data and developed the specific markers. HW contributed to the development of the materials. ML and YB wrote the manuscript. All authors approved the final version of the manuscript.

FUNDING

This project was supported by the National Natural Science Foundation of China (No. 32071998), the National Key Research and Development Program of China (No. 2016YFD0102004-02) and Key R&D Program of Shandong Province (Major Science and Technology Innovation Project) (2021LZGC009).

ACKNOWLEDGMENTS

We sincerely thank Prof. Yilin Zhou at the Institute of Plant Protection, Chinese Academy of Agricultural Sciences, for providing powdery mildew assessment at the seedling stage.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.943669/full#supplementary-material

REFERENCES

Braun, H. J., Atlin, G., and Payne, T. (2010). "Multi-location testing as a tool to identify plant response to global climate change," in *Climate Change and Crop Production*, ed M. P. Reynolds (CABI Climate Change Series), Oxfordshire: CABI, 115–138. doi: 10.1079/9781845936334.0115

Cao, X., Zhou, Y., Duan, X., Song, Y., He, W., Ding, K., et al. (2010). Postulation of wheat powdery mildew resistance genes in 101 wheat cultivars (lines) from major wheat regions in China. *J. Triticeae Crops* 30, 948–953. doi: 10.7606/j.issn.1009-1041.2010.05.030

Chen, Q., Conner, R. L., Laroche, A., and Thomas, J. B. (1998). Genome analysis of *Thinopyrum intermedium* and *Th. ponticum* using genomic *in situ* hybridization. *Genome* 41, 580–586. doi: 10.1139/g98-055

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

Fu, S., Lv, Z., Qi, B., Guo, X., Li, J., Liu, B., et al. (2012). Molecular cytogenetic characterization of wheat-*Thinopyrum elongatum* addition, substitution and translocation lines with a novel source of resistance to wheat Fusarium Head Blight. *J. Genet. Genomics* 39, 103–110. doi: 10.1016/j.jgg.2011.11.008

Gao, H., Xu, X., Ai, P., Luo, F., Guo, P., and Ma, P. (2022). Identification of the powdery mildew resistance in Chinese wheat cultivar Heng 4568 and its evaluation in marker-assisted selection. *Front. Genet*. 13, 819844. doi: 10.3389/fgene.2022.819844

Hart, G. E., Mcmillin, D. E., and Sears, E. R. (1976). Determination of the chromosomal location of a glutamate oxaloacetate transaminase structural gene using *Triticum-Agropyron* translocations. *Genetics* 83, 49–61. doi: 10.1093/genetics/83.1.49

Hawkesford, M. J., Araus, J. L., Park, R., Calderini, D., Miralles, D., Shen, T. M., et al. (2013). Prospects of doubling global wheat yields. *Food Energy. Secur*. 2, 34–48. doi: 10.1002/fes3.15

He, F., Xu, J., Qi, X., Bao, Y., Li, X., Zhao, F., et al. (2013). Molecular cytogenetic characterization of two partial wheat *Elytrigia* elongate amphiploids resistant to powdery mildew. *Plant Breed* 132, 553–557. doi: 10.1111/pbr.12104

He, H., Liu, R., Ma, P., Du, H., Zhang, H., Wu, Q., et al. (2021). Characterization of *Pm68*, a new powdery mildew resistance gene on chromosome 2BS of Greek durum wheat TRI 1796. *Theor. Appl. Genet*. 134, 53–62. doi: 10.1007/s00122-020-03681-2

He, H., Zhu, S., Zhao, R., Jiang, Z., Ji, Y., Ji, J., et al. (2018). *Pm21*, encoding a typical CC-NBS-LRR protein, confers broad-spectrum resistance to wheat powdery mildew disease. *Mol. Plant* 11, 879–882. doi: 10.1016/j.molp.2018.03.004

Hou, L., Jia, J., Zhang, X., Li, X., Yang, Z., Ma, J., et al. (2016). Molecular mapping of the stripe rust resistance gene *Yr69* on wheat chromosome 2AS. *Plant Dis*.100, 1717–1724. doi: 10.1094/PDIS-05-15-0555-RE

Huang, X., Zhu, M., Zhuang, L., Zhang, S., Wang, J., Chen, X., et al. (2018). Structural chromosome rearrangements and polymorphisms identified in Chinese wheat cultivars by high-resolution multiplex oligonucleotide FISH. *Theor. Appl. Genet*. 131, 1967–1986. doi: 10.1007/s00122-018-3126-2

Jia, M., Xu, H., Liu, C., Mao, R., Li, H., Liu, J., et al. (2020). Characterization of the powdery mildew resistance gene in the elite wheat cultivar Jimai 23 and its application in marker-assisted selection. *Front. Genet*. 11, 241. doi: 10.3389/fgene.2020.00241

Kato, A., Lamb, J. C., and Birchler, J. A. (2004). Chromosome painting using repetitive DNA sequences as probes for somatic chromosome identification in maize. *Proc. Natl. Acad. Sci. U.S.A*. 101, 13554–13559. doi: 10.1073/pnas.0403659101

Li, H., Boshoff, W., Pretorius, Z. A., Zheng, Q., Li, B., and Li, Z. (2019). Establishment of wheat-*Thinopyrum ponticum* translocation lines with resistance to *Puccinia graminis* f. sp. tritici Ug99. *J. Genet. Genomics* 46, 405-407. doi: 10.1016/j.jgg.2019.07.005

Li, H., and Durbin, R. (2010). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, H., and Wang, X. (2009). *Thinopyrum ponticum* and *Th. intermedium*: the promising source of resistance to fungal and viral diseases of wheat. *J. Genet. Genomics* 36, 557–565. doi: 10.1016/S1673-8527(08)60147-2

Li, H., Zheng, Q., Pretorius, Z. A., Li, B., Tang, D., and Li, Z. (2016). Establishment and characterization of new wheat-*Thinopyrum ponticum* addition and translocation lines with resistance to Ug99 races. *J. Genet. Genomics* 43, 573–575. doi: 10.1016/j.jgg.2016.07.004

Li, H. J., Wang, X. M., Song, F. J., Wu, C. P., Wu, X. F., Zhang, N., et al. (2011). Response to powdery mildew and detection of resistance genes in wheat cultivars from China. *Acta Agron. Sin.* 37, 943–954. doi: 10.3724/SP.J.1006.2011.00943

Li, M., Wang, Y., Liu, X., Li, X., Wang, H., and Bao, Y. (2021). Molecular cytogenetic identification of a novel wheat-*Thinopyrum ponticum* 1J^{S} (1B) substitution line resistant to powdery mildew and leaf rust. *Front. Plant. Sci.* 12, 727734. doi: 10.3389/fpls.2021.727734

Li, Z. S., Li, B., Zheng, Q., and Li, H. (2015). "Review and new progress in wheat wide hybridization for improving the resistance to biotic and abiotic stresses," in *Advances in Wheat Genetics: From Genome to Field*, eds Y. Ogihara, S. Takumi, H. Handa, et al. (Tokyo: Springer), 377–385. doi: 10.1007/978-4-431-55675-6_43

Liu, C., Han, R., Wang, X. L., Gong, W. P., Cheng, D. G., Cao, X. Y., et al. (2020). Research progress of wheat wild hybridization, disease resistance genes transfer and utilization. *Sci. Agric. Sin.* 53, 1287–1308. doi: 10.3864/j.issn.0578-1752.2020.07.001

Liu, W., Koo, D. H., Xia, Q., Li, C., Bai, F., Song, Y., et al. (2017). Homoeologous recombination-based transfer and molecular cytogenetic mapping of powdery mildew-resistant gene *Pm57* from *Aegilops searsii* into wheat. *Theor. Appl. Genet.* 130, 841–848. doi: 10.1007/s00122-017-2855-y

Luo, Q., Zheng, Q., Hu, P., Yang, G., Li, H., Li, B., et al. (2021). Mapping QTL for agronomic traits under two levels of salt stress in a new constructed RIL wheat population. *Theor. Appl. Genet.* 134, 171–189. doi: 10.1007/s00122-020-03689-8

Ma, S., Wang, M., Wu, J., Guo, W., Chen, Y., Li, G., et al. (2021).WheatOmics: a platform combining multiple omics data to accelerate functional genomics studies in wheat. *Mol. Plant* 14: 1965–1968. doi: 10.1016/j.molp.2021.10.006

Mago, R., Bariana, H. S., Dundas, I. S., Spielmeyer, W., Lawrence, G. J., Pryor, A. J., et al. (2005). Development of PCR markers for the selection of wheat stem rust resistance genes *Sr24* and *Sr26* in diverse wheat germplasm. *Theor. Appl. Genet.* 111, 496–504. doi: 10.1007/s00122-005-2039-z

Niu, Z., Klindworth, D. L., Yu, G. L., Friesen, T., Chao, S., et al. (2014). Development and characterization of wheat lines carrying stem rust resistance gene *Sr43* derived from *Thinopyrum ponticum*. *Theor. Appl. Genet.* 127, 969–980. doi: 10.1007/s00122-014-2272-4

Procunier, J. D., Townley-Smith, T. F., Fox, S., Prashar, S., Gray, M., Kim, W. K., et al. (1995). PCR-based RAPD/DGGE markers linked to leaf rust resistance genes *Lr29* and *Lr25* in wheat (*Triticum aestivum* L.). *J. Genet. Breed.* 49, 92–97.

Qu, Y., Wu, P., Hu, J., Chen, Y., Shi, Z., Qiu, D., et al. (2020). Molecular detection of the powdery mildew resistance genes in winter wheats DH51302 and Shimai 26. *J. Integr. Agric* 19, 931–940. doi: 10.1016/S2095-3119(19)62644-4

Ren, T., Tang, Z., Fu, S., Yan, B., Tan, F., Ren, Z., et al. (2017). Molecular cytogenetic characterization of novel wheat-rye T1RS●1BL translocation lines with high resistance to diseases and great agronomic traits. *Front. Plant. Sci.* 8, 799. doi: 10.3389/fpls.2017.00799

Sarma, D., and Knott, D. R. (1966). The transfer of leaf-rust resistance from *Agropyron* to *Triticum* by irradiation. *Genome* 8, 137–143. doi: 10.1139/g66-018

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, a5435. doi: 10.1126/science.aba5435

Wang, S., Wang, C., Wang, Y., Wang, Y., Chen, C., and Ji, W. (2019). Molecular cytogenetic identification of two wheat-*Thinopyrum ponticum* substitution lines conferring stripe rust resistance. *Mol. Breed.* 39, 143. doi: 10.1007/s11032-019-1053-9

Wang, Z., Li, L., He, Z., Duan, X., Zhou, Y., Chen, X., et al. (2005). Seedling and adult plant resistance to powdery mildew in Chinese bread wheat cultivars and lines. *Plant Dis.* 89, 457–463. doi: 10.1094/PD-89-0457

Whelan, E. D. P., and Lukow, O. M. (1990). The genetics and gliadin protein characteristics of a wheat-alien translocation that confers resistance to colonization by the wheat curl mite. *Genome* 33, 400–404. doi: 10.1139/g90-061

Xing, L., Hu, P., Liu, J., Witek, K., Zhou, S., Xu, J., et al. (2018). *Pm21* from *Haynaldia villosa* encodes a CC-NBS-LRR protein conferring powdery mildew resistance in wheat. *Mol. Plant* 11, 874–878. doi: 10.1016/j.molp.2018.02.013

Yang, G., Boshoff, W. H. P., Li, H., Pretorius, Z. A., Luo, Q., Li, B., et al. (2021). Chromosomal composition analysis and molecular marker development for the novel Ug99-resistant wheat-*Thinopyrum ponticum* translocation line WTT34. *Theor. Appl. Genet.* 134, 1587–1599. doi: 10.1007/s00122-021-03796-0

Yang, G., Tong, C., Li, H., Li, B., Li, Z., and Zheng, Q. (2022). Cytogenetic identification and molecular marker development of a novel wheat-*Thinopyrum ponticum* translocation line with powdery mildew resistance. *Theor. Appl. Genet.* doi: 10.1007/s00122-022-04092-1

Zhan, H., Li, G., Zhang, X., Li, X., Guo, H., Gong, W., et al. (2014). Chromosomal location and comparative genomics analysis of powdery mildew resistance gene Pm51 in a putative wheat-*Thinopyrum ponticum* introgression line. *PLoS ONE* 9, e113455. doi: 10.1371/journal.pone.0113455

Zhang, R., Xiong, C., Mu, H., Yao, R., Meng, X., Kong, L., et al. (2020). Pm67, a new powdery mildew resistance gene transferred from *Dasypyrum villosum* chromosome 1V to common wheat (*Triticum aestivum* L.). *Crop J.* 9, 882–888. doi: 10.1016/j.cj.2020.09.012

Zhao, Z., Sun, H., Song, W., Lu, M., Huang, J., Wu, L., et al. (2013). Genetic analysis and detection of the gene MlLX99 on chromosome 2BL conferring resistance to powdery mildew in the wheat cultivar Liangxing 99. *Theor. Appl. Genet.* 126, 3081–3089. doi: 10.1007/s00122-013-2194-6

Zheng, Q., Luo, Q., Niu, Z., Li, H., Li, B., Xu, S. S., et al. (2015). Variation in chromosome constitution of the Xiaoyan series partial amphiploids and its relationship to stripe rust and stem rust resistance. *J. Genet. Genomics* 42, 657–660. doi: 10.1016/j.jgg.2015.08.004

Zhou, S., Zhang, J., Che, Y., Liu, W., Lu, Y., Yang, X., et al. (2018). Construction of *Agropyron* Gaertn. genetic linkage maps using a wheat 660K SNP array reveals a homoeologous relationship with the wheat genome. *Plant Biotechnol. J.* 16, 818–827. doi: 10.1111/pbi.12831

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

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

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 20 September 2022
DOI 10.3389/fpls.2022.1006510

OPEN ACCESS

EDITED BY
Xiaoli Fan,
Chengdu Institute of Biology (CAS), China

REVIEWED BY
Junming Li,
Chinese Academy of Sciences, China
Zaijun Yang,
China West Normal University, China

*CORRESPONDENCE
Xiujin Lan
lanxiujin@163.com
Jian Ma
jianma@sicau.edu.cn

†These authors have contributed equally to this work

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

RECEIVED 29 July 2022
ACCEPTED 15 August 2022
PUBLISHED 20 September 2022

CITATION
Luo W, Zhou J, Liu J, Liu Y, Mu Y, Tang H, Xu Q, Deng M, Jiang Q, Chen G, Qi P, Wang J, Jiang Y, Chen Z, Zheng Z, Wei Y, Zheng Y, Lan X and Ma J (2022) Fine mapping of the *Hairy glume (Hg)* gene in a chromosome variation region at the distal terminus of 1AS.
Front. Plant Sci. 13:1006510.
doi: 10.3389/fpls.2022.1006510

Fine mapping of the *Hairy glume (Hg)* gene in a chromosome variation region at the distal terminus of 1AS

Wei Luo[1,2,3†], Jieguang Zhou[1,2†], Jiajun Liu[1,2], Yanlin Liu[2], Yang Mu[2], Huaping Tang[2], Qiang Xu[1,2], Mei Deng[2], Qiantao Jiang[1,2], Guoyue Chen[1,2], Pengfei Qi[1,2], Jirui Wang[1,2], Yunfeng Jiang[1,2], Zhongxu Chen[4], Zhi Zheng[5], Yuming Wei[1,2], Youliang Zheng[1], Xiujin Lan[1,2*] and Jian Ma[1,2*]

[1]State Key Laboratory of Crop Gene Exploration and Utilization in Southwest China, Sichuan Agricultural University, Wenjiang, Chengdu, Sichuan, China, [2]Triticeae Research Institute, Sichuan Agricultural University, Chengdu, Sichuan, China, [3]College of Life Science and Technology, Xinjiang University, Ürümqi, Xinjiang, China, [4]Department of Life Science, Chengdu Tcuni Technology, Chengdu, Sichuan, China, [5]Agriculture and Food, Commonwealth Scientific and Industrial Research Organisation, St. Lucia, QLD, Australia

Trichomes are differentiated epidermal cells and exist on above-ground organs of nearly all land plants with important roles in resistance to a wide range of biotic and abiotic stresses. We attempted to obtain candidate gene (s) for *Hairy glume* (*Hg*), responsible for the trichome on wheat glume, by using bulked segregant exome capture sequencing (BSE-Seq), while *Hg* was only mapped in 0.52–3.26Mb of 1AS. To further fine map this gene and identify candidate genes in this region, a near isogenic line-derived population consisting of 2,050 F_2 lines was generated in the present study. By analyzing this population, *Hg* was fine mapped into a 0.90 cM region covering a physical distance of ~825.03 Kb encompassing 6 high- and 23 low-confidence genes in the reference genome of Chinese Spring. A presence-absence variation was identified in the fine mapping region through analyses of sequence-tagged sites markers and genome sequences of the hairy glume parent of the near isogenic lines. The results presented here will be useful for further cloning *Hg* in wheat.

KEYWORDS

fine mapping, BSE-Seq, hairy glume, wheat, presence-absence variation

Introduction

Trichome (hair or pubescence) is a kind of unicellular or multicellular appendage with diverse functions on the plant epidermis in different organs (Werker, 2000). Previous studies showed that trichomes play important roles in resistance to a wide range of both biotic and abiotic stresses (Bickford, 2016; Galdon-Armero et al., 2018) as well as in plant defense

regard to phytophagous insects (Handley et al., 2005). Additionally, hairy leaf sheath was significantly and positively correlated with increased grain yield, grain weight and grain weight per spike (Wan et al., 2015). However, glume trichome is undesirable based on some recent studies as it may be harmful to human health. It has been reported glume trichome is associated with the increased risk of developing esophageal squamous cell carcinoma following the consumption of contaminated wheat flour (Lian et al., 2019). In addition, it generates dust during harvesting and grain processes causing acute and chronic irritation to eyes, skin, and respiratory systems (Zhang et al., 2012).

QTL mapping has been routinely used to identify initial genomic regions affecting specific traits. In wheat, eight QTL conferring glume trichome have been identified. A genome-wide association study for 404 Indian bread wheat genotypes identified six putative QTL conferring hairy glume on 1A (at 292 and 511 Mb), 1B and 2B (Sheoran et al., 2019). Wu et al. (2021) reported a hairy glume gene on 2BL which was named *Hg2*. *Hg* has been long recognized as the dominant gene for glume hairiness (Kadam, 1936; Tsunewaki, 1961; Dubcovsky and Dvorák, 1995; Khlestkina et al., 2006) and it was initially mapped in a 3.3 cM region in our previous study (Luo et al., 2016). However, QTL mapping only provides limited resolution (Paterson et al., 1988) and molecular markers obtained from such studies can not be reliably used to tag targeted loci. The heterogeneity in genetic background in such mapping populations also makes accurate phenotyping of many quantitative traits difficult (Jiang et al., 2019). Therefore, fixing genetic backgrounds in regarding to a targeted locus becomes efficient for accurate phenotyping and can be achieved with near isogenic lines (NILs) or populations derived from NILs.

To further characterize *Hg*, a lot of efforts have been made including developing NILs and generating RNA sequences. In our previous study, seven pairs of NILs targeting *Hg* have been developed and difference between the two isolines was only detected on glume hairiness (Luo et al., 2020). Taken advantage of the unique feature of NILs where phenotypic difference between the two isolines of NILs mainly depends on the difference between their genomes at target locus, we also adopted multi-NIL RNA-Seq approach (Habib et al., 2017; Gao et al., 2019) to generate and analyze RNA sequences against two of the seven NIL pairs. This approach allowed us to reduce the number of candidate genes and identified 37 differentially expressed genes (DEGs) and 39 SNPs in the target region (Luo et al., 2020).

Cost of high throughput sequencing rapidly decreasing made genomic data across hundreds or even thousands of genotypes available. For example, Guo et al. (2020) released draft genome sequence of Tibetan semi-wild wheat (*Triticum aestivum* ssp. *tibetanum* Shao) and resequencing data of 245 wheat accessions. Genomic data of 641 wheat accessions were available on SnpHub (Wang et al., 2020). A recent study generated ten chromosome pseudomolecule and five scaffold assemblies of hexaploid wheat (Walkowiak et al., 2020). These data would be greatly helpful to explore large panels of high-quality genomic variation data and provide valuable resources for designing markers, identifying trait-related genes, and molecular breeding. Glume trichome in wheat has been studied for a century (Wu et al., 2021), while *Hg* has not been fine mapped. It is an opportunity to analyze it with the explosive sequencing data. Together with the available transcriptomic and genomic data, we also explored candidate region of *Hg* by bulked segregant exome capture sequencing (BSE-Seq) analysis. By developing and exploiting a NIL-derived population consisting of 2,050 F_2 lines, we aimed to fine map *Hg* in the present study.

Materials and methods

Plant materials

Firstly, a population of 260 F_2 plants segregating at *Hg* generated by crossing a hairy wheat '3B1A' (Figure 1A) with a non-hairy wheat accession 'Baiying1' (Figure 1B) was used

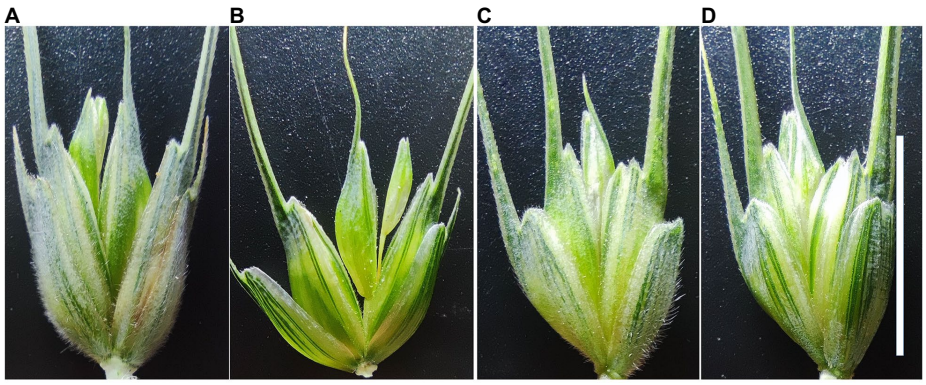


FIGURE 1
Phenotypes of the glume with or without trichome. The four glumes are **(A)** '3B1A', **(B)** 'Baiying1', **(C)** hairy iso-line of NIL1, and **(D)** non-hairy iso-line of NIL1, respectively. Scale bar=1cm.

for genetic analysis of *Hg*. These F_2 plants were harvested and further cultivated (20 seeds of each line), and the trichome trait for each plant was investigated for confirming the phenotype of each line.

Secondly, fifty hairy and fifty non-hairy lines were randomly selected from a population of recombinant inbred lines (RILs) derived from the cross between Tibetan semi-wild wheat (*Triticum aestivum* ssp. *tibetanum* Shao) 'Q1028' (hairy glume) and variety 'Zhengmai9023' (non-hairy glume) from a previous study (Luo et al., 2016) to confirm the initial location of *Hg* and refine makers flanking the targeted region in this study. In addition, seven pairs of NILs about glume trichome were also developed from the RIL population in a previous study (Luo et al., 2020). In the present study, an F_2 population consisting of 2,050 plants was generated by crossing the two isolines of NIL1 (Figures 1C,D) to fine map *Hg*.

Phenotyping

The 260 F_2 plants of '3B1A' × 'Baiying1' materials were planted in 1.5 m rows with 0.3 m space between rows in Wenjiang experimental farm of the Triticeae Research Institute, Sichuan Agricultural University, China. The 2,050 F_2 plants derived from the crossing of the two isolines of NIL1 were cultivated in the 5 × 10 plug tray with 50 cm in length × 28 cm in width × 10 cm in height. The tray was placed in the greenhouse (constant temperature at 22°C, light for 16 h and darkness for 8 h) of Triticeae Research Institute, Sichuan Agricultural University. The phenotypes of wheat glumes were investigated and recorded as hairy or non-hairy.

BSE-Seq analysis

The approximately equal leaf of 32 homozygous hairy and 32 non-hairy F_3 lines of '3B1A' × 'Baiying1' were sampled and pooled separately. DNA of the two pools and parents were extracted using the cetyltrimethylammonium bromide method. Exome capture sequencing (Illumina HiSeq Nova platform) and analysis for the four libraries ('3B1A', 'Baiying1', hairy pool, and non-hairy pool) were performed by Tiancheng Weilai Technology[1] (Chengdu, China) using a wheat exome capture panel. It contained 268.9 Mb capture space and covered 107,400 high confidence (HC) genes and 132,688 HC transcripts of International Wheat Genome Sequencing Consortium (IWGSC) Annotation v1.1 (Dong et al., 2020).

Both SNP-index and G' methods were used to screen the significant region for bulk segregant analysis, and R package QTLseqr (Mansfeld and Grumet, 2018) was used to calculate SNP-index and G' based on sliding window. Allelic frequency with the significance of linearity was performed using Euclidean distance (ED; Hill et al., 2013). ED value of each locus was calculated, the 4th power of the original ED was set as correlation value to eliminate background noise, and Loess was used to fit the ED value (Ji et al., 2021). varBScore was also used for BSE-Seq analysis (Dong et al., 2020).

Data collection and molecular marker development

The resequencing data of 641 wheat accessions were collected and analyzed on the database of SnpHub (Wang et al., 2020). In our previous study, 83 of the 641 accessions including 24 hairy and 59 non-hairy genotypes were observed (Supplementary Table S1; Luo et al., 2019). Resequencing data of these 83 genotypes were downloaded and used to detect the InDel and SNP variations in the candidate region.

The flanking genomic sequences of InDel (more than 14 base pairs) were used as templates to develop InDel markers. HC genes from IWGSC Annotation v1.1 identified from the initial mapping region were used to detect the sequence variants between the two isolines of the NIL1. The variations between the NIL1 were used for developing PCR-based markers, including KASP (Kompetitive Allele-Specific PCR) and CAPS (cleaved amplified polymorphic sequence) markers. Some HC and low confidence (LC) genes in the fine mapping region were used as templates to develop sequence-tagged sites (STS) markers. Primers for CAPS and KASP markers were designed on websites[2,3] and those for the others were designed using WheatOmics 1.0 (Ma et al., 2021).

Primer synthesis and gene sequencing were completed by Tsingke Biotechnology Co., Ltd. (Chengdu, China[4]). The amplification for KASP markers were carried out with KASP-TF V4.0 2× Master Mix (LGC Genomics, Hoddeson, United Kingdom) using CFX96 Touch Real-Time PCR Detection System (Bio-Rad, Hercules, United States) according to the manufacturer's instructions. Other markers were amplified by 2× Taq Master Mix (P213; Vazyme Biotech Co., Ltd., Nanjing, China).

The detection of chromosome variation in the targeted region

In order to detect the chromosome variation in the fine mapping region, the STS markers were amplified in NIL1. The amplification products were analyzed by 2% agarose gel electrophoresis. Furthermore, the chromosome variation was

1 http://www.tcuni.cn

2 https://galaxy.triticeaetoolbox.org/

3 http://www.polymarker.info/

4 https://tsingke.com.cn/

TABLE 1 Chi-square test for hairy:non-hairy glume.

Population	Total F_2 plants	Hairy plants	Non-hairy plants	χ^2
3B1A × Baiying1	260	195	65	0.01
Hairy × non-hairy isoline	2,050	1,510	540	1.90

$\chi^2_{0.05} = 3.84$

preliminarily analyzed using the genomic sequence of the hairy glume parent of NIL1.

The 30-fold wheat whole-genome data of 'Q1028' was provided by Triticeae Research Institute, Sichuan Agricultural University (unpublished). The leaves of 'Q1028'were used to extract genomic DNA using Plant Genomic DNA Kit of Tiangen Biotech Co., Ltd. (Beijing, China). Genomic DNA of 'Q1028' was sent to BerryGenomics (Beijing, China) for resequencing using 10× Genomics. After quality control, sequencing reads were mapped to IWGSC Reference Sequences (RefSeq) v1.0 (IWGSC, 2018) using Resequencing Analysis function of Toolbox on CLC Genomics Workbench 12.0.[5] The parameters for mapping resequencing reads of 'Q1028' to reference genomes were shown in Supplementary Table S2. Then the Whole Genome Coverage Analysis was performed in the Resequencing Analysis function of Toolbox on CLC Genomics Workbench 12.0 with the following parameters: P -Value threshold = 0.0001, Minimum length = 50, Create regions = Yes, Create report = Yes.

Results

Genetic analysis of trichome gene

Chi-square test demonstrated that the ratio of hairy:non-hairy glume fit the 3:1 segregation ratio in the F_2 populations of '3B1A' × 'Baiying1' (Table 1). The 2,050 NIL-derived F_2 plants derived from the cross between the NIL1 also displayed a 3:1 segregation ratio for the hairy and the non-hairy glume (Table 1). These results indicated that hairy glume trait was controlled by a dominant gene in '3B1A' and hairy isoline of NIL1.

Chromosome localization of trichome gene in BSE-Seq analysis

BSE-Seq generated 29, 26, 62 and 57 Gb raw data from '3B1A', 'Baiyin1', the hairy and non-hairy pools, respectively, and their data size reached about 109, 101, 234 and 214 times to the exome capture panel (Supplementary Table S3). The Q30 of the library was over 91% (Supplementary Table S3). Furthermore, the percentage of properly paired reads was more than 98% in each library (Supplementary Table S3). A total of 13,895 nonsymmetrical variations were obtained from the two bulked pools, and half of these variations were concentrated at chromosome 1A (Figure 2A; Supplementary Table S4). The variations at the distal terminus of 1AS was the densest, and several regions with no SNPs were detected on this interval, for example, 1.34–2.16 (only 21 sites were detected around 1.7 Mb in this gap) Mb (Figure 2B; Supplementary Table S4).

In the varBScore analysis, the peak scoring points were distributed in three regions on chromosome 1A, namely, 1.20–1.21, 6.56–6.57 and 10.72–12.59 Mb (Supplementary Figure S1). The results of tricube-smoothed Δ(SNP-index) showed that only one region existed on 1AS (0.52–3.26 Mb) with a confidence interval of larger than 99%, indicating that a single locus for glume trichome was defined in the genome (Figure 3). A similar result was also detected in G' analysis, and the peak values of tricube-smoothed Δ(SNP-index) and G' were both at 521,162 bp on 1AS (Supplementary Figure S2). In the result of ED analysis, the top 1% fitted ED values were distributed in 0.52–14.08 Mb (Supplementary Figure S3), and the peak value was also at 521,162 bp. These results suggested that the candidate gene of *Hg* may be located at 0.52–3.26 Mb of 1AS, and a gap, 1.34–2.16 Mb, existed in the region.

Identification of the genomic region containing *Hg*

A total of 27,607 polymorphic sites (Figure 4A) among the 83 resequencing accessions were obtained including 22,046 SNPs and 5,561 InDels. Of them, 163 InDel markers (Supplementary Table S5) and 19 KASP markers (Supplementary Table S6) were developed and used to screen against the NIL1, and eight markers were detected polymorphism between the two isolines.

The newly developed markers were then used to screen the 100 RILs to confirm the initial location of *Hg*. *Hg* was located in the region between InDel2 and InDel5 (Figure 4B). After that, *TraesCS1A02G002700* (one of the closest gene to InDel2) was isolated based on the chromosome 1A specific primers (Supplementary Table S7) and sequenced in NIL1, and a 122 bp sequence was detected to be inserted in the hairy isoline. An InDel marker, InDel2700, was designed and mapped closer to *Hg* in the RIL population (Figure 4B).

Fine mapping of *Hg*

Based on the positions and sequences of the two flanking markers, twelve genes (Supplementary Table S7) were sequenced from the two isolines of NIL1 and three polymorphic markers

5 https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/analysis-and-visualization/qiagen-clc-genomics-workbench/

FIGURE 2
The nonsymmetrical variations of BSE-Seq analysis. The variable sites in the genome **(A)** and 1–10Mb of chromosome 1A **(B)**.

(CAPS3400, KASP3700, and KASP4900; Supplementary Table S8) were generated and used to genotype the population and construct a high-density genetic map covering the targeted region. Finally, *Hg* was placed into a 0.90 cM region between InDel2700 and CAPS3400 covering a physical region of ~825.03 Kb (1,337,248-2,162,275 bp; Figure 4C). By assessing the whole population of 2,050 lines with the two flanking markers (InDel2700 and CAPS3400), 30 recombinants were identified (Figure 4D). Phenotyping of these recombinants found that 23 were hairy and the other 7 were non-hairy (Figure 4D). Six HC and 23 LC genes (Figure 4E; Supplementary Table S8) were annotated in this fine mapping region according to IWGSC Annotation v1.1.

A chromosome variation in the fine mapping region

In the fine mapping region, some genes can be easily amplified in non-hairy isoline of NIL1 while no bands or target sequences were obtained in the hairy isoline. For instance, the target sequence for the homolog *TraesCS1A02G003000*, was obtained in the non-hairy isolines of NIL1, while the amplified sequence from hairy isolines was more similar to the homoeolog on 1B (Supplementary Figure S4).

STS markers were then designed for the conserved region of 6 HC and 5 LC genes (Figure 4E; Supplementary Table S9). As expected, the 11 STS and 2 InDel markers were detected in

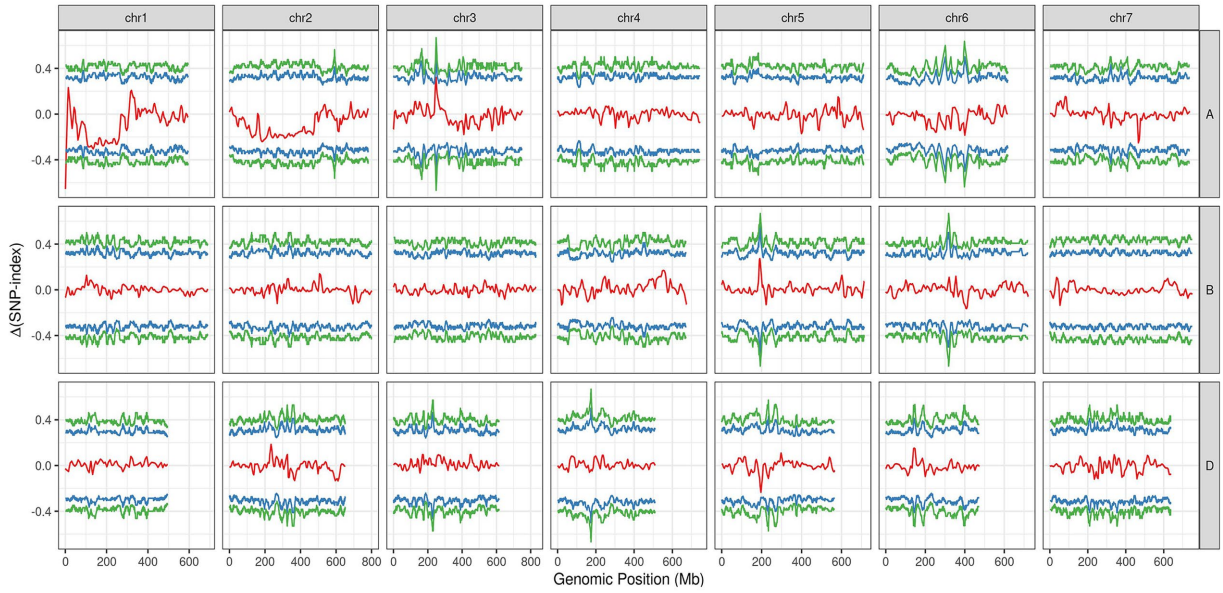


FIGURE 3
The tricube-smoothed Δ(SNP-index; red line) of BSE-Seq analysis. The blue and green lines indicated for confidence intervals of 95 and 99%, respectively.

non-hairy isoline of NIL1 (Figure 4F). However, the markers between *TraesCS1A02G003500LC* and *TraesCS1A02G005300LC* were not detected in hairy isoline of NIL1 (Figure 4F). These results indicated that a presence-absence variation (PAV) might exist in the fine mapping region.

A gap was detected in the region from ~1.40 Mb to ~2.10 Mb when mapping the resequencing reads of Q1028 to IWGSC RefSeq v1.0. There was a low coverage region with a few reads and most of these reads showed huge difference relative to the reference sequence. For example, a gap was shown in 1.82–1.90 Mb, and a normal region was shown in 2.17–2.25 Mb (Figure 5). It was found that the number of abnormal reads in fine mapping region were larger than in the flanking regions (Figure 6). The abnormal high and low coverage reads were quite less and larger than the flanking regions (Figure 6), respectively. Taken together, the PAV was further identified through resequencing of the donor of *Hg*.

Discussion

Hg as the dominant gene for glume trichome has been studied for many years. However, candidate genes underlying this locus remain unknown. In the present study, we developed and assessed a large NIL-derived population targeting this locus and finally placed this gene into a 0.90 cM genetic region corresponding to a ~825.03 Kb physical region (1,337,248–2,162,275 bp) on chromosome 1AS. A PAV was identified in the candidate region according to amplification results of STS markers and sequencing data analysis of Q1028. This PAV at the mapping region may hamper the following work of gene cloning of *Hg*.

It has been reported that prediction of candidate genes *via* fine mapping was impeded by chromosome variation, so alternative approaches have been arisen to solve the problem. Nsabiyera et al. (2019) suggested resequencing of the parents of mapping population and *de novo* assemble of the target region was the ideal way to obtain the functional genes. Lang et al. (2021) used the same method and discovered the gene for pre-harvest sprouting resistance gene from synthetic hexaploid in a PAV region on 3D. Hewitt et al. (2021) isolated the functional gene conferring powdery mildew resistance by single chromosome enrichment sequencing for 7A. These reports provided us clues to explore the candidate genes from a chromosome variation region.

Chromosome variants might be involved in crop domestication and adaptation (Huang et al., 2021). Besides the distal terminus of 1AS, chromosome variation has been reported at the distal terminus of other chromosomes. For example, a heat stress tolerance gene was mapped in the PAV region at the distal terminus of 4AL (Zhai et al., 2021). Bunt resistance gene *Bt9* was located on the distal terminus of 6DL, and PAV markers for *Bt9* selection had been identified (Morris and Beecher, 2012). Furthermore, there was a rearrangement at the distal region of chromosome 7AL (Hewitt et al., 2021). These variants indicated that chromosome extremities were the most variational regions in wheat (De Oliveira et al., 2020), which was proved by wheat-rice comparative genomics, and gene evolution appeared preferentially at chromosome extremities due to the high frequency of recombination duplication and divergence (See et al., 2006).

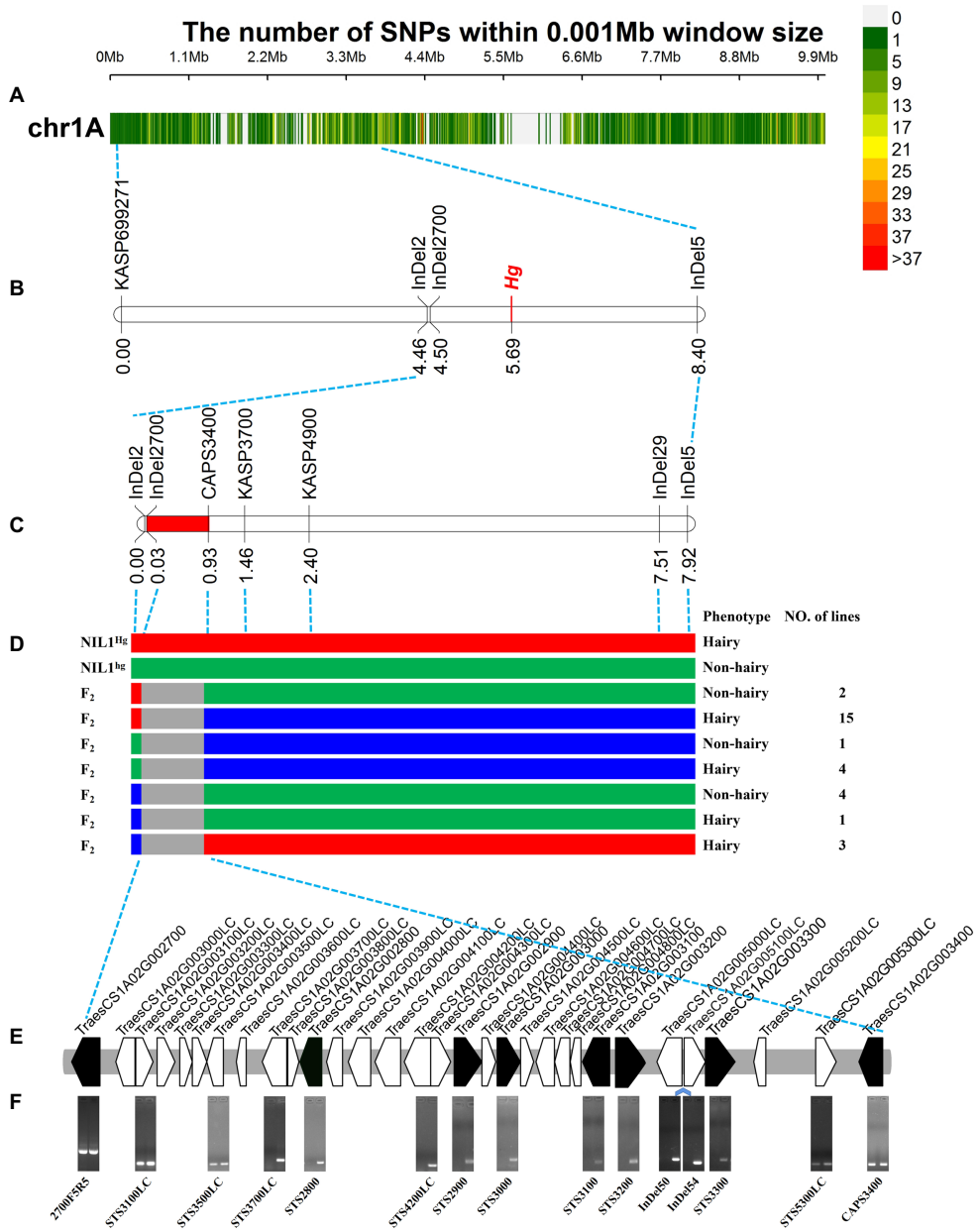


FIGURE 4
Fine mapping of *Hg*. **(A)** The variations in 0–10Mb among the 83 released resequencing wheat accessions. **(B)** The primary mapping of *Hg* with the markers based on the resequencing data. **(C)** The fine mapping region of *Hg*. **(D)** The recombinants from the fine mapping region, the three colors indicated the genotypes of hairy (red), heterozygous (blue), and non-hairy lines (green), respectively. **(F)** The genes in fine mapping region from IWGSC Annotation v1.1. **(D)** The amplification of the primers for the genes in the mapping region. The bands of *TraesCS1A02G002700* were the amplification products using sequencing primers of 2700F5 and 2700R5; STS4200LC was designed based on the overlapping sequence of *TraesCS1A02G004200LC* and *TraesCS1A02G004300LC*; InDel50 and InDel54 were two InDel markers based on the re-sequencing data in the region between *TraesCS1A02G005000LC* and *TraesCS1A02G005100LC*; the bands of *TraesCS1A02G003400* were the amplification products of CAPS3400F and CAPS3400R (Enzyme digestion has not been carried out). The left and right lanes were the hairy and non-hairy isolines, respectively.

We found that the percentage of hairy glume in released wheat cultivars is much less than that of wild or semi-wild wheat (Luo et al., 2019). Furthermore, 35.7% of the genes in bread wheat were dispensable (Montenegro et al., 2017). It was reported that trichome is not necessary for the growth and development of the plant itself although it provides resistance to biotic and abiotic stress (Zhang et al., 2020). Trichome on wheat glume may be a dispensable structure controlled by dispensable genes, while it may provide clues for studying the evolution and domestication of wheat.

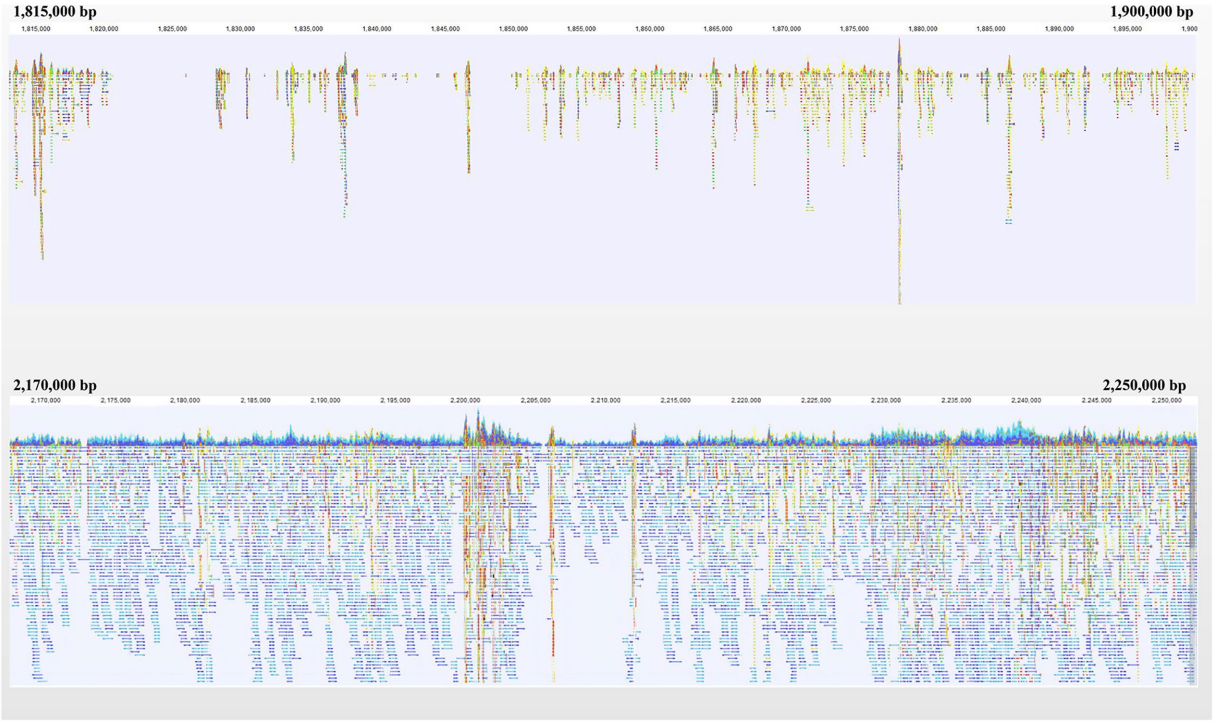


FIGURE 5
Navigation overview of two reads mapping regions on Chromosome 1A. Each color block represented a read. Each color block represents a read. The closer it is to blue or red, the more similar or different the read is to the reference, respectively.

Conclusion

In summary, *Hg* was fine mapped in a ~ 825.03 Kb physical region on chromosome 1A, and a PAV was identified in this region. Assembling of resequencing data may be an effective way to explore *Hg*.

Data availability statement

The data presented in the study are deposited in the National Genomics Data Center (NGDC) repository, accession number PRJCA010932.

Author contributions

WL finished the study and wrote this manuscript. JZ participated in glasshouse work, phenotype investigation, and data analysis. JL, YL, and YM helped phenotype investigation. HT helped data analysis and maker development. QX and MD collected and analyzed data. QJ, GC, and PQ helped with data analysis and revision. JW and YJ analyzed resequencing data. ZC performed BSE-Seq analysis. ZZ and YW revised the manuscript. YZ and XL discussed results. JM designed the experiments, guided the entire study, and extensively revised this manuscript. All authors participated in the research and approved the final manuscript.

Funding

This research was funded by the National Natural Science Foundation of China (31970243 and 31971937), Sichuan Science and Technology Program (2022YFH0053, 2022NSFSC1729, and 2021YFH0083), and the Basic Research Project of Science and Technology Plan of Guizhou Province (ZK[2021] General 131). We thank the anonymous referees for critical reading and revising this manuscript.

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the

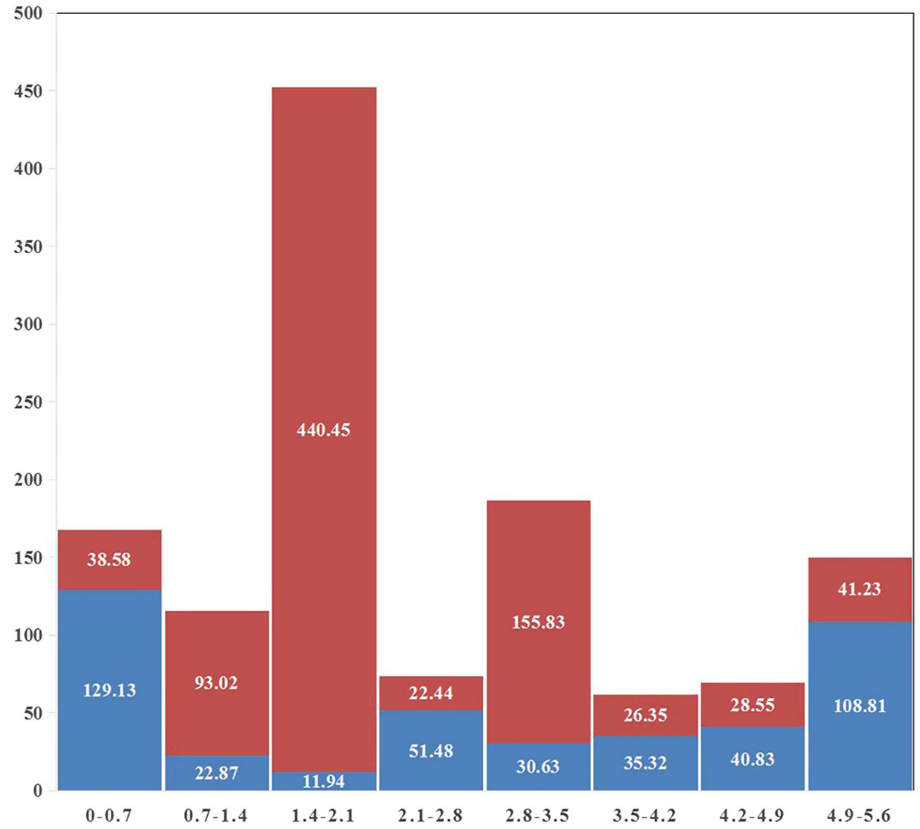


FIGURE 6
Statistics of abnormal reads coverage around fine mapping interval. Red and blue indicated low and high coverage, respectively. The abscissa was the physical map of 1A (MB) and the ordinate was the total length of abnormal reads (KB).

reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1006510/full#supplementary-material

SUPPLEMENTARY FIGURE S1
Result of varBScore analysis for BSE-Seq. The above and below figures were drawn based on the varBScore of the variable sites of the whole genome and chromosome 1A, respectively.

SUPPLEMENTARY FIGURE S2
Tricube-smoothed G' value for BSE-Seq.

SUPPLEMENTARY FIGURE S3
Result of ED analysis for BSE-Seq. The above and below figures were drawn based on the fitted curve and original value, respectively.

SUPPLEMENTARY FIGURE S4
Sequence comparison of amplified *TraesCS1A02G003000* in NIL1.

References

Bickford, C. P. (2016). Ecophysiology of leaf trichomes. *Funct. Plant Biol.* 43, 807–814. doi: 10.1071/FP16095

De Oliveira, R., Rimbert, H., Balfourier, F., Kitt, J., Dynomant, E., Vrana, J., et al. (2020). Structural variations affecting genes and transposable elements of chromosome 3B in wheats. *Front. Genet.* 11, 891. doi: 10.3389/fgene.2020.00891

Dong, C., Zhang, L., Chen, Z., Xia, C., Gu, Y., Wang, J., et al. (2020). Combining a new exome capture panel with an effective varBScore algorithm accelerates BSA-based gene cloning in wheat. *Front. Plant Sci.* 11, 1249. doi: 10.3389/fpls.2020.01249

Dubcovsky, J., and Dvorák, J. (1995). Ribosomal RNA multigene loci: nomads of the Triticeae genomes. *Genetics* 140, 1367–1377. doi: 10.1093/genetics/140.4.1367

Galdon-Armero, J., Fullana-Pericas, M., Mulet, P. A., Conesa, M. A., Martin, C., and Galmes, J. (2018). The ratio of trichomes to stomata is associated with water use efficiency in *Solanum lycopersicum* (tomato). *Plant J.* 96, 607–619. doi: 10.1111/tpj.14055

Gao, S., Zheng, Z., Powell, J., Habib, A., Stiller, J., Zhou, M., et al. (2019). Validation and delineation of a locus conferring *Fusarium crown* rot resistance on 1HL in barley by analysing transcriptomes from multiple pairs of near isogenic lines. *BMC Genomics* 20, 650. doi: 10.1186/s12864-019-6011-8

Guo, W., Xin, M., Wang, Z., Yao, Y., Hu, Z., Song, W., et al. (2020). Origin and adaptation to high altitude of Tibetan semi-wild wheat. *Nat. Commun.* 11, 5085. doi: 10.1038/s41467-020-18738-5

Habib, A., Powell, J. J., Stiller, J., Liu, M., Shabala, S., Zhou, M., et al. (2017). A multiple near isogenic line (multi-NIL) RNA-seq approach to identify candidate

genes underpinning QTL. *Theor. Appl. Genet.* 131, 613–624. doi: 10.1007/s00122-017-3023-0

Handley, R., Ekbom, B., and Ågren, J. (2005). Variation in trichome density and resistance against a specialist insect herbivore in natural populations of *Arabidopsis thaliana*. *Ecol. Entomol.* 30, 284–292. doi: 10.1111/j.0307-6946.2005.00699.x

Hewitt, T., Müller, M. C., Molnár, I., Mascher, M., Holušová, K., Šimková, H., et al. (2021). A highly differentiated region of wheat chromosome 7AL encodes a *Pm1a* immune receptor that recognizes its corresponding *AvrPm1a* effector from *Blumeria graminis*. *New Phytol.* 229, 2812–2826. doi: 10.1111/nph.17075

Hill, J. T., Demarest, B. L., Bisgrove, B. W., Gorsi, B., Su, Y. C., and Yost, H. J. (2013). MMAPPR: mutation mapping analysis pipeline for pooled RNA-seq. *Genome Res.* 23, 687–697. doi: 10.1101/gr.146936.112

Huang, Y., Huang, W., Meng, Z., Braz, G. T., Li, Y., Wang, K., et al. (2021). Megabase-scale presence-absence variation with Tripsacum origin was under selection during maize domestication and adaptation. *Genome Biol.* 22, 237. doi: 10.1186/s13059-021-02448-2

IWGSC (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361:eaar7191. doi: 10.1126/science.aar7191

Ji, G., Xu, Z., Fan, X., Zhou, Q., Yu, Q., Liu, X., et al. (2021). Identification of a major and stable QTL on chromosome 5A confers spike length in wheat (*Triticum aestivum* L.). *Mol. Breed.* 41:56. doi: 10.1007/s11032-021-01249-6

Jiang, Y., Habib, A., Zheng, Z., Zhou, M., Wei, Y., Zheng, Y.-L., et al. (2019). Development of tightly linked markers and identification of candidate genes for Fusarium crown rot resistance in barley by exploiting a near-isogenic line-derived population. *Theor. Appl. Genet.* 132, 217–225. doi: 10.1007/s00122-018-3209-0

Kadam, B. (1936). Genetics of the Bansi wheat of the Bombay-Deccan and a synthetic Khapli-part I. *Proc. Indian As. B.* 4, 357–369.

Khlestkina, E. K., Pshenichnikova, T. A., Roder, M. S., Salina, E. A., Arbuzova, V. S., and Borner, A. (2006). Comparative mapping of genes for glume colouration and pubescence in hexaploid wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 113, 801–807. doi: 10.1007/s00122-006-0331-1

Lang, J., Fu, Y., Zhou, Y., Cheng, M., Deng, M., Li, M., et al. (2021). *Myb10-D* confers *PHS-3D* resistance to pre-harvest sprouting by regulating *NCED* in ABA biosynthesis pathway of wheat. *New Phytol.* 230, 1940–1952. doi: 10.1111/nph.17312

Lian, C., Zuo, X., and Tian, L. (2019). A possible role of biogenic silica in esophageal cancer in North China? *Environ. Sci. Pollut. R.* 26, 8340–8343. doi: 10.1007/s11356-019-04332-w

Luo, W., Liu, J., Ding, P., Li, C., Liu, H., Mu, Y., et al. (2020). Transcriptome analysis of near-isogenic lines for glume hairiness of wheat. *Gene* 739:144517. doi: 10.1016/j.gene.2020.144517

Luo, W., Liu, J., Ding, P., Zou, Y., Li, T., Xie, Q., et al. (2019). Domestication of glume hairiness in wheat. *J. Sichuan Agr. Univ.* 37, 743–754. doi: 10.16036/j.issn.1000-2650.2019.06.001

Luo, W., Ma, J., Zhou, X.-h., Jiang, Y.-f., Sun, M., Yang, Y.-j., et al. (2016). Genetic analysis of glume hairiness (*Hg*) gene in bread wheat (*Triticum aestivum* L.). *Genet. Resour. Crop Evol.* 63, 763–769. doi: 10.1007/s10722-016-0393-0

Ma, S., Wang, M., Wu, J., Guo, W., Chen, Y., Li, G., et al. (2021). WheatOmics: A platform combining multiple omics data to accelerate functional genomics studies in wheat. *Mol. Plant* 14, 1965–1968. doi: 10.1016/j.molp.2021.10.006

Mansfeld, B. N., and Grumet, R. (2018). QTLseqr: an R package for bulk segregant analysis with next-generation sequencing. *Plant Genome* 11:180006. doi: 10.3835/plantgenome2018.01.0006

Montenegro, J. D., Golicz, A. A., Bayer, P. E., Hurgobin, B., Lee, H., Chan, C. K., et al. (2017). The pangenome of hexaploid bread wheat. *Plant J.* 90, 1007–1013. doi: 10.1111/tpj.13515

Morris, C. F., and Beecher, B. S. (2012). The distal portion of the short arm of wheat (*Triticum aestivum* L.) chromosome 5D controls endosperm vitreosity and grain hardness. *Theor. Appl. Genet.* 125, 247–254. doi: 10.1007/s00122-012-1830-x

Nsabiyera, V., Baranwal, D., Qureshi, N., Kay, P., Forrest, K., Valarik, M., et al. (2019). Fine mapping of *Lr49* using 90K SNP chip array and flow-sorted shromosome sequencing in wheat. *Front. Plant Sci.* 10, 1787. doi: 10.3389/fpls.2019.01787

Paterson, A. H., Lander, E. S., Hewitt, J. D., Peterson, S., Lincoln, S. E., and Tanksley, S. D. (1988). Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. *Nature* 335, 721–726. doi: 10.1038/335721a0

See, D. R., Brooks, S., Nelson, J. C., Brown-Guedira, G., Friebe, B., and Gill, B. S. (2006). Gene evolution at the ends of wheat chromosomes. *Proc. Natl. Acad. Sci. U. S. A.* 103, 4162–4167. doi: 10.1073/pnas.0508942102

Sheoran, S., Jaiswal, S., Kumar, D., Raghav, N., Sharma, R., Pawar, S., et al. (2019). Uncovering genomic regions associated with 36 agro-morphological traits in Indian spring wheat using GWAS. *Front. Plant Sci.* 10, 527. doi: 10.3389/fpls.2019.00527

Tsunewaki, K. (1961). Monosomic and conventional gene analyses in common wheat IV. Glume hairiness and ear density. *JPN. J. Genet.* 36, 55–62.

Walkowiak, S., Gao, L., Monat, C., Haberer, G., Kassa, M. T., Brinton, J., et al. (2020). Multiple wheat genomes reveal global variation in modern breeding. *Nature* 588, 277–283. doi: 10.1038/s41586-020-2961-x

Wan, H., Yang, Y., Li, J., Zhang, Z., and Yang, W. (2015). Mapping a major QTL for hairy leaf sheath introgressed from *Aegilops tauschii* and its association with enhanced grain yield in bread wheat. *Euphytica* 205, 275–285. doi: 10.1007/s10681-015-1457-5

Wang, W., Wang, Z., Li, X., Ni, Z., Hu, Z., Xin, M., et al. (2020). SnpHub: an easy-to-set-up web server framework for exploring large-scale genomic variation data in the post-genomic era with applications in wheat. *GigaScience* 9:giaa060. doi: 10.1093/gigascience/giaa060

Werker, E. (2000). Trichome diversity and development. *Adv. Bot. Res.* 31, 1–35. doi: 10.1016/S0065-2296(00)31005-9

Wu, P., Yang, L., Guo, G., Hu, J., Qiu, D., Li, Y., et al. (2021). Molecular mapping and identification of a candidate gene for new locus *Hg2* conferring hairy glume in wheat. *Plant Sci.* 307:110879. doi: 10.1016/j.plantsci.2021.110879

Zhai, H., Jiang, C., Zhao, Y., Yang, S., Li, Y., Yan, K., et al. (2021). Wheat heat tolerance is impaired by heightened deletions in the distal end of 4AL chromosomal arm. *Plant Biotechnol. J.* 19, 1038–1051. doi: 10.1111/pbi.13529

Zhang, Y., Song, H., Wang, X., Zhou, X., Zhang, K., Chen, X., et al. (2020). The roles of different types of trichomes in tomato resistance to cold, drought, whiteflies, and botrytis. *Agronomy* 10, 411. doi: 10.3390/agronomy10030411

Zhang, H., Wu, K., Wang, Y., Peng, Y., Hu, F., Wen, L., et al. (2012). A *WUSCHEL*-like homeobox gene, *OsWOX3B* responses to *NUDA/GL-1* locus in rice. *Rice* 5, 30. doi: 10.1186/1939-8433-5-30

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 19 October 2022
DOI 10.3389/fpls.2022.1042399

OPEN ACCESS

EDITED BY
Xiaoli Fan,
Chengdu Institute of Biology, Chinese Academy of Sciences (CAS), China

REVIEWED BY
Jindong Liu,
Institute of Crop Sciences, (CAAS), China
Liqiang Song,
Hebei Agricultural University, China

*CORRESPONDENCE
Meicheng Zhao
mczhao@sjziam.ac.cn
Diaoguo An
andiaoguo@163.com

†These authors have contributed equally to this work

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

RECEIVED 12 September 2022
ACCEPTED 06 October 2022
PUBLISHED 19 October 2022

CITATION
Liu H, Han G, Gu T, Jin Y, Shi Z, Xing L, Yan H, Wang J, Hao C, Zhao M and An D (2022) Identification of the major QTL *QPm.cas-7D* for adult plant resistance to wheat powdery mildew. *Front. Plant Sci.* 13:1042399. doi: 10.3389/fpls.2022.1042399

Identification of the major QTL *QPm.cas-7D* for adult plant resistance to wheat powdery mildew

Hong Liu[1†], Guohao Han[1†], Tiantian Gu[1†], Yuli Jin[1], Zhipeng Shi[1], Lixian Xing[1], Hanwen Yan[1], Jing Wang[1], Chenyang Hao[2], Meicheng Zhao[1*] and Diaoguo An[1,3*]

[1]Center for Agricultural Resources Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Shijiazhuang, China, [2]The National Key Facility for Crop Gene Resources and Genetic Improvement, Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing, China, [3]Innovative Academy of Seed Design, Chinese Academy of Sciences, Beijing, China

Developing effective and durable host plant resistance is crucial for controlling powdery mildew, a devastating disease caused by *Blumeria graminis* f. sp. *tritici* (*Bgt*). In the present study, we dissected the genetic basis of the adult plant resistance to powdery mildew using a recombinant inbred line (RIL) composed of 176 F_9 RILs population derived from a cross between PuBing 3228 (P3228) and susceptible cultivar Gao 8901. P3228 exhibits stable adult-plant resistance to powdery mildew in the field over consecutive years. We identified two QTLs on chromosomes 7DS (*QPm.cas-7D*) and 1AL (*QPm.cas-1A*) contributed by P3228, and one QTL on 3DS (*QPm.cas-3D*) contributed by Gao 8901, which could explain 65.44%, 3.45%, and 2.18% of the phenotypic variances, respectively. By analyzing the annotated genes in the 1.168 Mb physical interval of the major QTL *QPm.cas-7D*, we locked a previously cloned adult-plant resistance gene *Pm38* that was most probably the candidate gene of *QPm.cas-7D*. Sequence alignment analysis revealed that the candidate gene of *QPm.cas-7D* in P3228 was identical to the reported *Pm38* sequence. Two haplotypes *QPm-7D-R* and *QPm-7D-S* were identified in the whole *Pm38* genomic regions between P3228 and Gao 8901. To apply *QPm.cas-7D* in wheat breeding, we developed a kompetitive allele-specific PCR (KASP) marker *Kasp5249* that is closely linked with these haplotypes. It is worth mentioning that the *QPm-7D-R* haplotype significantly decreased TKW and underwent negative selection for higher yields in China wheat breeding. In this study, we identified a major QTL *QPm.cas-7D* and revealed the relationship between its resistance and yield, which could be beneficial for further applications in wheat disease resistance and high-yield breeding.

KEYWORDS

adult plant resistance, powdery mildew, *Pm38*, haplotypes, molecular markers, common wheat

Introduction

Common wheat (*Triticum aestivum*) is an important contributor to national food security and sustains one-third of humankind (IWGSC, 2018). With an estimated global population of more than nine billion over the next 30 years, wheat production is facing an approximately 70% growth challenge to meet the food demands (IWGSC, 2014). However, wheat powdery mildew, a globally epidemic wheat disease caused by the biotrophic fungus *Blumeria graminis* f. sp. *tritici* (*Bgt*), can severely reduce wheat yields and affect grain quality (Yahiaoui et al., 2004; Singh et al., 2016). In recent decades, the planting area of winter wheat in China affected annually by powdery mildew has reached 6 m ha, resulting in 300,000 tons of yield loss each year (Jia et al., 2020).

Developing effective and durable host plant resistance is crucial for controlling powdery mildew epidemics. Resistance to disease in crops is typically classified into two main patterns: qualitative resistance and quantitative resistance (Spielmeyer et al., 2005; Lillemo et al., 2008; McIntosh et al., 2019; He et al., 2021). Qualitative resistance is mostly race-specific where resistance (*R*) gene based and confers strong and life-long immunity at all stages (Kang et al., 2020; Sánchez-Martín et al., 2021). However, *Bgt* isolates have complex and highly variable virulence structures, so their constant evolution causes the constant breakdown of *R* genes, particularly in areas where *R* genes were widely used (An et al., 2019). Different from qualitative resistance, quantitative resistance which conferred by polygenes is mostly non-race-specific, of which adult plant resistance (APR) is one of the main types and exhibits effectiveness at the post-seedling stages. APR usually cannot display complete immunity, to great extent, which reduces selection pressure on pathogens (Li et al., 2014). Together, these two forms of resistance have provided the genetic basis of powdery mildew resistance in wheat.

To date, more than 100 powdery mildew (*Pm*) genes/alleles at 63 loci (*Pm1-Pm68*, noting that *Pm8* = *Pm17*, *Pm18* = *Pm1c*, *Pm22* = *Pm1e*, *Pm23* = *Pm4c*, and *Pm31* = *Pm21*) have been found from common wheat and its wild relatives (He et al., 2018; McIntosh et al., 2020). Most of the 68 formally designated *Pm* genes provided qualitative resistance which showed all stage resistance (ASR), only *Pm38* (Lagudah et al., 2009), *Pm39* (Lillemo et al., 2008), *Pm46* (Moore et al., 2015), and *Pm62* (Zhang et al., 2018) showed APR. Despite numerous *Pm* genes having been reported, most of them cannot be directly applied in wheat production due to unexpected linkage drag or longer breeding cycles required for genes that were from wheat relatives or landraces. For instance, the broad-spectrum powdery mildew resistance gene *Pm16* led to 15% yield loss when it was introgressed into the wheat backgrounds (Summers and Brown, 2013). In the current wheat breeding programs in China, only a few *Pm* genes have been extensively used in wheat improvement (Jin et al., 2021), which are facing huge selective pressure. It is therefore necessary for resistance durability to unceasingly identify and rationally deploy various types of *Pm* genes/alleles from various germplasm resources.

Once the effective gene was identified, the next challenge is its accurate and rapid transfer in breeding programs. Compared with the conventional breeding, marker-assisted selection (MAS) is more accurate for it also combines genotypic identification. The targeted genes could be selected or excluded in fewer generations by using the powerful diagnostic markers (Jiang et al., 2016). Therefore, cloning of target genes/loci and their tightly-linked molecular markers are key points for MAS. Recent progress in whole-genome sequencing and data-processing strategies have greatly promoted the isolation of the resistance genes. Up to now, 13 *Pm* genes, including *Pm1a* (Hewitt et al., 2021), *Pm2* (Sánchez-Martín et al., 2016), *Pm3* (Yahiaoui et al., 2004), *Pm4b* (Sánchez-Martín et al., 2021), *Pm5e* (Xie et al., 2020), *Pm8* (Hurni et al., 2013), *Pm17* (Singh et al., 2018), *Pm21* (He et al., 2018; Xing et al., 2018), *Pm24* (Lu et al., 2020), *Pm38/Yr18/Lr34/Sr57* (Krattinger et al., 2009), *Pm41* (Li et al., 2020), *Pm46/Yr46/Lr67/Sr55* (Moore et al., 2005) and *Pm60* (Zou et al., 2018), have been cloned by multiple strategies. Many diagnostic markers based on variations in the functional gene sequence have been consequently developed, such as the functional kompetitive allele-specific PCR (KASP) marker *Pm5e-KASP* for *Pm5e* (Xie et al., 2020), and *STS-Pm24* for *Pm24* (Lu et al., 2020). Those markers have no recombination with the target genes, which highlighted the extremely precise and efficient genotyping results.

PuBing3228 (P3228) is a wheat-*Agropyron cristatum* introgression line (Liu et al., 2020). It exhibits stable resistance to powdery mildew in wheat-growing regions over consecutive years, indicating that it should be a promising resource for durable powdery mildew resistance. To better clarify and use the resistance against powdery mildew in P3228, the objectives of this study were to (i) assess its resistance to powdery mildew, (ii) map the major QTL for powdery mildew resistance and predict candidate gene(s), (iii) reveal the relationship between *Pm* gene (s) and yield traits, (iv) determine the geographical distribution of the major *Pm* gene(s), and (v) develop KASP marker of the candidate gene(s) for MAS breeding.

Materials and methods

Plant materials and field trials

A mapping population composed of 176 F_9 RILs derived from 'P3228 × Gao 8901' was developed by single seed descent method. The wheat line P3228 exhibited resistance to powdery mildew at the adult stage whereas Gao 8901 was highly susceptible (Figure 1). The 176 RILs, and two parents were grown at the Center for Agricultural Resources Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences during 2021-2022 growing seasons (2022SJZ). Hill-drop (20 seeds per hill) were used and all

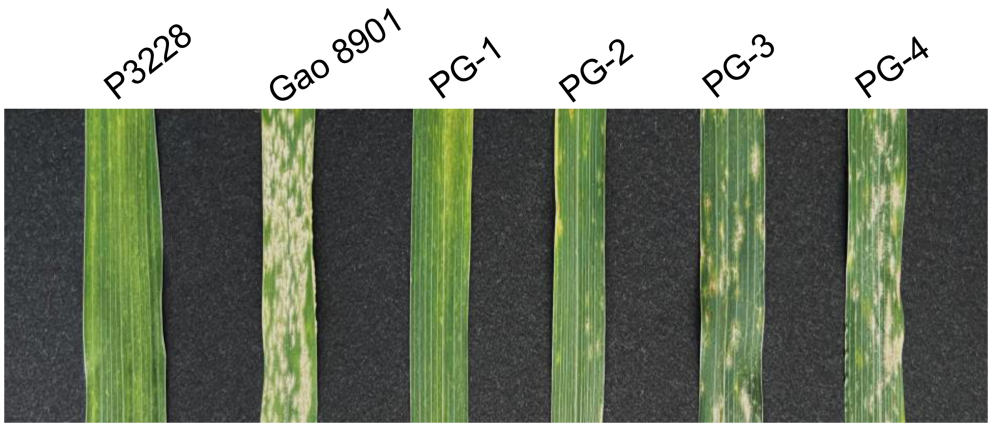


FIGURE 1
Phenotypic characterization of resistance to powdery mildew at the adult stage in two parents and some representative RILs. P3228, PuBing 3228; PG-1, PG-2, PG-3, and PG-4 are representative lines in PG-RIL population.

trials used a randomized block design with three replicates. The water, fertilizer and other management of all field trials were carried out in accordance with local standard practices. Two natural populations including 157 landraces of the Chinese wheat mini-core collection, and 348 modern cultivars from ten ecological zones of China were used for marker screening and association analysis as previously described (Zhao et al., 2019; Liu et al., 2020).

Phenotypic assessment of powdery mildew

Seedling stage reactions of P3228 and Gao 8901 to virulent *Bgt* isolates E09, E11, and E20 were separately tested with three replicates at a greenhouse as previously described (Qie et al., 2019). Heng 4399 was used as the susceptible control. For each line, 20 seeds were planted in rectangular trays (54 cm × 28 cm × 4 cm) with 128 wells (3 cm × 3 cm × 4 cm). After 15 days after inoculation, when sporulation was observed on the first leaf of Heng 4399, the tested plants were scored using a 0-4 scale, in which infection types (ITs) 0-2 considered resistant, while ITs 3-4 considered susceptible (Si et al., 1992).

At the jointing stage, the plants were inoculated with a mixture of *Bgt* isolates E09, E11, and E20. Adult plant reactions to powdery mildew were scored with mean maximum disease severity (MDS). When the susceptible control (Heng 4399) reached 80%, the MDS scores were calculated based on the Cobb scale (Peterson et al., 1948). According to the actual percentage of powdery mildew covered area (0-100%), the severity of infection in the secondary leaves (leaf below flag leaf) of five randomly selected plants in each hill was scored. Disease severities of five selected plants in each line were averaged to obtain a mean severity for each line. Each plant was assessed twice for confirmation.

QTL mapping

A whole-genome genetic map of the PG-RIL population to analysis the genetics of MDS was previously constructed from Wheat 660 K SNP array data (Liu et al., 2020). QTL mapping was conducted in IciMapping v4.1 software by the inclusive composite interval mapping of additive and dominant QTL (ICIM-ADD, Meng et al., 2015). The logarithm of odds (LOD) score≥2.5 (Sun et al., 2013). The MapChart 2.2 was used to draw the genetic map (Voorrips, 2002). The QTLs were named based on McIntosh et al. (2019) and 'cas' represents the Chinese Academy of Sciences.

Comparison of the identified QTLs with the known powdery mildew resistance genes

The physical position of the QTL was identified using the flanking SNP markers sequence of QTL to BLAST against the genome sequences of Chinese Spring v2.1 (Zhu et al., 2021). Candidate genes in the QTL interval was acquired based on coding sequence of Chinese Spring v2.1 and gene function annotations were manually using NCBI Non-redundant protein sequences.

Development of KASP markers

To develop markers that can efficiently trace the powdery mildew resistance genes in P3228 in MAS, KASP marker *Kasp5249* (QPm-7D-FAM: gaaggtgaccaagttcatgctATGGGAGCAT TATTTTTTTCCATCT, QPm-7D-HEX: gaaggtcggagtcaacg

gattATGGGAGCATTATTTTTTTCCATCA, QPm-7D-R: TGCTCATCTCTGGTATGCCATTTAA) was developed based on the distinctive insertions/deletions (InDels) in the targeted interval. KASP assays were performed in 96-well format in 5 μl mixture comprising 2.81 μl of 2 × KASP mix (LGC Genomics, UK), 1μl of DNA template, 1.11 μl of ddH_2O and 0.08 μl of primer mixture. KASP reactions were carried out on an Applied Biosytems™ Veriti™ 96 PCR system (Thermo Fisher, USA). PCR amplification procedure was performed as previously described (Liu et al., 2020). The fluorescence value was read using FLUOstar Omega SNP (LGC Genomics, UK). The KASP genotyping results were read using KlusterCaller genotyping software (LGC Genomics, UK).

Phenotypic evaluation of agronomic traits

Phenotypic traits of 157 landraces and 348 modern cultivars, including thousand kernel weight (TKW), kernel number per spike (KNS), spikelet number per spike (TSS), spike length (SL), effective tiller number (ETN) and plant height (PH), were investigating from plants grown in 2002 and 2005 at Luoyang, Henan province and 2010 at Shunyi, Beijing.

Statistical analyses

The frequency distribution of powdery mildew responses and analysis of variance (ANOVA) was calculated in performed with SPSS Statistics v20.0 software (SPSS, USA). The broad-sense heritability (H^2) was calculated using the QGAStation 2.0 (http://ibi.zju.edu.cn/software/qga/v2.0/indexc.htm) and the following formula $H^2 = VG/VP$; where VG and VP are the genetic variance and phenotypic variance, respectively. Two-tailed t test was performed with SPSS Statistics v20.0 software (SPSS, USA).

Results

Evaluation of powdery mildew resistance and correlation analysis

At the seedling stage, both P3228 and Gao 8901 developed abundant sporulation on the leaves with an IT 4 when inoculated with *Bgt* isolate E09, E11 and E20, respectively. At the adult plant stage, when the MDS of the susceptible control Heng 4399 ranged from 80% to 100%, the P3228 and Gao 8901 showed 1.00% and 67.67%, respectively, showing significant differences on MDS (Table 1). For the RIL population, the frequency distributions of MDS showed continuous variation (Supplemental Figure 1). The MDS scores showed broad-sense heritability (H^2) at 0.63 (Table 1).

QTL mapping

Based on the results of powdery mildew reaction evaluation, two QTLs from P3228 on chromosomes 1A and 7D, and one from Gao 8901 on chromosomes 3D, respectively, were detected in 2022SJZ environment (Table 2 and Figure 2), and were designated as *QPm.cas-1A*, *QPm.cas-3D*, and *QPm.cas-7D*, respectively. The QTL *QPm.cas-1A* was located in the marker interval *AX-109816727–AX-10877999* on the short arm of chromosome 1A and explained 3.45% of the phenotypic variance with an additive effect of -10.58 (Table 2 and Figure 2). The QTL *QPm.caas-3DS* was mapped on chromosome 3DS and flanked by markers *AX-94989783* and *AX-109499958*, which accounted for 2.18% of the phenotypic variance with an additive effect of 5.33 (Table 2 and Figure 2). The major QTL *QPm.cas-7D* was mapped on marker interval *AX-111197303–AX-89471347* on the short arm of chromosome 7D and explained 65.44% of the phenotypic variance with an additive effect of -29.38 (Table 2 and Figure 2).

Predicting of candidate gene *Pm38* for QTL *QPm.cas-7D*

The peak interval of the major QTL *QPm.cas-7D* was collocated between the markers *AX-111197303* and *AX-89471347*. Combined with the physical position of Bin markers *Bin-AX-111197303* (including *AX-111197303* and *AX-89378255*) and *Bin-AX-89471347* (including *AX-89471347*) based on the Chinese Spring reference genome v2.1 (Zhu et al., 2021), the QTL *QPm.cas-7D* was mapped to the 48.917–50.085 Mb position on chromosome arm 7DS. We found 17 high confidence annotated genes in the 1.168 Mb region using Chinese Spring reference genome v2.1 (Table 3). Among them, *Pm38* (*TraesCS7D03G0183600*), a previously cloned adult-plant resistance gene to powdery mildew, was considered as the preferred candidate gene for *QPm.cas-7D*. Then, we analyzed the genomic sequence of *Pm38* from Gao 8901 and P3228 using two

TABLE 1 Phenotypes of the parents and PG-RIL population in this study.

Trait	Parents		PG-RILs					
	P3228	Gao 8901	Minimum	Maximum	Mean	SD	CV(%)	*H*
MDS	1.00	66.67	0.20	100.00	46.61	35.78	76.76	0.63

MDS, maximum disease severities.

TABLE 2 QTLs for maximum disease severities (MDS) in the PG-RIL population.

Trait	QTL	Markers Interval	Genetic Interval (cM)	Physical Interval (Mb)	LOD	PVE%	Add
MDS	*QPm.cas-1A*	*AX-109816727–AX-108779994*	2.932–3.173	6.64–7.63	4.75	3.45	-10.58
	QPm.cas-3D	*AX-94989783–AX-109499958*	51.373–54.408	58.18–65.88	3.13	2.18	5.33
	QPm.cas-7D	*AX-111197303–AX-89471347*	75.65–76.066	48.92–50.09	47.39	65.43	-29.38

LOD, threshold log-of-odds; PVE, phenotypic variance explained; Add, additive effect.

pairs of genome-specific primers as reported to amplify the gene separately from the start codon to exon 14 (ExpF1 and Cssfr6-MR1), and from exon 11 to the stop codon (Cssfr6-MF2 and Lr34-ExpR1) (Fang et al., 2020). The results showed that the *Pm38* allele in P3228 was identical to the previously reported *Pm38*. Moreover, two SNPs (A1654T and C5597T) and two InDels (1-bp InDels at 4996th and 3-bp InDels at 5249th) were identified between P3228 and Gao 8901 in the whole *Pm38* genomic regions, which formed two haplotypes: *QPm-7D-R* (resistant haplotype) and *QPm-7D-S* (susceptible haplotype) (Figure 3A). Meanwhile, sequence alignment revealed that the 1-bp deletion at 4996th caused frameshift mutation, resulting in a loss-of-function *Pm38* protein in Gao 8901 (Figure 3A).

Development of KASP markers and analysis for *Pm38* alleles

Based on the 3-bp InDel on the *Pm38*- homologous sequence between P3228 and Gao 8901, we developed a KASP marker *Kasp5249* (Figure 3B). After screening the PG-RIL population using marker *Kasp5249*, a two-tailed *t* test was performed between the InDel of *Kasp5249* and MDS. The results showed that *Kasp5249* was significantly correlated with MDS in the PG-RIL population (Figure 3C). These results further demonstrated that the candidate gene for *QPm.cas-7D* was most likely *Pm38*.

Association analysis of *QPm-7D-R* haplotype with yield-related traits in common wheat

After screening 157 landraces of the Chinese wheat mini-core collection and 348 Chinese modern cultivars using the diagnostic marker *Kasp5249*, we performed haplotype association analysis of six agronomic traits (TKW, KNS, TSS, SL, ETN, PH) in multiple environments. The resistance haplotype *QPm-7D-R* was significantly correlated with TKW and SL in the 157 landraces of the Chinese wheat mini-core collection (Figures 4A-F). The mean

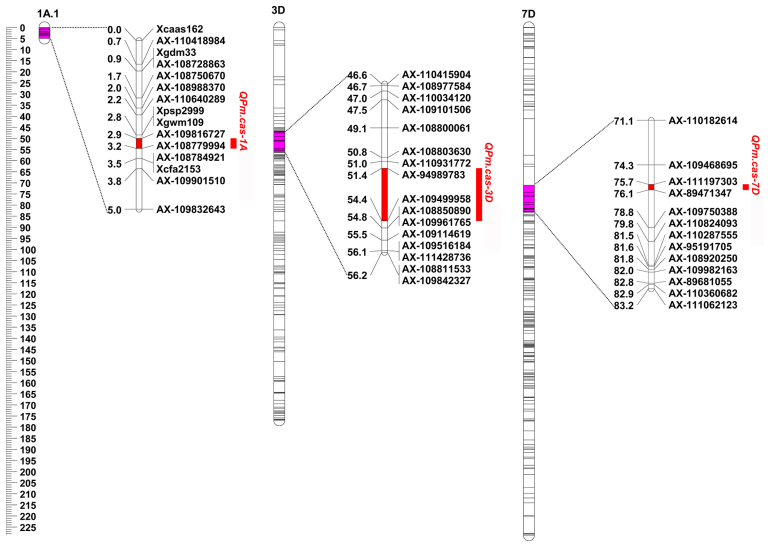


FIGURE 2
Genetic locations of QTL intervals associated with maximum disease severities detected in 2022SJZ environment. Uniform centimorgan (cM) scales are shown on the left. QTLs are indicated on the right side of each chromosome.

TABLE 3 Candidate genes identified for QTL *QPm.cas-7D* with putative functions of interest and their functional annotation.

Gene Name	Blast-hit-accession	Description
TraesCS7D03G0183500	tr|B9R9W6|B9R9W6_RICCO	Sugar transporter
TraesCS7D03G0183600	tr|W0TSU1|W0TSU1_ACAMN	Pleiotropic drug resistance ABC transporter
TraesCS7D03G0183700	tr|C0JSA9|C0JSA9_WHEAT	Cytochrome P450
TraesCS7D03G0183800	tr|B8XSM8|B8XSM8_WHEAT	Lectin receptor kinase
TraesCS7D03G0183900	tr|B8XSM9|B8XSM9_WHEAT	Lectin receptor kinase
TraesCS7D03G0184100	tr|B8XSN0|B8XSN0_WHEAT	Cytochrome P450
TraesCS7D03G0184300	AT1G65040.4	RING/U-box superfamily protein
TraesCS7D03G0184400	tr|A0A165XS50|A0A165XS50_DAUCA	UDP-glycosyltransferase
TraesCS7D03G0184600	tr|A0A165XS50|A0A165XS50_DAUCA	UDP-glycosyltransferase
TraesCS7D03G0184900	AT5G19400.5	Telomerase activating protein Est1
TraesCS7D03G0185400	tr|A0A061DYZ5|A0A061DYZ5_THECC	Ubiquitin-conjugating enzyme 23 isoform 1
TraesCS7D03G0185500	tr|A0A061DYZ5|A0A061DYZ5_THECC	Ubiquitin-conjugating enzyme 23 isoform 1
TraesCS7D03G0185700	tr|K7TR03|K7TR03_MAIZE	Basal layer antifungal peptide
TraesCS7D03G0185800	AT4G22640.2	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
TraesCS7D03G0186600	tr|A0A0K9NQW8|A0A0K9NQW8_ZOSMR	Tyrosine-tRNA ligase
TraesCS7D03G0187100	tr|A0A103XL20|A0A103XL20_CYN	Elongation factor 4
TraesCS7D03G0187200	tr|A0A059PZI7|A0A059PZI7_9POAL	Carboxypeptidase

TKW of *QPm-7D-R* plants was significantly lower than those of the *QPm-7D-S* plants (4.47 g lower in 2002, 5.71 g lower in 2005, and 3.55 g lower in 2010) (Figure 4A). Similar results were found in 348 Chinese modern cultivars. Significant differences were also detected in TKW and PH between *QPm-7D-R* and *QPm-7D-S* haplotypes (Figures 5A-F). The mean TKW of *QPm-7D-R* plants in three environments was also significantly lower than that of *QPm-7D-S* plants (5.28 g lower in 2002, 4.58 g lower in 2005, and 3.49 g lower in 2010) (Figure 5A). In summary, the above results indicate that *QPm-7D-R* is the lower TKW haplotype than *QPm-7D-S*.

QPm.cas-7D-R haplotype underwent negative selection during Chinese wheat breeding

The geographic distribution of the *QPm-7D-R* and *QPm-7D-S* haplotypes was evaluated in both landraces and modern cultivars from ten ecological zones of China. The frequency of the *QPm-7D-R* haplotype declined in the modern cultivars relative to the landraces in the major Chinese production zones (Figures 6A, B). By contrast, the frequency of the *QPm-7D-S* haplotype with high TKW was increased during the transition from landraces to modern cultivars (Figures 6A, B). These results suggested that *QPm-7D-R* underwent negative selection in China wheat breeding for higher yields.

Discussion

P3228 is a valuable wheat germplasm line carrying multiple and stable QTLs related to yield traits (Wang et al., 2011; Liu et al., 2020). In the present study, we found that P3228 was susceptible to *Bgt* isolates E09, E11, and E20 at the seedling stage but highly resistant to a mixture of those *Bgt* isolates at the adult plant stage, suggesting that P3228 conferred APR against powdery mildew (Figure 1). Subsequently, using a RIL population of 'P3228 × Gao 8901', we identified a major QTL *QPm.cas-7D* contributed by P3228 in an interval flanked by *AX-111197303–AX-89471347* on the short arm of chromosome 7D, which explained 65.44% of the phenotypic variance with an additive effect of -29.38 (Table 2 and Figure 2). Meanwhile, a minor QTL *QPm.cas-1A* also from P3228 was detected in the marker interval *AX-109816727–AX-10877999* on the short arm of chromosome 1A and explained 3.45% of the phenotypic variance with an additive effect of -10.58 (Table 2 and Figure 2). *Pm38* (*TraesCS7D03G0183600*), a previously cloned *Pm* gene conferring APR against powdery mildew (Krattinger et al., 2009) was included in the targeted physical interval of the major QTL *QPm.cas-7D*. Sequencing showed that P3228 had an identical sequence to the reported *Pm38* (Krattinger et al., 2009), while Gao 8901 had two different SNPs and two InDels in the genomic regions, resulting a frameshift mutation, the findings indicated that *QPm.cas-7D* was most likely *Pm38* (Figure 3A). Of course, we cannot completely exclude the possibility that *QPm.cas-7D* might be one of the other 17 candidate genes that cooperated with *Pm38* in the targeted physical interval. Further fine mapping and functional validation of the candidate genes should be performed to confirm the *Pm38* as the caused gene for *QPm.cas-7D* in the future.

Due to the ease of selection and phenotypic evaluation, most disease resistance studies have focused on ASR genes, which are known as qualitative or race-specific resistance

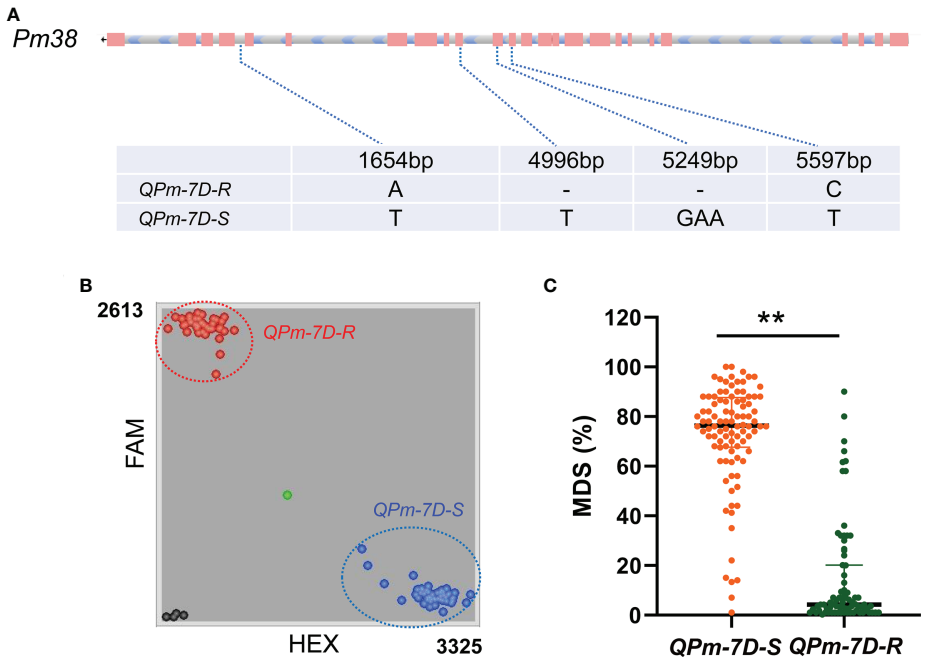


FIGURE 3
Haplotypes analysis with maximum disease severities (MDS) of *QPm.cas-7D* in PG-RIL population. **(A)** Haplotypes of *QPm.cas-7D* based on its genome regions in P3228 and Gao 8901. **(B)** Genotyping results of the diagnostic marker *Kasp5249* developed for *QPm-7D-R* and *QPm-7D-S* haplotypes in the PG-RIL population. **(C)** Haplotypes analysis with MDS of *Pm38* in the PG-RIL population. $^{**}P < 0.01$ (two-tailed *t* test).

genes (Wu et al., 2022). However, overuse of single race-specific resistance genes can lead to the rapid evolution of new virulent *Bgt* races and consequent massive economic losses (Singh et al., 2016). In this case, developing durable resistance conferred by APR genes is emphasized. Many APR genes appear to provide broad-spectrum resistance to one or multiple diseases at the same locus, such as *Pm38/Lr34/Yr18/Sr57* (Krattinger et al., 2009), *Pm39/Lr46/Yr29/Sr58* (Singh et al., 2013), *Pm46/Lr67/Yr46/Sr55* (Moore et al., 2015), and *Pmx/Lr27/Yr30/Sr2* (Mago et al., 2011). Previous studies revealed that the pyramiding of multiple APR genes enabled near-immunity of the plants (Singh and Trethowan, 2007). Therefore, the QTL for powdery mildew resistance identified in P3228 could enrich the available wheat genetic resources in breeding for durable and multiple resistance.

The balance of resistance and yield is the main concern for breeders when using a resistance gene in wheat breeding programs, but often, disease resistance is at the expense of some agronomic traits and reduces plant adaptation (Deng et al., 2017; Han et al., 2022; Mu et al., 2022). For example, *mildew resistance locus O* (*MLO*), is a durable and broad-spectrum resistance to powdery mildew in various plant species including common wheat, however, it also leads to growth penalties and yield losses, thereby limiting its widespread use in wheat breeding (Li et al., 2022). So, investigation of the corresponding yield traits tends to be an important index for evaluating the use of resistance gene(s). Recent research shows that *Pm5e* has no yield penalty by investigating the agronomic performance in a pair of near-isogenic lines H962R with *Pm5e* and H962S without *Pm5e* (Qiu et al., 2022). In this study, we found that *QPm.cas-7D* significantly decreased TKW in the 157 landraces of the Chinese wheat mini-core collection and 348 Chinese modern cultivars (Figures 4A-F and Figures 5A-F). An additional interesting phenomenon was also observed that the frequency of *QPm.cas-7D* in wheat landraces was higher than in breeding lines, which might attribute to the artificial selection of high TKW trait in wheat major production regions. Thus, *QPm.cas-7D* could be designed to transfer into the various wheat cultivars exhibiting desirable performance on TKW.

To better transfer *QPm.cas-7D* in MAS, we developed a KASP marker *Kasp5249* (Figure 3B) based on the differences in the *Pm38*-homologous sequence between P3228 and Gao 8901. It is worth mentioning that *Kasp5249* was developed according to the 3-bp deletion that causes the loss of protein function, which could also be used as the functional marker of *Pm38*. We believe that after the suitable selection for resistance and agronomic performance, *QPm.cas-7D* will release its full potential in wheat breeding programs.

Conclusion

We performed QTL analysis using the PG-RIL population for MDS, and three QTLs *QPm.cas-1A*, *QPm.cas-3D*, and *QPm.cas-7D* were identified in 2022SJZ environment (Table 2

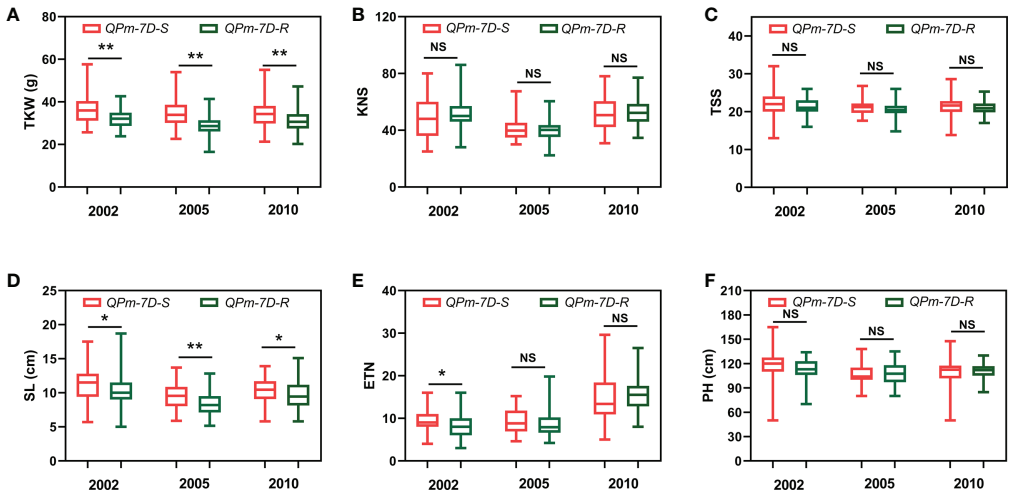


FIGURE 4
Haplotypes analysis with agronomic traits of *QPm.cas-7D* in the landraces of the Chinese wheat mini-core collection. Comparison analysis of *QPm.cas-7D* haplotypes with the TKW **(A)**, KNS **(B)**, TSS **(C)**, SL **(D)**, ETN **(E)**, and PH **(F)** of the landraces of the Chinese wheat mini-core collection in three environments. $^{**}P < 0.01$ and $^{*}P < 0.05$ (two-tailed *t* test) indicates a significant difference to the two haplotypes. NS, no significant difference.

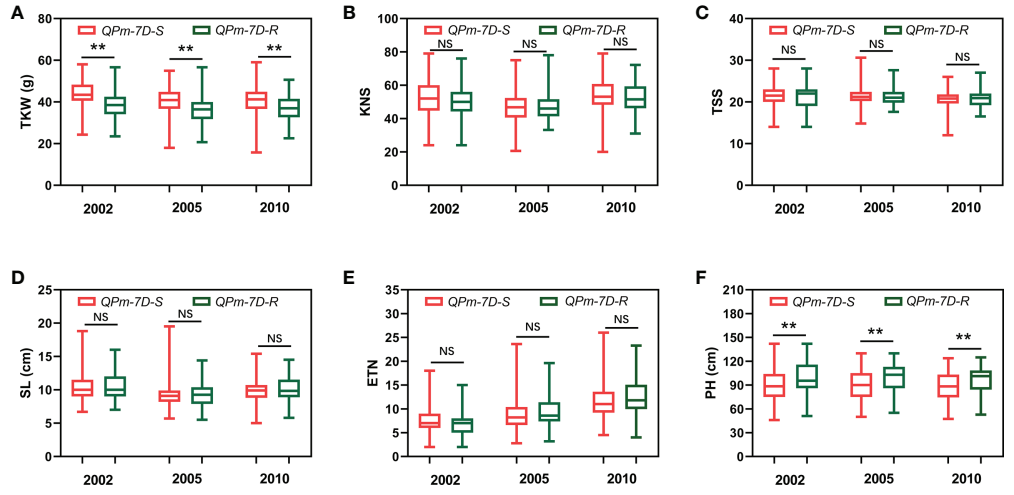


FIGURE 5
Haplotypes analysis with agronomic traits of *QPm.cas-7D* in the Chinese modern cultivars. Comparison analysis of *QPm.cas-7D* haplotypes with the TKW **(A)**, KNS **(B)**, TSS **(C)**, SL **(D)**, ETN **(E)**, and PH **(F)** of the Chinese modern cultivars in three environments. $^{**}P < 0.01$ and $^{*}P < 0.05$ (two-tailed *t* test) indicates a significant difference to the two haplotypes. NS, no significant difference.

and Figure 2). Notably, the major QTL *QPm.cas-7D* contributed by P3228, could explain 65.44% of the phenotypic variances (Table 2). Furthermore, the QTL *QPm.cas-7D* was delimited to the physical interval of approximately 1.168 Mb, and *Pm38* was considered as the candidate gene (Table 3). Based on a 3-bp InDel of *Pm38* genomic sequence between the two parents, a KASP marker *Kasp5249* of *Pm38* allele was developed and verified by PG-RIL population (Figures 3B, C). Furthermore, the *QPm-7D-R* haplotype significantly decreased TKW and underwent negative selection in China wheat breeding for

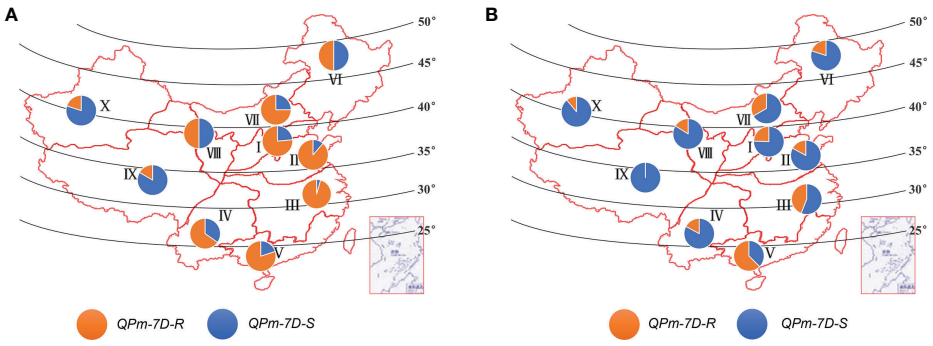


FIGURE 6
Geographic distribution of the *QPm.cas-7D* haplotypes among ten Chinese ecological regions. Distribution of *QPm.cas-7D* haplotypes in landraces **(A)** and modern cultivars **(B)** among ten Chinese ecological regions. I, northern winter wheat region; II, Yellow and Huai River valley winter wheat region; III, low and middle Yangtze River valley winter wheat region; IV, southwestern winter wheat region; V, southern winter wheat region; VI, northeastern spring wheat region; VII, northern spring wheat region; VIII, northwestern spring wheat region; IX, Qinghai–Tibet spring–winter wheat region; X, Xinjiang winter–spring wheat region.

higher yields (Figures 4A-F and Figures 5A-F). Our finding identified a major QTL *QPm.cas-7D* and analyzed its effects for yield-related traits, which could be helpful in improving wheat disease resistance and high-yield breeding.-

Data availability statement

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

Author contributions

DA and MZ conceived the study. TG, YJ, LX, HY, JW, and CH evaluated the phenotype. HL, GH and ZS carried out QTL mapping, predicted candidate gene, and developed the KASP markers. HL, GH and TG analyzed the data and wrote the manuscript. DA and MZ supervised and revised the writing of the article. All authors approved the final manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA24030102), the National Natural Science Foundation of China (32101686), and the Hebei Province Key Research and Development Program (22326306D). The funding bodies were not involved in the design of the study, and collection, analysis, and interpretation of data, and manuscript writing.

Acknowledgments

The authors are grateful to Prof. Xueyong Zhang for providing 157 landraces of the Chinese wheat mini-core collection and 348 Chinese modern cultivars.

Conflict of interests

The reviewer JL declared a shared affiliation with the author CH to the handling editor at the time of review.

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1042399/full#supplementary-material

References

An, D. G., Ma, P. T., Zheng, Q., Fu, S. L., Li, L. H., Han, F. P., et al. (2019). Development and molecular cytogenetic identification of a new wheat-rye 4R chromosome disomic addition line with resistances to powdery mildew, stripe rust and sharp eyespot. *Theor. Appl. Genet.* 132, 257–272. doi: 10.1007/s00122-018-3214-3

Deng, Y. W., Zhai, K. R., Xie, Z., Yang, D. Y., Zhu, X. D., Liu, J. Z., et al. (2017). Epigenetic regulation of antagonistic receptors confers rice blast resistance with yield balance. *Science* 355, eaai8898. doi: 10.1126/science.aai8898

Fang, T., Lei, L., Li, G., Powers, C., Hunger, R. M., Carver, B. F., et al. (2020). Development and deployment of KASP markers for multiple alleles of *Lr34* in wheat. *Theor. Appl. Genet.* 133, 2183–2195. doi: 10.1007/s00122-020-03589-x

Han, G. H., Yan, H. W., Gu, T. T., Cao, L. J., Zhou, Y. L., Wei, L., et al. (2022). Identification of a wheat powdery mildew dominant resistance gene in the *Pm5* locus for high-throughput marker-assisted selection. *Plant Dis*. doi: 10.1094/PDIS-07-22-1545-RE

He, H., Liu, R., Ma, P., Du, H., Zhang, H., Wu, Q., et al. (2021). Characterization of *Pm68*, a new powdery mildew resistance gene on chromosome 2BS of Greek durum wheat TRI 1796. *Theor. Appl. Genet.* 134, 53–62. doi: 10.1007/s00122-020-03681-2

Hewitt, T., Müller, M. C., Molnár, I., Mascher, M., Holušová, K., Šimková, H., et al. (2021). A highly differentiated region of wheat chromosome 7AL encodes a *Pm1a* immune receptor that recognizes its corresponding *AvrPm1a* effector from *Blumeria graminis*. *New Phytol.* 229, 2812–2826. doi: 10.1111/nph.17075

He, H. G., Zhu, S. Y., Zhao, R. H., Jiang, Z. N., Ji, Y. Y., Ji, J., et al. (2018). *Pm21*, encoding a typical CC-NBS-LRR protein, confers broad spectrum resistance to wheat powdery mildew disease. *Mol. Plant* 11, 879–882. doi: 10.1016/j.molp.2018.03.004

Hurni, S., Brunner, S., Buchmann, G., Herren, G., Jordan, T., Krukowski, P., et al. (2013). Rye *Pm8* and wheat *Pm3* are orthologous genes and show evolutionary conservation of resistance function against powdery mildew. *Plant J.* 76, 957–996. doi: 10.1111/tpj.12345

IWGSC (2014). A chromosome-based draft sequence of the hexaploid bread wheat (*Triticum aestivum*) genome. *Science* 345, 1251788. doi: 10.1126/science.1251788

IWGSC (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361, eaar7191. doi: 10.1126/science.aar7191

Jia, M. S., Xu, H. X., Liu, C., Mao, R. X., Li, H. S., Liu, J. J., et al. (2020). Characterization of the powdery mildew resistance gene in the elite wheat cultivar Jimai 23 and its application in marker-assisted selection. *Front. Genet.* 11, 241. doi: 10.3389/fgene.2020.00241

Jiang, Y., Schulthess, A. W., Rodemann, B., Ling, J., Plieske, J., Kollers, S., et al. (2016). Validating the prediction accuracies of marker-assisted and genomic selection of fusarium head blight resistance in wheat using an independent sample. *Theor. Appl. Genet.* 130, 471–482. doi: 10.1007/s00122-016-2827-7

Jin, Y. L., Shi, F. Y., Liu, W. H., Fu, X. Y., Gu, T. T., Han, G. H., et al. (2021). Identification of resistant germplasm and detection of genes for resistance to powdery mildew and leaf rust from 2,978 wheat accessions. *Plant Dis.* 105, 3900–3908. doi: 10.1094/PDIS-03-21-0532-RE

Kang, Y. C., Zhou, M. X., Merry, A. M., and Barry, K. M. (2020). Mechanisms of powdery mildew resistance of wheat-a review of molecular breeding. *Plant Pathol.* 69, 601–617. doi: 10.1111/ppa.13166

Krattinger, S. G., Lagudah, E. S., Spielmeyer, W., Singh, R. P., Huerta-Espino, J., McFadden, H., et al. (2009). A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. *Science* 323, 1360–1363. doi: 10.1126/science.1166453

Lagudah, E. S., Krattinger, S. G., Herrera-Foessel, S., Singh, R. P., Huerta-Espino, J., Spielmeyer, W., et al. (2009). Gene-specific markers for the wheat gene *Lr34/Yr18/Pm38* which confers resistance to multiple fungal pathogens. *Theor. Appl. Genet.* 119, 889–898. doi: 10.1007/s00122-009-1097-z

Li, M., Dong, L., Li, B., Wang, Z., Xie, J., Qiu, D., et al. (2020). A CNL protein in wild emmer wheat confers powdery mildew resistance. *New Phytol.* 228, 1027–1037. doi: 10.1111/nph.16761

Li, Z., Lan, C., He, Z., Singh, R. P., Rosewarne, G. M., Chen, X., et al. (2014). Overview and application of QTL for adult plant resistance to leaf rust and powdery mildew in wheat. *Crop Sci.* 54, 1907–1925. doi: 10.2135/cropsci2014.02.0162

Li, S., Lin, D., Zhang, Y., Deng, M., Chen, Y., Lv, B., et al. (2022). Genome-edited powdery mildew resistance in wheat without growth penalties. *Nature* 602, 455–460. doi: 10.1038/s41586-022-04395-9

Lillemo, M., Asalf, B., Singh, R. P., Huerta-Espino, J., Chen, X. M., He, Z. H., et al. (2008). The adult plant rust resistance loci *Lr34/Yr18* and *Lr46/Yr29* are important determinants of partial resistance to powdery mildew in bread wheat line Saar. *Theor. Appl. Genet.* 116, 1155–1166. doi: 10.1007/s00122-008-0743-1

Liu, H., Li, H., Hao, C., Wang, K., Wang, Y., Qin, L., et al. (2020). *TaDA1*, a conserved negative regulator of kernel size, has an additive effect with *TaGW2* in common wheat (*Triticum aestivum l.*). *Plant Biotechnol. J.* 18, 1330–1342. doi: 10.1111/pbi.13298

Liu, H., Zhang, X., Xu, Y., Ma, F., Zhang, J., Cao, Y., et al. (2020). Identification and validation of quantitative trait loci for kernel traits in common wheat (*Triticum aestivum* l.). *BMC Plant Biol.* 20, 529. doi: 10.1186/s12870-020-02661-4

Lu, P., Guo, L., Wang, Z., Li, B., Li, J., Li, Y., et al. (2020). A rare gain of function mutation in a wheat tandem kinase confers resistance to powdery mildew. *Nat. Commun.* 11, 680. doi: 10.1038/s41467-020-14294-0

Mago, R., Tabe, L., McIntosh, R. A., Pretorius, Z., Kota, R., Paux, E., et al. (2011). A multiple resistance locus on chromosome arm 3BS in wheat confers resistance to stem rust (*Sr2*), leaf rust (*Lr27*) and powdery mildew. *Theor. Appl. Genet.* 123, 615–623. doi: 10.1007/s00122-011-1611-y

McIntosh, R. A., Dubcovsky, J., Rogers, W. J., Xia, X. C., and Raupp, W. J. (2019) *Catalogue of gene symbols for wheat 2019 supplement*. Available at: https://wheat.pw.usda.gov/GG3/WGC.

Jia, M. S., Xu, H. X., Liu, C., Mao, R. X., Li, H. S., Liu, J. J., et al (2020). Characterization of the powdery mildew resistance gene in the elite wheat cultivar Jimai 23 and its application in marker-assisted selection. *Front. Genet.* 11, 241. doi: 10.3389/fgene.2020.00241

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

Moore, J. W., Herrera-Foessel, S., Lan, C., Schnippenkoetter, W., Ayliffe, M., Huerta-Espino, J., et al (2015). A recently evolved hexose transporter variant confers resistance to multiple pathogens in wheat. *Nat. Genet.* 47, 1494–8. doi: 10.1038/ng.3439

Moore, J. W., Herrera-Foessel, S., Lan, C., Schnippenkoetter, W., Ayliffe, M., Huerta-Espino, J., et al. (2015). A recently evolved hexose transporter variant confers resistance to multiple pathogens in wheat. *Nat. Genet.* 47, 1494–1498. doi: 10.1038/ng.3439

Mu, Y. J., Gong, W. P., Qie, Y. M., Liu, X. Q., Li, L. Z., Sun, N. N., et al. (2022). Identification of the powdery mildew resistance gene in wheat breeding line yannong 99102-06188 *via* bulked segregant exome capture sequencing. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1005627

Peterson, R. F., Campbell, A. B., and Hannah, A. E. (1948). A diagrammatic scale for estimating rust intensity of leaves and stem of cereals. *Can. J. Res. Sect. C.* 26, 496–500. doi: 10.1139/CJR48C-033

Qie, Y., Sheng, Y., Xu, H., Jin, Y., Ma, F., Li, L., et al. (2019). Identification of a new powdery mildew resistance gene *pmDHT* at or closely linked to the *Pm5* locus in the Chinese wheat landrace dahongtou. *Plant Dis.* 103, 2645–2651. doi: 10.1094/PDIS-02-19-0401-RE

Qiu, D., Huang, J., Guo, G. H., Hu, J. H., Li, Y. H., Zhang, H. J., et al. (2022). The *Pm5e* gene has no negative effect on wheat agronomic performance: evidence from newly established near-isogenic lines. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.918559

Sánchez-Martín, J., Steuernagel, B., Ghosh, S., Herren, G., Hurni, S., Adamski, N., et al. (2016). Rapid gene isolation in barley and wheat by mutant chromosome sequencing. *Genome Biol.* 17, 221. doi: 10.1186/s13059-016-1082-1

Sánchez-Martín, J., Widrig, V., Herren, G., Wicker, T., Zbinden, H., Gronnier, J., et al. (2021). Wheat *Pm4* resistance to powdery mildew is controlled by alternative splice variants encoding chimeric proteins. *Nat. Plants* 7, 327–341. doi: 10.1038/s41477-021-00869-2

Singh, R. P., Herrera-Foessel, S. A., Huerta-Espino, J., Lan, C. X., Basnet, B. R., Bhavani, S., et al. (2013). Pleiotropic gene *Lr46/Yr29/Pm39/Ltn2* confers slow rusting, adult plant resistance to wheat stem rust fungus. proceedings borlaug global rust initiative. *New Delhi India: 2013 Tech. Workshop*, 17.1.

Singh, S. P., Hurni, S., Ruinelli, M., Brunner, S., Sanchez-Martin, J., Krukowski, P., et al. (2018). Evolutionary divergence of the rye *Pm17* and *Pm8* resistance genes reveals ancient diversity. *Plant Mol. Biol.* 98, 249–260. doi: 10.1007/s11103-018-0780-3

Singh, R. P., Singh, P. K., Rutkoski, J., Hodson, D. P., He, X., Jørgensen, L. N., et al. (2016). Disease impact on wheat yield potential and prospects of genetic control. *Annu. Rev. Phytopathol.* 54, 303–322. doi: 10.1146/annurev-phyto-080615-095835

Singh, R. P., and Trethowan, R. (2007). *Breeding major food staples*. Eds. M. Kang and P. M. Priyadarshan (Ames, IA: Blackwell), 109–140, 349-355.

Si, Q. M., Zhang, X. X., Duan, X. Y., Sheng, B. Q., and Zhou, Y. L. (1992). On gene analysis and classification of powdery mildew (*Erysiphe graminis* f. sp. *tritici*) resistant wheat varieties. *Acta Phytopathol. Sinica.* 22, 349–355. doi: 10.13926/j.cnki.apps.1992.04.021

Spielmeyer, W., McIntosh, R. A., Kolmer, J., and Lagudah, E. S. (2005). Powdery mildew resistance and *Lr34/Yr18* genes for durable resistance to leaf and stripe rust cosegregate at a locus on the short arm of chromosome 7D of wheat. *Theor. Appl. Genet.* 111, 731–735. doi: 10.1007/s00122-005-2058-9

Summers, R. W., and Brown, J. K. M. (2013). Constraints on breeding for disease resistance in commercially competitive wheat cultivars. *Plant Phothol.* 62, 115–121. doi: 10.1111/ppa.12165

Sun, Z., Li, H., Zhang, L., and Wang, J. (2013). Properties of the test statistic under null hypothesis and the calculation of LOD threshold in quantitative trait loci (QTL) mapping. *Acta Agronom. Sin.* 39, 1–11. doi: 10.3724/SP.J.1006.2013.00001

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

Wang, J., Liu, W., Wang, H., Li, L., and Gao, A. (2011). QTL mapping of yield-related traits in the wheat germplasm 3228. *Euphytica* 177, 277–292. doi: 10.1007/s10681-010-0267-z

Wu, L. R., Zhu, T., He, H. G., Cao, X. Y., Li, H. S., Xu, H. X., et al. (2022). Genetic dissection of the powdery mildew resistance in wheat breeding line LS5082 using BSR-seq. *Crop J.* doi: 10.1016/j.cj.2021.12.008

Xie, J., Guo, G., Wang, Y., Hu, T., Wang, L., Li, J., et al. (2020). A rare single nucleotide variant in *Pm5e* confers powdery mildew resistance in common wheat. *New Phytol.* 228, 1011–1026. doi: 10.1111/nph.16762

Xing, L., Hu, P., Liu, J., Witek, K., Zhou, S., Xu, J., et al. (2018). *Pm21* from *Haynaldia villosa* encodes a CC-NBS-LRR protein conferring powdery mildew resistance in wheat. *Mol. Plant* 11, 874–878. doi: 10.1016/j.molp

Yahiaoui, N., Srichumpa, P., Dudler, R., and Keller, B. (2004). Genome analysis at different ploidy levels allows cloning of the powdery mildew resistance gene *Pm3b* from hexaploid wheat. *Plant J.* 37, 528–538. doi: 10.1046/j.1365-313x.2003.01977.x

Zhang, R. Q., Fan, Y. L., Kong, L. N., Wang, Z. J., Wu, J. Z., Xing, L. P., et al. (2018). *Pm62*, an adult-plant powdery mildew resistance gene introgressed from *Dasypyrum villosum* chromosome arm 2VL into wheat. *Theor. Appl. Genet.* 131, 2613–2620. doi: 10.1007/s00122-018-3176-5

Zhao, J., Wang, Z., Liu, H., Zhao, J., Li, T., Hou, J., et al. (2019). Global status of 47 major wheat loci controlling yield, quality, adaptation and stress resistance selected over the last century. *BMC Plant Biol.* 19, 5. doi: 10.1186/s12870-018-1612-y

Zhu, T., Wang, L., Rimbert, H., Rodriguez, J. C., Deal, K. R., De Oliveira, R., et al. (2021). Optical maps refine the bread wheat *Triticum aestivum* cv Chinese spring genome assembly. *Plant J.* 107, 303–314. doi: 10.1111/tpj.15289

Zou, S., Wang, H., Li, Y., Kong, Z., and Tang, D. (2018). The NB-LRR gene *Pm60* confers powdery mildew resistance in wheat. *New Phytol.* 218, 298–309. doi: 10.1111/nph.14964

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 26 October 2022
DOI 10.3389/fpls.2022.1054673

OPEN ACCESS

EDITED BY
Pengtao Ma,
Yantai University, China

REVIEWED BY
Wenxuan Liu,
Henan Agricultural University, China
Tianya Li,
Shenyang Agricultural University, China

*CORRESPONDENCE
Hongfei Yan
hongfeiyan2006@163.com
Qingfang Meng
qingfangmeng500@126.com

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

RECEIVED 27 September 2022
ACCEPTED 11 October 2022
PUBLISHED 26 October 2022

CITATION
Zhang L, Zhao X, Liu J, Wang X, Gong W, Zhang Q, Liu Y, Yan H, Meng Q and Liu D (2022) Evaluation of the resistance to Chinese predominant races of *Puccinia triticina* and analysis of effective leaf rust resistance genes in wheat accessions from the U.S. National Plant Germplasm System. *Front. Plant Sci.* 13:1054673. doi: 10.3389/fpls.2022.1054673

Evaluation of the resistance to Chinese predominant races of *Puccinia triticina* and analysis of effective leaf rust resistance genes in wheat accessions from the U.S. National Plant Germplasm System

Lin Zhang[1,2], Xuefang Zhao[1], Jingxian Liu[3], Xiaolu Wang[4,5,6], Wenping Gong[4,5,6], Quanguo Zhang[7], Yuping Liu[7], Hongfei Yan[1*], Qingfang Meng[1*] and Daqun Liu[1]

[1]College of Plant Protection, Hebei Agricultural University, Technological Innovation Center for Biological Control of Crop Diseases and Insect Pests of Hebei Province, Baoding, China, [2]School of Landscape and Ecological Engineering, Hebei Engineering University, Handan, China, [3]College of Agronomy, Shandong Agricultural University, Tai'an, China, [4]Crop Research Institute, Shandong Academy of Agricultural Sciences, Shandong Wheat Technology Innovation Center, Jinan, China, [5]National Engineering Laboratory of Wheat and Maize, Shandong Wheat Technology Innovation Center, Jinan, China, [6]Key Laboratory of Wheat Biology and Genetic Improvement in the North HuangHuai River Valley of Ministry of Agriculture, Shandong Wheat Technology Innovation Center, Jinan, China, [7]Institute of Cereal and Oil Crops, Hebei Academy of Agriculture and Forestry Sciences, Shijiazhuang, China

Puccinia triticina, which is the causative agent of wheat leaf rust, is widely spread in China and most other wheat-planting countries around the globe. Cultivating resistant wheat cultivars is the most economical, effective, and environmentally friendly method for controlling leaf rust-caused yield damage. Exploring the source of resistance is very important in wheat resistance breeding programs. In order to explore more effective resistance sources for wheat leaf rust, the resistance of 112 wheat accessions introduced from the U.S. National Plant Germplasm System were identified using a mixture of pathogenic isolates of THTT, THTS, PHTT, THJT and THJS which are the most predominant races in China. As a result, all of these accessions showed high resistance at seedling stage, of which, ninety-nine accessions exhibited resistance at adult plant stage. Eleven molecular markers of eight effective leaf rust resistance genes in China were used to screen the 112 accessions. Seven effective leaf rust resistance genes *Lr9*, *Lr19*, *Lr24*, *Lr28*, *Lr29*, *Lr38* and *Lr45* were detected, except *Lr47*. Twenty-three accessions had only one of those seven effective leaf rust resistance gene. Eleven accessions carried *Lr24+Lr38*, and 7 accessions carried *Lr9+Lr24+Lr38*, *Lr24+Lr38+Lr45*, *Lr24+Lr29+Lr38*

and *Lr19+Lr38+Lr45* respectively. The remaining seventy-one accessions had none of those eight effective leaf rust resistance genes. This study will provide theoretical guidance for rational utilization of these introduted wheat accessions directly or for breeding the resistant wheat cultivars.

KEYWORDS

leaf rust, wheat accessions, resistance gene, molecular markers, races

Introduction

Wheat leaf rust, caused by *Puccinia triticina* Erikss., is a serious fungal disease of wheat which occurs in the majority of wheat-growing regions worldwide, especially in North Africa, Southeast and Central Asia, Eastern Europe, North and South America (Bolton et al., 2008). In China, leaf rust is a common disease threatening wheat production, especially in the North China Plain, the Middle and Lower Reaches of the Yangtze River, Southwest and Northwest Regions (Liu and Chen, 2012). Varying on wheat cultivars and disease period, 7% to 30% yield loss can be encountered and even more than 50% in severe cases (Huerta-Espino et al., 2011). In recent years, the occurrence of wheat leaf rust has been in ascendancy as a result of varying climatic conditions as evident in 2008, 2009, 2012, 2013 and 2015 in the whole country or some regions (Zhang et al., 2015; Zhang et al., 2020a; Zhang et al., 2020b).

The most economical, effective and environment-friendly method to control leaf rust is to cultivate resistant cultivars (Pink, 2002). However, the variation of virulence and the emergence of new races of *P. triticina* always leads to loss of the effective resistance of wheat cultivars, especially which carried single leaf rust resistance genes and large-scale planted, and increase the potential risk of leaf rust epidemic on wheat (Zhang et al., 2015). The THTT, THTS, PHTT, THJT, THJS, PHJT, and PHTS were predominant races of *P. triticina* in China from 2011-2015, of which, THTT and PHTT were also the predominant races in India (Zhang et al., 2020a, b;Bhardwaj et al., 2019). Most wheat cultivars in the major wheat-growing regions such as Henan, Shandong and Hebei province are susceptible to these races (Zhang et al., 2017a; Zhang et al., 2017b; Zhang et al., 2020a; Zhang et al., 2020b). Previous studies also revealed that many of the major wheat cultivars in China carry a few leaf rust resistance genes, such as *Lr1*, *Lr16*, *Lr26*, and *Lr37*, which have lost their effective resistance (Zhang et al., 2017a; Zhang et al., 2017b; Zhang et al., 2020a; Zhang et al., 2020b). So, it is necessary to explore and utilize the effective wheat resistance resources for the breeding of new, sustainable, and durable resistant wheat cultivars.

Gene postulation and molecular marker-assisted selection (MAS) are the most commonly used methods for identification and analysis of wheat leaf rust resistance genes (Zhang et al., 2017a). Gene postulation is a method for presupposing and identifying leaf rust resistance genes in wheat cultivars. This method uses a set of wheat leaf rust resistance near-isogenic lines or single gene lines, but it is easily influenced by genetic background, environmental conditions and human factors (Hu et al., 2014). In addition, the different virulent races of *P. triticina* to the differential lines are the key factors for gene postulation. Therefore, the high-resistance wheat cultivars carrying effective leaf rust resistance genes cannot be analyzed by gene postulation methods due to the lack of corresponding virulent races for these genes. For example, so far, there are no virulent races against the leaf rust resistance genes *Lr9*, *Lr19*, *Lr24*, *Lr28* and *Lr38* in China and many parts of the world (Zhang et al., 2020a; Zhang et al., 2020b). So these genes cannot be postulated in the wheat cultivars by gene postulation. MAS can effectively track corresponding genes by using the molecular markers closely linked to the leaf rust resistance genes (Ding et al., 2010). Most of the leaf rust resistence gene markers had been developed and successfully applied to identify the known leaf rust resistance genes in wheat cultivars and molecular breeding for disease resistance (Bassi et al., 2015; Wang et al., 2016; Gebrewahid et al., 2017; Beukert et al., 2020; Wu et al., 2020). For instance, MAS has been successfully applied to practical commercial wheat breeding for rust resistance genes *Lr34* and *Yr36* (Miedaner and Korzun, 2012). Therefore, the gap created by gene postulation methods can be bridged by MAS methods, which has high efficiency for the identification of effective resistance genes contained in wheat cultivars.

There are abundant wheat germplasm resources (more than 49,000) preserved in the National Germplasm Bank of China. However, according to previous studies, the proportion of Chinese wheat cultivars with high resistance to leaf rust is relatively low by identifying the resistance of the main or core wheat breeding materials (lines) to leaf rust in different regions, and the majority of Chinese main wheat cultivars(lines) carry only a few leaf rust resistance genes which have lost their effectiveness (Ding et al., 2010; Shi et al., 2011; Zhao et al., 2013; Zhang et al., 2017a; Zhang et al., 2017b; Gao et al., 2019; Zhang et al., 2019a). For example, only 14 of 182 wheat cultivars

(lines) in Huang-Huai-Hai river wheat region were resistant to leaf rust at seedling stage, and a few resistance genes, such as *Lr1*, *Lr26*, and *Lr37* which had lost their effectiveness in China, were detected in these tested cultivars(lines) (Gao et al., 2019). It is therefore demand-driven to increase the genetic resources of wheat leaf rust resistance and the appropriate supplement to the wheat parent material resource pool that can lay a resource foundation for the breeding of more resistance cultivars. In the previous study, we preliminarily identified the resistance of 359 introduced accessions from the United States National Plant Germplasm System at seedling stage, of which 112 resistant accessions were screened (Unpublished data). So this study aimed to further identify the resistance of these 112 wheat accessions to the Chinese predominant races of *P. triticina* and determine the effective leaf rust resistance genes by MAS, and provide new and excellent resistance sources for wheat resistance breeding program in China.

Materials and methods

Plant materials

One hundred and twelve wheat accessions used in this study were provided by Dr. Harold Bockelman, National Plant Germplasm System (NPGS), USDA-ARS, Aberdeen, Idaho, USA. The susceptible wheat Thatcher, Zhengzhou 5389 and 8 Thatcher near-isogenic lines with single resistance genes *Lr9*, *Lr19*, *Lr24*, *Lr28*, *Lr29*, *Lr38*, *Lr45* and *Lr47*, the effective resistance genes until now in China, were provided by Wheat Leaf Rust Research Center of Hebei Agricultural University.

Puccinia triticina isolates

Five predominant races of *P. triticina*, THTT, THTS, PHTT, THJT and THJS were used in this study. These races were collected and identified by Wheat Leaf Rust Research Center of Hebei Agricultural University in China in 2015.

Evaluation of resistance to leaf rust at seedling stage

In 2016 and 2017, 112 wheat accessions, Thatcher and Zhengzhou 5389 were planted in 30×16 cm plastic trays in the greenhouses of Hebei Agricultural University. Each line was represented by 5 to 10 seedlings. These wheat materials were inoculated with predominant *P. triticina* races as described by Zhang et al. (2020a). Urediniospores of five predominant races of *P. triticina* were mixed with talcum powder in a 1:10 proportion and subsequently bestrewed on the pre-moistened leaves of the experimental wheat seedling. The inoculated seedlings were then transferred to a closed humid container for incubation at 18 to 24°C in darkness for 16 to 24 h, and subsequently moved to a greenhouse at 20±5°C and a photoperiod regime of 12-14 h with fluorescent light supplementation. Evaluation of infection types (IT) were performed at 12 days post-inoculation (dpi) as described by Roelfs (1984) when the disease was fully developed on the susceptible control Thatcher and Zhengzhou 5389. The identification experiment were repeated at least three times.

Validation of adult plant resistance in field

All wheat accessions were tested and evaluated for their resistance at adult plant stage in the field nurseries at Baoding in Hebei province in 2016 and 2017. In mid-October 2015 and 2016, seeds of each wheat accession were sown in single rows according to the standard of row spacing of 30 cm and length 2 m per line. The susceptible control Zhengzhou 5389 were sown adjacent to and around the test rows. The spore suspension was prepared by mixing equal amounts of urediniospores of five predominant races and adding Tween-80 at a final concentration of 1%. The spore suspension was then sprayed on the wheat plants in mid-April (Tillering stage) of 2016 and 2017. The inoculated seedlings were covered with plastic film overnight to moisturize them. Disease investigation was carried out when the disease was fully developed about middle of May (Filling stage) of 2016 and 2017. The infection types to the mixed races were identified and recorded as described by Roelfs (1984).

DNA extraction and molecular marker detection

The genomic DNA of wheat accessions were extracted according to the modified CTAB method (Gill et al., 1991). Eleven STS and SCAR markers for eight effective leaf rust resistance genes in China, *viz. Lr9*, *Lr19*, *Lr24*, *Lr28*, *Lr29*, *Lr38*, *Lr45* and *Lr47*, were used to screen the identified resistant wheat accessions (Table 1). PCR procedure was performed as described by references in Table 1. PCR products were detected by 1.0% (w/v) agarose gel electrophoresis in 1×TAE buffer and visualized under UV transilluminator.

Results

Seedling resistance

In this study, the predominant races THTT, THTS, PHTT, THJT and THJS were used to identify the leaf rust resistance of 112 wheat accessions at seedling stage. The identification results

TABLE 1 Primers of molecular markers used to detect the wheat leaf rust resistance.

Lr gene	Marker type	Primer name	Sequence of primer (5'-3')	FragmentSize (bp)	Reference
Lr9	SCAR	SCS5-550F	TGCGCCTTCAAAGGAAG	550	Gupta et al., 2005
		SCS5-550R	TGCGCCCTTCTGAACTGTAT		
Lr9	STS	J13/1	TCCTTTTATTCCGCACGCCGG	1100	Schachermayr et al., 1994
		J13/2	CCACACTACCCCAAAGAGACG		
Lr19	SCAR	SCS265-F	GGCGGATAAGCAGAGCAGAG	512	Gupta et al., 2006a
		SCS265-R	GGCGGATAAGTGGGTTATGG		
Lr19	SCAR	SCS253-F	GCTGGTTCCACAAAGCAAA	736	Gupta et al., 2006a
		SCS253-R	GGCTGGTTCCTTAGATAGGTG		
Lr24	STS	J09/1	TCTAGTCTGTACATGGGGGC	310	Schachermayr et al., 1995
		J09/2	TGGCACATGAACTCCATACG		
Lr24	SCAR	$S1302_{609}$-F	CGCAGGTTCCAAATACTTTTC	607	Gupta et al., 2006b
		$S1302_{609}$-R	CGCAGGTTCTACCTAATGCAA		
Lr28	SCAR	$SCS421_{570}$-F	ACAAGGTAAGTCTCCAACCA	570	Cherukuri et al., 2005
		$SCS421_{570}$-R	AGTCGACCGAGATTTTAACC		
Lr29	SCAR	OPY10/1	GTGACCTCAGGCAATGCA	850	Tar et al., 2002
		OPY10/2	GTGACCTCAGAACCGATG		
Lr38	SCAR	$Y_{38}SCAR_{982}$-F	GCTGAATCTGCGTATCGTCCC	982	Yan et al., 2008
		$Y_{38}SCAR_{982}$-R	GACTTGTTCTTCGGCGTGTTG		
Lr45	SCAR	PSc20H23	CGACGATCGAATCT CGGGCAAG	750	Yan, 2009
		PSc20H24	GCGCCCTGCGTTGAGGAGAC		
Lr47	STS	PS10R	GCTGATGACCCTGACCGGT	282	Helguera et al., 2000
		PS10L	TCTTCATGCCCGGTCGGGT		

showed that these wheat accessions showed different degrees of resistance to leaf rust at seedling stage (Table 2). Seven out of the 112 wheat accession representing 6.25 % of the total accessions (PI601428, PI542975, PI601429, PI478892, PI639450, Citr15929, and Citr15082) exhibited immunity (IT "0"). Sixty-eight accessions showed high resistance with ITs ";" or "1", while 37 other accessions showed moderate resistance with ITs "X", such as ";1, 3" or ";, 3". These results indicated higher resistance rates of these wheat cultivars to Chinese *P. triticinia* race. The wheat accessions with ITs "X" may be due to the specific resistance to some of the isolates of *P. triticinia*.

Field resistance

To further characterize the resistance of 112 wheat accessions to leaf rust at adult plant stage, field nursery experiments were carried out in the wheat cropping seasons. Ninety-nine (99) of 112 wheat accessions were resistant at adult plant stage, which were consistent with the seedling stage (Table 2). The remaining 13 accessions (PI601452, PI591702, PI486145, PI542976, PI547082, PI17769, PI478892, PI516197, PI531197, PI468977, PI469272, PI475771, and PI566924) were susceptible to leaf rust at adult plant stage with ITs "3" indicating that the seedling resistance, so-called whole growth period resistance, may encountered a new phenotype. In addition, 17 wheat accessions (marked with asterisks in Table 2) with ITs "X" to leaf rust at seedling stage conferred higher resistance at adult plant stage, implying that those accessions may carry adult leaf rust resistance genes or the heat sensitive resistance gene which induced the resistance to *P. triticinia* at high temperature.

Detection of resistance genes

To further identify the leaf rust resistance genes carried by these wheat accessions, the STS and SCAR markers of eight effective leaf rust resistance genes in China were used to detect the leaf rust resistance genes of these accessions. Seven leaf rust resistance genes, *Lr9*, *Lr19*, *Lr24*, *Lr28*, *Lr29*, *Lr38* and *Lr45*, were detected in 41 of 112 wheat accessions (Figure 1-5, Table 2). No corresponding leaf rust resistance genes were detected in the remaining 71 accessions, indicating that other unknown or new effective leaf rust resistance genes at seedling stage existed in these resistant accessions. Based on marker analyses, the 41 resistant accessions carrying the tested gene can be divided into three categories: The first type consisted of wheat accessions that carried only a single leaf rust resistance gene. It was observed that, 9 accessions carried *Lr9*, one accession carried *Lr24*, two accessions carried *Lr28*, 6 accessions carried *Lr29*, and 5 accessions carried *Lr45*, which accounted for 22.0%, 2.4%, 4.9%, 14.6% and 12.2% of 41 resistant accessions respectively.

10.3389/fpls.2022.1054673

TABLE 2 Leaf rust resistance levels at seedling and adult stages and detection results of molecular markers.

No.	accessions	Seeding Infection type	Adult Infection type	*Lr* gene	No.	accessions	Seeding Infection type	Adult Infection type	*Lr* gene
1	PI 601428	0	;1	*Lr24, Lr38*	57	CItr 17723	;, 1	;1	*Lr29*
2	PI 601429	0	;	*Lr9*	58*	CItr 17831	;1, 3	;	—
3	PI 601465	;	;	*Lr9*	59	CItr 17856	;, 3	;1, 3	—
4	PI 601606	1	;1	*Lr24, Lr38*	60	CItr 17857	;1, 3	;1, 3	—
5	PI 595212	;	;	—	61	CItr 15075	;1	;1, 3	—
6	PI 17389	;1, 3	;1, 3	—	62	CItr 15082	0	;	*Lr9*
7	PI 17879	;1	;1	—	63	CItr 15290	;1	;	—
8	PI 17898	;1	;	—	64	PI 548844	;	;	*Lr9*
9	PI 486147	;	;	*Lr24, Lr38*	65	PI 548845	;1, 3	;1, 3	—
10	PI 486212	;	;	*Lr24, Lr38*	66	PI 548847	;	;	—
11*	PI 486349	;1, 3	;	—	67	PI 550697	;1, 3	;1, 3	—
12	PI 494101	;1	;	—	68	PI 552813	;	;1, 3	*Lr45*
13	PI 542975	0	;1, 3	—	69	PI 543893	;	;	*Lr24, Lr38, Lr45*
14	PI 542979	;	;	*Lr9, Lr24, Lr38*	70	PI 547262	;1	;1	—
15	PI 547264	;	;1	*Lr45*	71*	PI 547263	1, 3	;	—
16	PI 594102	1	;, 3	—	72	PI 555586	;	;1, 3	—
17	PI 497988	;	;1	—	73*	CItr 3780	;, 3	;1	—
18	PI 17729	;	;1	—	74	CItr 15375	;	;	*Lr9*
19	PI 601262	;1	;	—	75	CItr 17264	;1, 3	;1, 3	—
20	PI 601263	;	;	*Lr9*	76	PI 564700	;	;1, 3	—
21	PI 601366	;	;	—	77	PI 564851	;	;	*Lr24, Lr29, Lr38*
22	PI 601207	;	;	—	78	PI 566923	;1	;1	*Lr24*
23	PI 601203	;1	;	*Lr24, Lr38*	79	PI 577793	;	;	*Lr24, Lr38*
24	PI 598214	;	;1	—	80	PI 578213	;1	;	*Lr24, Lr38, Lr45*
25	PI 599987	1	;	—	81	PI 491396	;	;	*Lr9*
26*	PI 598209	;1, 3	;	—	82	PI 583676	;1	;1	*Lr29*
27	PI 598211	;1	;	—	83	PI 591560	;1	;1	—
28	PI 598212	;	;	—	84	PI 476974	;1	;1	*Lr29*
29*	PI 298213	;1, 3	;1	—	85	PI 476975	;1	;	*Lr28*
30	PI 486140	;1	;	*Lr45*	86	PI 483469	;1	;	*Lr24, Lr38, Lr45*
31*	PI 508288	;1, 3	;	—	87	PI 596335	;1	;	*Lr24, Lr38, Lr45*
32	PI 511307	;1	;1	*Lr24, Lr38*	88*	PI 601722	;1, 3	;	—
33	PI 511308	;1	;	*Lr24, Lr38*	89	PI 559376	;	;	—
34	PI 506407	;	;	*Lr24, Lr38*	90	PI 557537	;1	;	*Lr28*
35	PI 506405	;	;	*Lr29*	91	PI 557538	;1, 3	;1, 3	—
36	PI 531246	;	;	*Lr24, Lr38*	92	PI 561197	;	;1	—
37	Citr 17940	;1	;	—	93*	PI 561198	;1, 3	;1	—
38	PI 600974	;1, 3	;, 3	—	94	PI 641952	1	;1	—
39*	PI 601069	1, 3	;	—	95*	PI 561200	;1, 3	;	—
40*	PI 601070	;1, 3	;	—	96*	PI 562382	1^{+}	;	—
41	PI 601723	;	;	*Lr29*	97	PI 639450	0	;1	*Lr45*
42	PI 601806	;1	;1	*Lr29*	98	PI 564083	;1, 3	;1, 3	—
43	PI 601807	;1	;1	—	99*	PI 573003	;1, 3	;	—

(Continued)

TABLE 2 Continued

No.	accessions	Seeding Infection type	Adult Infection type	*Lr* gene	No.	accessions	Seeding Infection type	Adult Infection type	*Lr* gene
44	Citr 13684	;1	;1	—	100	PI 601452	;1, 3	3	—
45	Citr 13874	;	;1, 3	—	101	PI 591702	;1, 3	3	—
46	Citr 14048	;	;	*Lr19, Lr38, Lr45*	102	PI 486145	;1, 3	3	—
47	Citr 15229	;	;	*Lr45*	103	PI 542976	;1, 3	3	—
48	Citr 15288	;	;	*Lr9*	104	PI 547082	;1, 3	3	—
49	Citr 15929	0	;	*Lr9*	105	PI 17769	;1, 3	3	—
50	Citr 17262	;	;1	—	106	PI 478892	0	3	—
51*	PI 535454	;1, 3	;1	—	107	PI 516197	;1, 3	3	—
52	PI 518591	;	;	—	108	PI 531197	;1, 3	3	—
53*	PI 527480	;, 3	;	—	109	PI 468977	;1, 3	3	—
54	PI 531244	;	;1, 3	—	110	PI 469272	;1, 3	3	—
55	PI 532282	1	;	*Lr24, Lr38*	111	PI 475771	;1	3	—
56*	PI 532912	;1, 3	;	—	112	PI 566924	;1, 3	3	—

"0": No chlorotic flecks or uredinia; ";": No uredinia, but flecks or chlorosis; "1": Small uredinia with necrosis; "3": Moderate size uredinia with slight chlorosis; "+": uredinia somewhat larger than normal for the infection type; "—": No tested gene is detected. "*": wheat accessions with higher resistance at adult plant stage than at seedling stage.

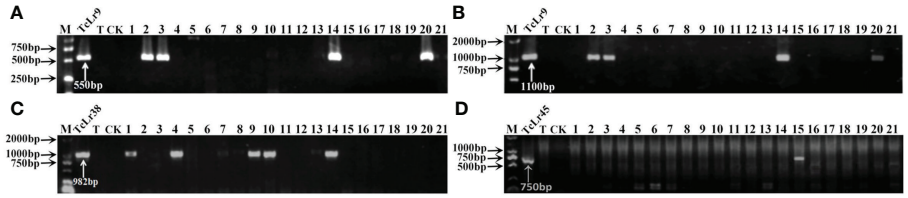


FIGURE 1
PCR amplifications results of molecular markers SCS5-550F/R **(A)** and J131/2 **(B)** for *Lr9*, $Y_{38}SCAR_{982}$-F/R for *Lr38* **(C)**, and PSc20H23/24 for *Lr45* **(D)** in part of accessions. M, DL2000 Marker; TcLr9, TcLr38 and TcLr45: Thatcher near-isogenic lines with single resistance genes *Lr9*, *Lr38* and *Lr45* (Positive control); T, Thatcher (Negative control); CK, ddH_2O, Lane 1-21: PI 601428, PI 601429, PI 601465, PI 601606, PI 595212, PI 17389, PI 17879, PI 17898, PI 486147, PI 486212, PI 486349, PI 494101, PI 542975, PI 542979, PI 547246, PI 594102, PI 497988, PI 17729, PI 601262, PI 601263, PI 601366.

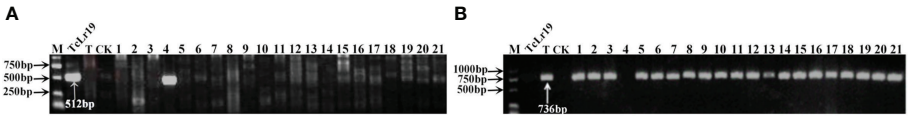


FIGURE 2
PCR amplifications results of molecular markers SCS265-F/R(Coupling) **(A)** and SCS253-F/R(Repulsion) **(B)** for *Lr19* in part of accessions. M, DL2000 Marker; TcLr19, Thatcher near-isogenic lines with single resistance genes *Lr19* (Positive control); T, Thatcher (Negative control); CK, ddH2O, Lane 1-21, PI 601807, Citr 13684, Citr 13874, Citr 14048, Citr 15229, Citr 15288, Citr 15929, Citr 17262, PI 535454, PI 518591, PI 527480, PI 531244, PI 532282, PI 532912, CItr 17723, CItr 17831, CItr 17856, CItr 17857, CItr 15075, CItr 15082, CItr 15290.

The second type was made up of *Lr24* and *Lr38* which existed in 11 wheat accessions representing 26.8%. The third type: *Lr9 +Lr24+Lr38* were detected in 1 accession, *Lr24+Lr38+Lr45* in 4 accessions, *Lr24+Lr29+Lr38* in 1 accession, and *Lr19+Lr38+Lr45* in 1 accession, which respectively accounted for 2.4%, 9.8%, 2.4% and 2.4% of 41 resistant accessions. The remaining 58 materials carried unknown effective leaf rust resistance genes, which accounted for 58.6% of 99 accessions with whole growth period resistance. These results indicated that the utilization ratios of *Lr45*, *Lr24* and *Lr38* were the highest among these

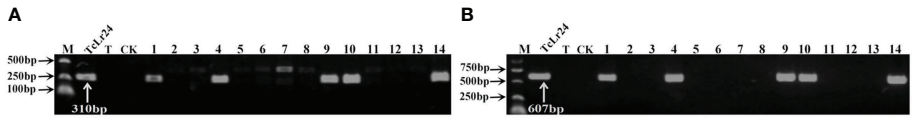


FIGURE 3
PCR amplifications results of molecular markers J09/1/2 **(A)** and $S1302_{609}$-F/R **(B)** for *Lr24* in part of accessions. M, DL2000 Marker; TcLr24, Thatcher near-isogenic lines with single resistance genes *Lr24* (Positive control); T, Thatcher (Negative control); CK, ddH2O, Lane 1-14, PI 601428, PI 601429, PI 601465, PI 601606, PI 595212, PI 17389, PI 17879, PI 17898, PI 486147, PI 486212, PI 486349, PI 494101, PI 542975, PI 542979.

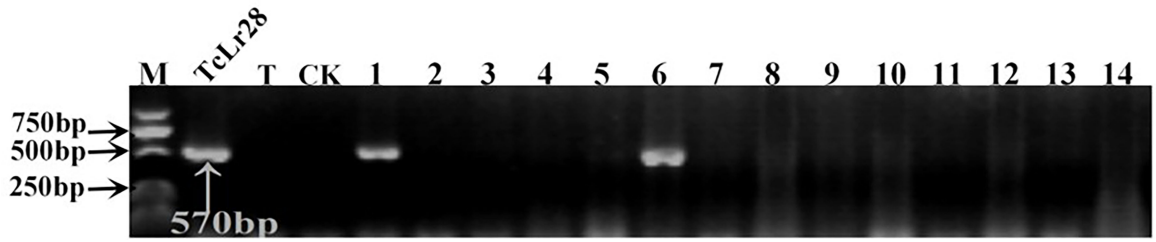


FIGURE 4
PCR amplifications results of molecular markers $SCS421_{570}$-F/R for *Lr28* in part of accessions. M, DL2000 Marker; TcLr28, Thatcher near-isogenic lines with single resistance genes *Lr28* (Positive control); T, Thatcher (Negative control); CK, ddH2O, Lane 1-14, PI 476975, PI 483469, PI 596335, PI 601722, PI 559376, PI 557537, PI 557538, PI 561197, PI 561198, PI 641952, PI 561200, PI 562382, PI 639450, PI 564083.

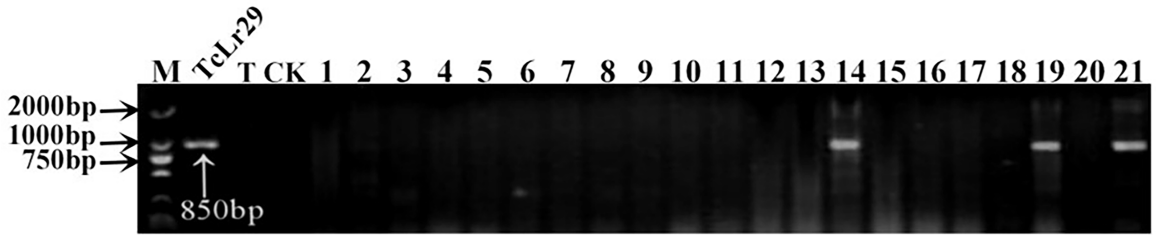


FIGURE 5
PCR amplifications results of molecular markers OPY10/1/2 for *Lr29* in part of accessions. M, DL2000 Marker; TcLr29, Thatcher near-isogenic lines with single resistance genes *Lr29* (Positive control); T, Thatcher (Negative control); CK, ddH2O, Lane 1-21, PI 548844, PI 548845, PI 548847, PI 550697, PI 552813, PI 543893, PI 547262, PI 547263, PI 555586, CItr 3780, CItr 15375, CItr 17264, PI 564700, PI 564851, PI 566923, PI 577793, PI 578213, PI 491396, PI 583676, PI 591560, PI 476974.

accessions, with *Lr45* accounted for 10.1%, *Lr24* for 18.2% and *Lr38* for 18.2% in 99 resistant accessions. In addition, *Lr47* was not detected in any of the tested wheat accessions in this study.

Discussion

Races of *P. triticina*, especially the predominant races THTT, THTS, PHTT, THJT and THJS from the wheat-growing regions of China, posed serious threat to wheat production in 2011-2015 due to high virulence to many cultivars and their widespread distribution (Zhang et al., 2020a; Zhang et al., 2020b). According to the field investigation and the identification of resistance to leaf rust, the majority of main wheat cultivars in the main wheat-growing regions were susceptible to leaf rust in China. For instance, at least 28 main wheat cultivars cultivated in the North China Plain, where is the largest wheat wheat-growing region with the highest wheat yield, were found to be susceptible to wheat leaf rust in recent years (Zhang et al., 2017a; Zhang et al., 2017b; Zhang et al., 2019b). Most of the Chinese

wheat cultivars, including the above mentioned, carried a few leaf rust resistance genes such as *Lr*1, *Lr16*, *Lr26*, *Lr37* among others (Gebrewahid et al., 2017; Zhang et al., 2017a; Zhang et al., 2017b; Gao et al., 2019; Zhang et al., 2019). Among these genes, *Lr1* and *Lr26* were the most used leaf rust resistance gene(s) in China. The proportions of *Lr1* and *Lr26* in 460 Chinese wheat accessions were 47.8% and 33.5% respectively (Zhang, 2015), while that of *Lr26* in 116 different wheat accessions in a study by Ren (2011) was observed to be as high as 37 %. While these genes have lost their effectiveness (Zhang et al., 2020a; Zhang et al., 2020b), which is key reasons for the poor resistance of wheat cultivars to leaf rust in China, so it is necessary to screen and identify more new sources of leaf rust resistance genes.

Wheat cultivars introduced from USA may confer different resistance sources compared with Chinese common wheat cultivars, due to the *P. triticina* predominant populations and the frequencies of virulence to leaf rust resistance genes are different (Kolmer and Hughes, 2017; Kolmer, 2019; Zhang et al., 2020a; Zhang et al., 2020b). For example, the wheat cultivars with the resistance genes *Lr9*, *Lr21*, *Lr24*, and *Lr39* have been released since the 1960s-1980s in the United States (Huerta-Espino et al., 2008; Kolmer et al., 2018), but these genes are rarely used in Chinese wheat cultivars. Some of these leaf rust resistance genes had lost their effectiveness, for instance, *Lr24* in the United State has begun to lose effectiveness to *P. triticinia* (Kolmer, 2019), but it is known to confer effective resistance in China and until now the virulent race of *P. triticinia* to *Lr24* is not be found. Against above background, we used the predominant races of Chinese *P. triticina* to identify the resistance of wheat cultivars from the United States for better and faster application in breeding. Due to the problem of hybridization incompatibility, it is more advantageous to select wheat resistant materials as parents for hybridization breeding compared with wild relatives of wheat or foreign gene introduction. Therefore, it is necessary to search for effective leaf rust resistance genes in known wheat cultivars or lines, especially those introduced cultivars which may have new potential resistance sources. In this study, 112 wheat accessions from the United States were resistant to Chinese predominant races of *P. triticina* at seedling stage, which indicated that the resistance resources of these wheat materials in the United States were abundant and may be a good source of resistance against wheat leaf rust in China. Moreover, the resistance of 99 out of the 112 wheat accessions also exhibited resistance to *P. triticinia* at adult plant stage, suggesting these accessions confers whole growth period resistance to leaf rust from seedling stage to adult plant stage. These cultivars were therefore subjected to effective leaf rust resistance gene analysis using molecular makers.

At present, 82 leaf rust resistance genes have been given gene designations (Bariana et al., 2022). The leaf rust resistance genes such as *Lr1*, *Lr2a*, *Lr2c*, *Lr3*, *Lr16*, *Lr26*, *Lr11*, *Lr17*, *LrB*, *Lr10*, *Lr14a*, *Lr2b*, *Lr3bg*, *Lr14b*, *Lr32*, *Lr33*, *Lr37*, and *Lr50* have lost the effectiveness in China from 2011 to 2015 (Zhang et al., 2020a; Zhang et al., 2020b). A few leaf rust resistance genes such as *Lr9*, *Lr19*, *Lr24*, *Lr28*, *Lr29*, *Lr38*, *Lr45* and *Lr47* still possessed their effective resistance to inhibit most of *P. triticina* isolates including those predominant races as mentioned above in China (Zhang et al., 2020a; Zhang et al., 2020b). These leaf rust resistance genes express effective resistance at both seedling and adult plant stages. No or very few race have been found to be virulent to these effective leaf rust resistance genes in China, so gene postulation method was difficult to be used for the analysis of these genes. While the molecular marker-assisted selection method is preferred and convenient for leaf rust resistance genes detection because of its rapidity and accuracy (Ding et al., 2010). Seven effective leaf rust resistance genes, *Lr9*, *Lr19*, *Lr24*, *Lr28*, *Lr29*, *Lr38* and *Lr45*, were detected in 41 of these accessions, which similar as the research reports that wheat cultivars with the leaf rust resistance genes *Lr9*, *Lr21*, *Lr24*, and *Lr39* have been released since the 1960s-1980s in the United States (Huerta-Espino et al., 2008; Kolmer et al., 2018). The exception to this assertion was *Lr21* and *Lr39* genes. The resistance of *Lr21* and *Lr39* to the Chinese predominant races of *P. triticina* were relatively low due to fact that they were losing their effectiveness (Zhang et al., 2020a; Zhang et al., 2020b). Research findings on the identification of wheat leaf rust resistance resources in China revealed that these effective leaf rust resistance genes are rarely distributed and accounted for a very low proportion in the wheat cultivars(lines) that have been in cultivation in China (Gao et al., 2019; Ren et al., 2012; Shi, 2010; Zhang, 2015). The tested leaf rust resistance genes were not be detected in some resistance accessions by the known leaf rust resistance gene markers, the main reason should be due to absent the correspondence resistance genes or maybe unknown leaf rust resistance gene in these wheat accessions. Although these tested effective leaf rust resistance genes were not present in the remaining 71 resistant accessions in this study, their high resistance phenotype indicated that these wheat accessions may carry others undetected, known or new leaf rust resistance genes. Therefore, all these resistant cultivars have certain potential as breeding materials in China.

In general, leaf rust resistance genes are broadly divided into two main categories: seedling resistance genes and adult plant resistance genes (Riaz, 2018). Most of the designated leaf rust resistance genes are the seedling resistance genes. These genes are usually detected at the seedling stage and remain effective throughout the growth stages of wheat. These genes are therefore known as all-stage resistance genes. Adult plant resistance genes are usually effective at the post-seedling stage. At present, among the designated 82 leaf rust resistance genes, only 16, *Lr12*, *Lr13*, *Lr22* (alleles a, and b), *Lr34*, *Lr35*, *Lr37*, *Lr46*, *Lr48*, *Lr49*, *Lr67*, *Lr68*, *Lr74*, *Lr75*, *Lr77* and *Lr78* specifically provide resistance at the adult plant stage (Kolmer et al., 2017). Some adult plant resistance genes are also known as slow rusting genes, such as *Lr34*, *Lr46*, *Lr67*, and *Lr68* because they can confer partial

resistance or slow rusting resistance (Parlevliet and Ommeren, 1985; Zhang et al., 2019). Thirteen of the 112 seedling resistance accessions were susceptible at adult plant stage which might have been caused by the temperature-sensitive genes. For example, the known temperature-sensitive leaf rust resistance genes *Lr11*, *Lr14a*, *Lr14b*, *Lr18*, *Lr20* and *Lr37* were noted to be more effective at lower temperatures (Mcintosh et al., 1995; Zhang et al., 2008; Wang et al., 2016). This phenomenon is worth further verification for conclusive establishment. Seventeen wheat accessions (marked with asterisks in Table 2) with moderate resistance at seedling stage exhibited high resistance at adult plant stage which indicated these wheat accessions may also carry polymeric genes with adult plant resistance gene (Kolmer, 2019).

The long-term cultivation of single resistance gene cultivars, especially those with single resistance gene cultivars are easy to lose resistance to leaf rust, while wheat cultivars with multiple resistance genes are more durable (Mcintosh, 1992). In this study, 18.2% of the resistant accessions carried polymeric genes, mainly including five types of polymeric genes, *Lr24+Lr38*, *Lr9 +Lr24+Lr38*, *Lr24+Lr38+Lr45*, *Lr24+Lr29+Lr38* and *Lr19+Lr38 +Lr45*, among which the wheat accessions that carried polymeric genes *Lr24+Lr38* were the most numerous. These wheat accessions showed high resistance at both seedling stage and adult plant stage, especially those accessions with polymeric genes were very valuable for breeding cultivars with resistance throughout the whole growth period. The rational utilization of polymeric genes can inhibit the predominant virulence race, stabilize the pathogen population by directional selection, and thus reduce the incidence and epidemic of leaf rust disease, and make the resistance of cultivars more durable (Staskawicz et al., 1995). Therefore, it is the future trend of wheat breeding to select polymeric leaf rust resistance genes with effectiveness, high resistance, and good comprehensive traits. In addition, the balance between yield traits and resistance traits is also important, and we need to pay attention to the coordination between them. In order to improve the level of leaf rust resistance of wheat cultivars, it is necessary for us to pyramid those effective leaf rust resistance genes into new cultivars without affecting other agronomic traits.

Conclusion

In the present study, we identified the seedling and adult plant resistance of 112 wheat accessions introduced from the U.S. National Plant Germplasm System using a mixture of Chinese predominant *P. triticinia* races THTT, THTS, PHTT, THJT and THJS. Seven effective resistance genes *Lr9*, *Lr19*, *Lr24*, *Lr28*, *Lr29*, *Lr38* and *Lr45* singly or in combination were found in 41 wheat accessions. Our study will provide theoretical guidance for rational using some of these wheat accessions as resistance material or variety to breeding program.

Data availability statement

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

Author contributions

HY, QM and DL designed the experiments. LZ and XZ carried out the experiments and wrote the manuscript. JL, XW, WG, QZ, and YL participated or assisted in some part of the study. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by Hebei Province Wheat Industry System (Grant No. HBCT2018010204), Key Research and Development Program of Hebei (Grant No. 21326508D).

Acknowledgments

We thank Dr. Harold Bockelman from National Plant Germplasm System (NPGS), USDA-ARS, Aberdeen, Idaho, USA for providing the wheat accessions.

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

Bariana, H. S., Babu, P., Forrest, K. L., Park, R. F., and Bansal, U. K. (2022). Discovery of the new leaf rust resistance gene Lr82 in wheat: Molecular mapping and marker development. *Genes (Basel)* 13, 964. doi: 10.3390/genes13060964

Bassi, F. M., Bentley, A. R., Charmet, G., Ortiz, R., and Crossa, J. (2015). Breeding schemes for the implementation of genomic selection in wheat (Triticum ssp.). *Plant Sci.* 242, 23–36. doi: 10.1016/j.plantsci.2015.08.021

Beukert, U., Thorwarth, P., Zhao, Y., Longin, C. F. H., Serfling, A., Ordon, F., et al. (2020). Comparing the potential of marker-assisted selection and genomic prediction for improving rust resistance in hybrid wheat. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.594113

Bhardwaj, S. C., Gangwar, O. P., Prasad, P., Kumar, S., Khan, H., and Gupta, N. (2019). Physiologic specialization and shift in puccinia triticina pathotypes on wheat in Indian subcontinent during 2013-2016. *Ind. J. Phytopathol.* 72, 23–34. doi: 10.1007/s42360-018-00110-9

Bolton, M. D., Kolmer, J. A., and Garvin, D. F. (2008). Wheat leaf rust caused by puccinia triticina. *Mol. Plant Pathol.* 9, 563–575. doi: 10.1111/j.1364-3703.2008.00487.x

Cherukuri, D. P., Gupta, S. K., Charpe, A. M., Koul, S., Prabhu, K., Singh, R. B., et al. (2005). Molecular mapping of aegilops speltoides derived leaf rust resistance gene Lr28 in wheat. *Euphytica* 143, 19–26. doi: 10.1007/s10681-005-1680-6

Ding, Y. H., Liu, H., Shi, L. H., Wen, X. L., Zhang, N., Yang, W. X., et al. (2010). Wheat leaf rust resistance in 28 Chinese wheat mini-core collections. *Acta Agron. Sin.* 36, 1126–1134. doi: 10.3724/SP.J.1006.2010.01126

Gao, Y., Zhao, N., Zhao, X. F., Yan, H. F., and Liu, D. Q. (2019). Identification of leaf rust resistance at seeding stage and analysis of lr genes in 182 Huang-Huai-Hai wheat cultivars(lines). *J. Plant Genet. Resour.* 20, 1213–1222. doi: 10.13430/j.cnki.jpgr.20190110001

Gebrewahid, T. W., Yao, Z. J., Yan, X. C., Gao, P., and Li, Z. F. (2017). Identification of leaf rust resistance genes in Chinese common wheat cultivars. *Plant Dis.* 101, 1729–1737. doi: 10.1094/PDIS-02-17-0247-RE

Gill, K. S., Lubbers, E. L., Gill, B. S., Raupp, W. J., and Cox, T. S. (1991). A genetic linkage map of triticum tauschii (DD) and its relationship to the d genome of bread wheat (AABBDD). *Genome* 34, 362. doi: 10.1139/g91-058

Gupta, S. K., Charpe, A., Koul, S., Haq, Q., and Prabhu, K. (2006b). Development and validation of SCAR markers Co-segregating with an agropyron elongatum derived leaf rust resistance gene Lr24 in wheat. *Euphytica* 150, 233–240. doi: 10.1007/s10681-006-9113-8

Gupta, S. K., Charpe, A., Koul, S., Prabhu, K. V., and Haq, Q. M. (2005). Development and validation of molecular markers linked to an aegilops umbellulata-derived leaf-rust-resistance gene, Lr9, for marker-assisted selection in bread wheat. *Genome* 48, 823–830. doi: 10.1139/g05-051

Gupta, S. K., Charpe, A., Prabhu, K. V., and Haque, Q. M. (2006a). Identification and validation of molecular markers linked to the leaf rust resistance gene Lr19 in wheat. *Theor. Appl. Genet.* 113, 1027–1036. doi: 10.1007/s00122-006-0362-7

Helguera, M., Khan, I. A., and Dubcovsky, J. (2000). Development of PCR markers for wheat leaf rust resistance gene Lr47. *Theor. Appl. Genet.* 100, 1137–1143. doi: 10.1007/s001220051397

Huerta-Espino, J., Singh, R. P., Germán, S., McCallum, B. D., Park, R. F., Chen, W. Q., et al. (2011). Global status of wheat leaf rust caused by puccinia triticina. *Euphytica* 179, 143–160. doi: 10.1007/s10681-011-0361-x

Huerta-Espino, J., Singh, R. P., and Reyna-Martinez, J. (2008). First detection of virulence to genes Lr9 and Lr25 conferring resistance to leaf rust of wheat caused by puccinia triticina in Mexico. *Plant Dis.* 92, 311. doi: 10.1094/PDIS-92-2-0311A

Hu, Y. Y., Sun, Y., Zhang, H. S., Wei, X. J., Du, D. D., Yang, W. X., et al. (2014). Wheat leaf rust resistance genes of eight wheat breeding parents. *J. Plant Genet. Resour.* 15, 802–809. doi: 10.13430/j.cnki. jpgr.2014.04.018

Kolmer, J. A. (2019). Virulence of puccinia triticina, the wheat leaf rust fungus, in the united states in 2017. *Plant Dis.* 103, 2113–2120. doi: 10.1094/PDIS-09-18-1638-SR

Kolmer, J. A., Bernardo, A., Hayden, M. J., and Chao, S. (2017). Adult plant leaf rust resistance derived from toropi wheat is conditioned by Lr78 and three minor QTL. *Phytopathology* 108, 246–253. doi: 10.1094/PHYTO-07-17-0254-R

Kolmer, J. A., and Hughes, M. E. (2017). Physiologic specialization of puccinia triticina on wheat in the united states in 2016. *Plant Dis.* 102, 1066–1071. doi: 10.1094/PDIS-11-17-1701-SR

Kolmer, J. A., Su, Z., Bernardo, A., Bai, G., and Chao, S. (2018). A backcross line of Thatcher wheat with adult plant leaf rust resistance derived from duster wheat has Lr46 and Lr77. *Phytopathology* 109, 127–132. doi: 10.1094/PHYTO-06-18-0184-R

Liu, T. G., and Chen, W. Q. (2012). Race and virulence dynamics of puccinia triticina in China during 2000-2006. *Plant Dis.* 96, 1601–1607. doi: 10.1094/PDIS-06-10-0460-RE

Mcintosh, R. A. (1992). Pre-emptive breeding to control wheat rusts. *Euphytica* 63, 103–113. doi: 10.1007/978-94-017-0954-5_9

Mcintosh, R., Wellings, C. R., and Park, R. (1995). *Wheat rusts: An atlas of resistance genes* (Dordrecht: Kluwer Academic Publishers).

Miedaner, T., and Korzun, V. (2012). Marker-assisted selection for disease resistance in wheat and barley breeding. *Phytopathology* 102, 560–566. doi: 10.1094/PHYTO-05-11-0157

Parlevliet, J. E., and Ommeren, (1985). Race-specific effects in major-genic and polygenic resistance of barley to barley leaf rust in the field and how to distinguish them. *Euphytica* 34, 689–695. doi: 10.1007/BF00035405

Pink, D. A. C. (2002). Strategies using genes for non-durable disease resistance. *Euphytica* 124, 227–236. doi: 10.1023/A:1015638718242

Ren, X. L. (2011). *Postulation and molecular detection of wheat leaf rust resistance genes of commercial wheat cultivars in China* (Beijing: Chinese Academy of Agricultural Sciences).

Ren, X. L., Liu, T. G., Liu, B., Gao, L., and Chen, W. Q. (2012). Multiplex PCR assay for detection of wheat leaf rust resistance genes Lr9-Lr26 and Lr19-Lr20 in 116 Chinese wheat cultivars (lines). *Plant Prot.* 38, 29–36. doi: 10.3969/j.issn.0529-1542.2012.02.006

Riaz, A. (2018). *Unlocking new sources of adult plant resistance to wheat leaf rust. Ph.D. thesis* (Brisbane, Australia: The University of Queensland).

Roelfs, A. P. (1984). Race specificity and methods of study. *Cereal Rusts* A4, 131–164. doi: 10.1016/B978-0-12-148401-9.50011-X

Schachermayr, G. M., Messmer, M. M., Feuillet, C., Winzeler, H., Winzeler, M., and Keller, B. (1995). Identification of molecular markers linked to the agropyron elongatum-derived leaf rust resistance gene Lr24 in wheat. *Theor. Appl. Genet.* 90, 982–990. doi: 10.1007/BF00222911

Schachermayr, G. M., Siedler, H., Gale, M. D., Winzeler, H., Winzeler, M., and Keller, B. (1994). Identification and localization of molecular markers linked to the Lr9 leaf rust resistance gene of wheat. *Theor. Appl. Genet.* 88, 110–115. doi: 10.1007/BF00222402

Shi, L. H. (2010). *Evaluation of wheat leaf rust resistance of 60 Chinese wheat cultivars in china. master's thesis* (Baoding: Hebei Agricultural University).

Shi, L. H., Zhang, N., Hu, Y. Y., Yang, W. X., and Liu, D. Q. (2011). Evaluation of wheat leaf rust resistance of 10 new wheat cultivars(lines). *Sci. Agric. Sin.* 44, 2900–2908. doi: 10.3864/j.issn.0578-1752.2011.14.006

Staskawicz, B. J., Ausubel, F. M., Baker, B. J., Ellis, J. G., and Jones, J. D. G. (1995). Molecular genetics of plant resistance. *Science* 268, 661–667. doi: 10.1126/science.7732374

Tar, M., Purnhauser, L., Csosz, L., Mesterházy, A., and Gyulai, G. (2002). Identification of molecular markers for an efficient leaf rust resistance gene (Lr29) in wheat. *Acta Biologica Szegediensis* 46, 133–134.

Wang, C. F., Yin, G. H., Xia, X. C., He, Z. H., and Zhang, P. P. (2016). Molecular mapping of a new temperature-sensitive gene LrZH22 for leaf rust resistance in Chinese wheat cultivar zhoumai 22. *Mol. Breed.* 36, 18. doi: 10.1007/s11032-016-0437-3

Wu, H., Kang, Z., Li, X., Li, Y., Li, Y., Wang, S., et al. (2020). Identification of wheat leaf rust resistance genes in Chinese wheat cultivars and the improved germplasms. *Plant Dis.* 104, 2669–2680. doi: 10.1094/PDIS-12-19-2619-RE

Yan, H. F. (2009). *Molecular marker for Lr38, Lr45 and their resistance-related analysis against wheat lesf rust. master's thesis* (Baoding: Hebei Agricultural University).

Yan, H. F., Yang, W. X., Chu, D., and Liu, D. Q. (2008). A new marker tagged to the leaf rust resistance gene Lr38. *Sci. Agric. Sin.* 41, 3604–3609. doi: 10.3864/j.issn.0578-1752.2008.11.021

Zhang, Y. Q. (2015). *Identifcation of wheat leaf rust resistance genes in 460 Chinese wheat cultivars and QTL mapping for adult-plant resistance gene to leaf rust in Italian wheat cultivar libellula. master's thesis* (Baoding: Hebei Agricultural University).

Zhang, P. P., Gebrewahid, T. W., Zhou, Y., Li, Q. L., Li, Z. F., and Liu, D. Q. (2019a). Seedling and adult plant resistance to leaf rust in 46 Chinese bread wheat landraces and 39 wheat lines with known lr genes. *J. Integr. Agr* 18, 60345–60352. doi: 10.1016/S2095-3119(19)62575-X

Zhang, L. Y., Meng, Q. F., Kang, J., Yang, W. X., Yan, H. F., and Liu, D. Q. (2015). Genetic diversity analysis of puccinia recondita by UP-PCR. *Mycosystema* 34, 215–226. doi: 10.13346/j.mycosystema.130283

Zhang, L., Shi, C. C., Li, L. R., Li, M., Meng, Q. F., Yan, H. F., et al. (2020a). Race and virulence analysis of puccinia triticina in China in 2014 and 2015. *Plant Dis.* 104, 455–464. doi: 10.1094/PDIS-05-19-1051-RE

Zhang, L., Wang, J., Zhang, M. Y., Xu, H. P., Yan, H. F., and Liu, D. Q. (2017a). Analysis of wheat leaf rust resistance genes in 16 main wheat cultivars

in henan. *J. Plant Genet. Resour.* 18, 546–554. doi: 10.13430/j.cnki.jpgr.2017.03.020

Zhang, L., Xiao, Y., Gao, Y., Zhao, N., An, Y. J., Yang, W. X., et al. (2020b). Race and virulence analysis of puccinia triticina in China during 2011 to 2013. *Plant Dis.* 104, 2095–2101. doi: 10.1094/PDIS-01-20-0047-RE

Zhang, L., Zhang, M. Y., Gao, Y., Xu, H. P., Liu, C., Liu, J. J., et al. (2017b). Analysis of leaf rust resistance in 12 main wheat cultivars (lines) in Shandong. *J. Plant Genet. Resour.* 18, 676–684. doi: 10.13430/j.cnki.jpgr.2017.04.010

Zhang, X. L., Zhang, H. H., Xu, X. Y., Yang, H. L., Duan, Z. Y., and Yao, Z. J. (2019b). Identification and analysis of leaf rust resistance genes in 12 main wheat cultivars(lines) from hebei province. *J. Plant Genet. Resour.* 20, 982–990. doi: 10.13430/j.cnki.jpgr.20181108001

Zhang, Y. C., Zhang, Z. Q., Yan, H. F., Li, Z. F., Liu, D. Q., and Yang, W. X. (2008). Postulation of temperature-sensitive genes for resistance to wheat leaf rust in wheat-wheatgrass no. 33 at seedling stage. *J. Triticeae Crops* 28, 150–153. doi: 10.7606/j.issn.1009-1041.2008.01.029

Zhao, L. N., Ren, X. D., Hu, Y. Y., Zhang, T., Zhang, N., Yang, W. X., et al. (2013). Evaluation of wheat leaf rust resistance of 23 Chinese wheat mini-core collections. *Sci. Agric. Sin.* 46, 441–450. doi: 10.3864/j.issn.0578-1752.2013.03.001

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 26 October 2022
DOI 10.3389/fpls.2022.1048252

OPEN ACCESS

EDITED BY
Xiaoli Fan,
Chengdu Institute of Biology, Chinese Academy of Sciences (CAS), China

REVIEWED BY
Huagang He,
Jiangsu University, China
Cheng Liu,
Crop Research Institute, Shandong Academy of Agricultural Sciences, China

*CORRESPONDENCE
He Zhao
hezhao311@163.com
Diaoguo An
dgan@sjziam.ac.cn

†These authors have contributed equally to this work

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

RECEIVED 19 September 2022
ACCEPTED 10 October 2022
PUBLISHED 26 October 2022

CITATION
Jin Y, Gu T, Li X, Liu H, Han G, Shi Z, Zhou Y, Fan J, Wang J, Liu W, Zhao H and An D (2022) Characterization of a new splicing variant of powdery mildew resistance gene *Pm4* in synthetic hexaploid wheat YAV249. *Front. Plant Sci.* 13:1048252. doi: 10.3389/fpls.2022.1048252

Characterization of a new splicing variant of powdery mildew resistance gene *Pm4* in synthetic hexaploid wheat YAV249

Yuli Jin[1†], Tiantian Gu[1†], Xiuquan Li[3], Hong Liu[1], Guohao Han[1], Zhipeng Shi[1], Yilin Zhou[4], Jieru Fan[4], Jing Wang[1], Wei Liu[4], He Zhao[2*] and Diaoguo An[1,5*]

[1]Center for Agricultural Resources Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Shijiazhuang, Hebei, China, [2]Institute of Biotechnology and Food Science, Hebei Academy of Agriculture and Forestry Science/the Key Laboratory of Plant Genetic Engineering of Hebei Province, Shijiazhuang, Hebei, China, [3]The National Key Facility for Crop Gene Resources and Genetic Improvement, Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing, China, [4]The State Key Laboratory for Biology of Plant Disease and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, China, [5]The Innovative Academy for Seed Design, Chinese Academy of Sciences, Beijing, China

Powdery mildew, caused by *Blumeria graminis* f. sp. *tritici* (*Bgt*), is a destructive fungal disease of wheat throughout the world. Utilization of effective powdery mildew resistance genes and cultivars is considered as the most economic, efficient, and environmental-friendly method to control this disease. Synthetic hexaploid wheat (SHW), which was developed through hybridization of diploid *Aegilops* and tetraploid wheat, is a valuable genetic resource for resistance to powdery mildew. SHW line YAV249 showed high levels of resistance to powdery mildew at both the seedling and adult stages. Genetic analysis indicated that the resistance was controlled by a single dominant gene, temporarily designated *PmYAV*. Bulked segregant analysis with wheat 660K single nucleotide polymorphism (SNP) array scanning and marker analysis showed that *PmYAV* was located on chromosome 2AL and flanked by markers *Xgdm93* and *Xwgrc763*, respectively, with genetic distances of 0.8 cM and 1.2 cM corresponding to a physic interval of 1.89 Mb on the Chinese Spring reference genome sequence v1.0. Sequence alignment analysis demonstrated that the sequence of *PmYAV* was consistent with that of *Pm4a* but generated an extra splicing event. When inoculated with different *Bgt* isolates, *PmYAV* showed a significantly different spectrum from *Pm4a*, hence it might be a new resistant resource for improvement of powdery mildew resistance. The flanked markers GDM93 and WGRC763, and the co-segregated markers BCD1231 and JS717/JS718 were confirmed to be easily performed in marker-assisted selection (MAS) of *PmYAV*. Using MAS strategy, *PmYAV* was transferred into the commercial cultivar Kenong 199 (KN199) and a

wheat line YK13 was derived at generation BC_3F_3 from the population of YAV249/4*KN199 due to its excellent agronomic traits and resistance to powdery mildew. In conclusion, an alternative splicing variant of *Pm4* was identified in this study, which informed the regulation of *Pm4* gene function.

KEYWORDS

***Blumeria graminis* f. sp. *tritici*, *PmYAV*, *Pm4a*, markers, marker-assisted selection**

Introduction

Powdery mildew, caused by *Blumeria graminis* f. sp. *tritici* (*Bgt*), is a destructive fungal disease of common wheat throughout the world, accounting for approximately 5% of the total yield losses caused by wheat pathogens and pests (Serge et al., 2019). Utilization of effective powdery mildew resistance genes (*Pm* genes) and resistant cultivars is the most economic, efficient, and environmental-friendly method to control this disease.

So far, more than 100 *Pm* genes/alleles from wheat and its relatives have been reported, including formally named *Pm1* to *Pm68* at 63 loci (noting that *Pm8=Pm17*, *Pm18=Pm1c*, *Pm22=Pm1e*, *Pm23=Pm4c*, and *Pm31=Pm21*) and 30 temporarily named genes (He et al., 2020; McIntosh et al., 2020). With the development of sequencing technologies and releases of reference genomes in succession, 13 *Pm* genes, including *Pm1a*, *Pm2*, *Pm3*, *Pm4b*, *Pm5e*, *Pm8*, *Pm17*, *Pm21*, *Pm24*, *Pm38*, *Pm41*, *Pm46* and *Pm60* have been cloned (Krattinger et al., 2009; Brunner et al., 2011; Hurni et al., 2013; Moore et al., 2015; Sánchez-Martín et al., 2016; Singh et al., 2018; Xing et al., 2018; Zou et al., 2018; Li et al., 2020; Lu et al., 2020; Xie et al., 2020; Hewitt et al., 2021; Sánchez-Martín et al., 2021). Because of the large-scale deployment of single resistance genes and the evolution of corresponding pathogens, many resistance genes have lost their effectiveness, such as *Pm1*, *Pm3*, *Pm7* and *Pm8*,either regionally or across the world (Jin et al., 2021). Therefore, it is essential to explore novel resistance genes/alleles of powdery mildew and transfer them into wheat cultivars to enrich the diversity of resistance sources and enhance the durability of disease resistance.

Gene *Pm4* has been one of the most widely used powdery mildew resistance genes and can be found in many cultivars all over the world (Ullah et al., 2018; Jin et al., 2021). Since *Pm4* allele *Pm4a* was first mapped on the long arm of wheat chromosome 2A in Khapli (The et al., 1979), five alleles (*Pm4a-Pm4e*) have been reported at that locus (Ullah et al., 2018). Recently, *Pm4b*, encoded a putative chimeric protein of a serine/threonine kinase with multiple C2 domains and transmembrane regions, has been cloned through Mutant Chromosome Flow Sorting and Sequencing (Sánchez-Martín et al., 2021). *Pm4b* consists of seven exons and contains two alternative transcripts, denoted *Pm4b_V1* and *Pm4b_V2*. Subsequently, three new *Pm4* alleles, tentatively denoted as *Pm4f*, *Pm4g* and *Pm4h* were discovered by the diagnostic functional markers JS717/JS718 (Sánchez-Martín et al., 2021). It is noted that other loci, for example, *Pm1*, *Pm2*, and *Pm5* also have multiple allelic variations because of the long term interactions with the co-evolving pathogen. It is reported that the allelic variations of the documented *Pm* genes played important roles in resistance breeding, for their different spectra of effectiveness to pathogen isolates. Therefore, identification of the allelic variations in detail is important, not only for increasing the diversity of resistance resource, but also for understanding the host-pathogen interaction mechanism and applying in resistance gene pyramiding breeding (Srichumpa et al., 2005; Koller et al., 2018).

Synthetic hexaploid wheat (SHW), artificially created hexaploid wheat, is usually developed from crosses of durum wheat, *Triticum turgidum* ssp. *durum* (2n = 4x = 28, AABB), and *Aegilops tauschii* (2n = 2x = 14, DD), followed by chromosome doubling of the F_1 hybrids (Kamali et al., 2014). SHW possesses valuable genes for wheat improvement including disease resistance, abiotic-stress tolerance and yield related genes, which can be transferred from tetraploid wheat and *Ae. tauschii* to common wheat cultivars as a bridge resource (Trethowan and Mujeeb-Kazi, 2008; Li et al., 2011). Since the late 1980s, the International Maize and Wheat Improvement Center (CIMMYT) has developed more than 1000 SHW lines (Das et al., 2016), these SHW lines are widely used to improve wheat quality, yield, and other important agronomic traits all over the world (Li et al., 2018). Some powdery mildew resistance genes from SHW have been reported. The SHW line "XX 194" was derived from the cross of susceptible cultivar *T. durum* "Moroccos 182" and the selected resistant accession *Ae. squarrosa* "AE 457/78". Lutz et al. (1995) revealed "XX 194" conferred the same response pattern to *Bgt* isolates as *Pm2*. The SHW line "XX 186" which was produced from a cross between susceptible cultivar *T. durum* "Santa Marta" and *Ae. squarrosa* "BGRC 1458" showed a different response pattern from all other

named *Pm* genes, being designated *Pm19* (Lutz et al., 1995). Hu et al. (2001) found that SHW line M81 (68.111/RGB-U//WARD/3/*Ae. tauschii* 452) showed resistance to powdery mildew at both the seedling and adult plant stages. Genetic analyses indicated that the resistance gene originated from *Ae. tauschii* (Hu et al., 2001).

YAV249 is a SHW line that showed effective resistance to powdery mildew in the field over multiple years. To make better use of its *Pm* gene(s), the major objectives of this study were to: (i) determine the inheritance of powdery mildew resistance and identify the *Pm* gene in YAV249; (ii) investigate the relationship between *PmYAV* and *Pm4*; (iii) evaluate diagnostic markers for marker-assisted selection (MAS); (iv) apply *PmYAV* in wheat powdery mildew resistance breeding.

Materials and methods

Plant materials

SHW line YAV249 was derived from the cross YAV-2/TEZ//*Ae. squarrosa* 249 (*T. durum* and *Ae. tauschii*) and is highly resistant to powdery mildew at all growth stages. Wheat cultivar Shixin 828 (SX828) is susceptible to the tested *Bgt* isolates and used as the susceptible parent to produce F_2 and $F_{2:3}$ populations for genetic analysis and mapping of *PmYAV* in this study. Wheat cultivar Mingxian 169, which does not carry any known *Pm* genes, was used as a susceptible control for phenotypic evaluation (Qie et al., 2019). The 47 wheat accessions with known *Pm* genes or gene combinations were tested with 24 different *Bgt* isolates collected from diseased wheat fields in different wheat growing areas in China to make comparison with the *Pm* gene(s) in YAV249 (Table 1) (An et al., 2019). Additionally, 192 wheat accessions including cultivars, landraces, advanced lines and introduced cultivars were determined by using the flanked or co-segregated markers of *Pm* gene(s) in YAV249 to validate the applicability for MAS (Supplementary Table 1). The susceptible commercial cultivar Kenong 199 (KN199) was used as the recurrent parent to cross and backcross with YAV249 to produce progenies.

Reaction-phenotyping to *Bgt* isolates

At the adult stage, YAV249 was inoculated with *Bgt* isolates mixture of E09, E11, E18 and E20 in the field nurseries with three replicates. The assessments were performed from 2019 to 2021 at Luancheng Agro-Ecological Experimental Station (37° 53′ 15″ N, 114° 40′ 47″ E), Chinese Academy of Sciences, Shijiazhuang, Hebei Province, China. YAV249 was planted with four rows, 30 seeds per row (1.5 m), and wheat cultivar Mingxian 169 was planted around the plot as susceptible control for each replicate. When Mingxian 169 showed severe disease symptoms, the phenotype reaction of YAV249 was assessed at least two times in weekly intervals with a 0-9 scale for ITs, ITs 0-2 were considered as highly resistant, ITs 3-4 as moderately resistant, ITs 5-6 as moderately susceptible and ITs 7-9 as highly susceptible (Sheng and Duan, 1991).At the seedling stage, YAV249 and 47 wheat accessions with known *Pm* genes or gene combinations were tested for response to 24 single-pustule-derived *Bgt* isolates to compare their resistance spectra (Table 1). *Bgt* isolate E09 prevalent in the main wheat producing regions of China (Zhou et al., 2005) was used to inoculate YAV249, SX828, KN199 and progenies of YAV249/SX828 and YAV249/KN199. At least 20 plants of each variety/progeny were used for phenotyping of *Bgt* isolates in three replicates. Response to powdery mildew of all materials was executed in a greenhouse. Materials were planted in the rectangular trays with 128 wells and inoculated at the one-leaf stage by dusting fresh conidia of a *Bgt* isolate, and Mingxian 169 was planted randomly in the trays as the susceptible control. After inoculation, the trays were treated in a high humidity environment at 22 ± 2°C and 10 h of darkness at 18°C. When the first leaf of Mingxian 169 showed full development of pustules about 14-15 days after inoculation, infection types (ITs) for each plant were scored using a 0-4 scale, and plants with ITs 0-2 were regarded as resistant and those with ITs 3 and 4 susceptible (Sheng, 1988; Wang et al., 2005).

Microscopic analyses of powdery mildew resistance reaction

Microscopic analyses of the plants were performed as previously described (Wang et al., 2014). Two cm leaf segments of YAV249 and SX828 at 7 days post-inoculation (dpi) inoculated with *Bgt* isolate E09 were fixed at 37°C for 24 h in 2 ml of Carnoy's Fluid (ethanol: acetic acid, 3:1, v/v), then stained with 2 ml of 0.6% (w/v) Coomassie blue solution for 3 min. Excess dye was rinsed off carefully with distilled water. Samples were observed under an Olympus BX-53 microscope (Olympus, Tokyo, Japan).

Bulked segregant analysis with the wheat 660K SNP array

Genomic DNA was extracted from young leaf tissues after evaluation of their powdery mildew reactions following the method of cetyletrimethylammonium bromide (CTAB) (Sharp et al., 1988). Resistant and susceptible bulks were made from equal amounts of DNA from 20 homozygous resistant and 20 homozygous susceptible $F_{2:3}$ families of YAV249/SX828 for Illumina wheat 660K single nucleotide polymorphism (SNP) array scanning by China Golden Marker Company (Beijing, China). Single nucleotide polymorphism genotype calling and clustering was performed with software Genome Studio

TABLE 1 Response spectra of YAV249 and wheat accessions with known powdery mildew (*Pm*) resistance genes to 24 different *Blumeria graminis* f. sp. *tritici* (*Bgt*) isolates at the seedling stage.

Cultivars/lines	*Pm* genes	Resistance/susceptible ratio	*Blumeria graminis* f. sp. *tritici* isolates																							
			E01	E05	E06	E07	E09	E11	E13	E15	E16	E17	E18	E20	E21	E23-(1)	E23-(2)	E26	E30-(1)	E30-(2)	E31	E32	E49	E50	E60	E69
YAV249	*PmYAV*	16/8	0	0	0	4	0;	2	1	4	2	0	4	4	2	0	0	0	0	4	4	4	0	4	2	2
Khapli/8cc	*Pm4a*	14/10	2	0	0;	4	0;	0;	0	4	0;	0	4	4	0;	0;	0;	0	0	4	4	4	0	4	3	3
Armada	*Pm4b*	18/6	2	0;	0;	3	0;	0;	0	0;	0;	0;	0	3	0	0	0;	0;	0	4	4	3	0	4	2	0;
81-7241	*Pm23=Pm4c*	23/1	0;	0;	0;	0	0;	0;	0	0;	0;	0;	0	0;	0;	0;	0;	0	0;	3	0;	0;	0	0	0;	0;
Axminster/8cc	*Pm1a*	2/22	4	4	4	3	4	4	0;	4	4	4	4	4	4	2	4	3	4	4	4	4	4	4	4	4
Ulka/8cc	*Pm2*	18/6	0;	0;	0;	0;	0;	0;	1	0;	0	0;	4	4	0;	0;	4	0;	0;	4	0;	0;	0;	4	0;	4
Asosan/8cc	*Pm3a*	5/19	4	4	3	3	4	3	3	0	4	1	3	4	3	4	4	4	0;	2	4	0;	4	4	4	4
Chul/8cc	*Pm3b*	4/20	4	4	0;	3	4	3	0;	0	4	4	4	4	3	3	4	4	3	4	0;	4	4	3	3	3
Sonora/8cc	*Pm3c*	1/23	4	4	4	4	4	4	4	0	3	4	3	4	3	3	4	4	3	4	4	3	4	4	4	4
Kolibri	*Pm3d*	17/7	0	0;	0;	4	0;	3	1	0;	4	0	0	0	0	0;	0;	0	0;	4	0;	0;	4	3	2	3
Mich.Amber/8cc	*Pm3f*	0/24	4	4	4	4	3	4	4	4	4	4	4	4	3	4	4	4	4	4	4	4	4	4	4	4
W150	*Pm3e*	0/24	4	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	3	4	4	4	4
Hope/8cc	*Pm5a*	0/24	4	4	4	4	4	3	4	4	4	4	4	3	4	3	3	4	3	4	3	3	4	4	3	3
Aquila	*Pm5b*	20/3	0;	0;	0;	0	0;	0;	0	0	0;	0;	4	3	0	0;	0;	0	0;	–	0;	0;	1	4	2	1
Fuzhuang 30	*Pm5e*	15/9	2	1	3	4	0;	3	4	0;	0;	4	1	3	1	3	4	3	1	2	0;	0;	2	1	2	1
Timgalen	*Pm6*	0/24	4	4	4	4	4	4	4	4	4	4	4	4	3	3	3	3	4	4	4	4	3	4	4	4
Coker 747	*Pm6*	4/20	4	4	3	3	3	3	4	4	2	4	4	3	3	3	0;	4	2	4	3	2	3	4	3	4
CI14189	*Pm7*	0/24	4	4	4	4	4	3	4	4	4	4	4	4	4	4	4	4	4	3	4	3	4	4	4	4
Kavkaz	*Pm8*	1/23	4	3	4	4	4	3	3	4	4	4	4	4	3	3	3	3	3	3	4	4	3	4	0	4
Wembley	*Pm12*	24/0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
R4A	*Pm13*	22/2	0;	0;	0;	0	0;	0;	0	0;	0;	0	2	1	0	3	0;	0	0;	3	0;	0;	0	0;	1	0;
Brigand	*Pm16*	24/0	0	0;	0	0	0;	0	0	0	0;	0	1	0;	0;	0	0;	0	0	0	0;	0	0	0;	0	0
Amigo	*Pm17*	10/14	4	0;	1	4	3	0;	0	0	0;	2	3	4	0;	0;	0;	4	3	4	4	3	4	4	3	4
MIN	*Pm18=Pm1c*	24/0	0	0;	0;	0	0;	0;	0;	0;	0;	0;	0;	0;	0;	0;	0;	0;	0;	0;	0;	0;	0	0	0	0;
XX186	*Pm19*	0/24	4	4	4	4	4	3	4	4	4	4	3	4	4	4	4	4	4	4	4	4	4	4	4	4
TAM104/Thatcher	*Pm20*	17/7	4	0;	0	0	3	0;	0	0	3	0	2	2	2	4	0;	0	0;	2	4	4	3	0;	0	0;
Nannong 9918	*Pm21*	24/0	0;	0;	0;	0	0;	0;	0;	0;	0;	0;	0	0;	0;	0;	0;	0	0	0;	0;	0;	0	0;	0;	0;
Virest	*Pm22=Pm1e*	22/2	3	0	0	0	0;	0;	0	0;	0;	0	0	0	0	0	0;	0	0;	3	0;	0;	0	0	0	0
Chiyacao	*Pm24*	22/2	0	2	0;	1	0;	0;	0;	4	0;	0;	0;	0;	0	0;	0;	3	0;	0;	0;	0;	0	0;	0;	0;
NCA5	*Pm25*	20/4	0	0	0;	0	0;	0;	1	0;	0;	3	3	0;	0	0;	0;	0	0;	3	2	0;	0	3	0;	0;

(Continued)

10.3389/fpls.2022.1048252

TABLE 1 Continued

Cultivars/ lines	*Pm* genes	Resistance/susceptible ratio	*Blumeria graminis* f. sp. *tritici* isolates																						
			E01	E05	E06	E07	E09	E11	E13	E15	E16	E17	E18	E20	E21	E23-(1)	E23-(2)	E26	E30-(1)	E30-(2)	E31	E32	E49	E50	E60	E69
5P27	*Pm30*	17/7	2	0;	0;	4	0;	0;	0;	0;	0;	0;	3	4	0;	0;	0;	0	0;	4	3	3	0;	4	0;	0;
NCA7	*Pm34*	14/10	0	2	2	0;	3	0;	3	2	2	0	2	3	3	3	2	4	0;	3	4	0;	3	0	0;	3
NCA3	*Pm35*	20/4	0;	0	3	3	1	0;	0;	1	0;	0;	0;	1	0;	3	0;	0;	0;	0;	3	0;	0;	0	0;	0;
Pm36	*Pm36*	18/6	0;	0;	0;	4	0;	0;	0;	0;	0;	0;	4	3	0	0;	0;	0;	0;	3	0;	3	0;	4	2	0;
GRY19	*Pm40*	16/8	3	1	0;	4	0;	0;	0;	0;	0;	0;	4	4	0;	0;	0;	0;	0;	4	4	4	0;	4	2	0;
Tabasco	*Pm46*	24/0	0	0	0	0	0	0;	0	0	0;	0;	1	0	0	0;	0;	0	0;	2	0;	0;	0	0;	0;	0;
CH7086	*Pm51*	21/3	0;	0;	0;	0;	0;	0;	0	0;	0;	0;	3	0;	1	0;	0;	0;	0	4	0;	0;	0	4	0;	2
Normandie	*Pm1+2+9*	3/21	4	3	3	4	3	3	3	3	0;	4	3	4	3	3	4	4	2	4	4	4	4	0;	3	4
Maris Huntsman	*Pm2+6*	21/3	0	0;	0	0;	0;	0;	0	0;	0;	0	3	3	0	0;	0;	0	1	3	0;	0;	0	1	0	0;
Maris Dove	*Pm2+Mld*	22/2	0;	0	0	0	0;	0;	0	0	0;	0	0	4	0	0;	0;	0	0;	4	0;	0;	0	0	0	0;
Mission	*Pm4b+5b*	18/6	3	0;	0;	4	0;	0;	0;	0;	0;	0;	0;	4	0;	0;	0;	0	0;	4	3	4	0;	0;	0;	0;
Bianmian 3	*Pm4+8*	17/7	2	0;	0;	4	0;	0;	0;	0;	0;	1	4	4	0;	0;	0;	0	0;	3	4	4	0;	3	0;	0;
Coker 983	*Pm5+6*	3/21	4	4	3	3	3	3	3	3	3	4	4	3	4	0;	2	4	3	4	4	2	3	3	3	3
Xiaobaidongmai	*Pm"XBD"*	15/9	0;	2	3	3	0;	4	3	0;	0;	3	0;	0;	0;	4	4	0;	0;	0;	4	0;	1	1	0;	4
NCV4	*Non-Pm1*	13/11	4	2	0	2	3	0;	0	2	3	3	4	3	3	0;	3	4	0;	0	2	0;	3	2	3	2
Chancellor	–	0/24	4	4	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
Funo	–	0/24	4	4	4	4	4	4	4	3	4	4	4	4	4	4	4	3	4	4	4	4	4	3	4	4
Era	–	15/9	1	0;	0	0;	0;	4	0;	0;	3	0;	4	4	0;	3	0;	0	0;	4	0;	4	0;	4	0;	4

Infection types 0-2 were resistant and 3-4 were susceptible as described by Sheng, 1988; Wang et al., 2005.

10.3389/fpls.2022.1048252

Polyploid Clustering v1.0 (Illumina, http://www.illumina.com). The monomorphic and poor quality SNP markers were excluded from further analyses. Then, the distribution of polymorphic SNPs between the resistant and susceptible DNA bulks on wheat chromosomes was analyzed. An enriched peak of differential SNPs was considered as the candidate interval of the *Pm* gene(s) in YAV249.

Molecular markers analysis

Based on the predicted interval of 660K SNP array scanning, 52 documented SSR molecular markers in the interval, including 40 SSR markers published on the GrainGenes website (http://wheat.pw.usda.gov) and 12 SSR markers developed by Ullah et al. (2018), were used to test the polymorphisms between the resistant and susceptible parents and bulks. The polymorphic markers between the parents and the bulks were used to genotype the $F_{2:3}$ families of YAV249/SX828. Then, the markers were aligned on the Chinese Spring reference genome sequence v1.0 (International Wheat Genome Sequencing Consortium (IWGSC), 2018) to confirm their physical locations.

PCR procedure was performed as previously described (Ma et al., 2015). The amplification products of marker BCD1231 were detected by agarose gel electrophoresis with a concentration of 1.5%, and stained with Super GelRed, while the remaining amplification products were separated in 8% non-denaturing polyacrylamide gels with 1×TBE buffer, and visualized by silver staining (Santos et al., 1993).

Map construction

A Chi squared (χ^2) test was carried out to investigate deviations of the observed phenotypic data of the $F_{2:3}$ families from the theoretically expected segregation ratios for analyzing the goodness-of-fit. Then, the linkage map of *PmYAV* was constructed using MAPMAKER 3.0 and the Kosambi function as reported previously (Kosambi, 1944; Lander et al., 1987).

Cloning and analysis of the *Pm4* homologous sequence in YAV249

Since *PmYAV* was mapped to the *Pm4* interval, the homologous sequence of YAV249 was cloned and assembled by using a homology-based cloning strategy according to the previous report on cloning *Pm4* (Sánchez-Martín et al., 2021). Then, the homologous sequence of YAV249 was compared with the cloned *Pm4* sequence.

RNA extraction and gene expression analysis

The wheat leaf samples of YAV249 were collected at 0, 2, 4, 8, 12, 24, 48 and 72 hours post inoculation (hpi) after inoculating *Bgt* isolate E09 and with three biological replicates. Total RNA was extracted using RNAiso Plus (TaKaRa, Shiga, Japan). For each sample, 2 μg of RNA was used for reverse transcription with a FastQuant RT Kit (Tiangen, Beijing, China). The transcript levels were detected by qRT-PCR on Bio-Rad CFX 96 with TB Green Premix Ex Taq™ II (TaKaRa, Shiga, Japan). The procedure included an initial step at 95°C for 30 s followed by 40 cycles of 95°C for 15 s, 60°C for 30 s, and 72°C for 10 s. The expression levels of target genes were normalized to that of *TaActin*. All qRT-PCR assays were performed in three independent replications.

Evaluation of the markers for MAS

The 192 wheat accessions including cultivars, landraces, advanced lines and introduced cultivars were tested by using the flanked or co-segregated markers of *PmYAV* and a specific marker of *Pm4* to evaluate their applicability for MAS breeding (Supplementary Table 1). If the polymorphic band(s) of a marker were same in both YAV249 and the tested accessions, it was considered not suitable for MAS. However, when the markers that amplified products in YAV249 that were different from the products in the tested accessions, this indicated that the marker could be used to detect *PmYAV* when it was transferred into these accessions and considered to be applicable for MAS breeding.

Application of *PmYAV* in wheat powdery mildew breeding

To transfer *PmYAV* into susceptible wheat cultivars to improve the resistance to powdery mildew, YAV249 was crossed and backcrossed with wheat cultivar KN199 that is widely grown in Huanghuai region (Zhang et al., 2011). Combined with identification of powdery mildew resistance and evaluation of agronomic traits, the flanked markers and co-segregated marker were used to trace *PmYAV* in the BC_1F_1, BC_2F_1, BC_3F_1, BC_3F_2 and BC_3F_3 populations.

The homozygous resistant progeny line YK13 derived from BC_3F_3 generation was selected to evaluate agronomic performance along with the parents YAV249 and KN199. These materials were planted at Luancheng Agro-ecosystem Experimental Station (37° 53′ 15″ N, 114° 40′ 47″ E) from

2019 to 2021 with three replicates. Each line was planted in five rows (1.5 m length and 0.25 m between rows) and 20 seeds per row. For each parent/line, 10 plants in the middle three internal rows were randomly sampled to investigate the plant height (PH), spike numbers per plant (SNPP), spike length (SL), spikelet numbers per spike (SNS), sterile spikelet numbers per spike (SSNS), kernel numbers per spike (KNS), thousand-kernel weight (TKW) and kernel related traits. PH and SNPP were assessed based on the mean of 10 plants. SL, SNS, SSNS and KNS were identified based on the mean of the main spike of 10 plants. TKW and kernel related traits (kernel length (KL), kernel width (KW), kernel length/kernel width (KL/W)) of at least 500 kernels randomly were measured three times using the rapid SC-G grain appearance quality image analysis system (WSeen Detection, Hangzhou, China). Analysis of variance (ANOVA) and least significant difference (LSD) test were performed with SPSS statistics v20.0 software (SPSS, Chicago, USA) to test the significance of differences between YK13 and its parents YAV249 and KN199 for each agronomic trait.

Results

Inheritance of powdery mildew resistance in YAV249

From 2019 to 2021, YAV249 was highly resistant with ITs 0-1 to the *Bgt* isolate mixture of E09, E11, E18 and E20 at the adult stage. When inoculated with *Bgt* isolate E09 at the seedling stage, YAV249 was highly resistant with IT 0, while SX828 was highly susceptible with IT 4 (Figure 1A). The results of microscopic observation indicated that many mycelia were produced in SX828 but not in YAV249 (Figure 1B). Twenty F_1 plants of the cross YAV249/SX828 were resistant with IT 0-1, indicating that the resistance of YAV249 to *Bgt* isolate E09 was controlled by dominant *Pm* gene(s). The F_2 population from the cross of YAV249/SX828 segregated with 231 resistant and 88 susceptible plants, which fit the theoretical ratio of 3:1 for monogenic segregation of a dominant gene (χ^2 = 1.14; P = 0.29). All 319 F_2 plants were transplanted in the field to produce $F_{2:3}$ families. The 250 $F_{2:3}$ families segregated with 64 homozygous resistant (RR), 119 segregating (Rr) and 67 homozygous susceptible (rr). The result further confirmed the ratio of monogenic inheritance of a single dominant powdery mildew resistance gene (χ^2 = 0.65; P = 0.72). Therefore, it was concluded that the resistance to *Bgt* isolate E09 in YAV249 was controlled by a single dominant gene, tentatively designated *PmYAV*.

Bulked segregant analysis with wheat 660K SNP array

To map *PmYAV*, the resistant and susceptible DNA bulks were genotyped with the 660K SNP array. A total of 1,901 SNPs which was distributed across 21 wheat chromosomes were homozygous polymorphic between the two bulks. Among them, 1,653 SNPs were enriched on chromosome 2A, and 248 SNPs were distributed on other chromosomes (Figure 2A and Supplementary Table 2). Of the 1,653 SNPs located on chromosome 2A, 1,179 SNPs were enriched in an interval of 29 Mb (734 Mb to 763 Mb) on chromosome 2AL (Figure 2B), indicating that *PmYAV* was mapped in or near this region.

Molecular mapping of *PmYAV*

Based on the mapping interval of the 660K SNP array, 52 SSR molecular markers located on chromosome 2AL were used to screen between resistant parent YAV249, susceptible parent SX828, resistant bulk and susceptible bulk. Among them, 14 markers PSP3039, GPW4474, GWM356, GWM311, GWM382, GDM93, GWM526, GPW4456, BCD1231, WGRC763, WGRC872, WGRC1096, WGRC869 and WGRC883 showed consistent polymorphism between parents and bulks. Subsequently, these markers were used to genotype the 250 $F_{2:3}$ families and confirm linkage with *PmYAV*. The result

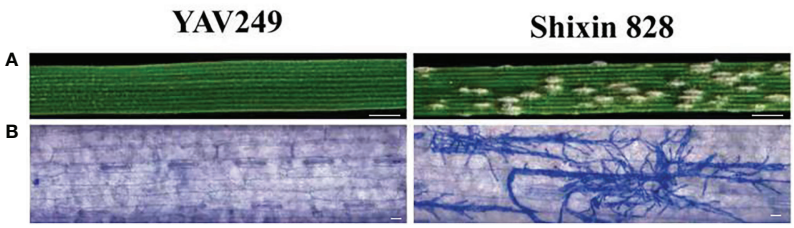


FIGURE 1
Macroscopic and microscopic characterization of YAV249 and Shixin 828 after inoculating isolate E09 of *Blumeria graminis* (f) sp. *tritici* (*Bgt*).
(A) Macroscopic view of the infected representative leaf segments from YAV249 and Shixin 828 at 10 d post-inoculation (dpi). Bar, 5 mm.
(B) Microscopic view of leaves from YAV249 and Shixin 828 at 7 dpi to visualize fungal mycelia. Bar, 100 μm.

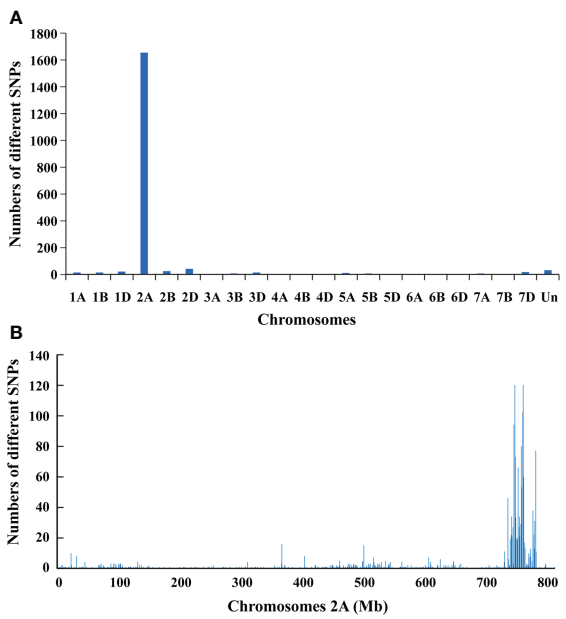


FIGURE 2
(A) Distribution of polymorphic SNP loci between the resistant and susceptible DNA bulks from the $F_{2:3}$ families of YAV249/Shixin 828 on 21 chromosomes. **(B)** Density of polymorphism SNP loci on chromosome 2A with 1 Mb intervals.

demonstrated that *PmYAV* co-segregated with five markers *Xwgrc883*, *Xwgrc969*, *Xwgrc1096*, *Xwgrc872* and *Xbcd1231* and was flanked by *Xwgrc763* and *Xgdm93* with genetic distances 0.8 cM and 1.2 cM corresponding to a physic interval of 1.89 Mb (760.58-762.47 Mb) on the Chinese Spring reference genome sequence v1.0 (Figure 3). There were four reported and formally designated *Pm* genes including *Pm4*, *Pm33*, *Pm50* and *Pm65* on chromosome 2AL (Li et al., 2019), and *PmYAV* mapped closed to *Pm4*.

Cloning and analysis of the *Pm4* homologous sequence in YAV249

To further explore the relationship between *PmYAV* and *Pm4*, we genotyped all the 250 $F_{2:3}$ families from YAV249/SX828 with the functional molecular marker JS717/JS718 of *Pm4*. JS717/JS718 showed positive amplification in YAV249 and co-segregated with *PmYAV* in the $F_{2:3}$ families. Then, the *Pm4* homologous coding sequence in YAV249 was cloned using nested PCR, and the result showed that *Pm4* homologous sequence in YAV249 was totally consistent with the *Pm4a* sequence. However, *PmYAV* generated an extra splicing event compared to the previously reported *Pm4a_V1* and *Pm4a_V2* (Sánchez-Martín et al., 2021), suggesting that there was an extra 67 bp-intron in *PmYAV*. The intron was located on the second exon in the serine/threonine kinase domain of *Pm4a_V1* (Figure 4).

Gene expression analysis of *PmYAV*

To investigate the expression patterns of *PmYAV*, we performed qRT-PCR analysis. *Pm4a_V1* and *Pm4a_V2* in YAV249 showed similar expression patterns and were upregulated after *Bgt* isolate E09 infection, peaked at 24 hpi, and then reduced at later stages between 48 and 72 hpi. The level of *Pm4a_V2* was higher significantly than that of *Pm4a_V1* at most of the time points (Figure 5). We speculated that *Pm4a_V2* had a greater contribution than *Pm4a_V1* to resistance in YAV249.

The difference of powdery mildew resistance spectrum between YAV249 and documented *Pm4a* stock

When inoculated with 24 *Bgt* isolates, *Pm4a* was resistant to 14 out of 24 *Bgt* isolates, including E01, E05, E06, E09, E11, E13, E16, E17, E21, E23-1, E23-2, E26, E30-1 and E49, while *PmYAV* was also resistant to E60 and E69 besides those 14 *Bgt* isolates (Table 1). Therefore, *PmYAV* showed a different reaction to two *Bgt* isolates compared to *Pm4a*.

Evaluation of the markers of *PmYAV* for MAS

To investigate the applicability of the markers linked or co-segregated with *PmYAV* in MAS, the flanked markers GDM93 and WGRC763 and the co-segregating markers BCD1231 and JS717/JS718 of *PmYAV* were used to test 192 wheat accessions (Supplementary Table 1). Markers GDM93 and WGRC763 amplified specific bands that differed from YAV249 in 147, and 140 out of 192 accessions respectively. The same as JS717/JS718, BCD1231 amplified different bands between YAV249 and 137 out of 192 accessions. The results indicated that these four markers could be used in MAS for detecting *PmYAV* in those wheat genetic backgrounds (Figure 6 and Supplementary Table 1). However, this was not the case for all the marker alleles in wheat cultivars Aikang 58, Xinmai 1998, ROANE, Safha 3 and Taixue 12 due to the same bands between YAV249 and them (Figure 6 and Supplementary Table 1).

Application of *PmYAV* in wheat powdery mildew breeding

To develop a resistance pre-breeding resource, we transferred *PmYAV* into the commercial susceptible cultivar KN199 by MAS, combined with evaluation of powdery mildew resistance and analysis of agronomic traits. An advanced

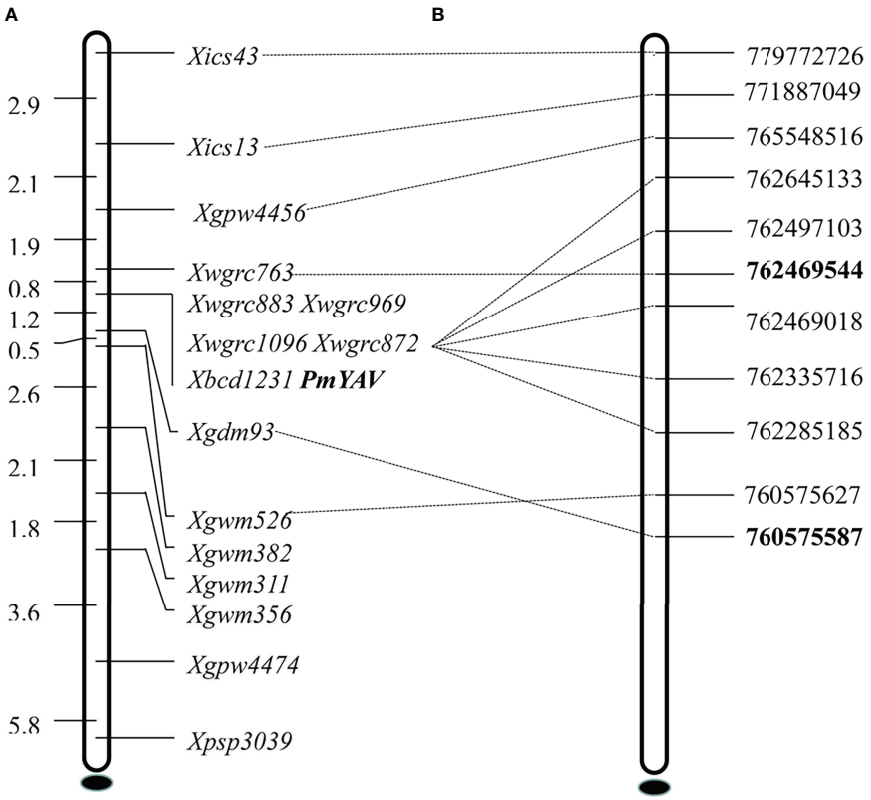


FIGURE 3
Genetic linkage map **(A)** and physical map **(B)** of powdery mildew resistance gene *PmYAV* using the $F_{2:3}$ families of YAV249/Shixin 828.

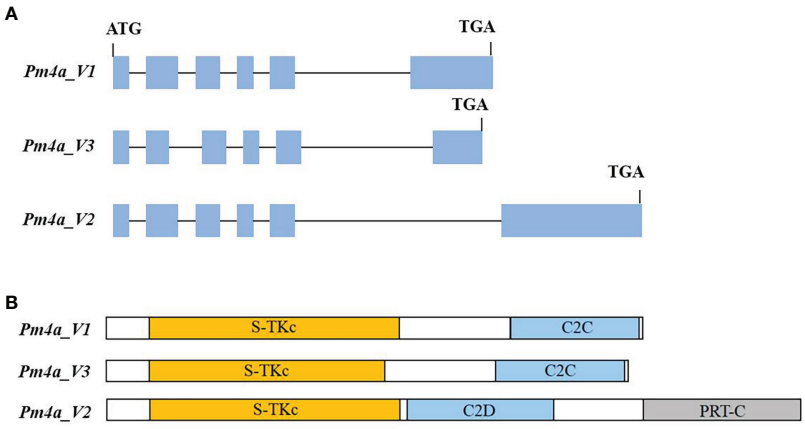


FIGURE 4
(A) Schematic representation of gene structures of *PmYAV* splicing variants. Gene exons and introns are shown by blue boxes and black lines, respectively. **(B)** *Pm4a_V1*, *Pm4a_V2* and *Pm4a_V3* protein isoforms with domains indicated by colours: yellow, serine/threonine kinase (S-TKc); light blue, C2; grey, phosphoribosyl transferase C-terminal (PRT-C).

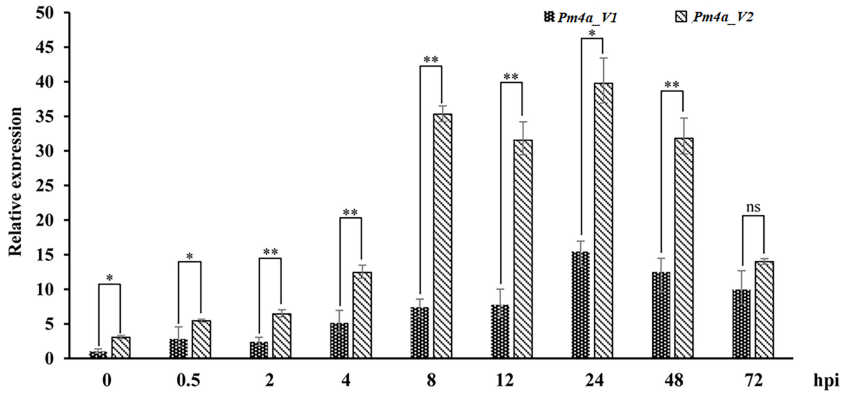


FIGURE 5
Expression patterns of the *Pm4a_V1* and *Pm4a_V2* splice variants in YAV249 during a 72-hour time course. Error bars represent SD based on three independent repeats. Asterisks indicate significant differences (*t*-tests) between *Pm4a_V1* and *Pm4a_V2* at each time point (*$P < 0.05$, **$P < 0.01$, ns: not significant). *TaActin* was used as the internal control.

breeding line YK13 was selected from a BC_3F_3 population of YAV249/4*KN199 due to its desirable agronomics traits and resistance to powdery mildew. During 2019-2021, analysis of agronomic traits indicated that the *PmYAV* donor parent YAV249 had a greater SL, TKW and KL but a poorer PH than the recurrent parent KN199. Compared with KN199, YK13 showed improved SL, TKW and KL in one or two years. Moreover, there was no significant difference on PH, SNPP, SNS, KNS and KW between YK13 and its recurrent parent KN199 (Table 2).

Discussion

SHW lines possess important resistance genes effective against various diseases, therefore, regular screening of SHW genotypes is important for identifying new sources of resistance to powdery mildew that could eventually be used in wheat disease resistance breeding programs. In this study, the SHW line YAV249 derived from the cross YAV-2/TEZ//*Ae. squarrosa* 249 was identified to be resistant to powdery mildew at all stages. Genetic analysis indicated that its resistance was conferred by a

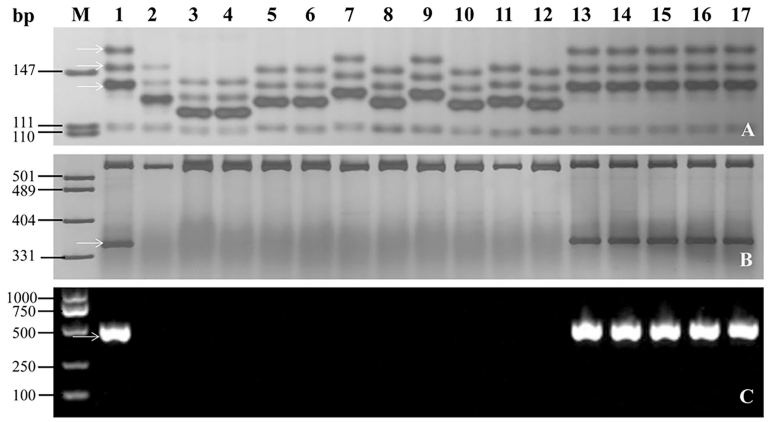


FIGURE 6
Amplification patterns of the flanked markers GDM93 **(A)** and WGRC763 **(B)** and the co-segregated marker BCD1231 **(C)** of *PmYAV* in YAV249, Shixin 828 and selected wheat cultivars/lines. M: pUC19/*Msp*I; 1: YAV249; 2: Shixin 828; 3: Shi 4185; 4: Shiyou 17; 5: Shimai 18; 6: Shixin 633; 7: Ji 5265; 8: Kenong 199; 9: Zhengyumai 9989; 10: Zhou 16; 11: Jinmai 41; 12: Gaoyou 503; 13: Xinmai 1998; 14: ROANE; 15: Safha 3; 16: Aikang 58; 17: Taixue 12. The white arrows indicate the specific amplified fragments.

TABLE 2 The agronomic trait evaluation of the advanced breeding line YK13 with its donor parent YAV249 and recurrent parent Kenong 199.

Year	Cultivars/Lines	PH (cm)	SL (cm)	SNPP	SNS	KNS	TKW (g)	KL (mm)	KW (mm)
2019-2020	YAV249	92.2 ± 1.7^{a}	13.0 ± 0.2^{a}	16.6 ± 1.1^{a}	19.0 ± 0.5^{a}	49.0 ± 3.1^{a}	38.4 ± 1.2^{a}	8.2 ± 0.1^{a}	3.0 ± 0.1^{a}
	Kenong 199	73.2 ± 1.5^{b}	8.0 ± 0.1^{c}	18.4 ± 1.2^{a}	21.0 ± 1.1^{a}	46.0 ± 2.7^{a}	37.2 ± 2.0^{a}	5.9 ± 0.1^{c}	3.3 ± 0.1^{a}
	YK13	73.8 ± 1.8^{b}	9.0 ± 0.1^{b}	17.6 ± 0.8^{a}	20.6 ± 1.5^{a}	52.0 ± 3.5^{a}	37.9 ± 1.8^{a}	6.7 ± 0.1^{b}	3.2 ± 0.1^{a}
2020-2021	YAV249	94.0 ± 1.4^{a}	12.7 ± 0.3^{a}	12.5 ± 0.7^{a}	19.5 ± 0.7^{a}	65.5 ± 7.1^{a}	53.9 ± 3.0^{a}	8.6 ± 0.1^{a}	3.4 ± 0.1^{a}
	Kenong 199	75.8 ± 1.5^{b}	8.1 ± 0.3^{c}	10.8 ± 1.3^{a}	20.8 ± 1.0^{a}	63.5 ± 0.7^{a}	43.0 ± 2.1^{c}	6.0 ± 0.1^{c}	3.5 ± 0.1^{a}
	YK13	75.3 ± 1.9^{b}	8.8 ± 0.5^{b}	10.8 ± 1.7^{a}	21.5 ± 1.7^{a}	77.0 ± 11.7^{a}	48.5 ± 2.0^{b}	6.8 ± 0.1^{b}	3.5 ± 0.1^{a}

Values with the same letters in the same column were not significantly different at the $P < 0.05$ according to the least significant difference test. PH, plant height; SL, spike length; SNPP, spike numbers per plant; SNS, spikelet numbers per spike; KNS, kernel numbers per spike; TKW, thousand-kernel weight; KL, kernel length; KW: kernel width.

single dominant gene, temporarily designated *PmYAV*. Bulked segregant analysis by using wheat 660K SNP array scanning and marker analysis showed that *PmYAV* was mapped in the *Pm4* interval on chromosome 2AL.

Up to now, eight alleles (*Pm4a-Pm4h*) have been reported at the *Pm4* locus (Sánchez-Martín et al., 2021), which makes it multiple allelic loci as previously known *Pm1* (Hsam et al., 1998), *Pm2* (Jin et al., 2018), *Pm3* (Bhullar et al., 2010) and *Pm5* (Xie et al., 2020) loci. *Pm4a* was derived from tetraploid wheat (*T. turgidum*) cultivars Yuma and Khapli (Briggle, 1966). However, *Pm4b* was originally reported in three bread wheat cultivars Weihenstephan M1, TP29 and ELS (Wolfe 1965; The et al., 1979). Subsequently, *Pm4b* was postulated in French cultivar VPM1 (Bariana and McIntosh, 1994). Three alleles at *Pm4* locus, including *Pm4c* (*Pm4c* = *Pm23*), *Pm4d*, and *Pm4e*, were identified in wheat line 81-7241, Tm27d2, and landrace D29, respectively (Hao et al., 2008; Schmolke et al., 2012; Li et al., 2017). *Pm4f*, *Pm4g* and *Pm4h* which were susceptible to most of the tested *Bgt* isolates were discovered by using the haplotype-specific marker JS717/JS718 in a global wheat collection of 512 accessions (Sánchez-Martín et al., 2021). In this study, *PmYAV* was resistant to 16 of 24 *Bgt* isolates at the seedling stage which was different from the tested *Pm4a* and *Pm4b* (Table 1). Meanwhile, the donor of *PmYAV* was also different from those of *Pm4a-Pm4h*. Therefore, *PmYAV* might be a new allele at the *Pm4* locus. The identification of *PmYAV* enriched the genetic diversity of *Pm4* locus, which was available for increasing the durable resistance conferred by *Pm4* locus in wheat powdery mildew resistance breeding since it has a resistance spectrum different from that of *Pm4a*, *Pm4b* and *Pm4c*, but unknown with that of *Pm4d* to *Pm4h* which were not tested in the phenotyping study.

Recently, *Pm4b*, which consisted of seven exons and contained two alternative transcripts, was cloned by using the strategy of MutChromSeq (Sánchez-Martín et al., 2021). Subsequently, all the alleles at *Pm4* locus were cloned with the full-length amplification and Sanger sequencing and indicated *Pm4* alleles contain SNPs and/or combinations of shared SNPs which have main effects on the kinase domain. What is even more interesting is that most of the SNP cause amino acid changes in the transmembrane and S_TKc domains. *Pm4* generated two isoforms which were both required for resistance (Sánchez-Martín et al., 2021). In this study, using homology-based cloning, we found that the *Pm4* homologous sequence in YAV249 was consistent with *Pm4a* sequence, while *PmYAV* generated an extra splicing event for the reported *Pm4a_V1* and *Pm4a_V2*, suggesting that there was an extra 67 bp-intron in *PmYAV* (Figure 4). It may cause different resistance reactions between *PmYAV* and other alleles. Then, the function of the extra splicing needs to be further investigated in future by other methods, such as virus-induced gene silencing and transgenic assay analysis. On the other hand, Sánchez-Martín et al. (2021) reported that the expression of *Pm4b_V1* and *Pm4b_V2* on the wild-type *Pm4b* wheat genotype Fed-Pm4b after infection with powdery mildew did not significantly differ from each other. Similar expression levels of both transcripts indicate that *Pm4b_V1* and *Pm4b_V2* have a similar contribution to powdery mildew resistance. However, in this study, when inoculated *Bgt* isolate E09, the expression level of *Pm4b_V2* in YAV249 was higher significantly than that of *Pm4b_V1* at most of the time points (Figure 5) which might cause the different resistance spectra of *PmYAV* from other alleles at *Pm4* locus.

When a valuable resistance gene or accession was identified, its rational utilization was the first factor to be considered in wheat disease resistance breeding. MAS is a high priority in wheat breeding programs to efficiently transfer *R* gene to susceptible cultivars or pyramid with other *R* gene(s). Sánchez-Martín et al. (2021) reported a dominant diagnostic marker JS717/JS718 which could detect the *Pm4* in different genetic background. In this study, YAV249 showed a high level of resistance to powdery mildew and possessed desirable agronomic traits, such as higher TKW and KL/KW. Two dominant markers WGRC763 closely linked to *PmYAV* and BCD1231 and JS717/718 co-segregated with *PmYAV* and can be used for MAS in many genetic backgrounds. More importantly, the co-dominant marker GDM93 which was closely linked to *PmYAV* could be used for selection of homozygotes during MAS breeding. In fact, an advanced breeding line YK13, which possessed not only desirable agronomic traits but excellent resistance to powdery mildew, was selected in BC_3F_3 population of YAV249/KN199 by MAS.

SHW is a valuable genetic resource for improving disease resistance in wheat cultivars. In our study, we identified an extra splicing variant of *Pm4a* from SHW line YAV249 and selected a

new wheat breeding line YK13 with high resistance to powdery mildew and excellent agronomic traits. The results will add diversity to the available genes for powdery mildew resistance and provide the foundation for analyzing alternative splicing mechanism in regulating the *Pm4* gene function.

Conclusion

To sum up, in the present study, we identified a dominant powdery mildew resistance gene *PmYAV* in synthetic hexaploid wheat line YAV249 using bulked segregant analysis with wheat 660K single nucleotide polymorphism (SNP) array scanning and marker analysis. Sequence alignment analysis demonstrated that the sequence of *PmYAV* was totally consistent with that of *Pm4a* but generated an extra splicing. The flanked markers GDM93 and WGRC763, and the co-segregated markers BCD1231 and JS717/JS718 were confirmed to be easily performed in marker-assisted selection (MAS) of *PmYAV*. Using MAS strategy, *PmYAV* was transferred into the commercial cultivar Kenong 199 (KN199) and a wheat line YK13 was selected in BC_3F_3 from the population of YAV249/KN199 due to its excellent agronomic traits and resistance to powdery mildew. Our study can be valuable for enhancing the genetic diversity of powdery mildew resistance in breeding.

Data availability statement

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

Author contributions

HZ and DA conceived the research. YJ, TG, XL, GH and ZS performed the experiments. TG, XL, YZ, JF, JW and HL developed the experimental materials. YZ, JF, JW and WL performed the phenotypic assessment. YJ and TG wrote the manuscript. DA and HZ supervised and revised the writing of the article. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by the National Key Research and Development Program of China No. 2021YFD1200600, and the National Natural Science Foundation of China No. 32272105.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1048252/full#supplementary-material

SUPPLEMENTARY TABLE 1
Validation of the flanked markers GDM93 and WGRC763 and the co-segregated markers BCD1231 and JS717/JS718 of *PmYAV* on 192 wheat accessions in marker-assisted selection (MAS). '+' amplified marker allele differs from that in YAV249, indicating that the marker might be useful in MAS, and '–' marker is similar to that in YAV249, indicating that the marker could not be used in MAS.

References

An, D. G., Ma, P. T., Zheng, Q., Fu, S. L., Li, L. H., Han, F. P., et al. (2019). Development and molecular cytogenetic identification of a new wheat-rye 4R chromosome disomic addition line with resistances to powdery mildew, stripe rust and sharp eyespot. *Theor. Appl. Genet.* 132, 257–272. doi: 10.1007/s00122-018-3214-3

Bariana, H. S., and McIntosh, R. A. (1994). Characterization and origin of rust and powdery mildew resistance genes in VPM1 wheat. *Euphytica* 76, 53–61. doi: 10.1007/bf00024020

Bhullar, N. K., Zhang, Z. Q., Wicker, T., and Keller, B. (2010). Wheat gene banks as a source of new alleles of the powdery mildew resistance gene *Pm3*: a large scale allele mining project. *BMC Plant Biol.* 10, 88. doi: 10.1186/1471-2229-10-88

Briggle, L. W. (1966). Transfer of resistance to *Erysiphe graminis* f. sp. *tritici* from khapli emmer and yuma durum to hexaploid wheat. *Crop Sci.* 6, 459–461. doi: 10.2135/cropsci1966.0011183x000600050020x

Brunner, S., Hurni, S., Herren, G., Kalinina, O., von Burg, S., Zeller, S. L., et al. (2011). Transgenic *Pm3b* wheat lines show resistance to powdery mildew in the field. *Plant Biotechnol. J.* 9, 897–910. doi: 10.1111/j.1467-7652.2011.00603.x

Das, M. K., Bai, G., Mujeeb-Kazi, A., and Rajaram, S. (2016). Genetic diversity among synthetic hexaploid wheat accessions (*Triticum aestivum*) with resistance to several fungal diseases. *Genet. Resour. Crop Evol.* 63, 1285–1296. doi: 10.1007/s10722-015-0312-9

Hao, Y. F., Liu, A. F., Wang, Y. H., Feng, D. S., Gao, J. R., Li, X. F., et al. (2008). *Pm23*: a new allele of *Pm4* located on chromosome 2AL in wheat. *Theor. Appl. Genet.* 117, 1205–1212. doi: 10.1007/s00122-008-0827-y

He, H. G., Liu, R. K., Ma, P. T., Du, H. N., Zhang, H. H., Wu, Q. H., et al. (2020). Characterization of *Pm68*, a new powdery mildew resistance gene on chromosome 2BS of Greek durum wheat TRI 1796. *Theor. Appl. Genet.* 134, 53–62. doi: 10.1007/s00122-020-03681-2

Hewitt, T., Müller, M. C., Molnár, I., Mascher, M., Holušová, K., Šimková, H., et al. (2021). A highly differentiated region of wheat chromosome 7AL encodes a *Pm1a* immune receptor that recognizes its corresponding *AvrPm1a* effector from *Blumeria graminis*. *New Phytol.* 229, 2812–2826. doi: 10.1111/nph.17075

Hsam, S. L. K., Huang, X. Q., Ernst, F., Hartl, L., and Zeller, F. J. (1998). Chromosomal location of genes for resistance to powdery mildew in common wheat (*Triticum aestivum* l. em thell.). 5. alleles at the *Pm1* locus. *Theor. Appl. Genet.* 96, 1129–1134. doi: 10.1007/s001220050848

Hurni, S., Brunner, S., Buchmann, G., Herren, G., Jordan, T., Krukowski, P., et al. (2013). Rye *Pm8* and wheat *Pm3* are orthologous genes and show evolutionary conservation of resistance function against powdery mildew. *Plant J.* 76, 957–969. doi: 10.1111/tpj.12345

Hu, Y. K., Xin, Z. Y., Chen, X., Zhang, Z. Y., and Duan, X. Y. (2001). Genetic analysis and gene deduction of powdery mildew resistance in *T. durum-ae. squarrosa* amphiploid. *Acta Genet Sin.* 28, 152–157. doi: 10.1088/0256-307x/18/11/313

International Wheat Genome Sequencing Consortium (IWGSC) (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361, eaar7191. doi: 10.1126/science.aar7191

Jin, Y. L., Shi, F. Y., Liu, W. H., Fu, X. Y., Gu, T. T., Han, G. H., et al. (2021). Identification of resistant germplasm and detection of genes for resistance to powdery mildew and leaf rust from 2,978 wheat accessions. *Plant Dis.* 105, 3900–3908. doi: 10.1094/pdis-03-21-0532-re

Jin, Y. L., Xu, H. X., Ma, P. T., Fu, X. Y., Song, L. P., Xu, Y. F., et al. (2018). Characterization of a new *Pm2* allele associated with broad-spectrum powdery mildew resistance in wheat line subtil. *Sci. Rep.* 8, 475. doi: 10.1038/s41598-017-18827-4

Kamali, S., Majidi, E., Khodarahmi, M., and Mehrabi, R. (2014). Evaluation of CIMMYT synthetic hexaploid wheats for resistance to septoria tritici blotch. *Crop Breed. J.* 4, 23–33. doi: 10.22092/cbj.2014.109669

Koller, T., Brunner, S., Herren, G., Hurni, S., and Keller, B. (2018). Pyramiding of transgenic *Pm3* alleles in wheat results in improved powdery mildew resistance in the field. *Theor. Appl. Genet.* 131, 861–871. doi: 10.1007/s00122-017-3043-9

Kosambi, D. D. (1944). The estimation of map distances from recombination values. *Ann. Eugen.* 12, 172–175. doi: 10.1007/978-81-322-3676-4_16

Krattinger, S. G., Lagudah, E. S., Spielmeyer, W., Singh, R. P., Huerta-Espino, J., McFadden, H., et al. (2009). A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. *Science* 323, 1360–1363. doi: 10.1126/science.1166453

Lander, E. S., Green, P., Abrahamson, J., Barlow, A., Daley, M., Lincoln, S., et al. (1987). Mapmaker: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* 1, 174–181. doi: 10.1016/0888-7543(87)90010-3

Li, G. Q., Cowger, C., Wang, X. W., Carver, B. F., and Xu, X. Y. (2019). Characterization of *Pm65*, a new powdery mildew resistance gene on chromosome 2AL of a facultative wheat cultivar. *Theor. Appl. Genet.* 132, 2625–2632. doi: 10.1007/s00122-019-03377-2

Li, M. M., Dong, L. L., Li, B. B., Wang, Z. Z., Xie, J. Z., Qiu, D., et al. (2020). A CNL protein in wild emmer wheat confers powdery mildew resistance. *New Phytol.* 228, 1027–1037. doi: 10.1111/nph.16761

Li, N., Jia, H. Y., Kong, Z. X., Tang, W. B., Ding, Y. X., Liang, J. C., et al. (2017). Identification and marker-assisted transfer of a new powdery mildew resistance gene at the *Pm4* locus in common wheat. *Mol. Breed.* 37, 79. doi: 10.1007/s11032-017-0670-4

Li, A., Liu, D., Yang, W., Kishii, M., and Mao, L. (2018). Synthetic hexaploid wheat: yesterday, today, and tomorrow. *Engineering* 4, 552–558. doi: 10.1016/j.eng.2018.07.001

Li, J., Wei, H. T., Hu, X. R., Li, C. S., Tang, Y. L., Liu, D. C., et al. (2011). Identification of a high-yield introgression locus from synthetic hexaploid wheat in chuanmai 42. *Acta Agronom. Sin.* 37, 255–262. doi: 10.3724/sp.j.1006.2011.00263

Lu, P., Guo, L., Wang, Z. Z., Li, B. B., Li, J., Li, Y. H., et al. (2020). A rare gain of function mutation in a wheat tandem kinase confers resistance to powdery mildew. *Nat. Commun.* 11, 680. doi: 10.1038/s41467-020-14294-0

Lutz, J., Hsam, S., Limpert, E., and Zeller, F. J. (1995). Chromosomal location of powdery mildew resistance genes in *Triticum aestivum* l. (common wheat). 2. genes *Pm2* and *Pm19* from *Aegilops squarrosa* l. *Heredity* 74, 152–156. doi: 10.1038/hdy.1995.22

Ma, P. T., Xu, H. X., Xu, Y. F., Li, L. H., Qie, Y. M., Luo, Q. L., et al. (2015). Molecular mapping of a new powdery mildew resistance gene *Pm2b* in Chinese breeding line KM2939. *Theor. Appl. Genet.* 128, 613–622. doi: 10.1007/s00122-015-2457-5

McIntosh, R. A., Dubcovsky, J., Rogers, W. J., Xia, X. C., and Raupp, W. J. (2020). *Catalogue of gene symbols for wheat: 2020 supplement*. Available at: https://wheat.pw.usda.gov/GG3/WGC.

Moore, J. W., Herrera-Foessel, S., Lan, C. X., Schnippenkoetter, W., Ayliffe, M., Huerta-Espino, J., et al. (2015). A recently evolved hexose transporter variant confers resistance to multiple pathogens in wheat. *Nat. Genet.* 47, 1494–1498. doi: 10.1038/ng.3439

Qie, Y. M., Sheng, Y., Xu, H. X., Jin, Y. L., Ma, F. F., Li, L. H., et al. (2019). Identification of a new powdery mildew resistance gene *PmDHT* at or closely linked to the *Pm5* locus in the Chinese wheat landrace dahongtou. *Plant Dis.* 103, 2645–2651. doi: 10.1094/pdis-02-19-0401-re

Sánchez-Martín, J., Steuernagel, B., Ghosh, S., Herren, G., Hurni, S., Adamski, N., et al. (2016). Rapid gene isolation in barley and wheat by mutant chromosome sequencing. *Genome Biol.* 17, 221. doi: 10.1186/s13059-016-1082-1

Sánchez-Martín, J., Widrig, V., Herren, G., Wicker, T., Zbinden, H., Gronnier, J., et al. (2021). Wheat *Pm4* resistance to powdery mildew is controlled by alternative splice variants encoding chimeric proteins. *Nat. Plants.* 7, 327–341. doi: 10.1038/s41477-021-00869-2

Santos, F. R., Pena, S. D. J., and Epplen, J. T. (1993). Genetic and population study of a y-linked tetranucleotide repeat DNA polymorphism with a simple non-isotopic technique. *Hum. Genet.* 90, 655–656. doi: 10.1007/bf00202486

Schmolke, M., Mohler, V., Hartl, L., Zeller, F. J., and Hsam, S. L. K. (2012). A new powdery mildew resistance allele at the *Pm4* wheat locus transferred from einkorn (*Triticum monococcum*). *Mol. Breed.* 29, 449–456. doi: 10.1007/s11032-011-9561-2

Sharp, P. J., Kreis, M., Shewry, P. R., and Gale, M. D. (1988). Location of β-amylase sequences in wheat and its relatives. *Theor. Appl. Genet.* 75, 286–290. doi: 10.1007/bf00303966

Sheng, B. Q. (1988). Wheat powdery mildew was recorded using infection type in seedling stage. *Plant Prot.* 14, 49. Available at: https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CJFD&dbname=CJFD8589&filename=ZWBH198801035&uniplatform=NZKPT&v=EleQOgszeUXu-PqypFjom3Acylv32Ikwf0r2_3oTXJpIMQ4n8xGbEPM7PZ-6Sx1q

Sheng, B. Q., and Duan, X. Y. (1991). Improvement of scale 0-9 method for scoring adult plant resistance to powdery mildew of wheat. Beijing. *Agr. Sci.* 1, 38–39. Available at: https://kns.cnki.net/kcms/detail/detail.aspx?filename=BJNK199101011&dbcode=CJFD&dbname=CJFD1991&v=Rxg8UkEIGV3B5TxBLD9dSq7XlRDsCXbtegphAfhso0TfBjG8a6XZCKAo6ZUWwcFN

Singh, S. P., Hurni, S., Ruinelli, M., Brunner, S., Sanchez-Martin, J., Krukowski, P., et al. (2018). Evolutionary divergence of the rye *Pm17* and *Pm8* resistance genes reveals ancient diversity. *Plant Mol. Biol.* 98, 249–260. doi: 10.1007/s11103-018-0780-3

Srichumpa, P., Brunner, S., Keller, B., and Yahiaoui, N. (2005). Allelic series of four powdery mildew resistance genes at the *Pm3* locus in hexaploid bread wheat. *Plant Physiol.* 139, 885–895. doi: 10.1104/pp.105.062406

The, T. T., McIntosh, R. A., and Bennett, F. G. A. (1979). Cytogenetical studies in wheat. IX. monosomic analyses, telocentric mapping and linkage relationships of genes *Sr21*, *Pm4*, and *Mle*. *Aust. J. Biol. Sci.* 32, 115–125. doi: 10.1071/bi9790115

Trethowan, R. M., and Mujeeb-Kazi, A. (2008). Novel germplasm resources for improving environmental stress tolerance of hexaploid wheat. *Crop Sci.* 48, 1255–1265. doi: 10.2135/cropsci2007.08.0477

Ullah, K. N., Li, N., Shen, T., Wang, P. S., Tang, W. B., Ma, S. W., et al. (2018). Fine mapping of powdery mildew resistance gene *Pm4e* in bread wheat (*Triticum aestivum* l.). *Planta* 248, 1329–1329. doi: 10.1007/s00425-018-2990-y

Wang, Y. P., Cheng, X., Shan, Q. W., Zhang, Y., Liu, J. X., Gao, C. X., et al. (2014). Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat. Biotechnol.* 32, 947–951. doi: 10.1038/nbt.2969

Wang, Z. L., Li, L. H., He, Z. H., Duan, X. Y., Zhou, Y. L., Chen, X. M., et al. (2005). Seedling and adult plant resistance to powdery mildew in Chinese bread wheat cultivars and lines. *Plant Dis.* 89, 457–463. doi: 10.1094/pd-89-0457

Wolfe, M. S. (1965). Physiologic specialization of *Erysiphe graminis* f. sp. *tritici* in the united kingdom. *Trans. Bril. Mycol. Soc* 48, 315–326. doi: 10.1016/S0007-1536(65)80099-7

Xie, J. Z., Guo, G. H., Wang, Y., Hu, T. Z., Wang, L. L., Li, J. T., et al. (2020). A rare single nucleotide variant in *Pm5e* confers powdery mildew resistance in common wheat. *New Phytol.* 228, 1011–1026. doi: 10.1111/nph.16762

Xing, L. P., Hu, P., Liu, J. Q., Witek, K., Zhou, S., Xu, J. F., et al. (2018). *Pm21* from *Haynaldia villosa* encodes a CC-NBS- LRR protein conferring powdery mildew resistance in wheat. *Mol. Plant* 11, 874–878. doi: 10.1016/j.molp.2018.02.013

Zhang, W., Wang, J., Ji, J., Wang, Z. G., An, D. G., Zhang, X. Q., et al. (2011).). development of "Kn199" new winter wheat variety and its cultivation in China. *Chin. J. Eco. Agric.* 19, 1215–1219. doi: 10.3724/sp.j.1011.2011.01215

Zhou, R. H., Zhu, Z. D., Kong, X. Y., Huo, N. X., Tian, Q. Z., Li, P., et al. (2005). Development of wheat near-isogenic lines for powdery mildew resistance. *Theor. Appl. Genet.* 110, 640–648. doi: 10.1007/s00122-004-1889-0

Zou, S. H., Wang, H., Li, Y. W., Kong, Z. S., and Tang, D. Z. (2018). The NB-LRR gene *Pm60* confers powdery mildew resistance in wheat. *New Phytol.* 218, 298–309. doi: 10.1111/nph.14964

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 14 November 2022
DOI 10.3389/fpls.2022.1062575

OPEN ACCESS

EDITED BY
Pengtao Ma,
Yantai University, China

REVIEWED BY
Long Li,
Chinese Academy of Agricultural Sciences (CAAS), China
Jihu Li,
SAAS, China

*CORRESPONDENCE
Junming Li
ljm@sjziam.ac.cn
Liqiang Song
songliqiang@hebau.edu.cn

†These authors have contributed equally to this work

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

RECEIVED 06 October 2022
ACCEPTED 24 October 2022
PUBLISHED 14 November 2022

CITATION
Liu J, Zhi L, Zhang N, Zhang W, Meng D, Batool A, Ren X, Ji J, Niu Y, Li R, Li J and Song L (2022) Transcriptomic analysis reveals the contribution of *QMrl-7B* to wheat root growth and development. *Front. Plant Sci.* 13:1062575. doi: 10.3389/fpls.2022.1062575

Transcriptomic analysis reveals the contribution of *QMrl-7B* to wheat root growth and development

Jiajia Liu[1,2†], Liya Zhi[1,2†], Na Zhang[1], Wei Zhang[1], Deyuan Meng[1,2], Aamana Batool[1,2], Xiaoli Ren[1,2], Jun Ji[1], Yanxiao Niu[3], Ruiqi Li[4], Junming Li[1,3*] and Liqiang Song[4*]

[1]Center for Agricultural Resources Research, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Shijiazhuang, Hebei, China, [2]The College of Life Science, University of Chinese Academy of Sciences, Beijing, China, [3]Ministry of Education Key Laboratory of Molecular and Cellular Biology, Hebei Collaboration Innovation Center for Cell Signaling, Hebei Key Laboratory of Molecular and Cellular Biology, College of Life Sciences, Hebei Normal University, Shijiazhuang, China, [4]State Key Laboratory of North China Crop Improvement and Regulation, College of Agronomy, Hebei Agricultural University, Baoding, Hebei, China

Roots are the major organs for water and nutrient acquisition and substantially affect plant growth, development and reproduction. Improvements to root system architecture are highly important for the increased yield potential of bread wheat. *QMrl-7B*, a major stable quantitative trait locus (QTL) that controls maximum root length (MRL), essentially contributes to an improved root system in wheat. To further analyze the biological functions of *QMrl-7B* in root development, two sets of *Triticum aestivum* near-isogenic lines (NILs), one with superior *QMrl-7B* alleles from cultivar Kenong 9204 (KN9204) named NIL^{KN9204} and another with inferior *QMrl-7B* alleles from cultivar Jing 411 (J411) named NIL^{J411}, were subjected to transcriptomic analysis. Among all the mapped genes analyzed, 4871 genes were identified as being differentially expressed between the pairwise NILs under different nitrogen (N) conditions, with 3543 genes expressed under normal-nitrogen (NN) condition and 2689 genes expressed under low-nitrogen (LN) condition. These genes encode proteins that mainly include NO_3^- transporters, phytohormone signaling components and transcription factors (TFs), indicating the presence of a complex regulatory network involved in root determination. In addition, among the 13524 LN-induced differentially expressed genes (DEGs) detected in this study, 4308 and 2463 were specifically expressed in the NIL^{KN9204} and NIL^{J411}, respectively. These DEGs reflect different responses of the two sets of NILs to varying N supplies, which likely involve LN-induced root growth. These

results explain the better-developed root system and increased root vitality conferred by the superior alleles of *QMrl-7B* and provide a deeper understanding of the genetic underpinnings of root traits, pointing to a valuable locus suitable for future breeding efforts for sustainable agriculture.

KEYWORDS

***Triticum aestivum*, root traits, RNA sequencing, differentially expressed gene, near-isogenic lines**

Introduction

Crops produce a large proportion of food and nutrients for humans in modern society. During the past 50 years, the application of chemical fertilizers and the Green Revolution have substantially increased crop production, maintaining food security for the increasingly growing population (Godfray et al., 2010). However, a continuous increase in crop output cannot sustainably be achieved by excessive amounts of fertilizer because of the limited nutrient use efficiency (NUE) of crop plants. Importantly, a large amount of wasted resources and environmental pollution resulting from excessive fertilizer applications cannot be ignored (Springmann et al., 2018). For these reasons, crop cultivars that have relatively high NUE are environmentally friendly and helpful to conserve resources for sustainable agriculture. Roots, the major organs for plant anchorage, nutrient and water absorption, and interactions with symbiotic organisms, have attracted increased amounts of attention in recent studies given the discovery of their correlations with nutrient use and crop yield (Gewin, 2010). In general, manipulation of root architecture traits, including root number, root length, root surface area (RA) and root volume (RV), especially the spatial configuration of the root system in the soil, is an essential approach to increase crop yields, which may represent an important component of a new green revolution (Lynch, 2007; Meister et al., 2014). Unfortunately, root traits have been chronically overlooked in the traditional breeding due to their lack of aboveground visibility (Maqbool et al., 2022).

Bread wheat (*Triticum aestivum* L.) is one of the most important crop species worldwide, providing ~20% of calories for human needs (Appels et al., 2018). Efforts have been made to explore the determinants of root traits that contribute to increase yield potential in wheat. In addition to traditional labor-intensive and time-consuming methods for root trait analysis (Kücke et al., 1995), artificial medium systems, such as hydroponic culture, have been used in many studies and revealed a series of quantitative trait loci (QTLs) that drive genetic variation in root traits (An et al., 2006; Cao et al., 2014; Kabir et al., 2015; Maccaferri et al., 2016; Merchuk-Ovnat et al., 2017; Fan et al., 2018; Zheng et al., 2019). However, it remains largely unclear of the mechanisms of these loci control root traits due to the complexity of the wheat genome. To date, a few genes regulating root morphology have been analyzed through reverse genetics studies in wheat. For example, the expressions of *TaRSL2* (Han et al., 2017a) and *TaLBD16* (Wang et al., 2018a) were shown to be positively correlated with root hair length and number of lateral roots, respectively. Heterogenous expressions of *TaMPK4* (Hao et al., 2015) and *TaMOR* (Li et al., 2016) demonstrated their positive effects on root development. Overexpression of the transcription factor (TF)-encoding genes *TaNAC2-5A* (He et al., 2015), *TaNFYA-B1* (Qu et al., 2015), *TaNAC69-1* (Chen et al., 2016) and *TaWRKY51* (Hu et al., 2018) resulted in enhanced root growth, while *TabZIP60* (Yang et al., 2019) had the opposite effects, implying a complex genetic basis of the formation of root traits. In addition, circular RNAs have been shown to participate in root length regulation (Xu et al., 2019).

At present, a holistic view of the root growth regulatory network and how it affects root traits and yield potential is strongly needed for molecular breeding. RNA sequencing (RNA-seq) makes it possible to explore the consistency between root morphology and its underlying mechanisms at the whole-plant level through DEG analysis. For example, the transcriptomic analysis revealed how the root growth of salt-tolerant bermudagrass was maintained under salinity stress; the salt tolerance was attributed to appropriate expression levels of genes involved in reactive oxygen species (ROS) homeostasis and cell wall loosening, which were regulated in part by a number of TFs linked to stress responses and growth regulation (Hu et al., 2015). DEGs detected in maize cultured under different nitrogen (N) levels highlighted the network of root growth in response to N deficiency, where the *LBD* TF-encoding gene participated in the balance between nutrient metabolism and the stress response (He et al., 2016). RNA-seq data suggested crosstalk occurs among the roots (in terms of

Abbreviations: NN, normal-nitrogen; LN, low-nitrogen; NIL, near-isogenic line; DEG, differentially expressed gene; QTL, quantitative trait locus; RNA-seq, RNA sequencing; NUE, nitrogen use efficiency; MRL, maximum root length; TRL, total root length; RA, root surface area; RV, root volume; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

growth promotion), gibberellin (GA) biosynthesis, auxin signaling, cell division and cell wall modification in uniconazole-treated soybean (Han et al., 2017b). These findings enriched our understanding of the molecular mechanisms behind the interplay of root traits and other agronomic characteristics, helping to generate strategies for crops with improved root systems, increased NUE and improved environmental adaptation.

Previous studies have shown that *Triticum aestivum* KN9204has a larger root system than does cultivar J411 (Wang et al., 2011). Using a RIL population derived from the cross of KN9204 and J411 (KJ-RIL), researchers identified *QMrl-7B* as a major stable QTL that controls maximum root length (MRL) (Fan et al., 2018), and its contribution to large root systems was dissected (Liu et al., 2021), prompting the question of how *QMrl-7B* regulates root architecture and N uptake. In the present study, transcriptomic analysis of the pairwise *QMrl-7B* NILs was performed to further analyze the function of *QMrl-7B* in root development. The results will help reveal the pathways controlling root development that are affected by *QMrl-7B* in wheat, providing insights into the genetic and physiological basis underlying desirable root traits for agriculture.

Materials and methods

Plant materials and root phenotyping

KN9204 is a representative cultivar released in 2003 with a prolific root system, and J411 is a major cultivar that has been planted in the Northern China Plain since 1990s (Wang et al., 2011). Two sets of *QMrl-7B* NILs, one with superior alleles from KN9204 (referred to as the NIL^{KN9204}) and another with inferior alleles from J411 (referred to as the NIL^{J411}) (Liu et al., 2021), were used in this study for transcriptomic analysis.

The root morphological traits of the pairwise *QMrl-7B* NILs were characterized in the experiments involving different N hydroponic cultures, with three replications. Seeds were imbibed in distilled water at 23°C for 24 h and then transferred to plastic pots containing moist vermiculite and nutrient-enriched soil (1:1, v/v) for ten days. Uniform seedlings (at the one- and half-leaf stages) were then chosen, after which their residual endosperm was removed; they were subsequently cultured in both normal-nitrogen (NN) and low-nitrogen (LN) nutrient solutions as described by Fan et al. (2018). On the fourteenth (four-leaf stage) and twenty-first (five-leaf stage) days in different N hydroponic cultures, five uniform seedlings of each NIL were evaluated for their average MRL (cm), total root length (TRL, cm), RA (cm^2) and RV (cm^3). MRL was measured using a ruler, and the remaining traits (TRL, RA and RV) were quantified by using a ScanMaker i800 Plus Scanner (600 DPI) and analyzed by LA-S software (Hangzhou Wanshen Detection Technology Co., Ltd., Hangzhou, China; https://www.wseen.com).

Moreover, the roots of the four-leaf-stage seedlings of the pairwise NILs were sampled and immediately frozen in liquid N for subsequent RNA extraction. The root samples were divided into four treatments: the NIL^{KN9204} under NN condition (NN_NIL^{KN9204}), the NIL^{J411} under NN condition (NN_NIL^{J411}), the NIL^{KN9204} under LN condition (LN_NIL^{KN9204}) and the NIL^{J411} under LN condition (LN_NIL^{J411}). Each sample pool comprised five individual seedlings, and all experiments were replicated three times.

Transcriptome analysis

RNA extraction and sequencing

Total RNA was extracted from the root samples using TRIzol Reagent (Plant RNA Purification Reagent for plant tissue, Invitrogen, USA) according to the manufacturer's instructions, and genomic DNA was removed using DNase I (TaKaRa, Beijing, China). The quantity, quality, and integrity of the extracted RNA were determined using an ND-2000 spectrophotometer (Thermo Fisher Scientific, MA, USA) and an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Only high-quality RNA samples (OD260/280 = 1.8~2.2, OD260/230 ≥ 2.0, RNA integrity number (RIN) ≥ 6.5, 28S:18S ≥ 1.0, RNA content > 1 μg) were used to construct a sequencing library. Messenger RNA (mRNA) was isolated from the total RNA using magnetic beads with oligo (dT) primers and then cut into short fragments in a fragmentation buffer. Complementary DNA (cDNA) was subsequently synthesized using a SuperScript Double-Stranded cDNA Synthesis Kit (Invitrogen, CA, USA), with random hexamer primers (Illumina). The synthesized cDNA was then subjected to end repair, phosphorylation and polyadenylation according to Illumina's library construction protocol. The libraries were size selected for cDNA target fragments of 300 bp on 2% low-range ultra-agarose followed by PCR amplification using Phusion DNA polymerase (NEB); fifteen PCR cycles were performed. After quantification by a TBS-380 fluorometer, the paired-end RNA-seq library was sequenced with an Illumina HiSeq Xten/NovaSeq 6000 sequencer (2 × 150 bp read length). The library was then sequenced using the Illumina HiSeq™ platform (Shanghai Majorbio Bio-pharm Technology Co., Ltd., Shanghai, China).

Analysis of DEGs

The data were analyzed on the free online platform of the Majorbio Cloud Platform (https://www.majorbio.com). To obtain high-quality reads, reads containing adapter sequences and reads of low quality were removed from the raw data to obtain clean reads in FASTQ format. Q20 and Q30 values and GC contents were determined according to conventional

methods. Afterwards, the remaining clean reads were mapped to the reference genome of Chinese Spring (CS) (http://plants.ensembl.org/Triticum_aestivum/Info/Index/) by HISAT2 software (http://ccb.jhu.edu/software/hisat2/index.shtml). Unigene expression levels were estimated according to their transcripts per million reads mapped (TPM) values. DEGs between the NIL^{KN9204} and the NIL^{J411} under both NN and LN conditions were then analyzed using the DESeq2 package (Love et al., 2014), and the expression results are presented in terms of the $\log_2$fold change (FC). $|\log_2 FC| \geq 1$ and p-adjusted< 0.05 were used as thresholds to judge the significance of gene expression differences. DEGs were also evaluated in response to LN stress for each NIL.

The DEGs of each one-to-one comparison were analyzed by Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis. GO functional enrichment and KEGG pathway analysis were carried out by Goatools (https://github.com/tanghaibao/Goatools) and KOBAS (http://kobas.cbi.pku.edu.cn/home.do), where GO terms and biological pathways with a p-adjusted< 0.05 were considered significantly enriched. The encoded products of the DEGs were subsequently annotated by the NCBI non-redundant protein (NR) database.

Quantitative real-time polymerase chain reaction validation

To confirm the reliability of the DEGs obtained from RNA-seq, eight DEGs on chromosome 7B were selected for qRT-PCR analysis. Gene-specific primers were designed using Primer Premier 5.0 software (Premier Biosoft International); the sequences are shown in Table S1. qRT-PCR was carried out in total volumes of 20 µl using a SYBR PCR kit (TaKaRa, Beijing, China) on an ABI 7500 real-time PCR system (Applied Biosystems, MA, USA) according to the manufacturers' instructions. Each sample was analyzed as three technical replicates, and the relative expression value of the DEGs was determined using the comparative cycle threshold (Ct) method after normalization to the expression of the GAPDH control.

Results

Root traits of the pairwise NILs

The root traits of the seedlings in hydroponic cultures with different N contents were evaluated. On the fourteenth day after the seedlings were transferred to the nutrient solutions, the average MRL, TRL, RA and RV of the four-leaf-stage NIL^{KN9204} seedlings were 22.28 cm, 255.77 cm, 25.47 cm^2 and 0.25 cm^3, respectively, under NN condition, which were significantly greater than those of the NIL^{J411} (19.85 cm, 203.47 cm, 20.64 cm^2 and 0.20 cm^3, respectively) under comparable conditions. Accordingly, the average MRL, TRL, RA and RV of the NIL^{KN9204} *vs* NIL^{J411} at the seedling stage were 28.07 *vs* 21.54 cm, 301.75 *vs* 282.12 cm, 27.28 *vs* 24.90 cm^2 and 0.25 *vs* 0.21 cm^3, respectively, under LN condition, indicating that the NIL^{KN9204} was superior to the NIL^{411} by 30.32%, 6.96%, 9.56% and 19.05%, respectively (p< 0.05). Notably, the MRL of the NIL^{KN9204} seedlings under LN condition was 25.99% greater than that under NN conditions, while only an 8.51% difference was recorded for the NIL^{411} under the hydroponic cultures with different N contents, indicating that LN more strongly induced the growth of primary roots for the NIL^{KN9204} than for the NIL^{411} (Figure 1). Similar change trends of MRL, TRL, RA and RV differences in the five-leaf-stage seedlings were also observed between the two sets of NILs at twenty-one days after the 10-day-old seedlings were transferred to the nutrient solutions (Figure S1). Taken together, the above results showed that, in comparison to the NIL^{411}, the NIL^{KN9204} had superior root architecture characteristics during the seedling stage, especially under LN condition.

Overview of transcriptome sequencing results

To explore the genetic and physiological basis of the superior root traits of the NIL^{KN9204}, roots of four-leaf-stage seedlings of the pairwise NILs cultured under both NN conditions and LN conditions were sampled for transcriptomic analysis. RNA-seq on the Illumina HiSeq™ platform (http://www.majorbio.com) generated 70576620~90524786 clean reads per sample, with a sequencing error of less than 0.3%. The average Q20, Q30 and GC content percentages were 97.06%, 92.27% and 53.98%, respectively. These findings attest to the fine quality of the sequencing results. Mapping of the reads with the Chinese Spring (CS) reference genome revealed an alignment in a range of 81.48~90.35%, among which 76.14~84.54% of the genes were uniquely mapped and 5.34~6.28% were multiply mapped (Table S2). The correlation coefficient of repeats ranged from 0.69~0.98 (Figure S2, Table S3). These data were deemed suitable for subsequent analysis (Table S4), in which the DEGs whose expression was upregulated ($\log_2 FC \geq 1$) and downregulated ($\log_2 FC \leq -1$) in each one-to-one comparison were statistically (p-adjusted< 0.05) analyzed.

Identification of DEGs between the pairwise NILs

The transcriptomic changes in the roots between the pairwise *QMrl-7B* NILs were compared under both NN and LN conditions. Under NN condition, a total of 3543 genes were differentially expressed between the two sets of NILs, of which

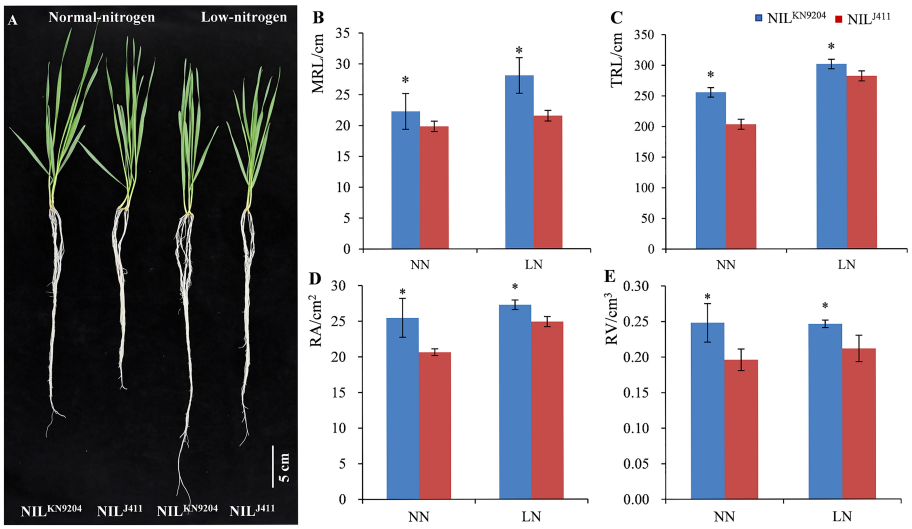


FIGURE 1
Phenotypes of four-leaf-stage seedlings **(A)** and the corresponding root traits **(B-E)** of the pairwise NILs cultured in nutrient solutions with different N contents for fourteen days NN and LN indicate normal- and low-nitrogen conditions, respectively. MRL, TRL, RA and RV indicate the maximum root length, total root length, root surface area and root volume, respectively. "*" indicates a significant difference at $p< 0.05$.

the expression of 1045 genes was upregulated and that of 2498 was downregulated in the NIL^{KN9204} (Figure 2A; Table S5). These genes were enriched in 320 subcategories according to GO analysis, of which 200, fourteen and 106 belonged to biological processes, cellular components and molecular functions, respectively, and the three most enriched biological processes were "biological process (GO:0008150)", "metabolic process (GO:0008152)" and "phosphate-containing compound metabolic process (GO:0006796)" (Table S6). "Phenylpropanoid biosynthesis (map00940)" was the most enriched pathway according to KEGG analysis (Table S7). Under the LN condition, 2689 genes were differentially expressed between the pairwise NILs, including 990 genes whose expressions were upregulated and 1699 genes whose expressions were downregulated in the NIL^{KN9204} (Figure 2A; Table S8). These DEGs were categorized into biological process (141), cellular component (12) and molecular function (77) categories by GO analysis. The three most enriched biological processes were "biological process (GO:0008150)", "metabolic process (GO:0008152)" and "phosphate-containing compound metabolic process (GO:0006796)" (Table S9). Similarly, "phenylpropanoid biosynthesis (map00940)" was the most enriched pathway according to the KEGG analysis (Table S10).

Notably, there were 315 DEGs whose expressions increased and 1034 DEGs whose expressions decreased in the NIL^{KN9204} compared to the NIL^{J411} simultaneously under both NN and LN conditions (Figure 2A). These genes might participate in translation, transcriptional regulation and cycle control according to GO and KEGG enrichment analysis (Figure 2B; Tables S11-S13), reflecting the permanent effects of the *QMrl-7B* locus that were not affected by N levels. Under NN condition, nine genes were expressed at higher levels in the NIL^{KN9204} than in the NIL^{J411}; however, these genes were expressed at a higher level in the NIL^{J411} under LN condition (Figure 2A; Table 1). In addition, three genes exhibited the opposite expression trend, implying their functional difference in response to N stress (Figure 2A; Table 1).

Identification of DEGs in response to N deficiency

The differences in DEGs in the root response to N deficiency between the NIL^{KN9204} and NIL^{J411} were analyzed. Compared with the NN condition, LN condition yielded 11073 DEGs (6341 whose expression increased and 4732 whose expression decreased) and 9228 DEGs (5072 whose expression increased and 4156 whose expression decreased) in the $NIL^{KN9204}$and NIL^{J411}, respectively (Figure 3A). A total of 2385 DEGs whose expressions increased and 1923 DEGs whose expressions decreased were detected only in the NIL^{KN9204} (Figure 3A; Table S14), whereas 1116 DEGs whose expressions increased and 1347 DEGs whose expressions decreased were detected specifically in the NIL^{J411} (Figure 3A; Table S17). Interestingly, there were more genes whose expressions were upregulated than genes whose expressions were downregulated in the NIL^{KN9204}, and the opposite trend occurred in the NIL^{J411}. GO analysis showed that these DEGs participate in multiple processes, and "phenylpropanoid biosynthesis (map00940)" was the most enriched process according to KEGG analysis (Figures 3B, C;

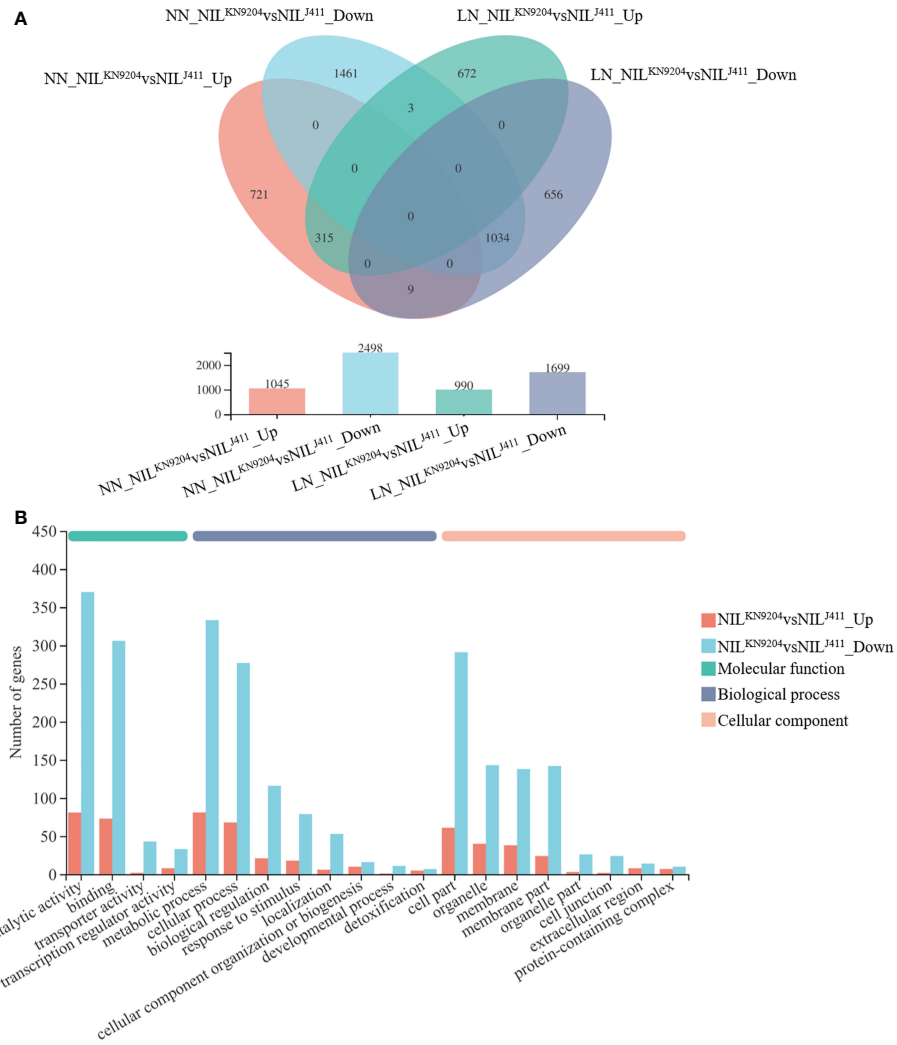


FIGURE 2
Identification of DEGs between the pairwise NILs NN and LN indicate normal- and low-nitrogen conditions, respectively. The expression levels between the NIL^{KN9204} and the NIL^{J411} were compared under each condition, with the former as a control. NIL^{J411}*vs*NIL^{KN9204}_Up and NIL^{J411}*vs*NIL^{KN9204}_Down indicate DEGs whose expression was upregulated and downregulated in the NIL^{J411} compared to the NIL^{KN9204} under the same culture conditions, respectively. **(A)** Venn diagram analysis of DEGs in each one-to-one comparison between the pairwise NILs under both NN and LN conditions. **(B)** GO annotation analysis of DEGs between the pairwise NILs that were simultaneously expressed under both NN and LN conditions with the first 20 terms as indicated.

Tables S14-S19). The expression of genes that differed specifically in the NIL^{KN9204} or the NIL^{J411} might give rise to a specific N response related to *QMrl-7B*, while a large number of common DEGs (3956 genes whose expression was upregulated and 2809 genes whose expression downregulated) in the two sets of NILs reflected the general response to N stress in wheat.

Quantitative real-time PCR verification

qRT-PCR analysis revealed that the expression levels of eight randomly selected DEGs on chromosome 7B were significantly different between the NIL^{KN9204} and the NIL^{J411} under NN conditions and LN conditions (Figure 4). These results nearly matched the expression profiles revealed by RNA-seq, confirming the results of the transcriptome sequencing of the pairwise NILs cultured under both NN and LN conditions.

Discussion

Improving plant roots is an effective way to increase NUE and minimize fertilizer application for sustainable agricultural development (Lynch, 2007). A number of genes/QTLs have been

TABLE 1 DEGs with opposite expression patterns between the pairwise NILs under NN and LN conditions.

Gene_ID	log_2FC (NN_NIL^{KN9204}/ NN_NIL^{J411})	*p*-adjusted	log_2FC (LN_NIL^{KN9204}/ LN_NIL^{J411})	*p*-adjusted	NR_description
TraesCS2A02G399900	1.07	0.002	-1.97	7.77E-20	MATE efflux family protein 9
TraesCS2B02G125500	-1.37	0.0098	1.08	1.25E-06	class III peroxidase
TraesCS2B02G626400	-5.11	2.88E-34	1.11	2.48E-05	probable 1-deoxy-D-xylulose-5-phosphate synthase 2, chloroplastic
TraesCS3B02G448000	1.74	1.07E-18	-1.35	1.98E-06	unnamed protein product
TraesCS3D02G094500	-2.06	5.54E-07	1.68	0.001	MYB-related protein 305-like
TraesCS3D02G225500	1.16	0.0002	-1.42	4.08E-14	uncharacterized protein LOC109770081
TraesCS3D02G408000	1.21	2.62E-07	-1.68	4.76E-16	BTB/POZ and TAZ domain-containing protein 2-like
TraesCS4A02G286600	1.32	4.98E-05	-1.73	2.50E-10	uncharacterized protein LOC109755115
TraesCS4B02G078800	1.85	3.67E-15	-1.07	5.58E-06	hypothetical protein TRIUR3_28321
TraesCS4D02G025500	1.18	0.0003	-1.97	2.85E-08	uncharacterized protein LOC109755115
TraesCS6B02G044300	1.15	0.001	-1.47	4.78E-08	high-affinity nitrate transporter TaNRT2
TraesCS6D02G035900	1.01	0.016	-1.02	0.002	high-affinity nitrate transporter 2.1-like

NN and LN indicate normal- and low-nitrogen conditions, respectively. AA and BB indicate the NIL^{KN9204} and the NIL^{J411}, respectively. The expression levels between the NIL^{KN9204} and the NIL^{J411} were compared under each condition, with the latter as a control. The expression ratios are presented as log_2FC values. Genes with a $log_2FC \leq -1$ (p-adjusted< 0.05, coloured blue) represent the genes whose expression was downregulated in the NIL^{KN9204}, and genes with $log_2FC \geq 1$ (p-adjusted< 0.05, coloured red) represent the genes whose expression was upregulated in the NIL^{KN9204}.

reported to be related to root growth and development; as such, these genes/QTLs have potential application in crop genetic improvement (Zhuang et al., 2021; Chen et al., 2022; Li et al., 2022). However, the genetic network explaining how these genes interact to influence root architecture still needs to be elucidated to assess the beneficial agronomic traits from an overall perspective and to guide the use of beneficial genes or pyramid of allelic variation in future breeding efforts. Compared with J411, KN9204 had a larger root system (Wang et al., 2011), with which the stable QTL *QMrl-7B* was associated (Fan et al., 2018). Further studies involving *QMrl-7B* NILs showed that, as compare to the NIL^{J411}s, the NIL^{KN9204}s had superior root traits, N-related traits and yield traits in field trials (Liu et al., 2021). The superiority of root traits of the NIL^{KN9204} at the seedling stage was also consistently proven in the present study. When these hydroponically cultured pairwise NILs were used for transcriptomic analysis, it was found that the DEGs detected by RNA-seq encoded proteins of various groups, such as transporters, enzymes, TFs, and components involved in hormone signaling. The important DEGs that might be involved in root trait regulation are discussed here.

The genes with opposite expression patterns between the pairwise NILs under the NN and LN conditions were very likely to be involved in N-related root growth (Tables 1, S5 and S8). For example, the expression level of the class III peroxidase gene (*TraesCS2B02G125500*) was higher in the NIL^{J411} under NN condition but was higher in the NIL^{KN9204} under LN condition. Class III peroxidase uses H_2O_2 as a substrate to produce OH^- radicals that enhance cell wall plastic extensibility and elongation by nonenzymatic cleavage of cell wall polysaccharides (Cosio and Dunand, 2009). Since root development is directly related to cell wall loosening and elongation, the upregulated expression of this gene by LN in the NIL^{KN9204} (Table S14) was proposed to participate in LN-promoted root growth by affecting these processes. Another most enriched process of the DEGs between the pairwise NILs was phenylpropanoid biosynthesis that might affect root enlargement (Muro-Villanueva et al., 2019; Liu et al., 2020). Interestingly, the expression levels of a series of phenylalanine ammonia-lyase and cinnamoyl-CoA reductase genes increased in the NIL^{J411} under both NN and LN conditions (Tables S5 and S8), implying that the biosynthesis of cinnamate and lignin monomers increased, which cannot directly explain the reason for the smaller roots.

Nitrate is substantially absorbed and transported by high-affinity (NRT2s) and low-affinity (NPFs) nitrate transporters for assimilation and storage (Xu et al., 2012; Wang et al., 2018b). N stress tolerant sorghum genotypes exhibit higher expression levels of high-affinity NRTs than do the sensitive genotypes under N-limiting conditions (Gelli et al., 2014). We previously compared the expression profiles of 44 NRT2 and 217 NPF genes in the roots of KN9204 and J411 cultivated hydroponically with and without N supply and found that up to 79.2% and 43.7% of the root-expressed NRT2 and NPF genes, respectively, displayed differential expression between KN9204 and J411. Under NN conditions, more than two-thirds of the differentially expressed NRT2 and NPF genes exhibited higher expression in KN9204 than in J411, but the proportion of NRT2 and NPF genes with higher expression levels in KN9204 decreased under LN treatment, as the upregulated expression levels of most nitrate transporter genes, NRT2 family in particular, were higher in J411 than in KN9204 (Shi et al.,

10.3389/fpls.2022.1062575

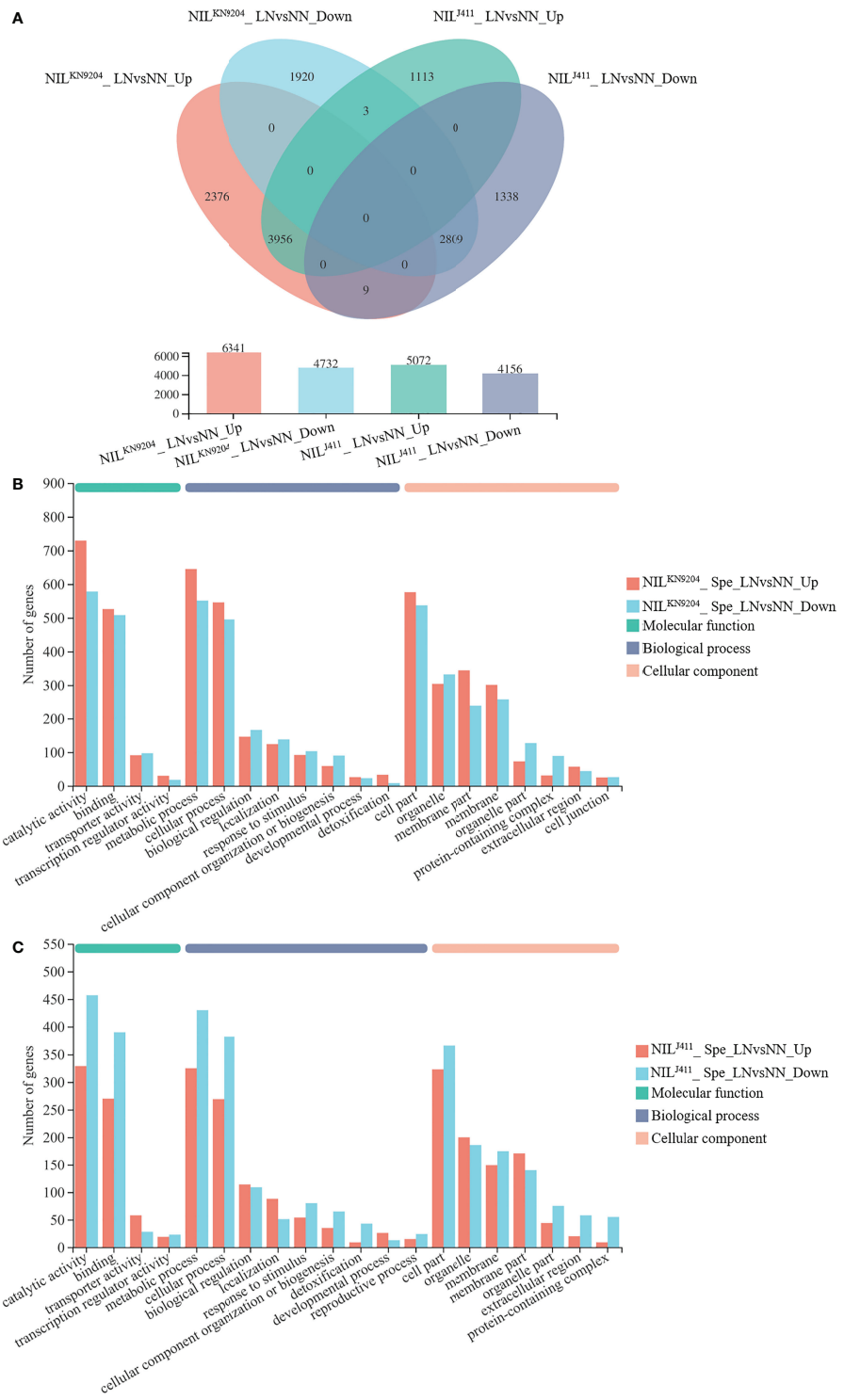


FIGURE 3
Identification of DEGs induced by LN NN and LN indicate normal- and low-nitrogen conditions, respectively. Spe indicates DEGs specifically expressed in the NIL^{KN9204} or the NIL^{J411}. The expression levels between NN and LN conditions were compared for each NIL, with NIL^{KN9204} as a control. LNvsNN_Up and LNvsNN_Down indicate DEGs whose expression was upregulated and downregulated under LN compared to NN for the same NIL. **(A)** Venn diagram analysis of the DEGs in each one-to-one comparison between NN and LN conditions for both the NIL^{KN9204} and NIL^{J411}. **(B, C)** GO annotation analysis of the genes whose expression differed specifically in the NIL^{KN9204}**(B)** and NIL^{J411} **(C)** with the first 20 terms as indicated.

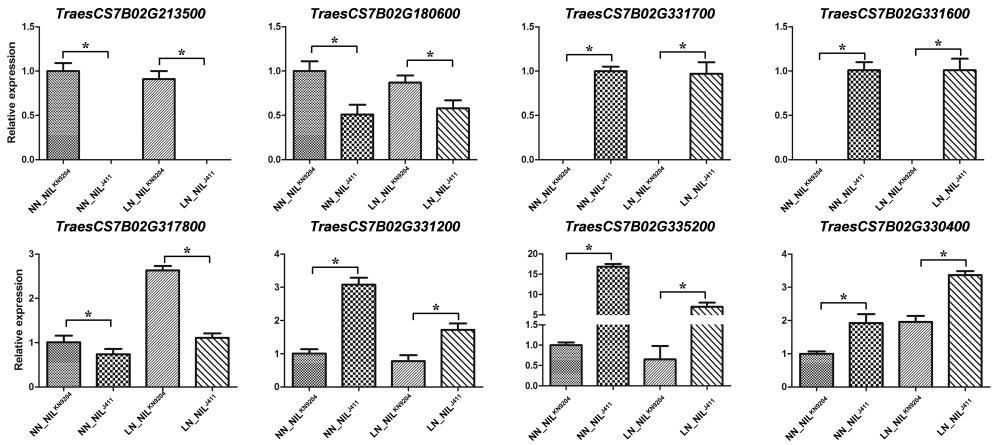


FIGURE 4
Validation of gene expression by qRT-PCR The relative mRNA expression of eight representative genes was checked by qRT-PCR. Each bar shows the mean ± standard error of three biological replicates. "*" indicates a significant difference at $p< 0.05$.

2022). Similarly, two high-affinity NRT genes (*TraesCS6B02G044300* and *TraesCS6D02G035900*) exhibited lower expression levels in the NIL^{J411} under NN condition, but higher expression levels in the NIL^{J411} than in the NIL^{KN9204} under LN condition (Table 1). In addition, the expressions of four high-affinity NRT genes (*TraesCS6A02G031000*, *TraesCS6A02G031100*, *TraesCS6B02G044100* and *TraesCS6B02G044500*) were induced specifically by LN in the NIL^{J411} but not in the NIL^{KN9204} (Table 2). Our present study showed that LN led to altere expression of a number of genes encoding both low-affinity and high-affinity NRTs (Table S4), in accordance with the viewpoint that different NRTs function at different N levels (Zhang and Forde, 1998).

Dual-affinity nitrate transporter 1 (NRT1) not only functions in N acquisition but also is a NO_3^- sensor; this action is independent of its uptake activity that controls root development (Ho et al., 2009). The cotransport of NO_3^-, peptides and hormones through the NRT1 family of transporters implies the occurrence of crosstalk between NO_3^- and other signaling pathways (Krouk et al., 2010; Tal et al., 2016). In our study, the expression levels of two and eleven putative NRT1 genes in the NIL^{KN9204} were higher and lower than those in the NIL^{J411} under both NN and LN conditions (Table 3), respectively, indicating that the differences in root traits may result mainly from NRT1 functional differentiation. Additionally, the higher expression level of the CBL-interacting protein kinase 23 (CIPK23) gene (*TraesCS2A02G251800*) in the NIL^{J411} compared to the NIL^{KN9204} was also independent of N status (Tables S5 and S8); this phenomenon may affect root formation together with NRT1 by phosphorylating T101 to regulate its primary NO_3^- response and high-affinity NO_3^- transport (Ho et al., 2009). Taken together, these results implied that a more sensitive response of roots to N acquisition occurred in the NIL^{J411} compared with the NIL^{KN9204}, which cannot be fully explained.

Phytohormones, together with their receptors and signaling components, constitute the main intrinsic regulatory pathways of root traits. Ethylene, whose biosynthesis is regulated by several key enzymes, such as 1-aminocyclopropane-1-carboxylic acid (ACC) synthase (ACS) and ACC oxidase (ACO), usually plays a negative regulatory role in root growth, inhibiting root enlargement and root meristem cell proliferation (Pattyn et al., 2020). In addition, ethylene regulates the expression of glutathione S-transferase (GST)-encoding genes. GST catalyses the conjugation of glutathione to various electrophilic compounds involved in the oxidative stress response, crosstalk with other hormones and drought-induced root growth (Dalal et al., 2018). Under both NN and LN conditions, significant increases in expression levels of genes encoding ACOs and GSTs were detected in the NIL^{J411}, with parallel increases in expression levels of genes encoding ethylene-responsive TFs (Table 4), implying that enhanced ethylene signaling together with glutathione metabolism negatively regulated root growth in the NIL^{J411}. GA, whose catabolism is catalysed by GA2-oxidase, is considered a negative regulator of lateral root formation (Gou et al., 2010). However, overexpression of GA2-oxidase-encoding genes led to shortened primary root length, and mutations in these kinds of genes led to somewhat longer roots (Li et al., 2017), indicating a negative correlation between GA2-oxidase-encoding genes and root length. Compared to the NIL^{KN9204}, the NIL^{J411} displayed upregulated expression levels of several GA2-oxidase-encoding genes under both NN and LN conditions (Table 4); these upregulated levels may have inhibitory effects on root development. Brassinosteroids (BRs) regulate both root meristem size and cell expansion in controlling root growth.

TABLE 2 *NRT* genes specifically expressed in the NIL^{KN9204} or NIL^{J411} in response to LN.

Gene_ ID	log_2FC (LN_NIL^{KN9204}/ NN_NIL^{KN9204})	*p*-adjusted	log_2FC (LN_NIL^{J411}/ NN_NIL^{J411})	*p*-adjusted	NR_description
TraesCS1A02G150200	1.03	0.004	0.14	0.84	Peptide transporter PTR3-A
TraesCS1A02G150400	1.18	2.29E-07	0.64	0.08	protein NRT1/PTR FAMILY 5.2-like
TraesCS1B02G168100	1.60	1.11E-12	0.90	0.01	protein NRT1/PTR FAMILY 5.2-like
TraesCS1B02G225000	-1.39	2.12E-06	-0.14	0.81	protein NRT1/PTR FAMILY 6.3-like
TraesCS1B02G267900	1.29	0.01	1.15	0.18	protein NRT1/PTR FAMILY 3.1-like
TraesCS1D02G201100	1.13	2.34E-11	0.81	0.0004	uncharacterized protein LOC109744571
TraesCS1D02G214300	-1.47	7.24E-08	0.20	0.72	protein NRT1/PTR FAMILY 6.3-like
TraesCS2A02G007500	0.60	0.08	1.05	1.54E-09	protein NRT1/PTR FAMILY 8.5-like
TraesCS2A02G074800	-1.89	0.003	-0.58	0.45	high-affinity nitrate transporter 2.1-like
TraesCS2A02G264500	0.99	8.86E-05	1.06	0.0006	protein NRT1/PTR FAMILY 4.5-like
TraesCS2B02G277600	0.71	0.003	1.27	6.40E-09	protein NRT1/PTR FAMILY 4.5-like
TraesCS2B02G626000	1.02	1.30E-05	0.43	0.19	protein NRT1/PTR FAMILY 5.10-like
TraesCS2B02G626700	1.76	1.07E-08	0.97	0.01	protein NRT1/PTR FAMILY 5.10-like
TraesCS2D02G413900	1.50	0.02	1.74	0.08	protein NRT1/PTR FAMILY 8.2-like
TraesCS2D02G583500	0.99	0.01	1.30	0.0005	protein NRT1/PTR FAMILY 5.10-like
TraesCS3A02G382100	-2.55	6.06E-20	-0.94	0.34	protein NRT1/PTR FAMILY 5.1-like isoform X1
TraesCS3A02G382200	-0.12	0.90	1.57	0.003	unnamed protein product
TraesCS3B02G095900	0.20	0.62	1.12	0.02	unnamed protein product
TraesCS3B02G414900	-3.35	5.34E-07	0.20	0.90	unnamed protein product
TraesCS3B02G415700	0.81	0.12	1.01	0.02	predicted protein
TraesCS3D02G056700	0.85	0.25	1.43	0.048	protein NRT1/PTR FAMILY 8.2-like isoform X1
TraesCS4A02G225400	-1.07	6.53E-06	-0.97	0.007	protein NRT1/PTR FAMILY 4.3-like
TraesCS4A02G283900	1.20	6.27E-15	0.93	9.83E-06	protein NRT1/PTR FAMILY 2.11-like
TraesCS4D02G026800	1.01	1.57E-08	0.85	1.76E-05	protein NRT1/PTR FAMILY 2.11-like
TraesCS4D02G087900	-1.39	2.77E-07	-0.997	0.002	protein NRT1/PTR FAMILY 4.3-like
TraesCS5B02G039100	3.41	0.02	2.02	0.22	protein NRT1/PTR FAMILY 2.11-like
TraesCS5B02G152000	1.42	0.31	3.72	0.04	putative nitrate excretion transporter 3
TraesCS5B02G245300	2.48	8.76E-05	1.89	0.09	predicted protein
TraesCS5B02G393100	1.46	0.39	4.38	0.03	protein NRT1/PTR FAMILY 4.5-like
TraesCS5B02G414000	0.79	0.008	1.14	2.22E-05	protein NRT1/PTR FAMILY 6.4 isoform X2
TraesCS5B02G498700	-0.82	0.02	-1.09	0.01	protein NRT1/PTR FAMILY 5.6-like
TraesCS5D02G419200	0.89	0.03	1.91	0.0007	protein NRT1/PTR FAMILY 6.4 isoform X1
TraesCS5D02G498700	-1.03	0.009	-0.95	0.0004	protein NRT1/PTR FAMILY 5.6-like
TraesCS6A02G030700	-2.93	7.26E-18	-0.63	0.17	high-affinity nitrate transporter 2.1-like
TraesCS6A02G030800	-2.84	1.38E-27	-0.38	0.50	high-affinity nitrate transporter 2.1-like
TraesCS6A02G030900	-2.75	1.36E-29	-0.43	0.33	High-affinity nitrate transporter 2.1
TraesCS6A02G031000	0.20	0.78	1.63	0.0002	High-affinity nitrate transporter 2.1
TraesCS6A02G031100	0.01	0.98	1.45	3.00E-05	high affinity nitrate transporter
TraesCS6A02G033100	-6.88	4.09E-06	-2.71	1	high-affinity nitrate transporter 2.1-like
TraesCS6A02G033200	-6.79	8.35E-06	-3.36	1	High affinity nitrate transporter 2.4

(Continued)

TABLE 2 Continued

Gene_ ID	log_2FC (LN_NIL^{KN9204}/ NN_NIL^{KN9204})	*p*-adjusted	log_2FC (LN_NIL^{J411}/ NN_NIL^{J411})	*p*-adjusted	NR_description
TraesCS6A02G267700	0.30	0.62	-1.13	4.90E-05	protein NRT1/PTR FAMILY 8.3-like
TraesCS6A02G369900	4.75	0.01	0.22	0.93	protein NRT1/PTR FAMILY 8.3-like isoform X1
TraesCS6B02G044100	0.06	0.90	1.51	4.09E-05	high affinity nitrate transporter
TraesCS6B02G044200	-1.21	2.49E-05	0.42	0.39	high-affinity nitrate transporter 2.1-like
TraesCS6B02G044400	-1.37	6.30E-07	0.93	0.01	high-affinity nitrate transporter 2.1-like
TraesCS6B02G044500	0.17	0.72	3.46	2.47E-36	high affinity nitrate transporter
TraesCS6B02G309200	-1.09	0.10	-2.04	0.008	protein NRT1/PTR FAMILY 7.3-like
TraesCS6D02G035900	-2.25	6.07E-16	-0.22	0.69	high-affinity nitrate transporter 2.1-like
TraesCS6D02G037800	-3.88	0.02	-3.59	0.23	high-affinity nitrate transporter 2.1-like
TraesCS6D02G132100	1.45	0.008	1.18	0.09	protein NRT1/PTR FAMILY 8.3-like
TraesCS7A02G121600	1.87	4.14E-10	0.60	0.24	protein NRT1/PTR FAMILY 2.3-like
TraesCS7A02G301700	-1.26	5.33E-12	-0.96	3.48E-08	protein NRT1/PTR FAMILY 6.3-like
TraesCS7B02G262200	0.99	2.59E-05	1.16	2.16E-08	protein NRT1/PTR FAMILY 4.6-like
TraesCS7B02G283400	0.93	0.0001	1.31	0.0001	protein NRT1/PTR FAMILY 8.3-like
TraesCS7D02G049300	0.71	0.16	1.80	0.004	protein NRT1/PTR FAMILY 2.3-like
TraesCS7D02G357300	1.11	2.26E-06	0.83	0.0003	protein NRT1/PTR FAMILY 4.6-like

NN and LN indicate normal- and low-nitrogen conditions, respectively. AA and BB indicate the NIL^{KN9204} and the NIL^{J411}, respectively. The expression levels between the NIL^{KN9204} and the NIL^{J411} were compared under each condition, with the latter as a control. The expression ratios are presented as log_2FC values. Genes with a $log_2FC \leq -1$ (p-adjusted< 0.05, coloured blue) represent the genes whose expression was downregulated under LN conditions, and genes with $log_2FC \geq 1$ (p-adjusted< 0.05, coloured red) represent the genes whose expression was upregulated under LN conditions.

Deficiency or insensitivity of BRs leads to reduced root growth and lateral root formation (Wei and Li, 2016). The decreased expression of receptor-like protein kinase BRI1-like 3 genes in the NIL^{J411} might impair BR signaling and inhibited root enlargement (Table 4). Auxin facilitates lateral root initiation and development (Fukaki and Tasaka, 2009), while cytokinin antagonistically negatively regulates root growth (Moubayidin et al., 2009). In rice, crown root primordium initiation was promoted by auxin signaling, whereas cytokinin signaling inhibited it (Neogy et al., 2021). Members of the *GH3* gene family respond to auxin and their encoding products use various amino acids as substrates to form indole-acetic acid (IAA)-amido conjugates for temporary storage and degradation (Staswick et al., 2005). Arabidopsis *GH3.17* gene regulates environment-induced hypocotyl elongation (Zheng et al., 2016) and it is also responsible for root elongation, whereas gh3 mutants accumulate more free-IAA at the expense of any IAA-Glu and IAA-Asp auxin conjugates which causes defects in root development, demonstrating the critical roles of *GH3* genes in auxin homeostasis and plant development (Guo et al., 2022). Cytokinin is activated by phosphoribohydrolase, which catalyses the hydrolysis of the bond between N^6-substituted bases and ribose 5'-monophosphates in cytokinin precursors during its biosynthesis process (Kurakawa et al., 2007). Rice *GH3-8* gene functions as a repressor of auxin-dependent developmental signaling, overexpression of *GH3-8* suppressed auxin action, which resulted in abnormal morphology similar to that reported in auxin-deficient plants i.e. shorter roots and fewer adventitious roots (Ding et al., 2008). Significant increases in expression levels were observed for genes encoding GH3.8 and phosphoribohydrolase in the NIL^{J411}, accompanied by decreased expression levels of a gene encoding an auxin-responsive protein (Table 4), implying the presence of weakened auxin signaling and accelerated cytokinin flux in the NIL^{J411}, which may be related to the differences in root traits of the pairwise NILs.

Several types of TFs play essential roles in root growth by regulating the expression of downstream genes that are usually closely associated with nutrient acquisition. Overexpression of *TaNFYA-B1* was shown to increase both primary root length and total lateral root length in wheat (Qu et al., 2015). The expression levels of genes encoding NF-Y subunit C-3-like proteins (*TraesCS1A02G354900*, *TraesCS1B02G366800* and *TraesCS1D02G355600*) were higher in the NIL^{KN9204} than in the NIL^{J411} under LN condition (Table 5), indicating their potential contribution to LN-induced root enlargement. NAC

TABLE 3 Differentially expressed *NRT* genes between the pairwise NILs.

Gene_ ID	log_2FC (NN_NIL^{KN9204}/ NN_NIL^{J411})	*p*-adjusted	log_2FC (LN_NIL^{KN9204}/ LN_NIL^{J411})	*p*-adjusted	NR_description
TraesCS2A02G309100	-1.43	9.78E-12	-1.24	2.60E-12	protein NRT1/PTR FAMILY 4.3-like
TraesCS2A02G572200	1.13	0.04	1.20	0.003	protein NRT1/PTR FAMILY 5.10-like
TraesCS5A02G485200	-1.03	0.006	-1.12	0.01	protein NRT1/PTR FAMILY 5.6-like
TraesCS5D02G498700	-2.07	4.20E-15	-2.15	1.62E-11	protein NRT1/PTR FAMILY 5.6-like
TraesCS6A02G267600	-1.44	0.0001	-1.92	0.002	protein NRT1/PTR FAMILY 8.3-like
TraesCS6A02G267700	-3.00	1.90E-24	-1.56	6.45E-06	protein NRT1/PTR FAMILY 8.3-like
TraesCS6A02G267800	-3.23	3.26E-05	-2.90	0.006	protein NRT1/PTR FAMILY 8.3-like
TraesCS6A02G268500	-1.09	7.42E-05	-1.56	0.005	protein NRT1/PTR FAMILY 8.3-like
TraesCS6B02G295000	-2.41	1.52E-13	-2.40	5.56E-08	protein NRT1/PTR FAMILY 8.3-like
TraesCS6B02G295600	-1.85	1.83E-12	-1.78	0.005	protein NRT1/PTR FAMILY 8.3-like
TraesCS6D02G247000	-4.10	4.91E-25	-2.03	0.0001	protein NRT1/PTR FAMILY 8.3-like
TraesCS6D02G247100	-1.45	9.32E-05	-1.88	5.83E-05	protein NRT1/PTR FAMILY 8.3-like isoform X1
TraesCS7D02G378300	1.53	0.0004	1.04	0.008	protein NRT1/PTR FAMILY 8.3-like isoform X1

NN and LN indicate normal- and low-nitrogen conditions, respectively. AA and BB indicate the NIL^{KN9204} and the NIL^{J411}, respectively. The expression levels between the $NIL^{KN9204}$and the NIL^{J411} were compared under each condition, with the latter as a control. The expression ratios are presented as log_2FC values. Genes with a $log_2FC \leq -1$ (p-adjusted< 0.05, coloured blue) represent the genes whose expression was downregulated in the NIL^{KN9204}, and genes with $log_2FC \geq 1$ (p-adjusted< 0.05, coloured red) represent the genes whose expression was upregulated in the NIL^{KN9204}.

TABLE 4 DEGs involved in phytohormone metabolism and signaling between the pairwise NILs.

Gene_ ID	log_2FC (NN_NIL^{KN9204}/ NN_NIL^{J411})	*p*-adjusted	log_2FC (LN_NIL^{KN9204}/ LN_NIL^{J411})	*p*-adjusted	NR_description
Ethylene					
TraesCS1A02G370400	-5.05	0.0004	-2.53	5.63E-05	ethylene-responsive transcription factor ERF109-like
TraesCS1A02G370600	-3.55	4.19E-17	-1.44	8.02E-05	AP2 domain containing protein
TraesCS1A02G370700	-5.28	1.08E-51	-2.08	8.05E-11	ethylene-responsive transcription factor ERF109-like
TraesCS1B02G117400	-1.32	5.13E-05	-1.44	6.19E-09	1-aminocyclopropane-1-carboxylate oxidase-like
TraesCS1B02G389700	-3.15	3.17E-20	-1.49	6.98E-06	ethylene responsive transcription factor 6
TraesCS1B02G389800	-3.64	1.61E-26	-1.78	1.94E-07	ethylene-responsive transcription factor ERF109-like
TraesCS1D02G098000	-1.80	2.67E-11	-1.77	5.13E-10	1-aminocyclopropane-1-carboxylate oxidase-like
TraesCS1D02G098100	-1.10	8.10E-06	-1.51	7.25E-09	1-aminocyclopropane-1-carboxylate oxidase-like
TraesCS1D02G376500	-4.11	8.13E-33	-1.79	7.68E-08	ethylene-responsive transcription factor ERF109-like
TraesCS1D02G376600	-3.26	2.22E-37	-1.33	0.0001	ethylene-responsive transcription factor ERF109-like
TraesCS1D02G376700	-3.34	1.30E-23	-1.36	2.22E-05	ethylene-responsive transcription factor ERF109-like
TraesCS1D02G376800	-4.79	1.53E-42	-2.08	0.0005	ethylene-responsive transcription factor ERF109-like
TraesCS2D02G397100	-5.47	0.0002	-2.92	4.59E-09	ethylene-responsive transcription factor ERF025-like
TraesCS3B02G158800	1.56	0.037	1.89	0.003	AP2 domain transcription factor TaDREB2
TraesCS3B02G412500	-2.39	4.76E-05	-1.17	0.046	F-box domain containing protein, expressed
TraesCS4A02G256500	-4.94	5.60E-08	-4.57	5.50E-19	1-aminocyclopropane-1-carboxylate synthase 1-like
TraesCS4B02G058200	-2.51	7.59E-09	-3.30	5.16E-23	1-aminocyclopropane-1-carboxylate synthase 1-like
TraesCS4D02G058200	-2.67	1.16E-12	-2.89	1.08E-23	1-aminocyclopropane-1-carboxylate synthase 1-like
TraesCS4D02G096400	-1.62	8.07E-10	-1.25	2.91E-11	lipase-like PAD4
TraesCS5B02G236900	-2.68	6.45E-17	-1.25	0.0009	ethylene-responsive transcription factor ERF109-like
TraesCS5D02G245300	-2.37	2.30E-10	-1.34	0.002	ethylene-responsive transcription factor ERF109-like
TraesCS6A02G235100	1.95	5.71E-07	1.08	0.004	predicted protein

(Continued)

TABLE 4 Continued

Gene_ ID	log_2FC (NN_NIL^{KN9204}/ NN_NIL^{J411})	*p*-adjusted	log_2FC (LN_NIL^{KN9204}/ LN_NIL^{J411})	*p*-adjusted	NR_description
TraesCS6A02G256600	-3.44	9.28E-14	-2.30	4.23E-12	ethylene-responsive transcription factor ERF027-like
TraesCS6A02G319500	-3.88	2.62E-11	-2.36	1.01E-09	ethylene-responsive transcription factor ERF109-like
TraesCS6A02G330500	-1.83	5.11E-16	-1.26	1.25E-06	ethylene-responsive transcription factor ERF014-like
TraesCS6B02G268700	-3.01	3.89E-17	-2.22	1.77E-19	ethylene-responsive transcription factor ERF027-like
TraesCS6B02G281000	-1.76	4.23E-09	-1.36	1.78E-07	ethylene-responsive transcription factor 1-like
TraesCS6B02G350100	-4.04	9.00E-12	-2.71	5.23E-10	ethylene-responsive transcription factor ERF109-like
TraesCS6B02G350400	-3.41	7.85E-42	-1.71	5.20E-07	ethylene-responsive transcription factor ERF109-like
TraesCS6B02G361400	-2.21	9.84E-22	-1.05	9.84E-08	ethylene-responsive transcription factor ERF014-like
TraesCS6D02G217800	1.72	1.14E-05	1.02	0.0004	ethylene-responsive transcription factor ERF054-like
TraesCS6D02G237800	-3.23	2.79E-19	-2.32	1.03E-21	ethylene-responsive transcription factor ERF027-like
TraesCS6D02G298700	-3.68	1.54E-06	-2.41	3.40E-09	ethylene-responsive transcription factor ERF109-like
TraesCS6D02G299000	-4.6422	0.007	-2.09	0.0022	ethylene-responsive transcription factor ERF109-like
TraesCS6D02G299100	-3.71	2.77E-37	-1.96	2.11E-06	ethylene-responsive transcription factor ERF109-like
TraesCS6D02G309600	-1.66	1.08E-15	-1.13	3.56E-05	ethylene-responsive transcription factor ERF014-like
Glutathione S-transferase					
TraesCS1B02G194100	-1.22	0.002	-1.11	0.0006	putative glutathione S-transferase GSTU6
TraesCS1B02G194500	-1.36	0.0005	-1.25	0.046	probable glutathione S-transferase GSTU6
TraesCS2A02G218700	-4.72	9.96E-19	-2.78	5.22E-20	probable glutathione S-transferase GSTU6
TraesCS2B02G244100	-3.87	3.27E-19	-2.39	6.06E-19	probable glutathione S-transferase GSTU6
TraesCS2B02G322800	-1.38	1.14E-06	-1.19	9.82E-05	glutathione S-transferase TCHQD
TraesCS2D02G044100	-3.99	0.03	-2.58	0.002	probable glutathione S-transferase GSTF1
TraesCS2D02G224700	-1.83	0.001	-1.72	5.55E-08	probable glutathione S-transferase GSTU6
TraesCS2D02G304500	-1.58	9.35E-06	-1.45	8.40E-12	glutathione S-transferase TCHQD
TraesCS3A02G309100	-1.92	5.18E-06	-1.14	4.15E-05	unnamed protein product
TraesCS3D02G130700	1.39	0.006	1.16	0.0006	probable glutathione S-transferase GSTU6
TraesCS3D02G486100	-1.15	9.69E-05	-1.07	3.18E-07	glutathione transferase GST 23-like
Gibberellin					
TraesCS1A02G334400	-1.81	2.73E-09	-2.32	3.52E-16	gibberellin 2-oxidase
TraesCS2A02G189600	-1.61	3.52E-11	-1.29	5.11E-09	chitin-inducible gibberellin-responsive protein 2-like
TraesCS2B02G217500	-1.59	8.03E-14	-1.32	1.55E-09	chitin-inducible gibberellin-responsive protein 2-like
TraesCS2B02G239400	-1.24	5.27E-08	-1.03	3.57E-05	chitin-inducible gibberellin-responsive protein 1
TraesCS2D02G198200	-1.36	5.50E-11	-1.46	1.23E-10	chitin-inducible gibberellin-responsive protein 2-like
TraesCS2D02G220200	-1.33	2.10E-10	-1.03	9.21E-07	chitin-inducible gibberellin-responsive protein 1
TraesCS3A02G122600	1.25	0.04	1.67	3.54E-07	gibberellin 3-oxidase 2-1
TraesCS3A02G294000	-1.73	2.71E-09	-1.45	9.93E-17	gibberellin 2-oxidase
TraesCS3B02G328700	-1.86	1.62E-08	-1.71	3.62E-07	gibberellin 2-oxidase 3 isozyme B1
TraesCS6A02G221900	-2.35	1.36E-13	-2.00	1.46E-10	gibberellin-2-oxidase-A9
TraesCS6D02G213100	-2.84	2.58E-08	-3.35	3.80E-17	gibberellin 2-beta-dioxygenase 8-like
Brassinosteroid					
TraesCS2A02G142300	1.42	4.85E-05	1.077	7.91E-05	probable cytochrome P450 313a4
TraesCS3B02G550900	-2.31	0.004	-1.70	0.0004	protein BRASSINOSTEROID INSENSITIVE 1-like isoform X1
TraesCS5B02G174400	1.24	9.18E-07	1.43	1.95E-07	receptor-like protein kinase BRI1-like 3
TraesCS5D02G181500	1.32	1.42E-07	1.56	4.46E-14	receptor-like protein kinase BRI1-like 3
Auxin					
TraesCS2A02G183900	-1.73	1.42E-12	-1.43	1.77E-13	probable indole-3-acetic acid-amido synthetase GH3.8
TraesCS5A02G378300	1.42	0.0498	1.12	3.10E-05	Auxin-responsive protein IAA30

(Continued)

TABLE 4 Continued

Gene_ ID	log_2FC (NN_NIL^{KN9204}/ NN_NIL^{J411})	*p*-adjusted	log_2FC (LN_NIL^{KN9204}/ LN_NIL^{J411})	*p*-adjusted	NR_description
TraesCS6B02G359400	-1.38	0.001	-1.07	0.005	predicted protein, partial
Cytokinin					
TraesCS1A02G156100	-3.19	2.83E-07	-2.22	3.79E-14	probable cytokinin riboside 5'-monophosphate phosphoribohydrolase LOGL10 isoform X1
TraesCS3D02G475800	-2.54	4.52E-07	-1.43	6.95E-05	cytokinin dehydrogenase 4-like

NN and LN indicate normal- and low-nitrogen conditions, respectively. AA and BB indicate the NIL^{KN9204} and the NIL^{J411}, respectively. The expression levels between the NIL^{KN9204} and the NIL^{J411} were compared under each condition, with the latter as a control. The expression ratios are presented as log_2FC values. Genes with a $log_2FC \leq -1$ (p-adjusted< 0.05, coloured blue) represent the genes whose expression was downregulated in the NIL^{KN9204}, and genes with $log_2FC \geq 1$ (p-adjusted< 0.05, coloured red) represent the genes whose expression was upregulated in the NIL^{KN9204}.

TABLE 5 Differentially expressed TF-encoding genes between the pairwise NILs.

Gene_ ID	log_2FC (NN_NIL^{KN9204}/ NN_NIL^{J411})	*p*-adjusted	log_2FC (LN_NIL^{KN9204}/ LN_NIL^{J411})	*p*-adjusted	NR_description
NTF					
TraesCS1A02G354900	0.70	1	2.16	0.02	nuclear transcription factor Y subunit C-3-like
TraesCS1B02G366800	0.64	0.89	1.83	0.0007	nuclear transcription factor Y subunit C-3-like
TraesCS1D02G355600	0.45	0.95	2.60	0.003	nuclear transcription factor Y subunit C-3-like
NAC					
TraesCS3A02G078400	1.48	0.004	1.37	3.04E-07	NAC transcription factor
TraesCS3A02G339600	-3.07	3.15E-22	-3.01	1.27E-27	NAC transcription factor 4
TraesCS3D02G078900	1.79	6.26E-06	1.39	5.36E-05	NAC transcription factor

NN and LN indicate normal- and low-nitrogen conditions, respectively. AA and BB indicate the NIL^{KN9204} and the NIL^{J411}, respectively. The expression levels between the NIL^{KN9204} and the NIL^{J411} were compared under each condition, with the latter as a control. The expression ratios are presented as log_2FC values. Genes with a $log_2FC \leq -1$ (p-adjusted< 0.05, coloured blue) represent the genes whose expression was downregulated in the NIL^{KN9204}, and genes with $log_2FC \geq 1$ (p-adjusted< 0.05, coloured red) represent the genes whose expression was upregulated in the NIL^{KN9204}.

TFs function in response to NO_3^-, depending on the NRT1.1 NO_3^- transport function (Vidal et al., 2014), to increase root biomass (He et al., 2015). Two genes encoding NAC TFs (*TraesCS3A02G078400* and *TraesCS3D02G078900*) exhibited higher expression levels in the NIL^{KN9204} than in the NIL^{J411} under both NN and LN conditions (Table 5), implying that they had similar effects on root growth. Ethylene responsive factors (ERFs) belong to the AP2/ERF (APETALA2/ERF) superfamily of transcription factors associated with various developmental processes and stress tolerance. It was reported that ERF transcription factor gene *TaSRL1* resulted in short root length when ectopically expressed in rice (Zhuang et al., 2021). In the present study, expressions of ERF-like genes were significantly down-regulated in the NIL^{KN9204} compared to the NIL^{J411} under both NN and LN conditions, which may contribute to longer root length in NIL^{KN9204}.

In conclusion, transcriptomic analysis in the present study revealed the landscape of the DEGs between the *QMrl-7B* NILs under both NN and LN conditions, reflecting the physiological basis of the large roots of the superior parent KN9204. This study used pairwise NILs for analysis to dampen the effects from different genetic backgrounds; as such, the DEGs are more likely to be related to *QMrl-7B*, either directly or indirectly. These results revealed many pathways regulated by genes that might interact genetically with *QMrl-7B*, which may provide a foundation for the application of *QMrl-7B* in molecular breeding. However, since the materials used for RNA-seq were seedlings and not mature plants, much work is still needed to elucidate the overall genetic network that controls ideal root traits of wheat plants at different stages.

Data availability statement

The data presented in the study are deposited in the NCBI repository, accession number PRJNA889686. The accession number PRJNA889686 (http://www.ncbi.nlm.nih.gov/bioproject/889686).

Author contributions

LS, JJL, LZ, NZ, WZ, DM, XR, JJ and JML conducted the hydroponic culture and field experiments. LS, JJL, LZ and AB analyzed the data. LS, JJL, YN and AB wrote the manuscript. RL and JML revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research was jointly supported by grants from the National Key Research and Development Program of China (2021YFF1000403), the National Natural Science Foundation of China (U22A6009), Hebei Natural Science Foundation (C2021205013) and China Agriculture Research System of MOF and MARA (CARS-03).

Acknowledgments

We are very grateful to Dr. Xigang Liu from Hebei Normal University, Shijiazhuang, China for his critical review of the manuscript.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1062575/full#supplementary-material

References

An, D., Su, J., Liu, Q., Zhu, Y., Tong, Y., Li, J., et al. (2006). Mapping QTLs for nitrogen uptake in relation to the early growth of wheat (Triticum aestivum l.). *Plant Soil.* 284, 73–84. doi: 10.1007/s11104-006-0030-3

Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361, eaar7191. doi: 10.1126/science.aar7191

Cao, P., Ren, Y., Zhang, K., Teng, W., Zhao, X., Dong, Z., et al. (2014). Further genetic analysis of a major quantitative trait locus controlling root length and related traits in common wheat. *Mol. Breeding.* 33, 975–985. doi: 10.1007/s11032-013-0013-z

Chen, D., Richardson, T., Chai, S., Lynne McIntyre, C., Rae, A. L., and Xue, G. P. (2016). Drought-UP-REGULATED TaNAC69-1 is a transcriptional repressor of TaSHY2 and TaIAA7, and enhances root length and biomass in wheat. *Plant Cell Physiol.* 57, 2076–2090. doi: 10.1093/pcp/pcw126

Chen, H., Wei, J., Tian, R., Zeng, Z., Tang, H., Liu, Y., et al. (2022). A major quantitative trait locus for wheat total root length associated with precipitation distribution. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.995183

Cosio, C., and Dunand, C. (2009). Specific functions of individual class III peroxidase genes. *J. Exp. Bot.* 60, 391–408. doi: 10.1093/jxb/ern318

Dalal, M., Sahu, S., Tiwari, S., Rao, A., and Gaikwad, K. (2018). Transcriptome analysis reveals interplay between hormones, ROS metabolism and cell wall biosynthesis for drought-induced root growth in wheat. plant physiol. *Biochem.* 130, 482–492. doi: 10.1016/j.plaphy.2018.07.035

Ding, X., Cao, Y., Huang, L., Zhao, J., Xu, C., Li, X., et al. (2008). Activation of the indole-3-acetic acid-amido synthetase GH3-8 suppresses expansin expression and promotes salicylate- and jasmonate-independent basal immunity in rice. *Plant Cell.* 20, 228–240. doi: 10.1105/tpc.107.055657

Fan, X., Zhang, W., Zhang, N., Chen, M., Zheng, S., Zhao, C., et al. (2018). Identification of QTL regions for seedling root traits and their effect on nitrogen use efficiency in wheat (Triticum aestivum l.). *Theor. Appl. Genet.* 131, 2677–2698. doi: 10.1007/s00122-018-3183-6

Fukaki, H., and Tasaka, M. (2009). Hormone interactions during lateral root formation. *Plant Mol. Biol.* 69, 437–449. doi: 10.1007/s11103-008-9417-2

Gelli, M., Duo, Y., Konda, A., Zhang, C., Holding, D., and Dweikat, I. (2014). Identification of differentially expressed genes between sorghum genotypes with contrasting nitrogen stress tolerance by genome-wide transcriptional profiling. *BMC Genomics* 15, 179. doi: 10.1186/1471-2164-15-179

Gewin, V. (2010). Food: An underground revolution. *Nature.* 466, 552–553. doi: 10.1038/466552a

Godfray, H. C., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., et al. (2010). Food security: the challenge of feeding 9 billion people. *Science.* 327, 812–818. doi: 10.1126/science.1185383

Gou, J., Strauss, S. H., Tsai, C. J., Fang, K., Chen, Y., Jiang, X., et al. (2010). Gibberellins regulate lateral root formation in populus through interactions with auxin and other hormones. *Plant Cell.* 22, 623–639. doi: 10.1105/tpc.109.073239

Guo, R., Hu, Y., Aoi, Y., Hira, H., Ge, C., Dai, X., et al. (2022). Local conjugation of auxin by the GH3 amido synthetases is required for normal development of roots and flowers in arabidopsis. *Biochem. Biophys. Res. Commun.* 589, 16–22. doi: 10.1016/j.bbrc.2021.11.109

Han, Y., Gao, Y., Shi, Y., Du, J., Zheng, D. , and Liu, G. (2017b). Genome-wide transcriptome profiling reveals the mechanism of the effects of uniconazole on root development in glycine max. *J. Plant Biol.* 60, 387–403. doi: 10.1007/s12374-017-0028-9

Han, H., Wang, H., Han, Y., Hu, Z., Xin, M. , Peng, H., et al. (2017a). Altered expression of the TaRSL2 gene contributed to variation in root hair length during allopolyploid wheat evolution. *Planta.* 246, 1019–1028. doi: 10.1007/s00425-017-2735-3

Hao, L., Wen, Y., Zhao, Y., Lu, W., and Xiao, K. (2015). Wheat mitogen-activated protein kinase gene TaMPK4 improves plant tolerance to multiple stresses through modifying root growth, ROS metabolism, and nutrient acquisitions. *Plant Cell Rep.* 34, 2081–2097. doi: 10.1007/s00299-015-1853-2

He, X., Ma, H., Zhao, X., Nie, S., Li, Y., Zhang, Z., et al. (2016). Comparative RNA-seq analysis reveals that regulatory network of maize root development controls the expression of genes in response to n stress. *PloS One* 11, e0151697. doi: 10.1371/journal.pone.0151697

He, X., Qu, B., Li, W., Zhao, X., Teng, W., Ma, W., et al. (2015). The nitrate-inducible NAC transcription factor TaNAC2-5A controls nitrate response and increases wheat yield. *Plant Physiol.* 169, 1991–2005. doi: 10.1104/pp.15.00568

Ho, C. H., Lin, S. H., Hu, H. C., and Tsay, Y. F. (2009). CHL1 functions as a nitrate sensor in plants. *Cell.* 138, 1184–1194. doi: 10.1016/j.cell.2009.07.004

Hu, L., Li, H., Chen, L., Lou, Y., Amombo, E., and Fu, J. (2015). RNA-Seq for gene identification and transcript profiling in relation to root growth of bermudagrass (Cynodon dactylon) under salinity stress. *BMC Genomics* 16, 575. doi: 10.1186/s12864-015-1799-3

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

Kabir, M. R., Liu, G., Guan, P., Wang, F., Khan, A. A., Ni, Z., et al. (2015). Mapping QTLs associated with root traits using two different populations in wheat (Triticum aestivum l.). *Euphytica.* 206, 175–190. doi: 10.1007/s10681-015-1495-z

Krouk, G., Lacombe, B., Bielach, A., Perrine-Walker, F., Malinska, K., Mounier, E., et al. (2010). Nitrate-regulated auxin transport by NRT1.1 defines a mechanism for nutrient sensing in plants. *Dev. Cell.* 18, 927–937. doi: 10.1016/j.devcel.2010.05.008

Kücke, M., Schmid, H., and Spiess, A. (1995). A comparison of four methods for measuring roots of field crops in three contrasting soils. *Plant Soil.* 172, 63–71. doi: 10.1007/BF00020860

Kurakawa, T., Ueda, N., Maekawa, M., Kobayashi, K., Kojima, M., Nagato, Y., et al. (2007). Direct control of shoot meristem activity by a cytokinin-activating enzyme. *Nature.* 445, 652–655. doi: 10.1038/nature05504

Li, B., Liu, D., Li, Q., Mao, X., Li, A., Wang, J., et al. (2016). Overexpression of wheat gene TaMOR improves root system architecture and grain yield in oryza sativa. *J. Exp. Bot.* 67, 4155–4167. doi: 10.1093/jxb/erw193

Li, H., Torres-Garcia, J., Latrasse, D., Benhamed, M., Schilderink, S., Zhou, W., et al. (2017). Plant-specific histone deacetylases HDT1/2 regulate GIBBERELLIN 2-OXIDASE2 expression to control arabidopsis root meristem cell number. *Plant Cell.* 29, 2183–2196. doi: 10.1105/tpc.17.00366

Liu, H., Wang, Z., Xu, W., Zeng, J., Li, L., Li, S., et al. (2020). Bacillus pumilus LZP02 promotes rice root growth by improving carbohydrate metabolism and phenylpropanoid biosynthesis. *Mol. Plant Microbe Interact.* 33, 1222–1231. doi: 10.1094/mpmi-04-20-0106-r

Liu, J., Zhang, Q., Meng, D., Ren, X., Li, H., Su, Z., et al. (2021). QMrl-7B enhances root system, biomass, nitrogen accumulation and yield in bread wheat. *Plants.* 10, 764. doi: 10.3390/plants10040764

Li, C., Wang, J., Li, L., Li, J., Zhuang, M., Li, B., et al. (2022). TaMOR is essential for root initiation and improvement of root system architecture in wheat. *Plant Biotechnol. J.* 20, 862–875. doi: 10.1111/pbi.13765

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

Lynch, J. (2007). Roots of the second green revolution. *Aust. J. Bot.* 55, 493–512. doi: 10.1071/BT06118

Maccaferri, M., El-Feki, W., Nazemi, G., Salvi, S., Canè, M. A., Colalongo, M. C., et al. (2016). Prioritizing quantitative trait loci for root system architecture in tetraploid wheat. *J. Exp. Bot.* 67, 1161–1178. doi: 10.1093/jxb/erw039

Maqbool, S., Hassan, M. A., Xia, X., York, L. M., Rasheed, A., and He, Z. (2022).). root system architecture in cereals: progress, challenges and perspective. *Plant J.* 110, , 23–, 42. doi: 10.1111/tpj.15669

Meister, R., Rajani, M. S., Ruzicka, D., and Schachtman, D. P. (2014). Challenges of modifying root traits in crops for agriculture. *Trends Plant Sci.* 19, 779–788. doi: 10.1016/j.tplants.2014.08.005

Merchuk-Ovnat, L., Fahima, T., Ephrath, J. E., Krugman, T., and Saranga, Y. (2017). Ancestral QTL alleles from wild emmer wheat enhance root development under drought in modern wheat. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.00703

Moubayidin, L., Mambro, R. D., and Sabatini, S. (2009). Cytokinin-auxin crosstalk. *Trends Plant Sci.* 14, , 557–, 562. doi: 10.1016/j.tplants.2009.06.010

Muro-Villanueva, F., Mao, X., and Chapple, C. (2019). Linking phenylpropanoid metabolism, lignin deposition, and plant growth inhibition. *Curr. Opin. Biotechnol.* 56, 202–208. doi: 10.1016/j.copbio.2018.12.008

Neogy, A., Singh, Z., Mushahary, K., and Yadav, S. (2021). Dynamic cytokinin signaling and function of auxin in cytokinin responsive domains during rice crown root development. *Plant Cell Rep.* 40, 1367–1375. doi: 10.1007/s00299-020-02618-9

Pattyn, J., Vaughan-Hirsch, J., and Poel, B. (2020). The regulation of ethylene biosynthesis: a complex multilevel control circuitry. *New Phytol.* 229, 770–782. doi: 10.1111/nph.16873

Qu, B., He, X., Wang, J., Zhao, Y., Teng, W., Shao, A., et al. (2015). A wheat CCAAT box-binding transcription factor increases the grain yield of wheat with less fertilizer input. *Plant Physiol.* 167, 411–423. doi: 10.1104/pp.114.246959

Shi, X., Cui, F., Han, X., He, Y., Zhao, L., Zhang, N., et al. (2022). Comparative genomic and transcriptomic analyses uncover the molecular basis of high nitrogen-use efficiency in the wheat cultivar kenong 9204. *Mol. Plant* 15, 1440–1456. doi: 10.1016/j.molp.2022.07.008

Springmann, M., Clark, M., Mason-D'Croz, D., Wiebe, K., Bodirsky, B. L., Lassaletta, L., et al. (2018). Options for keeping the food system within environmental limits. *Nature.* 562, 519–525. doi: 10.1038/s41586-018-0594-0

Staswick, P. E., Serban, B., Rowe, M., Tiryaki, I., Maldonado, M. T., Maldonado, M. C., et al. (2005). Characterization of an arabidopsis enzyme family that conjugates amino acids to indole-3-acetic acid. *Plant Cell.* 17, 616–627. doi: 10.1105/tpc.104.026690

Tal, I., Zhang, Y., Jorgensen, M. E., Pisanty, O., Barbosa, I. C. R., Zourelidou, M., et al. (2016). The arabidopsis NPF3 protein is a GA transporter. *Nat. Commun.* 7, 11486. doi: 10.1038/ncomms11486

Vidal, E. A., Alvarez, J. M., and Gutierrez, R. A. (2014). Nitrate regulation of AFB3 and NAC4 gene expression in arabidopsis roots depends on NRT1.1 nitrate transport function. *Plant Signal. Behav.* 9, e28501. doi: 10.4161/psb.28501

Wang, R. F., An, D. G., Hu, C. S., Li, L. H., Zhang, Y. M., Jia, Y. G., et al. (2011). Relationship between nitrogen uptake and use efficiency of winter wheat grown in the north China plain. *Crop Pasture Sci.* 62, 504–514. doi: 10.1071/CP10383

Wang, Y., Cheng, Y., Chen, K. , and Tsay, Y. (2018b). Nitrate transport, signaling, and use efficiency. *Annu. Rev. Plant Biol.* 69, 85–122. doi: 10.1146/annurev-arplant-042817-040056

Wang, H., Hu, Z., Huang, K., Han, Y., Zhao, A. , Han, H., et al. (2018a). Three genomes differentially contribute to the seedling lateral root number in allohexaploid wheat: evidence from phenotype evolution and gene expression. *Plant J.* 95, 976–987. doi: 10.1111/tpj.14005

Wei, Z., and Li, J. (2016). Brassinosteroids regulate root growth, development, and symbiosis. *Mol. Plant* 9, 86–100. doi: 10.1016/j.molp.2015.12.003

Xu, G., Fan, X., and Miller, A. (2012). Plant nitrogen assimilation and use efficiency. *Annu. Rev. Plant Biol.* 63, 153–182. doi: 10.1146/annurev-arplant-042811-105532

Xu, Y., Ren, Y., Lin, T., and Cui, D. (2019). Identification and characterization of CircRNAs involved in the regulation of wheat root length. *Biol. Res.* 52, 19. doi: 10.1186/s40659-019-0228-5

Yang, J., Wang, M., Li, W., He, X., Teng, W., Ma, W., et al. (2019). Reducing expression of a nitrate-responsive bZIP transcription factor increases grain yield and n use in wheat. *Plant Biotechnol. J.* 17, 1823–1833. doi: 10.1111/pbi.13103

Zhang, H., and Forde, B. G. (1998). An arabidopsis MADS box gene that controls nutrient-induced changes in root architecture. *Science.* 279, 407–409. doi: 10.1126/science.279.5349.40

Zheng, Z., Guo, Y., Novák, O., Chen, W., Ljung, K., Noel, J., et al. (2016). Local auxin metabolism regulates environment-induced hypocotyl elongation. *Nat. Plants.* 2, 16025. doi: 10.1038/nplants.2016.25

Zheng, X., Wen, X., Qiao, L., Zhao, J., Zhang, X., Li, X., et al. (2019). A novel QTL QTrl.saw-2D.2 associated with the total root length identified by linkage and association analyses in wheat (Triticum aestivum l.). *Planta.* 250, 129–143. doi: 10.1007/s00425-019-03154-x

Zhuang, M., Li, C., Wang, J., Mao, X., Li, L., Yin, J., et al. (2021). The wheat SHORT ROOT LENGTH 1 gene TaSRL1 controls root length in an auxin-dependent pathway. *J. Exp. Bot.* 72, 6977–6989. doi: 10.1093/jxb/erab357

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 07 December 2022
DOI 10.3389/fpls.2022.1057701

OPEN ACCESS

EDITED BY
Ling-Ling Chen,
Guangxi University, China

REVIEWED BY
Yi Wang,
Sichuan Agricultural University, China
Shulin Xue,
Henan University, China

*CORRESPONDENCE
Shuyu Liu
Shuyu.Liu@ag.tamu.edu

†PRESENT ADDRESS
Shuhao Yu,
Department of Horticulture and Landscape Architecture, Oklahoma State University, Stillwater, OK, United States
Frank Maulana,
California Cooperative Rice Research Foundation, Inc., Biggs, CA, United States
Wangqi Huang,
State Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan, Yunnan Agricultural University, Kunming, Yunnan, China
Xue-Feng Ma,
Forage Genetics International, West Salem, WI, United States

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

RECEIVED 29 September 2022
ACCEPTED 04 November 2022
PUBLISHED 07 December 2022

CITATION
Wang Z, Dhakal S, Cerit M, Wang S, Rauf Y, Yu S, Maulana F, Huang W, Anderson JD, Ma X-F, Rudd JC, Ibrahim AMH, Xue Q, Hays DB, Bernardo A, St. Amand P, Bai G, Baker J, Baker S and Liu S (2022) QTL mapping of yield components and kernel traits in wheat cultivars TAM 112 and Duster.
Front. Plant Sci. 13:1057701.
doi: 10.3389/fpls.2022.1057701

QTL mapping of yield components and kernel traits in wheat cultivars TAM 112 and Duster

Zhen Wang[1], Smit Dhakal[1], Mustafa Cerit[1], Shichen Wang[2], Yahya Rauf[1], Shuhao Yu[1†], Frank Maulana[3†], Wangqi Huang[3†], Joshua D. Anderson[3], Xue-Feng Ma[3†], Jackie C. Rudd[1], Amir M. H. Ibrahim[4], Qingwu Xue[1], Dirk B. Hays[4], Amy Bernardo[5], Paul St. Amand[5], Guihua Bai[5], Jason Baker[1], Shannon Baker[1] and Shuyu Liu[1*]

[1]Texas A&M AgriLife Research and Extension Center, Amarillo, TX, United States, [2]Genomics and Bioinformatics Service Center, Texas A&M AgriLife Research, College Station, TX, United States, [3]Noble Research Institute, Ardmore, OK, United States, [4]Department of Soil and Crop Science, Texas A&M University, College Station, TX, United States, [5]Central Small Grain Genotyping Lab and Hard Winter Wheat Genetics Research Unit, U.S. Department of Agriculture-Agricultural Research Service, Manhattan, KS, United States

In the Southern Great Plains, wheat cultivars have been selected for a combination of outstanding yield and drought tolerance as a long-term breeding goal. To understand the underlying genetic mechanisms, this study aimed to dissect the quantitative trait loci (QTL) associated with yield components and kernel traits in two wheat cultivars `TAM 112' and `Duster' under both irrigated and dryland environments. A set of 182 recombined inbred lines (RIL) derived from the cross of TAM 112/Duster were planted in 13 diverse environments for evaluation of 18 yield and kernel related traits. High-density genetic linkage map was constructed using 5,081 single nucleotide polymorphisms (SNPs) from genotyping-by-sequencing (GBS). QTL mapping analysis detected 134 QTL regions on all 21 wheat chromosomes, including 30 pleiotropic QTL regions and 21 consistent QTL regions, with 10 QTL regions in common. Three major pleiotropic QTL on the short arms of chromosomes 2B (57.5 - 61.6 Mbps), 2D (37.1 - 38.7 Mbps), and 7D (66.0 - 69.2 Mbps) colocalized with genes *Ppd-B1*, *Ppd-D1*, and *FT-D1*, respectively. And four consistent QTL associated with kernel length (KLEN), thousand kernel weight (TKW), plot grain yield (YLD), and kernel spike^{-1} (KPS) (*Qklen.tamu.1A.325*, *Qtkw.tamu.2B.137*, *Qyld.tamu.2D.3*, and *Qkps.tamu.6A.113*) explained more than 5% of the phenotypic variation. QTL Qklen.tamu.1A.325 is a novel QTL with consistent effects under all tested environments. Marker haplotype analysis indicated the QTL combinations significantly increased yield and kernel traits. QTL and the linked markers identified in this study will facilitate future marker-assisted selection (MAS) for pyramiding the favorable alleles and QTL map-based cloning.

KEYWORDS

wheat, QTL mapping, multiple-environment trials, yield components, kernel traits

Introduction

Bread wheat (*Triticum aestivum* L.) is one of the most common staple foods in the human daily diet, providing calories and proteins for one-third of the world population (Shewry and Hey, 2015). By the end of this century, the projected world population is expected to be 10.9 billion, leading to a significantly higher wheat consumption forecast (Enghiad et al., 2017; Roser et al., 2019). Additionally, droughts, diseases and pest prevalence increase concern about insufficient wheat production (Yue et al., 2019). To meet future demands, novel wheat cultivars with higher yield potential under stressed environments need to be developed (Trethowan and Mujeeb-Kazi, 2008).

Yield is a complex quantitative trait involving multiple QTL, epistatic, and genetic by environment interactions. Multiple environment QTL mapping analysis can reveal the genetic architecture and QTL stability, which could facilitate the effective utilization of genetic resources in breeding by marker-assisted selection (MAS) (Collard and Mackill, 2008). Yield consists of three main components, kernels spike^{-1} (KPS), spikes m^{-2} (SPM), and thousand kernel weight (TKW) (Cao et al., 2020). Most yield-related QTL may influence one or multiple yield components (Mason et al., 2013; Yu et al., 2018; Juliana et al., 2019; Li et al., 2019; Yang et al., 2020). Analysis of wheat breeding during practices from the past several decades show that yield improvement was mainly due to the continuous gains in KPS and SPM (Feng et al., 2018; Würschum et al., 2018). QTL for KPS have been reported on all wheat chromosomes from various sources (Assanga et al., 2017; Li et al., 2018; Liu et al., 2018; Yu et al., 2018; Muhammad et al., 2020; Pang et al., 2020). Several genes have been well-characterized and known for controlling KPS. For example, *FLOWERING LOCUS T1* gene (*TaFT1*) on chromosome group seven, also known as *VRN3*, which controls KPS, heading date (HD), as well as the integrator of several autonomous pathways (Yan et al., 2006; Lv et al., 2014). Deletion of *FT-B1* or *FT-D1* in hexaploid wheat increased the KPS and delayed the HD (Finnegan et al., 2018; Shaw et al., 2019; Glenn et al., 2021; Chen et al., 2022). Functional *WHEAT ORTHOLOG of APO1* (*WAPO1*) gene on chromosome group seven also increased KPS in wheat (Ikeda et al., 2007; Kuzay et al., 2019; Kuzay et al., 2022). For SPM, QTL have been identified on all wheat chromosomes (Assanga et al., 2017; Li et al., 2018; Ma et al., 2018; Yu et al., 2018; Li et al., 2020; Pang et al., 2020; Isham et al., 2021; Ren et al., 2021). Several genes associated with SPM have been characterized, including *TaD27* on chromosome group seven encoding a strigolactones biosynthesis enzyme that regulates the tiller number in many plant species, including wheat (Zhao et al., 2020). Additionally, *TaPIL1* on chromosome group five reduced the tiller number by interacting with *TaSPL3/17* and regulating *TaTB1* transcription (Zhang et al., 2021).

Thousand kernel weight is another important yield component, which has intensively selected during domestication and modern breeding (Golan et al., 2015; Qin et al., 2015). Similar to KPS and SPM, TKW QTL have been detected on all wheat chromosomes (Assanga et al., 2017; Sun et al., 2017; Li et al., 2018; Ma et al., 2018; Duan et al., 2020; Muhammad et al., 2020; Pang et al., 2020; Isham et al., 2021; Ren et al., 2021). *TaGL3* (Yang et al., 2019a), *TaGS1* (Guo et al., 2013), *TaGS3* (Yang et al., 2019b), *TaGS5* (Wang et al., 2015), *TaGW2* (Su et al., 2011), *TaGW7* (Wang et al., 2019), *TaGW8* (Yan et al., 2019), *TaTGW6* (Hanif et al., 2016), and *TaTGW7* (Hu et al., 2016), have been characterized for association with kernel size and TKW. Genes with pleiotropic effects on multiple yield-related traits of wheat were also reported, such as *Ppd1*, responsible for photoperiod and HD regulation. The copy number variations of *Ppd-B1* facilitated the adaption of wheat worldwide by modifying the HD (Würschum et al., 2015). Besides affecting HD, *Ppd-B1* also showed pleiotropic effects on KPS, SPM, and TKW in durum wheat (Arjona et al., 2018).

Although abundant QTL were characterized in previous studies, only a few were used in wheat breeding since the majority of them showed inconsistent or minor effects in different genetic backgrounds (Zheng et al., 2021). To avoid the unfavorable impact on QTL deployment, and concurrently meet the breeding purposes for the Southern Great Plains, we used the two most drought tolerant cultivars 'TAM 112' and 'Duster' in Texas and Oklahoma, respectively, to construct a mapping population. TAM 112 and Duster significantly differed in TKW, KPS, and kernel shape while both maintained outstanding yield performance in the Southern Great Plains region, where the extreme drought stresses frequently threaten the wheat production. In order to provide insight into the genomic regions contributing to high yield under water-limited conditions, this study evaluated the yield and agronomic traits of the mapping population in diverse environments. The consistent and pleiotropic QTL were identified using single nucleotide polymorphisms (SNPs), and their effects were assessed through haplotype analysis.

Material and methods

Plant material

The mapping population contains 182 F_6 recombinant inbred lines (RILs) derived from the cross between 'TAM 112' (PI 643143) and 'Duster' (PI 644016). TAM 112 is an awned, red glume, medium-early maturing, semidwarf hard red winter wheat released by Texas A&M AgriLife Research in 2005 (Rudd et al., 2014). Carrying *1AL.1RS* chromosome translocation from rye (*Secale cereale* L), TAM 112 has greenbug [*Schizaphis graminum* (Rondani)] and wheat curl mite (*Aceria tosichella* Keifer) resistance,

also the excellent grain yield potential, end-use quality, and drought tolerance. Duster is an awned, white glume, intermediate heading time, dual-purpose, semidwarf hard red winter wheat cultivar released by the Oklahoma Agricultural Experiment Station and the USDA-ARS in 2006 (Edwards et al., 2012). It shows resistance to *Soilborne wheat mosaic virus* (SBWMV), *Wheat spindle streak mosaic virus* (WSSMV), and Hessian fly (*Mayetiola destructor* Say).

Field experiments

The RIL population and both parents were planted under different irrigation regimes at seven locations across Texas and Oklahoma from 2017 to 2019 with a total of 13 environments (combination of treatments, years, and locations) (Table S1). The four Texas dryland environments are Bushland middle school (35° 06' N, 102° 27' W) in 2017, 2018 and 2019 (17BMS, 18BMS, 19BMS), and Chillicothe (34° 07' N, 99° 18' W) in 2017 (17CH). Six Texas irrigated environments are Texas A&M AgriLife Research stations in Bushland (35° 06' N, 102° 27' W) in 2018 (18BI), Bushland south pivot (35° 06' N, 102° 27' W) in 2017, 2018 and 2019 (17BSP100, 18BSP100, 19BSP100 at full irrigation, and 17BSP67 at 67% irrigation level) and Dumas (35° 51' N, 101° 58' W) in 2018 (18DMS). The three Oklahoma dryland environments include Noble Foundation Dupy Farm (34.29° N, 96.99° W) in 2018 (18NFD) and Red River Research and Demonstration Farm (33.88°N, 97.28°W) in 2017 and 2019 (17RRD, 19RRD).

The field experiments employed an Alpha Lattice design with incomplete blocks each containing five plots. All individual environments had two replications, except one replication for 18DMS and three replications for 18NFD and 19RRD. The plot size of all Oklahoma locations was 2.32 m^2. However, for Texas locations, the plot size was 4.64 m^2 for irrigated land and 6.96 m^2 for dryland. Standard management was employed for all field trials (Assanga et al., 2017).

Phenotyping and statistical analysis

A set of 18 traits from 13 individual environments (Table S1) were collected following the same procedures as described by Assanga et al. (2017). After maturity, a biomass sample was collected by manually harvesting a 0.5-meter-long evenly growing inner row in each plot, and the rest of plots were harvested by a combine harvester. The combine-harvested seeds were weighed for grain yield (YLD) and test weight (TW). Biomass samples were dried in a drying chamber for more than 72 h at 60°C and then weighed to calculate total biomass weight (BM). Spikes from each biomass sample were counted, weighed, and converted to SPM and single head dry weight (SHDW). The spike length (SL) was measured in 17BSP100 and 18BI. All spikes were threshed to get seeds and weighed to calculate the biomass yield (BMYLD).

The threshed seeds of biomass samples from 18BSP100, 18DMS, 18BI, 18BMS, 19BSP100, and 19BMS were further scanned for kernel traits with an HP G4010 (HP 11956A, Hewlett-Packard, Plao Alto, CA, USA) scanner, and seed weights of the scanned seeds were recorded. The scan resolution was 300 dots per inch (DPI). Seed images were analyzed using grain scan v1.0 software (Whan et al., 2014) with parameters set as minimum grain width = 20, minimum grain length = 20, and auto threshold sensitivity = 0.8. Kernel area (AREA), kernel perimeter (PERI), kernel length (KLEN), kernel width (KWID), and kernel number were automatically output from the software. Two methods were employed to estimate TKW: dividing the scanned grain weight by kernel number and multiplying by 1000 in these locations using the seed scanner, and by directly hand-counting and weighing 200 threshed seeds to estimate TKW for the rest of environments. Kernels $spike^{-1}$ was calculated using BMYLD, TKW, and SPM. Single culm weight (SCW) was computed from BM divided by SPM. Harvest Index (HI) was calculated using BMYLD divided by BM. Single head dry grain weight (SHGW) was computed using BMYLD and the number of spikes.

The raw phenotypic data was first filtered manually to remove the outliers that were beyond three standard deviations from the mean. Mega-environments were classified based on the principal component analysis (PCA). Analysis of variance (ANOVA) was conducted to estimate the significance of genotypes (G), environments (E), and genotype-by-environment interactions (G x E). The best linear unbiased estimate (BLUE) values were calculated for all phenotypic data by fitting the mixed linear model $Y_{ijkl} = \mu + Env_i + Rep_j(Env_i) + Block_k(Env_i\ Rep_i) + Gen_l + Env_i x Genl + \epsilon_{ijkl}$, where genotypes and environments were fixed and random effects, respectively, under both individual environments and mega-environments. In the mixed linear model, *Y* represents the phenotype, *Env*, *Rep*, *Block*, and ϵ are random effects that accounted for the effects of environment, replication, block, and error, respectively. *Gen* is fixed genotypic effect. The broad-sense heritability was calculated on entry-mean basis using the following formula:

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \sigma_{ge}^2/nLoc + \sigma_g^2/(nLoc \times nRep)}$$

where σ_g^2, σ_{ge}^2, and σ_g^2, are the estimated variances of genotype, genotype-by-environment, and error, while nLoc and nRep indicate the number of environments and replications, respectively. The Multiple Environment Trait Analysis with R (META-R) software (Alvarado et al., 2020) based on the R package lme4 (Bates et al., 2015) was used for PCA, ANOVA, and heritability analysis. Removal of outliers from the raw data as well as BLUE based Pearson correlation coefficients among variables were analyzed by JMP pro 16 (Sall et al., 2017).

Genotyping and linkage map construction

DNA of parents and 182 RILs was extracted from the young leaves using the modified cetyltrimethylammonium bromide (CTAB) protocol at 2 to 3 leaf stage (Liu et al., 2014). The quality of the extracted DNA was evaluated using agarose gel electrophoresis. About 200 ng genomic DNA per sample was used for genotyping-by-sequence (GBS) library construction in the USDA Central Small Grain Genotyping Center (CSGGC), Manhattan, KS (https://hwwgenotyping.ksu.edu/genotyping.html) following the protocol from Poland and Rife (2012). In brief, normalized genomic DNA was double-digested using restriction enzymes *PstI* and *MspI* (New England Biolabs Inc., Ipswich, MA) and ligated to a set of barcode adapters using T4 ligase (New England Biolabs Inc). The GBS library was size-selected for 200-300 bp fragments in an E-gel system (ThermoFisher Scientific). The selected polymerase chain reaction (PCR) fragments were purified with the GenCatch PCR purification kit (Epoch Life Science Inc., Sugar Land, TX), and quantified with the Qubit dsDNA HS assay kit (ThermoFisher Scientific) before sequencing using PI v3 chips and Hi-Q sequencing kits (ThermoFisher Scientific) in an Ion Torrent Proton sequencer (Life Technologies, Carlsbad, CA). The raw sequence reads were assigned to each sample based on attached barcodes, then trimmed to 64 bp DNA sequences including the 5' restriction site. SNPs were detected using a reference-based pipeline in TASSEL 5.0 (Bradbury et al., 2007) using the IWGSC RefSeq v1.1 genome (IWGSC et al., 2018) to assign the physical positions for each SNP.

The homozygous SNP genotypes of each RIL were converted to parental genotypes for 'A' as the female parent TAM 112 and 'B' as the male parent Duster. The heterozygous SNPs were considered missing values. SNPs with no calls (NN) from parental genotypes and missing value ratio of more than 20%, or A/B ratio outside of 0.5 to 2 were removed before linkage map construction. Filtered GBS data was inspected, and the false double crossovers were manually removed according to the physical location of the SNPs. Redundant and distorted makers were purged by JoinMap 4.0 software (Van Ooijen, 2006). Only one SNP was retained for linkage map construction when multiple SNPs have more than 99.5% similarity. SNPs significantly deviated from 1:1 Mendelian segregation ratios ($P< 0.01$) were also removed for the chi-square (χ^2) test. Finally, the remaining unique SNPs were grouped by setting independent logarithm of odds (LOD) threshold ranging from 3 to 30, with maximum likelihood (ML) as the mapping algorithm and Kosambi's regression mapping (Kosambi, 2016) as mapping function to construct the linkage map.

QTL and haplotype analysis

BLUE values from individual and mega-environments were used for QTL analysis using the inclusive composite interval mapping (ICIM) function in QTL IciMapping v 4.1 (Meng et al., 2015). Additive and additive by environment interaction (A x E) effects of QTL based on single allele were determined by ICIM additive mapping (ICIM-ADD) function. The threshold LOD score to declare a significant QTL with a type I error less than 0.05 was determined using one thousand permutations. The step length in QTL scanning was one centimorgan (cM), and the probability in stepwise regression (PIN) was 0.001.

The epistasis effects were analyzed using the epistatic mapping (ICIM-EPI) function. The threshold LOD score was 10 for reducing interaction number, step length was 5 cM, and PIN was 0.0001. QTL names were designated following procedure described by McIntosh et al. (2013) with slight modifications in the format as: *Qtrait.tamu.chrom.Mbps*, where 'Q' represents QTL, 'trait' is the trait name, 'tamu' is for the 'Texas A&M University', 'chrom' represent chromosome number, and 'Mbps' is the physical position aligned to the IWGSC Chinese Spring RefSeq v 1.0 (https://wheat-urgi.versailles.inra.fr/Seq-Repository/Assemblies). If multiple QTL were mapped in the same genomic region with overlapping flanking intervals, they were considered to belong to the same QTL region. A 'consistent QTL' was defined as the QTL mapped on the same physical position for the same trait that was significant in at least two individual environments. While 'pleiotropic QTL' was defined as a QTL mapped on the same physical position for different traits that were not highly correlated.

For each trait, the allelic effects of a QTL were estimated using the average trait BLUE values based on marker haplotype groups. The single peak marker which explained the maximum phenotypic variation was used to distinguish the allele source. RILs were grouped based on marker haplotypes for five traits, including YLD, KPS, TKW, KLEN, and KWID. The RILs carrying different alleles were analyzed based on their allelic combinations in the same individual environment. For statistical analysis, Tukey's honest significance difference (HSD) test was conducted in JMP pro 16 (Sall et al., 2017) to compare the trait BLUE values among haplotype groups at $P< 0.05$.

Results

Phenotype and heritability

Distributions of all traits across testing environments were depicted using boxplots (Figure S1). Grain yield averages ranged from 115.8 to 564.2 g m^{-2} while BM averages ranged from 964.7 to 2331.0 g m^{-2}. Biomass yield averages ranged from 312.3 to 705.8 g m^{-2}. The ranges of HI and TW were 16.3 to 39.5% and 706.0 to 794.8 Kg m^{-3}. The three yield components, TKW, SPM, and KPS, had average values ranging from 22.2 to 32.6 g, 534.5 to 1124.9 spikes m^{-2}, and 17.1 to 35.1 kernel spike^{-2} respectively. For means of kernel traits, KLEN ranged from 4.7 to 5.4 mm and KWID ranged from 2.3 to 2.7 mm. In addition, average HD

ranged from 112.4 to 134.4 days while average PH ranged from 54.3 to 84.0 cm.

The variance of genotype, environment, and G x E for all traits were significant ($p< 0.001$), except the variance of genotype for BM and G x E for SPM (Table S2). TAM 112 had higher values of BM, SHDW, SHGW, TKW and all four kernel traits, while Duster had higher values of KPS, YLD, HD and PH in the mean performance as well as the overall BLUE values across all individual environments. No significant difference was observed for BMYLD, KPS, SPM, HI, TW, and PH between the two parents. The entry mean based broad-sense heritability was moderate to high for all traits (0.51 – 0.95), except BM (0.10) (Table S2).

Correlations among traits

Significant positive correlations were observed between BMYLD and SPM (r = 0.43 – 0.77, $p< 0.001$), TKW (r = 0.21 – 0.46, $p< 0.01$), KPS (r = 0.26 – 0.59, $p< 0.001$) in most environments (Table S3). Consistent negative correlations were observed between KPS and TKW (r = -0.22 – -0.51, $p< 0.01$) and between SPM and TKW (r = -0.22 – -0.35, $p< 0.01$). However, the correlation between KPS and SPM was not significant (r = -0.18 – 0.04) in most of the environments, suggesting that KPS and SPM can be improved simultaneously. Both TW (r = 0.27 – 0.52, $p< 0.001$) and HI (r = 0.16 – 0.66, $p< 0.001$) were significantly and positively correlated with YLD in all environments, except in 19BMS for TW, and in 17BSP67 and 17BSP100 for HI. Heading date and PH had inconsistent correlations with YLD under individual environments, and the overall correlation value was negative. In addition, moderate to high positive correlations were observed between TKW and all four kernel size traits (r = 0.57 – 0.97, $p< 0.001$) from all six environments in 2018 and 2019.

Mega-environment classification

Mega environments were assigned based on the biplot data for each trait across multiple environments. Only the traits directly related to yield were analyzed for the mega-environment. YLD, SPM, and HI had three mega-environments (ME1 – ME3), while BMYLD, KPS, TKW, KLEN, and KWID had two mega-environments (ME1 – ME2) (Figure S2 and Table S4).

Genetic map construction

A total of 23,609 SNPs between TAM 112 and Duster were called and 7,139 polymorphic SNPs were retained for linkage map construction after removal of monomorphic makers or markers with missing datapoints >20% or heterozygotes >10%. Furthermore, 712 distorted and 1,346 redundant SNPs were removed from the final map. A total of 5,081 unique SNPs were assigned to 45 linkage groups corresponding to 21 chromosomes. The final map spanned 2,399.5 cM in genetic distance or 11,540.9 mega-base-pairs (Mbps) in physical distance, with an average marker density of 2.12 SNPs cM^{-1} and 0.44 SNPs $Mbps^{-1}$. The SNPs in A, B, and D genomes were 1,517, 2,581, and 983, respectively (Table S5). The SNPs density maps that based on genetic and physical positions were shown in Figure S3.

QTL analysis

A total of 134 QTL regions were identified by analyzing BLUE values of 16 traits from individual environments, across individual environments, within mega-environment, and across mega-environments (Figures 1, S4, S5 and Table S6). Among them, 21 consistent QTL for 12 traits on 10 chromosomes were identified in at least two individual environments (Table 1). A set of 30 QTL showing pleiotropic effects to at least two traits without strong correlations were mapped on 12 chromosomes. Ten QTL were in common between consistent and pleiotropic QTL (Table S6).

Yield and yield component traits

A set of 10 consistent QTL associated with yield and yield components was mapped on chromosomes 2B, 2D, 2A, 6A, 6B, and 7D (Table 1 and Figure 1). Among them, the QTL on 2D at 37.1 – 38.9 Mbps showed consistent effects on YLD, BMYLD, KPS, TKW, and HI. For YLD and BMYLD, *Qyld.tamu.2D.37* had LOD scores ranging 3.7 – 17.4, PVE values ranging 2.8 – 17.6%, and increased YLD up to 11.7 g m^{-2} in 17RRD. While *Qbmyld.tamu.2D.37* had LOD scores ranging 3.8 – 15.4, PVE ranging 9.0 – 26.5%, and increased BMYLD up to 28.3 g m^{-2} in 17BSP100. The favorable alleles that increased YLD and BMYLD were from Duster. For TKW and KPS, *Qtkw.tamu.2D.37* had LOD scores of 3.7 – 37.1, PVE values of 6.1 – 23.4%, and increased 1.4 g of TKW in 17BSP100 with the Duster allele. *Qkps.tamu.2D.37* had LOD scores of 4.3 – 20.8, PVE of 5.0 – 13.5%, and increased KPS up to 1.10 kernel $spike^{-1}$ in 19BMS with the TAM 112 allele. In addition, a consistent YLD QTL was also identified on chromosome 2D at 2.5 – 3.3 Mbps with the favorable allele from Duster that increased YLD up to 14.2 g m^{-2} in 17RDD, with LOD scores of 5.8 – 18.3, and PVE of 6.1 – 11.0%. Another consistent KPS QTL was identified on 6A at 113 Mbps with the favorable allele from TAM 112 that had LOD scores ranging 3.8 – 17.3, PVE ranging 7.1 – 11.6, and increased KPS up to 1.4 kernel $spike^{-1}$ in 17BMS. Moreover, another seven consistent QTL for TKW were identified on chromosomes 2B

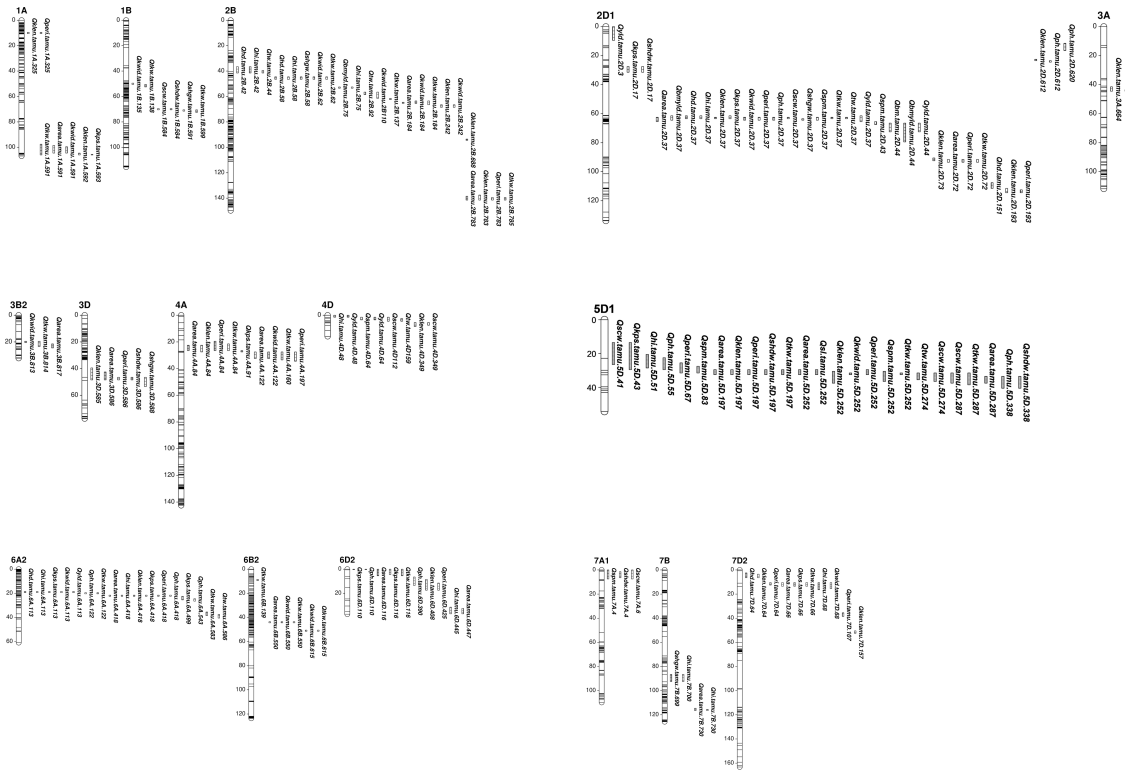


FIGURE 1
Consistent and pleiotropic QTL for yield related, agronomy and kernel traits identified with TAM112/Duster population by individual and multiple environment analysis. Traits include: 1) grain yield from whole plot (YLD, g m^{-2}); 2) dry weight of biomass sample (BM, g m^{-2}); 3) grain yield from biomass sample (BMYLD, g m^{-2}); 4) test weight (TW, Kg m^{-3}); 5) harvest index (HI, %); 6) kernel $spike^{-1}$ (KPS, kernel $spike^{-1}$); 7) spike m^{-2} (SPM, spike m^{-2}); 8) thousand kernel weight (TKW, g); 9) single head dry weight (SHDW, mg); 10) single head grain weight (SHGW, mg); 11) single culm weight (SCW, g); 12) kernel area (AREA, mm^2); 13) kernel perimeter (PERI, mm); 14) kernel length (KLEN, mm); 15) kernel width (KWID, mm); 16) spike length (SL, cm); 17) heading date (HD, days); 18) plant height (PH, cm).

(61.6, 136.9, and 784.7 Mbps), 2D (72.7 Mbps), 4A (83.8 – 90.5 Mbps), 6B (138.7 Mbps), and 7D (66 – 69.2 Mbps) that explained 2.8 – 23.4% of TKW variations with LOD scores of 3.5 – 37.1. Except for *Qtkw.tamu.4A.84*, all these TKW QTL had favorable alleles from TAM 112 that increased TKW up to 1.5 g. In summary, all YLD and BMYLD QTL were increased by Duster alleles, while all KPS QTL and most of TKW QTL were increased by TAM 112 alleles (Table 1). Although there were 11 QTL regions associated with SPM, none of them were consistent QTL (Table S6 and Figure 1).

Kernel size traits

Eleven kernel size related consistent QTL were identified on chromosomes 1A, 2B, 2D, 3A, 4A, 5D, and 7D (Table 1 and Figure 1). For AREA, five consistent QTL were located on chromosomes 2D (37.2 – 38.7 Mbps, 71.9 – 72.7 Mbps), 4A (83.8 – 90.5 Mbps), 5D (196.9 Mbps), and 7D (66 – 67.7 Mbps). They had LOD scores ranging 3.7 – 26.7, PVE ranging 2.0 – 12.9%, and increased AREA up to 0.34 mm^2. The beneficial alleles of *Qarea.tamu.2D.72* and *Qarea.tamu.7D.66* were from TAM 112 while the rest three QTL had beneficial alleles from Duster. Six KLEN QTL were mapped on chromosomes 1A (324.6 Mbps), 2B (688.4 Mbps, 781.2 – 784.7 Mbps), 3A (664.0 – 667.5 Mbps), 4A (83.8 – 90.5 Mbps), and 7D (156.8 Mbps). The LOD scores and PVE values of these QTL ranged 3.6 – 56.5 and 1.9 – 33.6, respectively. The additive effect of *Qklen.tamu.3A.664* for KLEN was up to 0.2 mm in 19BMS, with the favorable alleles from Duster. Only two consistent KWID QTL were mapped on chromosomes 2D (37.1 – 37.4 Mbps) and 7D (62.2 – 66.0 Mbps) that increased KWID up to 0.037 mm^2, with LOD scores of 3.5 – 17.2, and PVE values of 7.3 – 16.2%. The favorable alleles were from Duster and TAM 112, respectively. Moreover, six consistent QTL that increased PERI up to 0.23 mm were located on chromosomes 1A (324.6 Mbps), 2D (37.1 – 38.7 Mbps, 71.9 – 72.7 Mbps), 4A (83.8 – 90.5 Mbps), and 7D (62.2 – 66.0 Mbps, 107.0 – 109.8 Mbps) with LOD scores ranging 3.6 – 22.8 and PVE values ranging 1.5 – 15.7%. All of the PERI QTL had favorable alleles from TAM 112, except

TABLE 1 Consistent and pleiotropic QTL for yield related, agronomy and kernel traits identified with TAM112/Duster population by individual and multiple environment analysis.

QTL name	Chrom	Peak Position (Mb)[a]	Traits[b]	Threshold	LOD[c]	LOD (A)	LOD (AbyE)	PVE (%)[d]	PVE (A) (%)	PVE (AbyE)	Add effect[e]	Parental favorable allele	LeftMarker	Left SNP alleles	TAM 112 alleles	Consistent QTL	Pleiotropic QTL	Environments[f]
Qarea.tamu.2D.37	2D	37.2 - 38.7	AREA	3.4 - 6.1	4.6 - 22.2	1.7 - 11.8	4.8 - 10.5	2.0 - 12.9	0.6 - 4.5	1.3 - 8.0	-0.34 - -0.04	Duster	chr2D_35558557	G/A	G	y	y	18BI, 18BMS, 18DMS, 19BSP100, 19BMS, across 6 envs
Qarea.tamu.2D.72	2D	71.9 - 72.7	AREA	3.4 - 6.1	5.2 - 16.2	14.1	2.0	5.1 - 6.7	5.3	1.3	0.13 - 0.21	TAM 112	chr2D_71764015	C/T	C	y		18BI, 19BSP100, across6 envs
Qarea.tamu.4A.84	4A	83.8 - 90.5	AREA	3.4 - 6.1	4.1 - 14.6	9.2	5.3	5.5 - 6.3	3.4	2.8	-0.24 - -0.1	Duster	chr4A_30316423	G/C	G	y		18DMS, 19BSP100, across 6 envs
Qarea.tamu.5D.197	5D	196.9	AREA	3.4	3.7 - 6.4			7.3 - 8.4			-0.2 - -0.16	Duster	chr5D_70896874	G/A	G	y		18BSP100, 19BMS
Qarea.tamu.5D.252	5D	252.0	AREA	3.4 - 6.1	6.1 - 26.9	21.7	5.2	7.7 - 9.8	8.30	1.50	-0.23 - -0.16	Duster	chr5D_233797558	C/T	C		y	18BI, across 6 envs
Qarea.tamu.6D.116	6D	115.7	AREA	3.4 - 6.1	4.1 - 9.1	7.4	1.7	3.3 - 4.6	2.7	0.6	0.09 - 0.16	TAM 112	chr6D_97915036	G/A	G		y	19BMS, across 6 envs
Qarea.tamu.7D.66	7D	66 - 67.7	AREA	3.4 - 6.1	8.2 - 26.7	16.9	9.8	9.6 - 12	6.2	3.3	0.1 - 0.3	TAM 112	chr7D_64236400	G/T	G	y	y	18BI, 18BMS, across 6 envs
Qbm.tamu.2D.44	2D	43.7	BM	7.7	8.0	0.5	7.5	14.5	1.6	12.9	6.9	TAM 112	chr2D_42872584	A/C	A		y	across 9 envs
Qbmyld.tamu.2D.44	2D	43.7	BMYLD	3.4	3.7			9.2			-16.9	Duster	chr2D_42872584	A/C	A		y	18BSP100
Qbmyld.tamu.2D.37	2D	37.1 - 37.4	BMYLD	3.4 - 7.5	3.8 - 15.4	2.2 - 5.0	3.8 - 13.2	9.0 - 26.5	4.7 - 10.1	2.9 - 21.8	-28.3 - -6.2	Duster	chr2D_35558557	G/A	G	y	y	17BSP67, 17BSP100, ME2, ME12, across 9 envs
Qhd.tamu.2B.42	2B	41.8	HD	3.4 - 6.3	7.3 - 9.5	4.8	4.7	3.7 - 7.6	1.6	2.1	-0.92 - -0.27	Duster	chr2B_35320841	A/T	A		y	18BSP100, across 6 envs
Qhd.tamu.2B.58	2B	57.5	HD	3.4 - 6.3	4.5 - 10.7	4.1	4.9	3.3 - 7.3	1.4	2.0	-1.4 - -0.24	Duster	chr2B_57246450	A/C	A	y	y	17BMS, 18BI, 18BMS, across 6 envs
Qhd.tamu.2D.37	2D	37.4 - 38.7	HD	3.4 - 6.3	20.4 - 82.8	56.0	26.7	22.0 - 39.9	21.3	18.6	0.96 - 2.2	TAM 112	chr2D_35558557	G/A	G	y	y	18BI, 18BMS, 18DMS, 18BSP100, 19BSP100, across 6 envs
Qhd.tamu.6A.113	6A	113.0	HD	3.4 - 6.3	6.5 - 8.8	5.1	3.7	2.6 - 5.3	1.7	0.9	-0.65 - -0.27	Duster	chr6A_122456115	C/T	C		y	18BMS, across 6 envs
Qhd.tamu.7D.64	7D	64.4 - 67.7	HD	3.4 - 6.3	4.4 - 64.5	57.8	6.7	2.9 - 25.3	21.3	2.5	-1.7 - -0.7	Duster	chr7D_64236400	G/T	G	y	y	17BMS, 18BI, 18BMS, 18DMS, 18BSP100, 19BSP100, across 6 envs
Qhi.tamu.2B.42	2B	41.8	HI	3.4 - 4.0	4.2 - 4.5	2.5	2.0	3.7 - 5.9	3.20	0.6	0.59 - 0.86	TAM 112	chr2B_35320841	A/T	A		y	18BSP100, ME3
Qhi.tamu.2B.58	2B	57.5	HI	3.4 - 7.5	4.5 - 8.1	4.1 - 6.3	0.5 - 4.0	5.3 - 11.9	2.1 - 8.1	2.2 - 3.8	0.3 - 1.1	TAM 112	chr2B_57246450	A/C	A		y	17BSP100, ME2, ME123, across 9 envs
Qhi.tamu.2D.37	2D	37.4 - 38.7	HI	3.4 - 7.5	7.4 - 46.7	25.2	21.4	13.4 - 40.7	22.9	17.7	-2.2 - -0.74	Duster	chr2D_35558557	G/A	G	y	y	17BSP67, 18DMS, 18BSP100, 19BSP100, ME2, ME3, ME123, across 9 envs

(Continued)

10.3389/fpls.2022.1057701

TABLE 1 Continued

QTL name	Chrom	Peak Position (Mb)[a]	Traits[b]	Threshold	LOD[c]	LOD (A)	LOD (AbyE)	PVE (%)[d]	PVE (A) (%)	PVE (AbyE)	Add effect[e]	Parental favorable allele	LeftMarker	Left SNP alleles	TAM 112 alleles	Consistent QTL	Pleiotropic QTL	Environments[f]
Qhi.tamu.6A.113	6A	113.0	HI	3.4 - 7.5	4.5 - 10.7	3.0 - 8.0	1.5 - 2.7	6.1 - 12.4	4.0 - 5.9	2.1 - 6.5	0.4 - 0.94	TAM 112	chr6A_122456115	C/T	C		y	18BI, ME1, across 9 envs
Qhi.tamu.7D.68	7D	67.7	HI	3.4 - 7.5	7.9 - 11.8	3.7 - 5.9	4.6 - 5.9	5.5 - 11.6	3.0 - 4.7	2.1 - 2.6	0.3 - 1.2	TAM 112	chr7D_64236400	G/T	G		y	18BSP100, ME3, across 9 envss
Qklen.tamu.1A.325	1A	324.6	KLEN	3.4 - 6.3	6.3 - 63	11.4 - 59.9	0.05 - 5.7	5.7 - 19.5	4.5 - 18.0	0.02 - 1.2	0.05 - 0.10	TAM 112	chr1A_306143327	G/A	G	y		18BI, 18BSP100, 18DMS, 18BMS, 19BSP100, 19BMS, ME1, ME2, ME12, across 6 envs
Qklen.tamu.2B.688	2B	688.4	KLEN	3.4 - 6.3	3.6 - 12.1	4.0 - 8.1	0.4 - 4.0	1.9 - 6.5	1.8 - 4.3	0.3 - 1.0	-0.06 - -0.03	Duster	chr2B_690936220	A/G	A	y		18BMS, 19BMS, ME1, ME2, across 6 envs
Qklen.tamu.2B.783	2B	781.2 - 784.7	KLEN	3.4 - 6.3	5.6 - 22.4	6.1 - 19.8	0.8 - 2.6	3.7 - 6.2	3.2 - 5.2	0.0015 - 0.70	0.04 - 0.07	TAM 112	chr2B_779319734	C/T	C	y		18BMS, 18BSP100, 19BMS, ME1, ME2, ME12, across 6 envs
Qklen.tamu.2D.37	2D	37.4 - 38.7	KLEN	4.1 - 4.2	4.5 - 6.5	4.3 - 6.0	0.2 - 0.5	2.0 - 3.1	1.9 - 2.2	0.04 - 0.86	-0.04 - -0.03	Duster	chr2D_39423582	A/T	A		y	ME2, ME12
Qklen.tamu.3A.664	3A	664 - 667.3	KLEN	3.4 - 6.3	3.8 - 56.5	39.9	16.6	4.0 - 33.6	20.6	13.5	-0.2 - 0.05	Duster	chr3A_659529719	C/T	C	y		18BSP100, 19BMS, 19BSP100, ME2, ME12, across 6 envs
Qklen.tamu.4A.84	4A	83.8 - 90.5	KLEN	3.4 - 6.3	3.9 - 8.4	2.5 - 6.2	0.17 - 4.9	2.1 - 6.6	0.9 - 2.8	0.049 - 1.2	-0.06 - -0.02	Duster	chr4A_30316423	G/C	G	y		18DMS, 19BSP100, ME2, ME12, across 6 envs
Qklen.tamu.5D.252	5D	252.0	KLEN	3.4 - 6.3	4.4 - 11.9	7.7 - 10.9	0.3 - 1.7	3.3 - 6.5	2.6 - 6.5	0.04 - 1.9	-0.056 - -0.034	Duster	chr5D_233797558	C/T	C		y	18BMS, ME1, ME12, across 6 envs
Qklen.tamu.7D.157	7D	156.8	KLEN	3.4 - 6.3	6.1 - 23.9	7.3 - 15.4	0 - 12.6	4.7 - 8.5	1.9 - 7.6	0.16 - 4.1	0.029 - 0.088	TAM 112	chr7D_156192771	A/G	A	y		19BMS, 19BSP100, ME2, ME12, across 6 envs
Qklen.tamu.7D.64	7D	64.4 - 67.7	KLEN	3.4 - 4.1	4.2 - 6.9	4.0 - 5.1	0.2 - 1.3	3.2 - 10.3	1.9 - 3.2	0.8 - 1.3	0.03 - 0.08	TAM 112	chr7D_64236400	G/T	G		y	18BI, ME1, ME12
Qkps.tamu.2D.37	2D	37.4 - 38.7	KPS	3.4 - 7.6	4.3 - 20.8	2.4 - 14	1.4 - 10.9	5.0 - 13.5	2.8 - 10.6	1.7 - 3.4	0.32 - 1.1	TAM 112	chr2D_35558557	G/A	G	y	y	17BSP67, 18BI, 19BMS, ME1, ME2, ME12, across 9 envs
Qkps.tamu.6A.113	6A	113.0	KPS	3.4 - 7.6	3.8 - 17.3	5.8 - 11.8	3.2 - 5.5	7.1 - 11.6	4.7 - 6.7	2.5 - 4.9	0.4 - 1.4	TAM 112	chr6A_122456115	C/T	C	y	y	17BMS, 18BSP100, 19BSP100, ME1, across 9 envs
Qkps.tamu.6D.110	6D	110.4 - 115.7	KPS	3.4 - 7.6	5.4 - 13.9	4.2 - 9.2	0.2 - 4.7	5.3 - 10.9	3.6 - 8.1	0.3 - 2.1	-0.8 - -0.4	Duster	chr6D_97915036	G/A	G		y	19BSP100, ME1, ME12, across 9 envs
Qkps.tamu.7D.66	7D	66.0 - 67.7	KPS	3.4 - 7.6	4.3 - 11.9	5.3 - 8.2	0.6 - 3.8	4.1 - 6.7	3.2 - 5.3	0.4 - 1.2	-0.7 - -0.4	Duster	chr7D_64236400	G/T	G		y	18BI, ME1, ME12, across 9 envs
Qkwid.tamu.2B.62	2B	61.6	KWID	3.4	4.8			8.8			0.04	TAM 112	chr2B_57246450	A/C	A		y	18DMS
Qkwid.tamu.2D.37	2D	37.1 - 37.4	KWID	3.4 - 6.3	5.3 - 9.8	3.1 - 7	0.3 - 6.7	7.3 - 13.1	2.2 - 9.2	0.5 - 5.5	-0.036 - -0.009	Duster	chr2D_35558557	G/A	G	y	y	18BMS, 19BSP100, ME1, ME2, across 6 envs
Qkwid.tamu.5D.252	5D	252.0	KWID	4.1	4.4	3.0	1.4	1.9	1.3	0.5	-0.012	Duster	chr5D_233797558	C/T	C		y	ME12
Qkwid.tamu.6A.113	6A	113.0	KWID	3.4	3.6			4.1			-0.02	Duster	chr6A_122456115	C/T	C		y	18BSP100

(Continued)

10.3389/fpls.2022.1057701

TABLE 1 Continued

QTL name	Chrom	Peak Position (Mb)[a]	Traits[b]	Threshold	LOD[c]	LOD (A)	LOD (AbyE)	PVE (%)[d]	PVE (A) (%)	PVE (AbyE)	Add effect[e]	Parental favorable allele	LeftMarker	Left SNP alleles	TAM 112 alleles	Consistent QTL	Pleiotropic QTL	Environments[f]
Qkwid.tamu.7D.68	7D	62.2 - 66	KWID	3.4 - 6.3	3.5 - 17.2	12.1 - 13.2	0.6 - 4.2	7.6 - 16.2	5.9 - 13	0.8 - 2.3	0.018 - 0.037	TAM 112	chr7D_64236400	G/T	G	y	y	18BI, 18BMS, ME1, ME12, across 6 envs
Qperi.tamu.1A.325	1A	324.6	PERI	3.4 - 6.2	4.5 - 22.8	19.3	3.4	3.1 - 15.7	7.4	2.1	0.13 - 0.21	TAM 112	chr1A_306143327	G/A	G	y		18BI, 18BSP100, 19BMS, across 6 envs
Qperi.tamu.2D.37	2D	37.1 - 38.7	PERI	3.4 - 6.2	4.1 - 14.2	5.6	8.5	4.5 - 9.2	2.1	2.4	-0.02 - -0.07	Duster	chr2D_35558557	G/A	G	y	y	18BMS, 18DMS, 18BMS, 19BSP100, across 6 envs
Qperi.tamu.2D.72	2D	71.9 - 72.7	PERI	3.4 - 6.2	3.6 - 13.0	8.9	4.1	3.4 - 8.3	3.3	2.2	0.085 - 0.17	TAM 112	chr2D_71764015	C/T	C	y		18BMS, 18BSP100, 19BSP100, across 6 envs
Qperi.tamu.4A.84	4A	83.8 - 90.5	PERI	3.4 - 6.2	3.7 - 10.1	6.8	3.3	1.5 - 7.7	2.4	1.4	-0.178 - -0.072	Duster	chr4A_30316423	G/C	G	y		18DMS, 19BMS, 19BSP100, across 6 envs
Qperi.tamu.5D.252	5D	252.0	PERI	3.4	9.5			8.9			-0.17	Duster	chr5D_233797558	C/T	C		y	18BMS
Qperi.tamu.7D.107	7D	107 - 109.8	PERI	3.4 - 3.4	4.8 - 5.2			2.1 - 7.2			0.16 - 0.17	TAM 112	chr7D_110098182	T/C	T	y		19BMS, 19BSP100, across 6 envs
Qperi.tamu.7D.64	7D	62.2 - 66	PERI	3.4 - 6.2	3.9 - 14.2	8	6.3	4.2 - 12.8	2.9	2.6	0.08 - 0.23	TAM 112	chr7D_53634310	A/G	A	y	y	18BI, 18BMS, across 6 envs
Qph.tamu.2D.37	2D	37.1 - 37.4	PH	3.4 - 7.0	3.8 - 25.5	12.2	13.3	3.6 - 11.3	3.3	4.8	0.7 - 1.9	TAM 112	chr2D_35558557	G/A	G	y	y	18BI, 18DMS, 19BMS, 19BSP100, across 8 envs
Qph.tamu.2D.620	2D	619.7 - 620.2	PH	3.4 - 7.0	3.6 - 12.4	9.4	3.1	3.5 - 7.6	2.3	1.2	0.6 - 1.7	TAM 112	chr2D_620685649	T/G	T	y		18BI, 19BMS, across 8 envs
Qph.tamu.5D.338	5D	338 - 350.8	PH	3.4 - 7	4.4 - 14.3	9.0	5.3	4.9 - 8.3	2.3	2.8	-2.23 - -0.62	Duster	chr5D_269148535	G/A	G	y		18BI, 18DMS, across 8 envs
Qph.tamu.6A.122	6A	121.8	PH	3.4 - 7.0	4.2 - 59.6	45.7	13.8	8.0 - 21.8	12.9	8.5	-3.2 - -1.1	Duster	chr6A_115036912	G/A	G	y		18BI, 18BMS, 18DMS, 19BMS, 19BSP100, across 8 envs
Qph.tamu.6A.418	6A	418.4	PH	3.4 - 7.0	4.4 - 30.1	8.4	21.7	8.0 - 13.6	2.2	5.9	-3.2 - -0.6	Duster	chr6A_406554946	T/G	T	y		17BSP100, 18BSP100, across 8 envs
Qph.tamu.6D.110	6D	110.4	PH	3.4 - 7.0	4.8 - 9.0	5.8	3.2	2.0 - 10.4	1.5	0.5	0.5 - 1.2	TAM 112	chr6D_110425575	T/C	T		y	18BSP100, across 8 envs
Qscw.tamu.2D.37	2D	37.4	SCW	3.4 - 7.6	5.5 - 10.9	6.6	4.3	2.7 - 15.2	2.3	0.4	0.026 - 0.045	TAM 112	chr2D_37467471	A/T	A		y	18BSP100, across 9 envs
Qscw.tamu.5D.274	5D	273.8 - 299.5	SCW	3.4 - 7.6	4.0 - 22.8	15.9	6.8	8.3 - 15	5.4	2.8	-0.10 - -0.04	Duster	chr5D_269148535	G/A	G	y		18BMS, 18DMS, across 9 envs
Qscw.tamu.7A.6	7A	5.7 - 7.0	SCW	3.4 - 7.6	3.6 - 10.8	7.7	3.2	4.9 - 9.5	2.5	2.4	-0.097 - -0.027	Duster	chr7A_2579229	G/A	G	y		19BMS, 19BSP100, across 9 envs
Qshgw.tamu.2B.58	2B	57.5 - 61.6	SHGW	3.4 - 7.5	5.9 - 10.8	6.2	4.6	5.1 - 11.2	3.1	2.0	10.0 - 30.0	TAM 112	chr2B_57246450	A/C	A		y	17BSP100, across 9 envs
Qshgw.tamu.2D.37	2D	37.1	SHGW	3.5 - 7.5	3.6 - 8.3	0.9	7.4	4.2 - 7.3	0.5	3.7	-22.8 - -3.9	Duster	chr2D_37224692	C/T	C		y	17BSP67, across 9 envs

(Continued)

10.3389/fpls.2022.1057701

TABLE 1 Continued

QTL name	Chrom	Peak Position (Mb)[a]	Traits[b]	Threshold	LOD[c]	LOD (A)	LOD (AbyE)	PVE (%)[d]	PVE (A) (%)	PVE (AbyE)	Add effect[e]	Parental favorable allele	LeftMarker	Left SNP alleles	TAM 112 alleles	Consistent QTL	Pleiotropic QTL	Environments[f]
Qsl.tamu.5D.252	5D	252.0	SL	3.4 - 4.2	5.9 - 7.4	1.7	5.7	4.9 - 8.8	2.2	2.7	-0.18 - -0.10	Duster	chr5D_233797558	C/T	C		y	17BSP100, across 2 envs
Qspm.tamu.2D.43	2D	42.9	SPM	3.4 - 7.6	6.3 - 9.5	1.2 - 5.8	1.4 - 8.3	2.7 - 15.2	0.6 - 7.7	1.4 - 2.1	-30.8 - -6.1	Duster	chr2D_42872548	T/C	T		y	17BSP100, ME3, across 9 envs
Qspm.tamu.2D.37	2D	37.1 - 37.4	SPM	3.4 - 4.6	4.9 - 5.2	1.7	3.5	5.0 - 12.9	1.9	3.1	-33.0 - -7.5	Duster	chr2D_37467471	A/T	A		y	18BSP100, ME123
Qspm.tamu.4D.64	4D	64.2	SPM	4.6 - 7.5	10.0 -10.4	5.2 - 7.8	2.6 - 4.7	5.0 - 8.5	3.5 - 6.1	1.5 - 2.5	-13.4 - -3.6	Duster	chr4D_61853616	A/C	A		y	ME123, across 9 envs
Qspm.tamu.5D.252	5D	252.0	SPM	7.5	8.1	6.1	2.0	3.7	2.7	1.0	13.3	TAM 112	chr5D_233797558	C/T	C		y	across 9 envs
Qtkw.tamu.2B.62	2B	61.6	TKW	3.4 - 7.6	7.7 - 22.5	8.9 - 12.6	2.2 - 12.7	4.7 - 18.0	4.0 - 11.3	2.4 - 6.7	0.3 - 1.0	TAM 112	chr2B_57246450	A/C	A	y	y	17BSP67, 17BSP100, ME1, ME12, across 9 envs
Qtkw.tamu.2B.137	2B	136.9	TKW	3.4	3.5 - 4.4			5.9 - 9.2			0.75 - 0.81	TAM 112	chr2B_138085124	C/T	C	y		17BMS, 18DMS
Qtkw.tamu.2B.785	2B	784.7	TKW	3.4 - 7.6	5.8 - 18.7	5.9 - 8.8	0.4 - 4.7	4.4 - 12.9	3.6 - 5.8	0.6 - 2.9	0.3 - 1.5	TAM 112	chr2B_781532381	A/C	A	y		17BSP67, 17BSP100, ME1, ME12, across 9 envs
Qtkw.tamu.2D.37	2D	37.4 - 38.7	TKW	3.4 - 7.6	3.7 - 37.1	6.9 - 29.8	0.02 - 19.8	6.1 - 23.4	5.1 - 22.6	0.7 - 11.8	-1.4 - -0.4	Duster	chr2D_35558557	G/A	G	y	y	17BSP100, 18BI, 18DMS, 18BMS, 19BMS, ME1, ME2, ME12, across 9 envs
Qtkw.tamu.2D.72	2D	71.9 - 72.7	TKW	3.4 - 7.6	3.7 - 9.9	4.4 - 8.4	0.6 - 1.7	3.9 - 5.2	3.1 - 3.4	0.8 - 1.3	0.3 - 0.7	TAM 112	chr2D_71764015	C/T	C	y		18BI, 19BSP100, ME2, ME12, across 9 envs
Qtkw.tamu.4A.84	4A	83.8 - 90.5	TKW	3.4 - 7.6	3.7 - 19	5.5 - 12.1	0.5 - 7	2.8 - 8.7	4.2 - 5.0	0.7 - 3.7	-0.8 - -0.36	Duster	chr4A_30316423	G/C	G	y		17BSP67, 17BSP100, 18DMS, 19BMS, 19BSP100, ME1, ME2, across 9 envs
Qtkw.tamu.5D.252	5D	252.0	TKW	4.8	8.0	7.8	0.1	5.8	5.6	0.2	-0.5	Duster	chr5D_233797558	C/T	C		y	ME2
Qtkw.tamu.6A.583	6A	583.2 - 584.4	TKW	5.3 - 7.6	8.3 - 10.0	2.1 - 2.9	6.2 - 7.1	4.0 - 7.1	1.2 - 1.9	2.8 - 5.2	-0.18 - -0.22	Duster	chr6A_574253510	G/A	G		y	ME1, across 8 envs
Qtkw.tamu.6B.139	6B	138.7	TKW	3.4 - 4.3	3.6 - 7.4	5.4	2.0	2.8 - 5.4	3.5	1.9	0.4 - 0.7	TAM 112	chr6B_135720799	C/G	C	y		17BSP67, 17BSP100, ME12
Qtkw.tamu.6D.116	6D	115.7	TKW	4.3	4.3	4.3	0.04	2.8	2.7	0.1	0.4	TAM 112	chr6D_97915036	G/A	G		y	ME12
Qtkw.tamu.7D.66	7D	66 - 69.2	TKW	3.4 - 7.6	9.3 - 26.2	8.3 - 18	2.5 - 8.9	6.9 - 16.8	5.3 - 13.7	1.6 - 4.3	0.43 - 1.2	TAM 112	chr7D_64236400	G/T	G	y	y	18BI, 18BMS, ME2, ME12, across 9 envs
Qtw.tamu.2B.44	2B	43.8	TW	6.1	6.7	5.6	1.1	5.3	3.6	1.7	2.8	TAM 112	chr2B_42242271	A/G	A		y	across 6 envs
Qtw.tamu.2D.37	2D	37.4 - 38.7	TW	3.4 - 6.1	4.5 - 15.2	10.2	5.1	10.3 - 25.8	7	12.7	-8.8 - -0.6	Duster	chr2D_35558557	G/A	G	y	y	17RDD, 18NFD, 19RRD, 19BSP100, across 6 envs
Qtw.tamu.6A.586	6A	585.5 - 586.7	TW	3.5 - 6.1	3.5 - 8.3	5.2	3.1	5.5 - 6.2	3.4	2.8	-2.7 - -6.5	Duster	chr6A_583269008	A/G	A		y	19RDD, across 6 envs
Qyld.tamu.2D.44	2D	43.7	YLD	4.7 - 8.0	6.5 - 8.3	1.1 - 3.0	5.3	7.8 - 10.0	1.8	6.0 - 8.2	-2.6 - -2.1	Duster	chr2D_42872584	A/C	A		y	ME123, across 9 envs

(Continued)

10.3389/fpls.2022.1057701

TABLE 1 Continued

QTL name	Chrom	Peak Position (Mb)[a]	Traits[b]	Threshold	LOD[c]	LOD (A)	LOD (AbyE)	PVE (%)[d]	PVE (A) (%)	PVE (AbyE)	Add effect[e]	Parental favorable allele	LeftMarker	Left SNP alleles	TAM 112 alleles	Consistent QTL	Pleiotropic QTL	Environments[f]
Qyld.tamu.2D.3	2D	2.5 - 3.3	YLD	3.3 - 8.0	5.8 - 18.3	4.2 - 8.8	2.3 - 13.0	6.1 - 11.0	3.3 - 6.9	0.5 - 2.9	-14.2 - -3.6	Duster	chr2D_682149	G/T	G	y		17CH, 17RRD, 19RRD, ME2, ME123, across 10 envs
Qyld.tamu.2D.37	2D	37.1 - 38.7	YLD	3.4 - 8.0	3.7 - 17.4	0.6 - 6.8	0.8 - 13.4	2.8 - 17.6	0.4 - 9.9	1.9 - 7.8	-11.7 - -1.2	Duster	chr2D_35558557	G/A	G	y	y	17CH, 17RRD, 19RDD, ME2, ME3, ME123, across 10 envs
Qyld.tamu.4D.64	4D	64.2	YLD	3.4 - 8.0	10.6 - 15.7	5.4 - 7.8	5.6 - 10.3	10.4 - 15.8	3.4 - 8.8	4.7 - 7.1	-14.2 - -3.6	Duster	chr4D_61853616	A/C	A		y	19BSP100, ME1, across 9 envs
Qyld.tamu.6A.113	6A	113.0	YLD	3.4 - 4.7	3.7 - 5.1	2.9	2.2	9.3 - 15.6	4.7	10.9	3.3 - 12.6	TAM 112	chr6A_122456115	C/T	C		y	18BI, ME123

[a]Physical position based on IWGSC RefSeq v 1.0 Mega base pair position.

[b]Abbreviation of traits: grain yield form whole plot (YLD), thousand kernel weight (TKW), kernels spike^{-1} (KPS), spikes m^{-2} (SPM), harvest index (HI), dry weight of biomass sample (BM), grain yield from biomass sample (BMYLD), single head dry weight (SHDW), single head grain weight (SHGW), single culm weight (SCW), heading date (HD), plant height (PH), test weight (TW), kernel area (AREA), kernel perimeter (PERI), kernel length (KLEN), kernel width (KWID), spike length (SL).

[c]Logarithm of the odds value for overall, additive effect (A), and additive by environment interaction effect (AbyE).

[d]Phenotypic variance explained for overall, additive effect (A), and additive by environment interaction effect (AbyE).

[e]Additive effect value from single allele for positive value with favorable allele from TAM 112 and negative value with favorable allele from Duster.

[f]QTL mapping under individual environment, across all environments, within mega-environment, and across mega-environments.Abbreviation of environments: Texas A&M AgriLife Research stations in Bushland, TX, Bushland middle school dryland in 2017, 2018, 2019 (17BMS, 18BMS, 19BMS), Bushland irrigated in 2018 (18BI), and Bushland South Pivot with two irrigated levels (67% and 100%) in 2017, 2018, 2019 (17BSP67, 17BSP100, 18BSP100, 19BSP100), Red River Demonstration Farm, CO, in 2017, 2019 (17RRD, 19RRD), Noble Foundation Dupy Farm, CO, in 2018 (18NFD), Dumas, TX, in 2018 (18DMS), Chillicothe, TX, in 2017 (2017CH). Mega-environments (MEs) are as follow: BMYLD: ME1 (17BSP67, 18BSP100), ME2 (17BMS, 17BSP100, 19BSP100), KLEN: ME1 (19BMS, 19BSP100), ME2 (18BI, 18BMS), KWID: ME1 (19BMS, 19BSP100), ME2 (18BI, 18BMS), HI: ME1 (18BI, 19BMS), ME2 (17BSP67, 17BSP100, 19BSP100), ME3 (17BMS, 18BSP100), KPS: ME1 (17BMS, 17BSP67, 17BSP100, 19BSP100), ME2 (18BI, 18BMS, 19BMS), SPM: ME1 (19BMS, 19BSP100), ME2 (17BMS, 18BMS, 18BI), ME3 (17BSP67, 17BSP100), TKW: ME1 (17BMS, 17BSP67, 17BSP100, 19BSP100), ME2 (18BI, 18BMS, 19BMS), YLD: ME1 (18BI, 19BMS, 19BSP100), ME2 (17CH, 17BMS, 19RRD), ME3 (17RRD, 18BSP100, 18NFD).

10.3389/fpls.2022.1057701

Qperi.tamu.2D.37 and *Qperi.tamu.4A.84*. Among all of detected kernel size related QTL, TAM 112 and Duster each contributed half of the favorable alleles (Table 1).

In addition, colocalizations were observed for kernel size related consistent QTL. QTL on chromosome 1A at 324.6 Mbps was linked to KLEN and PERI. QTL on 2D at 37.1 – 38.7 Mbps and 7D at 62.2 – 67.7 Mbps were associated with AREA, KWID, PERI. QTL on 2D at 71.9 – 72.7 Mbps was linked to AREA and PERI while the QTL on 4A at 83.8 - 90.5 Mbps was associated with AREA, PERI, and KLEN (Table 1 and Figure 1).

Agronomy traits

In total, eight agronomic consistent QTL were detected on chromosomes 2B, 2D, 5D, 6A, 7A, and 7D (Table 1 and Figure 1). One major QTL on 2D at 37.1 – 38.7 Mbps consistently influenced HD, PH, TW, and increased these traits up to 2.2 days, 2.2%, 1.9 cm, and 8.8 Kg m^{-3}. For HD, *Qhd.tamu.2D.37* had LOD scores of 20.4 – 82.8 and PVE values of 22.0 – 39.9%. For PH, *Qph.tamu.2D.37* had LOD scores of 3.8 – 25.5 and PVE values of 3.6 – 11.3%. The LOD scores of Qhi.tamu.2D.37 and Qtw.tamu.2D.37 were ranged from 7.4 – 46.7, 4.5 – 15.2 while their PVE values were ranged from 13.4 – 40.7% and 10.3 – 25.8%, respectively. The increasing allele was from TAM 112 for HD and PH, and from Duster for TW. Furthermore, we detected four consistent PH QTL located on chromosomes 2D (619.7 – 620.2 Mbps), 5D (338 – 350.8 Mbps), 6A (121.8 Mbps, 418.4 Mbps) that increased PH up to 3.2 cm, explained 3.5 – 21.8% of phenotype variations, and with LOD scores of 3.6 – 59.6. All three QTL had favorable alleles from Duster except *Qph.tamu.2D.620*. Two HD QTL were consistently detected on chromosomes 2B (57.6 Mbps) and 2D (37.4 – 38.7 Mbps) that increased HD up to 1.7 days with favorable alleles from Duster. Their LOD score and PVE value ranged 4.4 – 64.5 and 2.9 – 25.3%, respectively. Moreover, we detected two consistent QTL for SCW on chromosomes 5D (273.8 – 299.5 Mbps) and 7A (5.7 – 7.0 Mbps) with favorable alleles from Duster. Two SCW QTL had LOD scores of 3.6 – 22.8 and PVE values of 4.9 – 15%, which increased SCW up to 0.1 g. Nearly all above mentioned QTL had increasing alleles from Duster, while only one QTL for HD and two QTL for PH had favorable alleles from TAM 112 (Table 1).

Pleiotropic QTL

The QTL analysis identified 30 pleiotropic QTL regions associated with at least two traits without strong correlations and 13 QTL regions in common with consistent QTL regions (Tables 1, S6). Pleiotropic QTL linked to yield and yield-related traits were detected on chromosomes 2B, 2D, 4D, 5D, 6A, 6D, and 7D. Among them, four QTL regions with large effects were classified as major pleiotropic QTL. The first QTL region was located on chromosome 2B between 57.5 and 61.6 Mbps with TAM 112 alleles that increased HI, TKW, and KWID, while HD was delayed by the Duster allele. The second QTL region was located on chromosome 2D at 37.1 – 38.7 Mbps, where the Duster alleles increased AREA, BMYLD, HI, KLEN, KWID, PERI, SPM, TKW, TW, and YLD, while TAM 112 allele increased HD, KPS, and PH. The third QTL region on chromosome 6A at 113.0 Mbps, where the TAM 112 alleles increased HI, KPS, and YLD, while Duster alleles increased HD and KWID. The fourth QTL region on chromosome 7D at 62.2 – 69.2 Mbps increased AREA, HI, KLEN, KWID, PERI, and TKW with TAM 112 alleles, while Duster alleles increased HD and KPS.

Moreover, six pleiotropic QTL regions with minor effects were also identified (Figure 1 and Tables 1, S6). The pleiotropic QTL on chromosome 2B at 41.8 – 43.8 Mbps was related to HD, HI, and TW. On chromosome 2D at 42.9 – 43.7 Mbps, a QTL region was associated with BM, SPM, BMYLD, and YLD. Another QTL controlling SPM and YLD was identified on chromosome 4D at 64.2 Mbps. The QTL region located on chromosome 5D at 252.0 Mbps increased AREA, SL, KLEN, KWID, PERI, and TKW with Duster alleles but TAM 112 alleles increased SPM. Another QTL region on chromosome 6A at 583.2 – 586.7 Mbps increased TKW and TW with alleles from Duster. Furthermore, the QTL region on 6D at 110.4 – 115.7 Mbps increased KPS with the Duster allele while TAM 112 alleles increased TKW, AREA, and PH (Figure 1 and Tables 1, S6).

Epistatic interactions

A total of 185 epistatic interactions (EI) with a LOD score ≥ 10 were identified for fifteen yield, yield-related, kernel and agronomic traits from individual environments, within, and across mega-environments (Figure S6 and Table S7). Among all EI, two interactions for AREA and one for each SPM, TW, and SCW explained 2.2 – 2.5% of the phenotypic variation. For YLD, a total of 16 interactions were identified, including two interactions involving the consistent QTL *Qyld.tamu.2D.3* that interacted with a locus on 2D at 605.0 Mbps. Its epistasis by environment interactions (AA x E) effects increased YLD by 9.0 g m^{-2} in 18BSP100. Meanwhile, this QTL also interacted with another locus on 7D at 598.7 Mbps and the AA x E effects increased YLD by 10.6 g m^{-2} in 18DMS. In addition, the interaction between SNPs on 1B at 171.7 Mbps and 7B at 60.5 Mbps had an epistatic effect which increased YLD by 4.6 g m^{-2}. AA x E effects increased yield by 4.8 g m^{-2} in 19RRD with the TAM 112 allele, while it increased yield by 8.4 g m^{-2} in 18BI with the Duster allele. Although four EI were detected for the KPS, none of them were involved with mapped major QTL. Within 68 identified EI for TKW, six interactions were involved with five detected QTL (*Qtkw.tamu.2B.62*, *Qtkw.tamu.4A.84*,

Qtkw.tamu.4A.160, Qtkw.tamu.6B.640, and *Qtkw.tamu.7D.69*). These interactions only explained 1.3 – 1.6% of the phenotypic variation and increased TKW up to 0.33 g. Among them, the maximum effects increased TKW up to 0.52 g with the Duster allele and up to 0.8 g with the TAM 112 allele. A total of 12 EI were identified associated with KLEN, including one interaction from the consistent QTL *Qklen.tamu.4A.84* and a QTL on 7A at 181.3 Mbps that increased KLEN by 0.033 mm. Moreover, only one KWID interaction was detected, with the QTL on 6B at 28.5 Mbps and 7B at 727.8 Mbps, but no major QTL was involved (Figure S6 and Table S7).

For HD, nine EI were detected involving three previously detected QTL, including *Qhd.tamu.2B.42*, *Qhd.tamu.2B.58*, and *Qhd.tamu.2D.151*, while only *Qhd.tamu.2B.58* was a consistent QTL. These three EI only explained 1.7 – 2.0% of the phenotypic variation with maximum effects less than 0.65 days. Among 24 HI EI, one involved with previously detected QTL *Qhi.tamu.3D.590*. This interaction increased HI by 0.43% in 17BSP100. The maximum effect only increased HI by 0.8%. Moreover, 28 interactions were detected for PH, including one interaction between major QTL *Qph.tamu.5D.398* and a marker on 5B at 601.4 Mbps. The maximum effect increased PH by 1.06 cm with the TAM 112 allele. Overall, not many major QTL were involved in the EI and the observed interaction effects were not significantly high (Figure S6 and Table S7).

Haplotype analysis

To assess the performance of different haplotypes, RILs were grouped according to their all-allelic combinations. Marker haplotypes identified 46 RILs carrying two QTL (*Qyld.tamu.2D.3*, *Qyld.tamu.2D.37*) for YLD in both 17RRD and 19RRD (Figures S7-A1, A2). Both QTL on chromosome 2D had insignificant effects on YLD when present alone in 17RDD. However, the QTL combination significantly increased YLD in both environments when compared with no QTL presence. For KPS, 46 RILs were also identified with two QTL (*Qkps.tamu.2D.37*, *Qkps.tamu.6A.113*) in both 17BSP67 and 18BI environments (Figures S7-B1, B2). The haplotype with only 6A QTL did not show significant effects by itself but it significantly increased KPS in combination with 2D QTL when compared to lines without QTL. There were four TKW QTL (*Qtkw.tamu.2B.62*, *Qtkw.tamu.2B.785*, *Qtkw.tamu.2D.37*, *Qtkw.tamu.4A.84*) mapped in 17BSP100 (Figure S7-C1) and two TKW QTL (*Qtkw.tamu.2D.37*, *Qtkw.tamu.7D.68*) mapped in 18BI (Figure S7-C2). In 17BSP100, at least two QTL in different combinations had significantly higher TKW than those without QTL, except when *Qtkw.tamu.2B.62* combined with *Qtkw.tamu.2B.785*, and *Qtkw.tamu.2B.785* combined with *Qtkw.tamu.4A.84*.

In 18BI, the mean TKW of those lines with both QTL on 2D and 7D was significantly higher than those lines with only 2D or without QTL. For KWID, 30 RILs had two QTL (*Qkwid.tamu.2D.37* and *Qkwid.tamu.7D.68*) in 18BMS (Figure S7-D1). Haplotypes with only one QTL or both QTL did not show significant differences on KWID. However, they were significantly higher than the haplotype without QTL. Only 12 RILs carried all four favorable alleles (*Qklen.tamu.1A.325*, *Qklen.tamu.2B.781*, *Qklen.tamu.3A.664*, and *Qklen.tamu.7D.157*) for KLEN in all three environments, 18BSP100, 19BSP100, and 19BMS (Figure S7-E1, E2, E3). Combining any three QTL significantly increased KLEN across all environments, except the combination of 2B, 3A, and 7D in 19BSP100.

Discussion

Phenotype under multiple environments

Phenotype evaluation under multiple environments is crucial for identifying consistent genomic regions linked to yield-related traits (Ogbonnaya et al., 2017). In the current study, both irrigated and dryland management was included in the field experiments, coupled with broad geographic regions including Texas High Plains, Texas Rolling Plains, and Oklahoma South Central Plains, which laid the foundation for exploring the stability of detected QTL regions (Tables S1, S6). Environmental variance showed significance and contributed a large proportion of the total phenotypic variation for all evaluated traits, which indicated the high environmental diversity among trials. Except for BM, the genetic variance for all traits was also significant, which indicated genetic stability (Table S2). A moderate to high entry-mean based broad-sense heritability (0.51 – 0.95) showed that most traits were selectable, except BM (0.10). The low heritability of BM was consistent with a previous study, which indicated BM was a low selectable trait (Yang et al., 2020). Since yield is a complex trait usually controlled by many minor effect genes and significantly influenced by environments (Li et al., 2007), the heritability of YLD (0.74) was comparatively high. The dissection of yield into yield components generally improves selection efficiency. Specific to this study, QTL for TKW, KPS, and SPM showed a higher additive effect proportion of the PVE compared to YLD and BMYLD (Figures S5A, F–H and Table S6). Correspondingly, the heritability of all three yield components was higher than the heritability of yield. The heritability estimates were highest for TKW followed by KPS and SPM, which agreed with a previous study (Guan et al., 2018). The further dissection of TKW into kernel traits had increased the heritability of KLEN (Table S2). The QTL for KWID and KLEN had higher additive PVE proportion than TKW (Figures S5H, N, O and Table S6).

The G x E is an essential source of phenotypic variation, providing important adaptability for plants in changing environments (Wang et al., 2014). Except for SPM, the observed G x E variances were highly significant for all traits

due to the diverse phenotyping field environments (Table S2), indicating more environmentally stable genomic regions may control SPM. Environment-specific QTL and related G x E effects were detected with the MET analysis (Table S6). For instance, the YLD QTL on 2D at 2.5 – 3.3 Mbps was only mapped in three dryland environments (17RRD, 17CH and 19RRD), which is an indication of a specific G x E interaction. Those environment-specific QTL will be valuable to breeders for environment targeted gene deployment.

Pyramiding favorable alleles for yield-related traits

Wheat yield can be partitioned into three main components including SPM, KPS, and TKW (Cao et al., 2020). Understanding the relationship among yield components is essential for better gene deployment or for pyramiding favorable alleles associated with different yield components. Generally, overall yield improvement requires increasing either kernel number per unit area (KN) or TKW (Sukumaran et al., 2018; Golan et al., 2019). However, KN and TKW are usually negatively correlated (Kuchel et al., 2007; Guan et al., 2018). The compensation effects between these two factors were also observed. For example, the pleiotropic QTL region on 2D at 37.1 – 38.7 Mbps had Duster alleles that increased TKW, SPM and YLD, while the TAM 112 allele increased KPS. Reversely, the pleiotropic QTL region on 7D at 62.2 – 69.2 Mbps had TAM 112 alleles increasing TKW, while Duster alleles increased KPS. The introgression of a different combination of these alleles may improve KN and/or TKW and relieve the unfavorable allele effect, which needs to be further verified. In addition, some QTL regions improved both yield and yield components without contrary favorable alleles from both parents, including the QTL region on chromosome 6A at 113 Mbps increased both YLD and KPS with the TAM 112 allele, and the QTL region on 2D at 42.9 – 43.7 Mbps increased both YLD and SPM with the Duster allele. Moreover, the improvement of YLD also can be achieved by increasing HI, such as the QTL region on 4D at 47.6 Mbps increased YLD and HI together with the Duster allele. These results indicated the potential to enhance YLD through synergistically improving TKW, KPS, SPM, and HI. In addition, we observed TKW colocalized with KWID on 1D at 135.3 – 137.8 Mbps, 2B at 61.6 and 184.3 Mbps, 3B at 813.4 – 814.4 Mbps, and 6B at 549.7 and 615.5 Mbps. TKW colocalized with KLEN on 2B at 781.2 – 784.7 Mbps, 2D at 71.9 – 72.7 Mbps, and 5D at 196.9 Mbps. Both KLEN and KWID colocalized with TKW on 1A at 590.8 – 592.1 Mbps, 2D at 37.1 – 38.7 Mbps, 5D at 252.0 Mbps, and 7D at 64.4 – 69.2 Mbps. These results indicated the potentially diverse mechanism for TKW determination.

The SNP haplotype analysis of yield and yield component traits estimated the potential effect of different allelic combinations at two, three and/or more loci. The SNP haplotype results suggested that more loci combinations had greater effects on traits including YLD, KPS, TKW, KLEN, and KWID and that gene pyramiding significantly improved performance as compared to the RILs without QTL. However, these QTL combinations need to be validated in other genetic backgrounds and diverse environments.

Agronomic traits related to adaptability

Heading date determined the adaption of wheat to broad geographic locations and diverse climatic environments, which was also a critical agronomic trait in dry climate (Zanke et al., 2014). In the U.S Southern High Plains, drought is major stress limiting wheat yield (Xue et al., 2014). The early heading was favorable for reducing the time exposed to drought stress during the sensitive flowering and post-anthesis grain filling periods, hence increasing the TKW, TW, and HI in dryland environments (Shavrukov et al., 2017). Similarly, we identified three QTL regions consistently associated with HD. All of them had alleles from the early heading parent that increased at least one of YLD, TKW, HI, and TW.

Plant height is another important agronomic trait correlated with yield. The green revolution introduced the dwarf allele of *Rht1* into wheat cultivars that significantly reduced PH and improved both the HI and YLD (Allan, 1989; Hedden, 2003). However, under drought and heat stress conditions, the taller wheat usually had higher grain yield due to more biomass, longer coleoptile, and higher growth vigor that were favorable for a better yield performance (Jatayev et al., 2020). In our study, PH was negatively correlated with yield and yield components under both dry and irrigated environments except in 18BSP100 and 17BSP100. Except for the QTL on 2D at 37.4 Mbps, most of the identified PH QTL were physically located far away from the QTL linked to yield and yield-related traits. The only exception, *Qph.tamu.2D.37* colocalized with yield-related traits, and may be explained by the pleiotropic effect of *Ppd-D1*. *Qph.tamu.2D.37* was physically close to *Ppd-D1*, which was reported for its strong pleiotropic effect on all yield-related traits and PH (Guo et al., 2010). The diagnostic marker analysis also proved that Duster carried the insensitive *Ppd-D1a* allele, and had early HD (Beales et al., 2007). Moreover, the identified PH loci other than *Qph.tamu.2D.37* may not affect yield, which is consistent with the recently cloned *Rht8* gene, which does not influence coleoptile elongation and yield components like KPS and TKW (Chai et al., 2022).

QTL comparisons

Yield-related traits are important in wheat breeding (Hu et al., 2020). Since 1900, yield-related traits were intensively

selected and improved for the higher KPS, TKW, and HI (Calderini et al., 1995). Previous studies characterized numerous QTL and genes associated with yield-related traits across multiple environments on all 21 wheat chromosomes (Assanga et al., 2017; Cao et al., 2020; Yang et al., 2020; Dhakal et al., 2021a). To evaluate the reliability and novelty of those QTL, we compared the physical positions of QTL identified in this study with previously reported QTL and genes based on the IWGSC reference genome version 1.0.

For yield-related traits, the genome-wide association study (GWAS) of CIMMYT spring bread wheat conducted by Sehgal et al. (2020) identified a yield QTL that shared the same physical location with *Qbmyld.tamu.2B.75*. On chromosome 2B, between 57.5 and 61.6 Mbps, we mapped a major pleiotropic QTL associated with TKW, KWID, HD, HI, and SHGW. This genomic region is at the proximal side of the photoperiod gene *Ppd-B1*, which is known for its influence on YLD, HD, and TKW (Mohler et al., 2004; Maphosa et al., 2014). The previously reported gene *TaDA1* located on chromosome 2D at 8.3 Mbps, had effects on TKW, KLEN, and KWID (Liu et al., 2020a). *Qyld.tamu.2D.3* was an environment-specific QTL only detected under dryland environments in Oklahoma and Chillicothe in Texas. Although the physical position of *Qyld.tamu.2D.3* is in the close distal side of *TaDA1*, the yield components and kernel traits were not evaluated under the abovementioned environments. Hence those traits still need to be explored in future studies to see if *Qyld.tamu.2D.3* also affects TKW and kernel traits. The *Qbmyld.tamu.2D.37* in the proximal side of photoperiod gene *Ppd-D1* at 33.9 Mbps. The diagnostic marker analysis also detected Duster has an insensitive *Ppd-D1a* allele. Previous studies revealed that different haplotypes of *Ppd-D1* under multiple environments influenced YLD, KPS, TKW, HI, HD, and PH (Guo et al., 2010; Chen et al., 2018), which were consistent with the result from this study. *Qyld.tamu.6A.113* was colocalized with *Qkwid.tamu.6A.113* and *Qkps.tamu.6A.113*. Several studies reported loci linked to YLD, KWID, and KPS located at similar physical positions, proving the reliability of this QTL region as a consistent and pleiotropic QTL (Assanga et al., 2017; Li et al., 2019a; Li et al., 2019b; Duan et al., 2020; Liu et al., 2020c; Dhakal et al., 2021b). In the Nongda3338/Jingdong6 doubled haploid (DH) population, Guan et al. (2018) identified a stable SPM QTL *QEp.cau-4D.1* which overlapped with the *Qspm.tamu.4D.64* identified in this study. The peak marker of *Qspm.tamu.5D.252* was only 2 Mbps proximal to the peak marker of the SPM QTL detected by Li et al. (2020). Moreover, the physical position of *Qspm.tamu.7A.4* overlapped with the locus linked to SPM in a population of 246 F_8 RILs derived from the cross of Zhou 8425B/Chinese Spring (Gao et al., 2015). In addition, we detected a stable QTL region related to KPS, HI, TKW, KLEN, KWID, AREA, PERI, and HD, which was located on 7D between 64.4 and 69.2 Mbps. This genomic region contains *FT-D1*, a homolog gene of Arabidopsis *FLOWERING LOCUS T* that regulates wheat flowering and has been frequently reported by many previous studies (Sun et al., 2017; Li et al., 2019; Ward et al., 2019; Chen et al., 2020; Hu et al., 2020; Liu et al., 2020b; Liu et al., 2020c; Yang et al., 2020; Dhakal et al., 2021a; Isham et al., 2021). Isham et al. (2021) reported that *FT-D1* had an effect on TKW and KPS, and a recent study by Chen et al. (2022) further proved its effects on HD and KPS through gene editing. Those results were consistent with the traits associated with the *FT-D1* locus detected in this study. The diagnostic marker Vrn-D3-KASP was used to test both parents, which showed TAM 112 carried early heading and Duster carried late heading allele. Moreover, the QTL region on 7D between 107 and 109.8 Mbps linked to KLEN and PERI co-localized with the KLEN QTL detected in three wheat RIL populations (Li et al., 2018).

For TKW and kernel traits, *Qklen.tamu.1A.517* and *Qkwid.tamu.1A.591* had the physical position identical to KLEN and KWID QTL mapped in 163 bread wheat lines (Sun et al., 2017). The TKW QTL located on 2B at 136.9 Mbps, TKW and KWID QTL located on 2B at 184.3 Mbps, AREA and KLEN QTL located on 2B at 647.6 Mbps, and TKW and KLEN QTL located on 2B at 781.2 – 784.7 Mbps shared similar physical locations with previously reported QTL (Assanga et al., 2017; Sun et al., 2017; Zhang et al., 2018; Ma et al., 2019; Rahimi et al., 2019; Corsi et al., 2021). Besides, on chromosome 2D between 71.9 and 72.7 Mbps, a QTL region controlling TKW, KLEN, PERI, and AREA was close to previously reported QTL associated with TKW, KLEN, and AREA (Ma et al., 2018; Ma et al., 2019; Liu et al., 2020c; Qu et al., 2021). The genomic region on chromosome 4A between 83.8 and 90.5 Mbps harbored a QTL cluster associated with TKW, KLEN, PERI, and AREA, which has also been reported for association with TKW and KLEN (Sun et al., 2017; Ren et al., 2021). QTL associated with KWID was identified on chromosome 4A at 122.3 Mbps and was physically close to a KWID-related QTL at 120.6 Mbps reported by Sun et al. (2017).

For PH, QTL on 5D at 338.0 Mbps and on 6A at 122 Mbps were also previously mapped (Li et al., 2019b). Another consistent QTL linked to PH on chromosome 6A at 418.4 Mbps was colocalized with the previously identified dwarf gene *Rht18* and *Rht24b*. Both genes affect the expression level of gibberellin (GA) 2-oxidase, TaGA2ox-A9 (Ford et al., 2018; Tian et al., 2021). We found that PH QTL also influenced KPS, HI, AREA, and KLEN. In addition, a locus linked to HD shared the exact physical location with *Qhd.tamu.2B.42* in TAM 112/TAM 111 population (Dhakal et al., 2021a).

After further comparing with the previously identified stable QTL and genes recently summarized by Cao et al. (2020), several novel genomic regions with stable effects were determined in this study. A major QTL region linked to KLEN on 1A at 324.6 Mbps was detected under all six evaluated individual environments, and across or within mega environments. The favorable allele was from TAM 112, consistent with TAM 112 having a higher KLEN than Duster. A major gene may underlie this QTL region that accounts for the consistent kernel shape difference between parents. Another

novel QTL region was on 5D at 196.9 Mbps and had pleiotropic effects on TKW, KLEN, PERI, AREA, and SHDW, with favorable alleles from Duster. Moreover, the QTL associated with TKW on 6B at 138.7 Mbps is also a novel locus, with favorable alleles from TAM 112.

Conclusion

In this study, a set of 18 yield-related, and agronomic traits were evaluated in the TAM 112/Duster RIL population which was tested in 13 environments for three crop years. Detected QTL included 21 consistent and 30 pleiotropic QTL regions with 10 in common. Three genes, on chromosomes 2B (*Ppd-B1*), 2D (*Ppd-D1*), and 7D (*FT-D1*), aligned with three major consistent and pleiotropic QTL regions, increasing yield and yield-related traits.

Six consistent QTL for YLD, TKW, KLEN, KPS, and PH have additive effects that constantly explained more than 5% of the phenotypic variations, including *Qyld.tamu.2D.3*, *Qtkw.tamu.2B.137*, *Qklen.tamu.1A.325*, *Qkps.tamu.6A.113*, *Qph.tamu.6A.122*, and *Qph.tamu.6A.418*. One novel KLEN QTL on 1A at 324.6 Mbps appeared in all testing environments, which might be a valuable locus for exploring the genetic mechanism of KLEN. These QTL could be potentially used for future breeding applications. Most identified pleiotropic QTL regions had significant compensation effects on yield components. The majority of consistent QTL showed a higher portion of additive effects which is selectable. In addition, no significant epistatic interactions were observed on major consistent and pleotropic QTL. These results will be useful on understanding the genetic basis of yield and agronomic traits, pyramiding the favorable alleles through MAS, and moving forward with any future fine-mapping studies. The RILs carrying most favorable alleles can be directly used in as germplasms in the breeding programs.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: Texas Data Repository, https://dataverse.tdl.org/dataset.xhtml?persistentId=doi:10.18738/T8/O8WIU3.

Author contributions

ZW, SL, YR prepared manuscript and analyzed data. ZW, SD, MC collected phenotype data. SL, YR, SY, GB revised manuscript. SL, JR, AI, QX, DH applied funding, gave suggestions on data analysis, and helped with field management. AB, PS, GB, SW conducted GBS library preparation, DNA sequencing, SNP calling, and genotype data filtering. FM, WH, JA, XM, JB, SB managed field trials. All authors contributed to the article and approved the submitted version.

Funding

This project is partly funded by Texas Wheat Producer Board, Texas A&M AgriLife Research, the National Research Initiative Competitive Grants 2017-67007-25939, 2019-67013-29172, 2021-67013-33940, and 2022-68013-36439 from the U.S. Department of Agriculture National Institute of Food and Agriculture.

Acknowledgments

We thank the technical support from Lisa Garza, Chenggen Chu, Kele Hui, Xiaoxiao Liu, and Kirk Joseph. We appreciated the review suggestions from Kyle Parker.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1057701/full#supplementary-material

References

Allan, R. (1989). Crop breeding, genetics, and cytology: Agronomic comparisons between Rht1 and Rht2 semidwarf genes in winter wheat. *Crop Sci.* 29, 1103–1108. doi: 10.2135/cropsci1989.0011183X002900050001x

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, 745–756. doi: 10.1016/j.cj.2020.03.010

Arjona, J. M., Royo, C., Dreisigacker, S., Ammar, K., and Villegas, D. (2018). Effect of ppd-A1 and ppd-B1 allelic variants on grain number and thousand kernel weight of durum wheat and their impact on final grain yield. *Front. Plant Sci.* 9, 888. doi: 10.3389/fpls.2018.00888

Assanga, S. O., Fuentealba, M., Zhang, G., Tan, C., Dhakal, S., Rudd, J. C., et al. (2017). Mapping of quantitative trait loci for grain yield and its components in a US popular winter wheat TAM 111 using 90K SNPs. *PloS One* 12, e0189669. doi: 10.1371/journal.pone.0189669

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 48. doi: 10.18637/jss.v067.i01

Beales, J., Turner, A., Griffiths, S., Snape, J., and Laurie, D. (2007). A pseudo-response regulator is misexpressed in the photoperiod insensitive ppd-D1a mutant of wheat (Triticum aestivum l.). *TAG Theor. Appl. Genet.* 115, 721–733. doi: 10.1007/s00122-007-0603-4

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, 2633–35. doi: 10.1093/bioinformatics/btm308

Calderini, D., Dreccer, M., and Slafer, G. (1995). Genetic improvement in wheat yield and associated traits. a re-examination of previous results and the latest trends. *Plant Breed.* 114, 108–112. doi: 10.1111/j.1439-0523.1995.tb00772.x

Cao, S., Xu, D., Hanif, M., Xia, X., and He, Z. (2020). Genetic architecture underpinning yield component traits in wheat. *Theor. Appl. Genet.* 133, 1811–1823. doi: 10.1007/s00122-020-03562-8

Chai, L., Xin, M., Dong, C., Chen, Z., Zhai, H., Zhuang, J., et al. (2022). A natural variation in ribonuclease h-like gene underlies Rht8 to confer" green revolution" trait in wheat. *Mol. Plant* 15 (3), 377–380 doi: 10.1016/j.molp.2022.01.013

Chen, Z., Cheng, X., Chai, L., Wang, Z., Bian, R., Li, J., et al. (2020). Dissection of genetic factors underlying grain size and fine mapping of QTgw.cau-7D in common wheat (Triticum aestivum l.). *Theor. Appl. Genet.* 133, 149–162. doi: 10.1007/s00122-019-03447-5

Chen, Z., Ke, W., He, F., Chai, L., Cheng, X., Xu, H., et al. (2022). A single nucleotide deletion in the third exon of FT-D1 increases the spikelet number and delays heading date in wheat (Triticum aestivum l.). *Plant Biotechnol. J* 20(5), 920–33. doi: 10.1111/pbi.13773

Chen, L., Yang, Y., Cui, C., Lu, S., Lu, Q., Du, Y., et al. (2018). Effects of vrn-B1 and ppd-D1 on developmental and agronomic traits in Rht5 dwarf plants of bread wheat. *Field Crops Res.* 219, 24–32. doi: 10.1016/j.fcr.2018.01.022

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. B: Biol. Sci.* 363, 557–572. doi: 10.1098/rstb.2007.2170

Corsi, B., Obinu, L., Zanella, C. M., Cutrupi, S., Day, R., Geyer, M., et al. (2021). Identification of eight QTL controlling multiple yield components in a German multi-parental wheat population, including Rht24, WAPO-A1, WAPO-B1 and genetic loci on chromosomes 5A and 6A. *Theor. Appl. Genet.* 134, 1435–1454. doi: 10.1007/s00122-021-03781-7

Dhakal, S., Liu, X., Chu, C., Yang, Y., Rudd, J. C., Ibrahim, A. M., et al. (2021a). Genome-wide QTL mapping of yield and agronomic traits in two widely adapted winter wheat cultivars from multiple mega-environments. *PeerJ* 9, e12350. doi: 10.7717/peerj.12350

Dhakal, S., Liu, X., Girard, A., Chu, C., Yang, Y., Wang, S., et al. (2021b). Genetic dissection of end-use quality traits in two widely adapted wheat cultivars 'TAM 111'and 'TAM 112'. *Crop Sci.* 61, 1944–1959. doi: 10.1002/csc2.20415

Duan, X., Yu, H., Ma, W., Sun, J., Zhao, Y., Yang, R., et al. (2020). A major and stable QTL controlling wheat thousand grain weight: Identification, characterization, and CAPS marker development. *Mol. Breed.* 40, 1–12. doi: 10.1007/s11032-020-01147-3

Edwards, J. T., Hunger, R. M., Smith, E. L., Horn, G. W., Chen, M.-S., Yan, L., et al. (2012). 'Duster' wheat: A durable, dual-purpose cultivar adapted to the southern great plains of the USA. *J. Plant Reg.* 6, 37–48. doi: 10.3198/jpr2011.04.0195crc

Enghiad, A., Ufer, D., Countryman, A. M., and Thilmany, D. D. (2017). An overview of global wheat market fundamentals in an era of climate concerns. *Int. J. Agron.* 2017, 3931897. doi: 10.1155/2017/3931897

Feng, F., Han, Y., Wang, S., Yin, S., Peng, Z., Zhou, M., et al. (2018). The effect of grain position on genetic improvement of grain number and thousand grain weight in winter wheat in north China. *Front. Plant Sci.* 9, 129. doi: 10.3389/fpls.2018.00129

Finnegan, E. J., Ford, B., Wallace, X., Pettolino, F., Griffin, P. T., Schmitz, R. J., et al. (2018). Zebularine treatment is associated with deletion of FT-B1 leading to an increase in spikelet number in bread wheat. *Plant Cell Environ.* 41, 1346–1360. doi: 10.1111/pce.13164

Ford, B. A., Foo, E., Sharwood, R., Karafiatova, M., Vrána, J., MacMillan, C., et al. (2018). Rht18 semidwarfism in wheat is due to increased GA 2-oxidaseA9 expression and reduced GA content. *Plant Physiol.* 177, 168–180. doi: 10.1104/pp.18.00023

Gao, F., Wen, W., Liu, J., Rasheed, A., Yin, G., Xia, X., et al. (2015). Genome-wide linkage mapping of QTL for yield components, plant height and yield-related physiological traits in the Chinese wheat cross zhou 8425B/Chinese spring. *Front. Plant Sci.* 6. doi: 10.3389/fpls.2015.01099

Glenn, P., Zhang, J., Brown-Guedira, G., DeWitt, N., Cook, J. P., Li, K., et al. (2021). Identification and characterization of a natural polymorphism in FT-A2 associated with increased number of grains per spike in wheat. *Theor. Appl. Genet.* 135(2), 679–92. doi: 10.21203/rs.3.rs-906395/v1

Golan, G., Ayalon, I., Perry, A., Zimran, G., Ade-Ajayi, T., Mosquna, A., et al. (2019). GNI-A1 mediates trade-off between grain number and grain weight in tetraploid wheat. *Theor. Appl. Genet.* 132, 2353–2365. doi: 10.1007/s00122-019-03358-5

Golan, G., Oksenberg, A., and Peleg, Z. (2015). Genetic evidence for differential selection of grain and embryo weight during wheat evolution under domestication. *J. Exp. Bot.* 66, 5703–5711. doi: 10.1093/jxb/erv249

Guan, P., Lu, L., Jia, L., Kabir, M. R., Zhang, J., Lan, T., et al. (2018). Global QTL analysis identifies genomic regions on chromosomes 4A and 4B harboring stable loci for yield-related traits across different environments in wheat (Triticum aestivum l.). *Front. Plant Sci.* 9, 529. doi: 10.3389/fpls.2018.00529

Guo, Y., Sun, J., Zhang, G., Wang, Y., Kong, F., Zhao, Y., et al. (2013). Haplotype, molecular marker and phenotype effects associated with mineral nutrient and grain size traits of TaGS1a in wheat. *Field Crops Res.* 154, 119–125. doi: 10.1016/j.fcr.2013.07.012

Guo, Z., Yanxia, S., Ronghua, Z., Zhenglong, R., and Jizeng, J. (2010). Discovery, evaluation and distribution of haplotypes of the wheat ppd-D1 gene. *New Phytol.* 185, 841–851. doi: 10.1111/j.1469-8137.2009.03099.x

Hanif, M., Gao, F., Liu, J., Wen, W., Zhang, Y., Rasheed, A., et al. (2016). TaTGW6-A1, an ortholog of rice TGW6, is associated with grain weight and yield in bread wheat. *Mol. Breed.* 36, 1. doi: 10.1007/s11032-015-0425-z

Hedden, P. (2003). The genes of the green revolution. *Trends Genet.* 19, 5–9. doi: 10.1016/S0168-9525(02)00009-4

Hu, J., Wang, X., Zhang, G., Jiang, P., Chen, W., Hao, Y., et al. (2020). QTL mapping for yield–related traits in wheat based on four RIL populations. *Theor. Appl. Genet.* 133, 917–933. doi: 10.1007/s00122-019-03515-w

Hu, M.-J., Zhang, H.-P., Liu, K., Cao, J.-J., Wang, S.-X., Jiang, H., et al. (2016). Cloning and characterization of TaTGW-7A gene associated with grain weight in wheat *via* SLAF-seq-BSA. *Front. Plant Sci.* 7, 1902. doi: 10.3389/fpls.2016.01902

Ikeda, K., Ito, M., Nagasawa, N., Kyozuka, J., and Nagato, Y. (2007). Rice ABERRANT PANICLE ORGANIZATION 1, encoding an f-box protein, regulates meristem fate. *Plant J.* 51, 1030–1040. doi: 10.1111/j.1365-313X.2007.03200.x

Isham, K., Wang, R., Zhao, W., Wheeler, J., Klassen, N., Akhunov, E., et al. (2021). QTL mapping for grain yield and three yield components in a population derived from two high-yielding spring wheat cultivars. *Theor. Appl. Genet.* 134(7), 2079–2095. doi: 10.1007/s00122-021-03806-1

IWGSC, R., Appels, K., Eversole, N., Stein, C., Feuillet, B., Keller, J., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361 (6403), eaar7191. doi: 10.1126/science.aar7191

Jatayev, S., Sukhikh, I., Vavilova, V., Smolenskaya, S. E., Goncharov, N. P., Kurishbayev, A., et al. (2020). Green revolution" stumbles" in a dry environment: Dwarf wheat with rht genes fails to produce higher grain yield than taller plants under drought. *Plant Cell Environ* 43 (10), 2355–2364. doi: 10.1111/pce.13819

Juliana, P., Poland, J., Huerta-Espino, J., Shrestha, S., Crossa, J., Crespo-Herrera, L., et al. (2019). Improving grain yield, stress resilience and quality of bread wheat using large-scale genomics. *Nat. Genet.* 51, 1530–1539. doi: 10.1038/s41588-019-0496-6

Kosambi, D. D. (1944). The estimation of map distances from recombination values. *DD Kosambi. Springer*, 125–130. doi: 10.1007/978-81-322-3676-4_16

Kuchel, H., Williams, K. J., Langridge, P., Eagles, H. A., and Jefferies, S. P. (2007). Genetic dissection of grain yield in bread wheat. i. QTL analysis. *Theor. Appl. Genet.* 115, 1029–1041. doi: 10.1007/s00122-007-0629-7

Kuzay, S., Lin, H., Li, C., Chen, S., Woods, D. P., Zhang, J., et al. (2022). WAPO-A1 is the causal gene of the 7AL QTL for spikelet number per spike in wheat. *PloS Genet.* 18, e1009747. doi: 10.1371/journal.pgen.1009747

Kuzay, S., Xu, Y., Zhang, J., Katz, A., Pearce, S., Su, Z., et al. (2019). Identification of a candidate gene for a QTL for spikelet number per spike on wheat chromosome arm 7AL by high-resolution genetic mapping. *Theor. Appl. Genet.* 132, 2689–2705. doi: 10.1007/s00122-019-03382-5

Li, S., Jia, J., Wei, X., Zhang, X., Li, L., Chen, H., et al. (2007). A intervarietal genetic map and QTL analysis for yield traits in wheat. *Mol. Breed.* 20, 167–178. doi: 10.1007/s11032-007-9080-3

Li, L., Mao, X., Wang, J., Chang, X., Reynolds, M., and Jing, R. (2019a). Genetic dissection of drought and heat-responsive agronomic traits in wheat. *Plant Cell Environ.* 42:2540–53. doi: 10.1111/pce.13577

Li, L., Peng, Z., Mao, X., Wang, J., Chang, X., Reynolds, M., et al. (2019b). Genome-wide association study reveals genomic regions controlling root and shoot traits at late growth stages in wheat. *Ann. Bot.* 124, 993–1006. doi: 10.1093/aob/mcz041

Liu, H., Li, H., Hao, C., Wang, K., Wang, Y., Qin, L., et al. (2020a). TaDA1, a conserved negative regulator of kernel size, has an additive effect with TaGW2 in common wheat (Triticum aestivum l.). *Plant Biotechnol. J.* 18, 1330–1342. doi: 10.1111/pbi.13298

Liu, H., Mullan, D., Zhang, C., Zhao, S., Li, X., Zhang, A., et al. (2020b). Major genomic regions responsible for wheat yield and its components as revealed by meta-QTL and genotype–phenotype association analyses. *Planta* 252, 65. doi: 10.1007/s00425-020-03466-3

Liu, S., Rudd, J. C., Bai, G., Haley, S. D., Ibrahim, A. M. H., Xue, Q., et al. (2014). Molecular markers linked to important genes in hard winter wheat. *Crop Sci.* 54, 1304–1321. doi: 10.2135/cropsci2013.08.0564

Liu, J., Xu, Z., Fan, X., Zhou, Q., Cao, J., Wang, F., et al. (2018). A genome-wide association study of wheat spike related traits in China. *Front. Plant Sci.* 9, 1584. doi: 10.3389/fpls.2018.01584

Liu, H., Zhang, X., Xu, Y., Ma, F., Zhang, J., Cao, Y., et al. (2020c). Identification and validation of quantitative trait loci for kernel traits in common wheat (Triticum aestivum l.). *BMC Plant Biol.* 20, 1–15. doi: 10.1186/s12870-020-02661-4

Li, F., Wen, W., He, Z., Liu, J., Jin, H., Cao, S., et al. (2018). Genome-wide linkage mapping of yield-related traits in three Chinese bread wheat populations using high-density SNP markers. *Theor. Appl. Genet.* 131, 1903–1924. doi: 10.1007/s00122-018-3122-6

Li, F., Wen, W., Liu, J., Zhang, Y., Cao, S., He, Z., et al. (2019). Genetic architecture of grain yield in bread wheat based on genome-wide association studies. *BMC Plant Biol.* 19, 168. doi: 10.1186/s12870-019-1781-3

Li, X., Xu, X., Liu, W., Li, X., Yang, X., Ru, Z., et al. (2020). Dissection of superior alleles for yield-related traits and their distribution in important cultivars of wheat by association mapping. *Front. Plant Sci.* 11, 175. doi: 10.3389/fpls.2020.00175

Lv, B., Nitcher, R., Han, X., Wang, S., Ni, F., Li, K., et al. (2014). Characterization of FLOWERING LOCUS T1 (FT1) gene in brachypodium and wheat. *PloS One* 9, e94171. doi: 10.1371/journal.pone.0094171

Maphosa, L., Langridge, P., Taylor, H., Parent, B., Emebiri, L. C., Kuchel, H., et al. (2014). Genetic control of grain yield and grain physical characteristics in a bread wheat population grown under a range of environmental conditions. *Theor. Appl. Genet.* 127, 1607–1624. doi: 10.1007/s00122-014-2322-y

Mason, R. E., Hays, D. B., Mondal, S., Ibrahim, A. M., and Basnet, B. R. (2013). QTL for yield, yield components and canopy temperature depression in wheat under late sown field conditions. *Euphytica* 194, 243–259. doi: 10.1007/s10681-013-0951-x

Ma, F., Xu, Y., Ma, Z., Li, L., and An, D. (2018). Genome-wide association and validation of key loci for yield-related traits in wheat founder parent xiaoyan 6. *Mol. Breed.* 38, 1–15. doi: 10.1007/s11032-018-0837-7

Ma, J., Zhang, H., Li, S., Zou, Y., Li, T., Liu, J., et al. (2019). Identification of quantitative trait loci for kernel traits in a wheat cultivar Chuannong16. *BMC Genet.* 20, 1–12. doi: 10.1186/s12863-019-0782-4

McIntosh, R., Dubcovsky, J., Rogers, W., Morris, C., Appels, R., Xia, X., et al. (2013). *Catalogue of gene symbols for wheat*, Vol. 2012, Eps 1:1DL-M11193.

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

Mohler, V., Lukman, R., Ortiz-Islas, S., William, M., Worland, A. J., Van Beem, J., et al. (2004). Genetic and physical mapping of photoperiod insensitive gene ppd-B1 in common wheat. *Euphytica* 138, 33–40. doi: 10.1023/B:EUPH.0000047056.58938.76

Muhammad, A., Hu, W., Li, Z., Li, J., Xie, G., Wang, J., et al. (2020). Appraising the genetic architecture of kernel traits in hexaploid wheat using GWAS. *Int. J. Mol. Sci.* 21, 5649. doi: 10.3390/ijms21165649

Ogbonnaya, F. C., Rasheed, A., Okechukwu, E. C., Jighly, A., Makdis, F., Wuletaw, T., et al. (2017). Genome-wide association study for agronomic and physiological traits in spring wheat evaluated in a range of heat prone environments. *Theor. Appl. Genet.* 130, 1819–1835. doi: 10.1007/s00122-017-2927-z

Pang, Y., Liu, C., Wang, D., St. Amand, P., Bernardo, A., Li, W., et al. (2020). High-resolution genome-wide association study identifies genomic regions and candidate genes for important agronomic traits in wheat. *Mol. Plant* 13, 1311–1327. doi: 10.1016/j.molp.2020.07.008

Poland, J. A., and Rife, T. W. (2012). Genotyping-by-sequencing for plant breeding and genetics. *Plant Genome* 5 (3), 92–102. doi: 10.3835/plantgenome2012.05.0005

Qin, X., Zhang, F., Liu, C., Yu, H., Cao, B., Tian, S., et al. (2015). Wheat yield improvements in China: Past trends and future directions. *Field Crops Res.* 177, 117–124. doi: 10.1016/j.fcr.2015.03.013

Qu, X., Liu, J., Xie, X., Xu, Q., Tang, H., Mu, Y., et al. (2021). Genetic mapping and validation of loci for kernel-related traits in wheat (Triticum aestivum l.). *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.667493

Rahimi, Y., Bihamta, M. R., Taleei, A., Alipour, H., and Ingvarsson, P. K. (2019). Genome-wide association study of agronomic traits in bread wheat reveals novel putative alleles for future breeding programs. *BMC Plant Biol.* 19, 1–19. doi: 10.1186/s12870-019-2165-4

Ren, T., Fan, T., Chen, S., Li, C., Chen, Y., Ou, X., et al. (2021). Utilization of a Wheat55K SNP array-derived high-density genetic map for high-resolution mapping of quantitative trait loci for important kernel-related traits in common wheat. *Theor. Appl. Genet.* 134, 807–821. doi: 10.1007/s00122-020-03732-8

Roser, M., Ritchie, H., and Ortiz-Ospina, E. (2019). *World population growth* (OurWorldInDataorg 2013).

Rudd, J. C., Devkota, R. N., Baker, J. A., Peterson, G. L., Lazar, M. D., Bean, B., et al. (2014). 'TAM 112' wheat, resistant to greenbug and wheat curl mite and adapted to the dryland production system in the southern high plains. *J. Plant Registrations* 8, 291–297. doi: 10.3198/jpr2014.03.0016crc

Sall, J., Stephens, M. L., Lehman, A., and Loring, S. (2017). JMP start statistics: a guide to statistics and data analysis using JMP. *Sas Institute*.

Sehgal, D., Rosyara, U., Mondal, S., Singh, R., Poland, J., and Dreisigacker, S. (2020). Incorporating genome-wide association mapping results into genomic prediction models for grain yield and yield stability in CIMMYT spring bread wheat. *Front. Plant Sci.* 11, 197. doi: 10.3389/fpls.2020.00197

Shavrukov, Y., Kurishbayev, A., Jatayev, S., Shvidchenko, V., Zotova, L., Koekemoer, F., et al. (2017). Early flowering as a drought escape mechanism in plants: How can it aid wheat production? *Front. Plant Sci.* 8, 1950. doi: 10.3389/fpls.2017.01950

Shaw, L. M., Lyu, B., Turner, R., Li, C., Chen, F., Han, X., et al. (2019). FLOWERING LOCUS T2 regulates spike development and fertility in temperate cereals. *J. Exp. Bot.* 70, 193–204. doi: 10.1093/jxb/ery350

Shewry, P. R., and Hey, S. J. (2015). The contribution of wheat to human diet and health. *Food Energy Secur.* 4, 178–202. doi: 10.1002/fes3.64

Su, Z., Hao, C., Wang, L., Dong, Y., and Zhang, X. (2011). Identification and development of a functional marker of TaGW2 associated with grain weight in bread wheat (Triticum aestivum l.). *Theor. Appl. Genet.* 122, 211–223. doi: 10.1007/s00122-010-1437-z

Sukumaran, S., Lopes, M., Dreisigacker, S., and Reynolds, M. (2018). Genetic analysis of multi-environmental spring wheat trials identifies genomic regions for locus-specific trade-offs for grain weight and grain number. *Theor. Appl. Genet.* 131, 985–998. doi: 10.1007/s00122-017-3037-7

Sun, C., Zhang, F., Yan, X., Zhang, X., Dong, Z., Cui, D., et al. (2017). Genome-wide association study for 13 agronomic traits reveals distribution of superior alleles in bread wheat from the yellow and huai valley of China. *Plant Biotechnol. J.* 15, 953–969. doi: 10.1111/pbi.12690

Tian, X., Xia, X., Xu, D., Liu, Y., Xie, L., Hassan, M. A., et al. (2021). Rht24b, an ancient variation of TaGA2ox-A9, reduces plant height without yield penalty in wheat. *New Phytol* 233 (2), 738–750. doi: 10.1111/nph.17808

Trethowan, R. M., and Mujeeb-Kazi, A. (2008). Novel germplasm resources for improving environmental stress tolerance of hexaploid wheat all rights reserved. No part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. permission for printing and for reprinting the material contained herein has been obtained by the publisher. *Crop Sci.* 48, 1255–1265. doi: 10.2135/cropsci2007.08.0477

Van Ooijen, J. (2006). *JoinMap® 4, software for the calculation of genetic linkage maps in experimental populations* Vol. 33 (Wageningen: Kyazma BV).

Wang, X., Pang, Y., Zhang, J., Zhang, Q., Tao, Y., Feng, B., et al. (2014). Genetic background effects on QTL and QTL× environment interaction for yield and its

component traits as revealed by reciprocal introgression lines in rice. *Crop J.* 2, 345–357. doi: 10.1016/j.cj.2014.06.004

Wang, W., Pan, Q., Tian, B., He, F., Chen, Y., Bai, G., et al. (2019). Gene editing of the wheat homologs of TONNEAU 1-recruiting motif encoding gene affects grain shape and weight in wheat. *Plant J.* 100, 251–264. doi: 10.1111/tpj.14440

Wang, S., Zhang, X., Chen, F., and Cui, D. (2015). A single-nucleotide polymorphism of TaGS5 gene revealed its association with kernel weight in Chinese bread wheat. *Front. Plant Sci.* 6, 1166. doi: 10.3389/fpls.2015.01166

Ward, B. P., Brown-Guedira, G., Kolb, F. L., Van Sanford, D. A., Tyagi, P., Sneller, C. H., et al. (2019). Genome-wide association studies for yield-related traits in soft red winter wheat grown in Virginia. *PloS One* 14, e0208217. doi: 10.1371/journal.pone.0208217

Whan, A. P., Smith, A. B., Cavanagh, C. R., Ral, J.-P. F., Shaw, L. M., Howitt, C. A., et al. (2014). GrainScan: A low cost, fast method for grain size and colour measurements. *Plant Methods* 10, 1–10. doi: 10.1186/1746-4811-10-23

Würschum, T., Boeven, P. H., Langer, S. M., Longin, C. F. H., and Leiser, W. L. (2015). Multiply to conquer: Copy number variations at ppd-B1 and vrn-A1 facilitate global adaptation in wheat. *BMC Genet.* 16, 1–8. doi: 10.1186/s12863-015-0258-0

Würschum, T., Leiser, W. L., Langer, S. M., Tucker, M. R., and Longin, C. F. H. (2018). Phenotypic and genetic analysis of spike and kernel characteristics in wheat reveals long-term genetic trends of grain yield components. *Theor. Appl. Genet.* 131, 2071–2084. doi: 10.1007/s00122-018-3133-3

Xue, Q., Rudd, J. C., Liu, S., Jessup, K. E., Devkota, R. N., and Mahano, J. R. (2014). Yield determination and water-use efficiency of wheat under water-limited conditions in the U.S. *South. High Plains Crop Sci.* 54, 34–47. doi: 10.2135/cropsci2013.02.0108

Yan, L., Fu, D., Li, C., Blechl, A., Tranquilli, G., Bonafede, M., et al. (2006). The wheat and barley vernalization gene VRN3 is an orthologue of FT. *Proc. Natl. Acad. Sci.* 103, 19581–19586. doi: 10.1073/pnas.0607142103

Yang, Y., Dhakal, S., Chu, C., Wang, S., Xue, Q., Rudd, J. C., et al. (2020). Genome wide identification of QTL associated with yield and yield components in two popular wheat cultivars TAM 111 and TAM 112. *PloS One* 15 (12), e0237293. doi: 10.1371/journal.pone.0237293

Yang, J., Zhou, Y., Wu, Q., Chen, Y., Zhang, P., Zhang, Y., et al. (2019a). Molecular characterization of a novel TaGL3-5A allele and its association with grain length in wheat (Triticum aestivum l.). *Theor. Appl. Genet.* 132, 1799–1814. doi: 10.1007/s00122-019-03316-1

Yang, J., Zhou, Y., Zhang, Y., Hu, W., Wu, Q., Chen, Y., et al. (2019b). Cloning, characterization of TaGS3 and identification of allelic variation associated with kernel traits in wheat (Triticum aestivum l.). *BMC Genet.* 20, 1–11. doi: 10.1186/s12863-019-0800-6

Yan, X., Zhao, L., Ren, Y., Dong, Z., Cui, D., and Chen, F. (2019). Genome-wide association study revealed that the TaGW8 gene was associated with kernel size in Chinese bread wheat. *Sci. Rep.* 9, 1–10. doi: 10.1038/s41598-019-38570-2

Yue, Y., Zhang, P., and Shang, Y. (2019). The potential global distribution and dynamics of wheat under multiple climate change scenarios. *Sci. Total Environ.* 688, 1308–1318. doi: 10.1016/j.scitotenv.2019.06.153

Yu, M., Mao, S. L., Hou, D. B., Chen, G. Y., Pu, Z. E., Li, W., et al. (2018). Analysis of contributors to grain yield in wheat at the individual quantitative trait locus level. *Plant Breed.* 137, 35–49. doi: 10.1111/pbr.12555

Zanke, C., Ling, J., Plieske, J., Kollers, S., Ebmeyer, E., and Korzun, V. (2014). Genetic architecture of main effect QTL for heading date in European winter wheat. *Front. Plant Sci.* 5. doi: 10.3389/fpls.2014.00217

Zhang, J., Gizaw, S. A., Bossolini, E., Hegarty, J., Howell, T., Carter, A. H., et al. (2018). Identification and validation of QTL for grain yield and plant water status under contrasting water treatments in fall-sown spring wheats. *Theor. Appl. Genet.* 131, 1741–1759. doi: 10.1007/s00122-018-3111-9

Zhang, L., He, G., Li, Y., Yang, Z., Liu, T., Xie, X., et al. (2021). PILs transcription factors directly interact with SPLs and repress tillering/branching in plants. *New Phytol* 233 (3), 1414–25. doi: 10.1111/nph.17872

Zhao, B., Wu, T. T., Ma, S. S., Jiang, D. J., Bie, X. M., Sui, N., et al. (2020). TaD27-b gene controls the tiller number in hexaploid wheat. *Plant Biotechnol. J.* 18, 513–525. doi: 10.1111/pbi.13220

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

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 14 December 2022
DOI 10.3389/fpls.2022.1042078

OPEN ACCESS

EDITED BY
Xiaoli Fan,
Chengdu Institute of Biology, Chinese Academy of Sciences (CAS), China

REVIEWED BY
Sajid Fiaz,
The University of Haripur, Pakistan
Jen-Tsung Chen,
National University of Kaohsiung, Taiwan
Parviz Heidari,
Shahrood University of Technology, Iran

*CORRESPONDENCE
Mingjian Ren
mxren@gzu.edu.cn
Yuanyuan Yuan
happyxinhai20047@163.com

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

RECEIVED 12 September 2022
ACCEPTED 02 November 2022
PUBLISHED 14 December 2022

CITATION
Li K, Liu X, He F, Chen S, Zhou G, Wang Y, Li L, Zhang S, Ren M and Yuan Y (2022) Genome-wide analysis of the *Tritipyrum* WRKY gene family and the response of *TtWRKY256* in salt-tolerance.
Front. Plant Sci. 13:1042078.
doi: 10.3389/fpls.2022.1042078

Genome-wide analysis of the *Tritipyrum* WRKY gene family and the response of *TtWRKY256* in salt-tolerance

Kuiyin Li[1,2], Xiaojuan Liu[1], Fang He[1], Songshu Chen[1], Guangyi Zhou[1], Yuhai Wang[3], Luhua Li[1], Suqin Zhang[1], Mingjian Ren[1]* and Yuanyuan Yuan[4,5]*

[1]Guizhou Subcenter of National Wheat Improvement Center, College of Agronomy, Guizhou University, Guiyang, China, [2]Anshun University, Anshun, China, [3]Zaozhuang University, Zaozhuang, China, [4]Jinan Academy of Agricultural Sciences, Jinan, China, [5]Yantai Academy of Agricultural Sciences, Yantai, China

Introduction: The transcription factor WRKY is widespread in the plant kingdom and plays a crucial role in diverse abiotic stress responses in plant species. Tritipyrum, an octoploid derived from an intergeneric cross between Triticum aestivum (AABBDD) and Thinopyrum elongatum (EE), is a valuable germplasm resource for introducing superior traits of Th. elongatum into T. aestivum. The recent release of the complete genome sequences of T. aestivum and Th. elongatum enabled us to investigate the organization and expression profiling of Tritipyrum WRKY genes across the entire genome.

Results: In this study, 346 WRKY genes, from TtWRKY1 to TtWRKY346, were identified in Tritipyrum. The phylogenetic analysis grouped these genes into three subfamilies (I-III), and members of the same subfamilies shared a conserved motif composition. The 346 TtWRKY genes were dispersed unevenly across 28 chromosomes, with 218 duplicates. Analysis of synteny suggests that the WRKY gene family may have a common ancestor. Expression profiles derived from transcriptome data and qPCR demonstrated that 54 TtWRKY genes exhibited relatively high levels of expression across various salt stresses and recovery treatments. Tel1E01T143800 (TtWRKY256) is extremely sensitive to salt stress and is on the same evolutionary branch as the salt-tolerant A. thaliana genes AtWRKY25 and AtWRKY33. From 'Y1805', the novel AtWRKY25 was cloned. The Pearson correlation analysis identified 181 genes that were positively correlated (R>0.9) with the expression of TtWRKY256, and these genes were mainly enriched in metabolic processes, cellular processes, response to stimulus, biological regulation, and regulation of biological. Subcellular localization and qRT-PCR analysis revealed that TtWRKY256 was located in the nucleus and was highly expressed in roots, stems, and leaves under salt stress.

Discussion: The above results suggest that TtWRKY256 may be associated with salt stress tolerance in plants and may be a valuable alien gene for improving salt tolerance in wheat.

KEYWORDS

Tritipyrum, WRKY, salt-tolerance, *TtWRKY256*, genome-wide, expression patterns

Introduction

Salinization of the soil is one of the principal abiotic stressors, as salt inhibits plant growth and development and affects food productivity (Van Zelm et al., 2020). On recent years, food production in high-yielding fields has slowed, but the vast majority of low- and medium-yielding farms have significant space for development. How to efficiently utilize saline and other low- and medium-yielding areas to boost total food production is a crucial issue that must be addressed. Salt stress can result in basic strains like as osmotic stress and ion toxicity, as well as secondary effects including oxidative stress and nutritional stress (Yang and Guo, 2018). Multiple pressures have an impact on cell growth and metabolic activities, which in turn affects seed germination, seedling growth, and crop output (Munns and Tester, 2008). Plants have evolved complex mechanisms to deal with salt stress at the morphological-structural, physiological-metabolic, and molecular levels, including reduced leaf number and area, stomatal closure, accumulation of osmoregulatory substances, Na^+ and Cl^- efflux and compartmentalization, scavenging of reactive oxygen species, and modifications in stress-responsive gene expression (Dai et al., 2018; Hou et al., 2018). The expression of stress-responsive genes determines the degree to which morphological structure and physiological metabolic levels improve, and transcription factors play a crucial role in regulating this expression (Liang et al., 2018).

WRKY are specific plant transcription factors whose members are involved in a wide range of biological processes, including morphogenesis, biotic and abiotic stresses, seed germination, and plant senescence (Chen et al., 2017). All members of the WRKY transcription factor family contain a WRKY domain (WD) consisting of 60 amino acid residues. The N-terminal part of the domain is the WRKYGQK sequence, which is associated with DNA binding activity, and the C-terminal part of the domain is the C_2H_2 ($C\text{-}X_{4\text{-}5}\text{-}C\text{-}X_{22\text{-}23}\text{-}H\text{-}X_1\text{-}H$) or C_2HC ($C\text{-}X_7\text{-}C\text{-}X_{23}\text{-}H\text{-}X_1\text{-}C$) zinc finger structure, which is involved in protein interactions and auxiliary DNA binding (Eulgem et al., 2000; Jiang et al., 2017b). The W-box is the shortest sequence necessary for the binding of WRKY transcription factors to DNA. The W-box has highly conserved C/TTGACT/C amino acid residues, of which TGAC is the most conserved region (Ülker and Somssich, 2004). Mutations in any of the nucleotides will affect the ability to bind WRKY transcription factors, which are also involved in the WRKY-W-box binding reaction through phosphorylation reactions, and the phosphorylation process requires the involvement of Zn^{2+} (Maeo et al., 2001; Duan et al., 2007). WRKY transcription factors are classified into three classes according to the number of structural domains and structural domain differences. Class I contains two WRKYGQK sequences, while classes II and III each contain one. Class I and class II have the same zinc finger structure as $C\text{-}X_{4\text{-}5}\text{-}C\text{-}X_{22\text{-}23}\text{-}H\text{-}X_1\text{-}H$, while class III has a zinc finger structure of $C\text{-}X_7\text{-}C\text{-}X_{23}\text{-}H\text{-}X_1\text{-}C$. Based on the differences in amino acid sequences, class II can be further subdivided into five subclasses: IIa, IIb, IIc, IId, and IIe (Yang et al., 1999).

A large number of salt-responsive WRKY transcription factors have been identified from a variety of plants, with subfamily IIc WRKY transcription factors showing key roles (Chen et al., 2017). Salt-responsive WRKY transcription factors drive or repress target gene transcription by interacting with W-box elements in promoters, participating in ABA, ethylene, and SOS signaling pathways, and acting as intermediate factors in the interaction of different signaling pathways. In *Arabidopsis*, overexpression of *AtWRKY25* and *AtWRKY33* genes can enhance plant salt tolerance by regulating the SOS pathway (Jiang and Deyholos, 2009; Rajappa et al., 2020), while expression of *AtWRKY8* is up-regulated under salt stress and enhances plant resistance to salt stress by activating the *RD29A* gene (Chen et al., 2013); In cotton, *GhWRKY17* can be induced by drought, salt, H_2O_2 and ABA, and its constitutive expression in tobacco suppresses the transcriptional levels of ABA-inducible genes (including AREB, DREB, NCED, ERD, LEA) and blocks the ABA signalling pathway (Yan et al., 2014). In response to salt stress, plants usually produce ROS, which mediates the response to salt stress through the ROS signaling process. *Sorghum bicolor SbWRKY50* can negatively regulate the salt response by reducing the expression level of the *Arabidopsis* Na^+/H^+ reverse transporter protein gene *AtSOS1* (Song et al., 2020).

Abbreviations: ABA, abscisic acid; AREB, ABA-responsive element binding; CaMV, Cauliflower mosaic virus; CAT, catalase; DREB, dehydration-responsive element binding; ERD, early responsive to dehydration; GFP, green fluorescent protein; GUS, β-glucuronidase; LEA, late embryogenesis-abundant protein; qPCR, quantitative real-time PCR; NCED, 9-cis-epoxycarotenoid dioxygenase; POD, peroxidase; ROS, reactive oxygen species; SOD, superoxide dismutase; SOS, salt overly sensitive.

Fortunella crassifolia FcWRKY40, on the other hand, can directly activate the expression of *FcSOS2*, a serine/threonine protein kinase gene in the SOS pathway, and indirectly regulate the expression of *FcSOS1* and *FcSOS3* genes to promote Na^+ efflux and positively regulate the response to salt stress. In addition, *FcWRKY40* can in turn be induced by ABA to be expressed as a target of the ABA response element binding factor *FcABF2*, and *FcWRKY40* may be a key transcription factor for the formation of cross-talk between the SOS and ABA pathways (Dai et al., 2018). In *GhWRKY17* overexpressing tobacco, the levels of ROS accumulation, electrolyte leakage, and malondialdehyde accumulation were significantly increased, and the expression levels of ROS scavenging genes, CAT and SOD genes, as well as proline content and enzyme activities were significantly decreased, thus reducing the tolerance of tobacco to salt stress (Yan et al., 2014). In *DgWRKY5* overexpressing chrysanthemum, the expression of ROS scavenging genes (including POD, CAT and SOD) was up-regulated, which improved the resistance of chrysanthemum to salt stress (Liang et al., 2017).

T. aestivum is a moderately salt-tolerant crop, with higher salt tolerance than rice but lower than barley, and is one of the major cultivated crops in saline lands. *Th. elongatum* is a close relative of *T. aestivum* and can grow in salt concentrations similar to seawater. The octoploid *Tritipyrum* obtained by intergeneric crosses between *T. aestivum* (AABBDD) and *Th. elongatum* (EE) is an important germplasm resource for introducing *Th. elongatum* salt tolerance genes into *T. aestivum*. The genomes of *T. aestivum* and T*h. elongatum* have been completely sequenced, giving a solid foundation for structural and functional investigations of the relevant genes (Mayer et al., 2014; Wang et al., 2020). In this study, the genomic structure characteristics, chromosomal positions, gene duplication, and evolutionary divergence of the WRKY gene family in *Tritipyrum* were investigated. In addition, the expression profiles of 54 *TtWRKY* genes were examined in response to salt stress. Finally, *TtWRKY256* was cloned and its subcellular location and expression levels during salt stress and recovery were determined. These results will pave the way for suggestions on how to improve salt tolerance in plants. These results will pave the way for suggestions on how to improve plant salt tolerance.

Material and methods

Plant material

Tritipyrum is a synthetic octoploid, which contains A, B, and D genomes from *Triticum aestivum* and a set of E genomes from *Thinopyrum elongatum*. Salt-tolerant 'Y1805' is a stable *Tritipyrum* octoploid from a wide cross between *T. aestivum* and *Th. elongatum*. The *Tritipyrum* protein and nucleic acid sequences used for identifying the WRKY genes in this study were obtained from *T. aestivum* genome database (http://plants.ensembl.org/Triticum_aestivum/Info/Index) and the *Th. elongatum* genome database (https://ngdc.cncb.ac.cn/gwh/Assembly/965/show). The genome sequences of *Arabidopsis thaliana*, *Hordeum vulgare*, *Oryza sativa*, *Zea mays*, and *Thinopyrum intermedium* were located in the Plant Genomics portal Phytozome13 (https://phytozome-next.jgi.doe.gov/). The *Secale cereal* genome sequences were obtained from the China National GeneBank DataBase (https://ngdc.cncb.ac.cn/gwh/Assembly/12832/show). Publicly transcriptome available datasets were analyzed in this study. This data can be found here: [https://dataview.ncbi.nlm.nih.gov/object/PRJNA769794?reviewer=7v7it5jpc4odu8mlni9frga27g/accessionnumber:PRJNA769794].

Genomic *in situ* hybridization

Seeds were germinated at 25°C on moist filter paper in Petri dishes, then kept at 4°C for about 24 hours before being transferred to 25°C. Roots 1 to 2 cm in length were cut and treated in ice water for approximately 24 h before fixation in Carnoy's solution. After fixation, the root tips were stained with carbol fuchsin, and their mitotic chromosomes were observed under a microscope. When the plants reached the flag leaf stage, spikes were sampled and anthers at metaphase I (MI) of meiosis were fixed in Carnoy's solution, dissociated in 1 M HCl at 60°C for 6 to 8 min, and homogenized in 1% acetocarmine. *Th. elongatum* DNA was labeled with fluorescein-12-dUTP by the nick translation method and used as probes. Sheared genomic DNA from Chinese Spring (AABBDD, 2n = 42) was used as blocking DNA. In the Vectashield mounting medium (Vector Laboratories, USA), the slides were counterstained with propidium iodide (PI, 0.25 mg/mL).

Identification of the WRKY genes in *Thinopyrum*

From the Pfam database (http://www.pfam.sanger.ac.uk/), the consensus protein sequences (PF03106) of the WRKY hidden Markov model (HMM) were downloaded. Additionally, a search file library was created from 68 reported AtWRKY sequences (Eulgem et al., 2000) that were acquired from the UniProt database (www.uniprot.org). With the published sequences of the WRKY proteins from *A. thaliana* and their WRKY domain as query sequences, we utilised the Basic Local Alignment Search Tool algorithm software (BLASTP) to look for *Thinopyrum* WRKY (*TtWRKY*) proteins. After culling duplicates, we used the Pfam database and the SMART tool (http://smart.embl-heidelberg.de/) to verify the validity of the remaining candidate sequences (Letunic et al., 2012). ExPASy (http://web.expasy.org/protparam/) was used to generate the genes' physical and chemical properties.

Phylogenetic analyses and conserved motif determination

The UniProt database was used to obtain *A. thaliana* WRKY protein sequences for phylogenetic trees. The Clustalx2.0 tool was used to align several amino acid sequences from discovered WRKY genes. The phylogenetic trees comparing *Tritipyrum* and *A. thaliana* were built using the NJ method, with the Poisson model and 1000 bootstrap replications as the particular parameters. The MEME online tool (http://meme.nbcr.net/meme/intro.html) was used to determine the conserved motifs in *TtWRKY* proteins, with the parameters set to the optimum mode width of 6 to 200 and the maximum number of motifs of 10. The phylogenetic tree was shown, edited, and coloured using FigTree software and iTOL (http://itol.embl.de/).

Chromosomal distribution and gene duplication of the *TtWRKY* genes

The method of mapping *TtWRKY* genes to the chromosomes of *Tritipyrum* was performed according to Liu et al. Analysis of the *TtWRKY* gene replication events was conducted using multiple collinear scanning toolkits (MCScanX) using default parameters. All-vs-all protein sequence comparisons necessary for MCScanX were performed using DIAMOND v0.8.25 (–max-target-seqs 5 –evalue 0.00001) (Buchfink et al., 2015).

Plant growth conditions and stress treatments

In a growth chamber (relative humidity of 75% and a 20/15°C light-dark photocycle), the seeds of 'Y1805' were germinated. The seedlings were seeded on a floater board in 1/2 Hoagland's solution with a 16/8 h light/dark cycle, 400 µmol $m^{-2}s^{-1}$ irradiance, and the same temperature and relative humidity as the germination chamber. Every three days, the culture solution was renewed. On the fourteenth day (two-leaf stage), salt stress treatments were initiated (1/2 Hoagland's solution supplemented with 250 mM NaCl). 5 hours after exposure to salt stress, the first root, stem, and leaf samples of uniform size were collected from *T. aestivum*. After 24 hours of salt stress, the materials were recovered (in 1/2 Hoagland's solution without NaCl). The second sample was taken 1 hour after recovery. CK1 and CK2 were employed as parallel controls, both consisting of normal (1/2 Hoagland's solution without NaCl) grown materials. All tissue samples were immediately frozen in liquid nitrogen and preserved at -80°C for qPCR, as well as gene cloning. Three biological replications were utilised, with at least ten seedlings per replicate being combined.

Expression analysis and qPCR validation of WRKY genes under salt stresses and recovery

These RNA-Seq data sets were utilised for the salt stress and recovery expression investigation of *Tritipyrum* WRKY genes. Using FeatureCounts v1.5.1, reads corresponding to WRKY genes were tallied (Liao et al., 2014). Calculated read abundance as transcripts per kilobase million (TPM). The R Circlize package was used to log transform and visualise the results (https://cran.r-project.org/web/packages/circlize.html). GO annotation assignment was utilized to achieve functional gene annotation by mapping GO terms utilizing GO (http://www.geneontology.org/), and KEGG (http://www.genome.jp/kegg/) databases in the BLAST2GO tool (Biobam Bioinformatics S.L., Valencia, Spain, http://www.blast2go.com/b2ghome/about-blast2go) with an E-value threshold of 10^{-6}. Primer 5.0 was used to design the primer sequences (Table S2). As an internal control, we utilized the *Actin* gene, which was consistently expressed at each growth stage in nearly all tissues. The *ACTIN* gene was utilized to calibrate the detection of three technical repeats of the three biological repeats, and technique $2^{-\Delta\Delta Ct}$ was employed to assess the expression. RNA reverse transcription, WRKY gene amplification, and plasmid construction

Utilizing a PrimeScript™ RT reagent Kit (Takara), RNA was reverse-transcribed into cDNA. The complete coding sequence of *TtWRKY256* was amplified from 'Y1805' cDNA using *Bsa* I restriction site-containing primers at the 5' and 3' ends of the amplified fragment. The amplification primers are presented in Table S2. Following the manufacturer's instructions, the amplified fragment was digested with *Sac*/*Spe* I and *BamH* I/*Kpn* I and then inserted into the pBI121 vector using T4-DNA ligase (Takara). This vector has been altered to include the gene for green fluorescence protein (GFP). The inserted sequence was driven by the 35S promoter of CaMV.

Sequence alignments and phylogenetic analysis of *TtWRKY256* gene

DNAMAN performed multiple alignments of the WRKY gene sequences. From Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html) and Ensembl (http://plants.ensembl.org/Triticumaestivum/Info/Index) servers, 21 DNA sequences encoding for WRKYs were retrieved. With MEGA X and the maximum likelihood method and 1000 bootstrap replications, a phylogenetic tree was constructed (Kumar et al., 2018).

Subcellular localization of *TtWRKY256*

The recombinant *TtWRKY256*-GFP and the vector pBI121-GFP as a negative control were infiltrated into tobacco epidermal

cells *via Agrobacterium tumefaciens*-mediated transformation. After incubation for 24 h, the green fluorescent protein (GFP) fluorescence signal was observed using a confocal microscope (FV1000 Olympus Corp., Tokyo, Japan).

Statistical analysis

SPSS software was used to conduct an analysis of variance (ANOVA) on the study's data (IBM Corporation). At a significance level of 0.05, mean values were compared using Fisher's least significant difference (LSD) test. The histograms were created with the Origin 8.0 program (OriginLab Corporation, Northampton, Massachusetts, USA).

Results

Identification of the *TtWRKY* genes in *Tritipyrum*

Cytogenetic analysis revealed that *Tritipyrum* 'Y1805' had 56 chromosomes, 42 of which were from *T. aestivum* (red) and 14 from *Th. elongatum* (green) (Figure 1A). After eliminating the duplicated sequences, 346 WRKY genes were retrieved from the *Tritipyrum* genome using the hidden Markov model (HMM) and BLASTp approaches in this study. WRKY genes were renamed based on where they were location on *Tritipyrum* chromosomes (Supplementary Table S1). The results showed that the 346 WRKY genes identified in *Tritipyrum* were mainly distributed in D subgenomes and five homologous groups (Figures 1B, C). The predicted WRKY proteins of *Tritipyrum* varied greatly in length and molecular weight (MV). The *Tritipyrum* WRKY genes encoded proteins ranging from 135 (TraesCS1A02G348600.1 and TraesCS1D02G351600.1) to 1482 (TraesCS5A02G344100.1) amino acids (aa) in length and from 14.93 (TraesCS1D02G351600.1) to 165.34 (FtPinG0000377600.01) kDa in MV. The isoelectric points (PIs) of the proteins ranged from 4.64 (TraesCS3D02G289500.1) to 10.86 (TraesCS5D02G070700.1) (Supplementary Table S1).

Phylogenetic analysis, and motif composition of the *TtWRKY* gene

According to the phylogenetic tree and motif composition, these gene families were divided into seven subfamilies

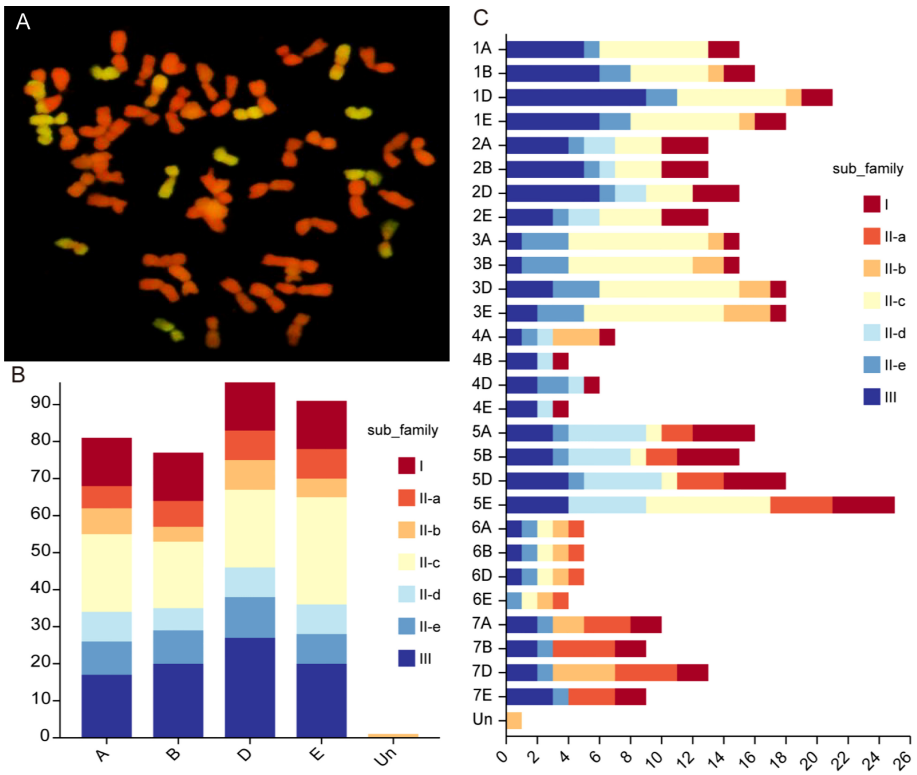


FIGURE 1
Chromosome configuration and distribution of WRKY genes in *Tritipyrum*. **(A)** Chromosome configuration of *Tritipyrum* 'Y1805'; **(B)** Subgenomes distribution of WRKY genes in *Tritipyrum*; **(C)** Chromosomal distribution of WRKY genes in *Tritipyrum*.

(Figure 2A; Figure S1). To investigate the evolutionary relationships and classify the WRKY family in *Tritipyrum*, a phylogenetic tree was constructed using 414 potential WRKY structural domains identified in *A. thaliana* and *Tritipyrum* (Figure 2A; Figure S1). According to the *AtWRKY* classification and primary amino acid sequence characteristics in the model organism *A. thaliana*, the WRKY family members in *Tritipyrum* could be divided into three major groups (Group I-III). Among them, 52 *TtWRKY* containing 2 WDs and 1 C2H2-type zinc finger structure belong to Group I; Group II contains 210 members with 1 WD and C2H2-type zinc finger structure in each sequence, and the family can be further divided into 5 subfamilies IIa, IIb, IIc, IId, and IIe, containing 29, 25, 89, 30, and 37 members; 84 *TtWRKY* contained 1 WD and C2HC-type zinc finger structure and belonged to Group III. In this study, Group II was the largest WRKY transcription factor family in *Tritipyrum*, accounting for 60.7% of the total, and subfamilies IIa and IIc and IId and IIe clustered to one branch, respectively, which was consistent with previous studies in *T. aestivum*. This is consistent with the results of previous studies in *T. aestivum*. The box plot indicates that the genetic distance I_vs_II is smaller than the genetic distance I_vs_III and II_vs_III, and that the genetic distance between Is is the smallest, indicating that Is and IIs are more closely related and that the relationship between Is is the closest (Figure 2B). In addition, the genetic distances of IIs and IIIs were comparable, but the range of IIs was the smallest, indicating that IIs had less sequence divergence than IIIs (Figure 2B).

A total of 10 different motifs, named motif 1-10 (Figure 2A; Figure S1), were identified using the MEME online program. Motif 1 and motif 6 both contain the characteristic sequence WRKYGQK of WRKY proteins, and almost all identified WRKY proteins contain at least one of these motifs. Motif anlysis indicates that WRKY proteins of the same family or subfamily in *Tritipyrum* have similar motif composition. For example, motif 9 is unique to Group I, while motif 8 is unique to Group IIa and Group IIb. The motifs specific to a family may be involved in plant-specific biological processes. Thus, each family or subfamily of WRKY genes may be associated with a specific biological process. Motif distribution patterns of members of subfamilies IIa and IIb are the same, indicating that their functions are similar. These two families were also clustered onto the same branch in the phylogenetic tree, and the same phenomenon was observed for the IId and IIe subfamilies. These results further validate the classification of WRKY in *Tritipyrum* and its phylogenetic relationships.

Chromosomal distribution, gene duplication and synteny analysis of the *TtWRKY* gene

Of the 346 *TtWRKY* genes, 345 were localized to 28 chromosomes, and only *TtWRKY255* was not shown in the chromosome localization map due to anchoring on scaffolds (Supplementary Table S1; Figure 3). Most of the *TtWRKY* genes were distributed on chromosomes 5E (25, 7.22%) and 1D (21, 6.06%), while chromosomes 4B, 4E and 6E contained only four *TtWRKY* genes (Supplementary Table S1; Figure 1, 3). Based on where they were on the chromosome, most of the *TtWRKY* genes were at the ends, and only a few were close to the near-central region (Figure 3) The above results indicated that the localization of WRKY genes on chromosomes was non-random and unevenly distributed. In this study, a total of 405 *TtWRKY* gene duplication pairs were identified, including 218 genes (Figure 3). This suggests

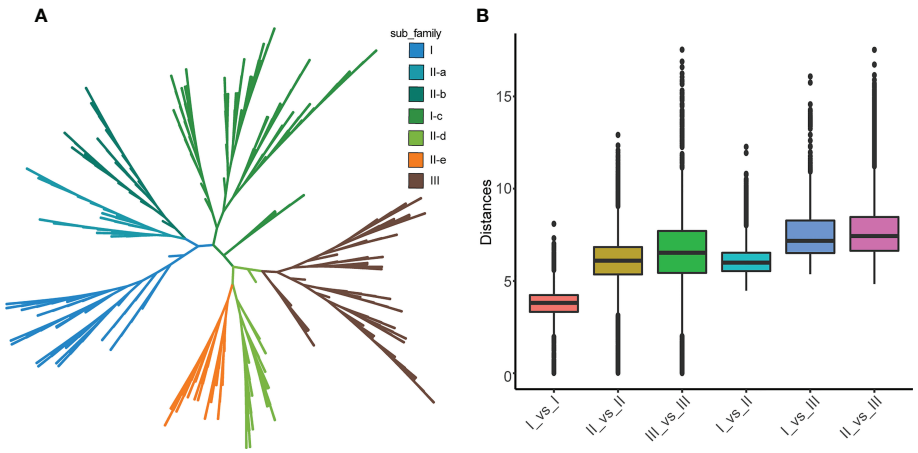


FIGURE 2
phylogenetic relationships and distance among the WRKY proteins from *Tritipyrum* and *A thaliana*. **(A)** Phylogenetic relationships among 414 WRKY proteins from *Tritipyrum* and *A thaliana*; **(B)** genetic distance among the different clades of WRKY genes. The box plot shows the median (black line), interquartile range (box), and maximum and minimum scores (whiskers) of each data set.

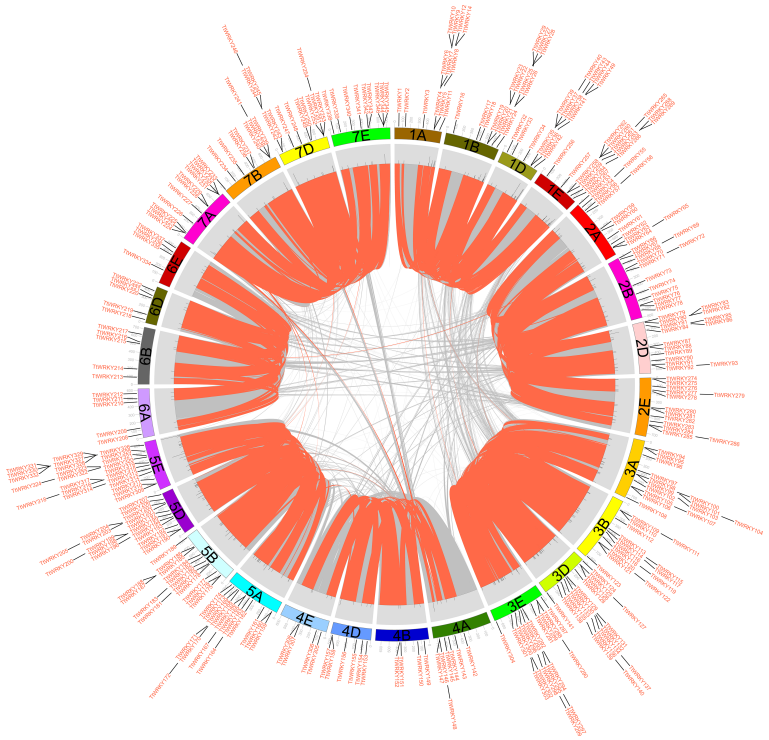


FIGURE 3
distribution, duplication and synteny analysis of WRKY genes in *Tritipyrum*. Collinear correlations of WRKY in *Tritipyrum* genomes are displayed by Circos. *Tritipyrum* chromosomes are colored according to the inferred ancestral chromosomes following an established convention. In the center, the relative map position of 345 WRKY genes is shown on each of the 28 *Tritipyrum* chromosomes.

that some genes contain more than one duplicate gene, probably due to multiple replication throughout *T. aestivum* genome. The tandem duplicated gene pairs were mainly distributed in the first, third, and fifth homologous groups, with the vast majority of homologous genes distributed on the same homologous groups and only a few in the fourth, fifth, and seventh homologous groups, which is consistent with the natural translocations generated during the formation and evolution of common *T. aestivum*.

Evolutionary analysis of the WRKY Families in several different species

To further deduce the evolutionary ties between members of the WRKY gene family in *Tritipyrum*, *H. vulgare*, *O. sativa*, *S. cereal*, *Th. intermedium*, and *Z. mays*, the syntenic relationships between the six species were studied. Five distinct classes of syntenic linkages were identified (Figure 4). 255 *TtWRKY* genes shared syntenic connections with those of *Th. intermedium*, followed by *Z. mays* (251), *O. sativa* (262), *H. vulgaris* (218), and *S. cereal* (185). (Figure 4). Intriguingly, collinear pairs were discovered between *Tritipyrum* and the other five species, suggesting that these orthologous pairs may have existed prior to the original separation. In addition, certain WRKY collinear gene pairs found between *Tritipyrum* and *H. vulgare* were anchored to the highly conserved syntenic blocks, which encompass over 500 collinear sites. Similar trends were also identified between *Tritipyrum* and *S. cereal*, which may be attributable to the evolutionary connection between *Tritipyrum* and the other five plant species. Significantly, several *TtWRKY* genes were discovered to be connected with at least three syntenic gene pairs, indicating that these genes may have played a crucial role in the evolution of the WRKY gene family. These results revealed that the *TtWRKY* gene family was highly conserved and that the *TtWRKY* genes were more closely related to those of *H. vulgare* than to those of *Z. mays*. The *TtWRKY* genes in several plants may have developed from a common ancestor.

Expression of *TtWRKY* genes under salt stresses and recovery

To confirm whether the expression of *TtWRKY* genes was influenced by different salt stresses and recovery treatments, the transcriptional levels of all 346 *TtWRKY* genes under different salt stresses and recovery treatments were investigated. Among

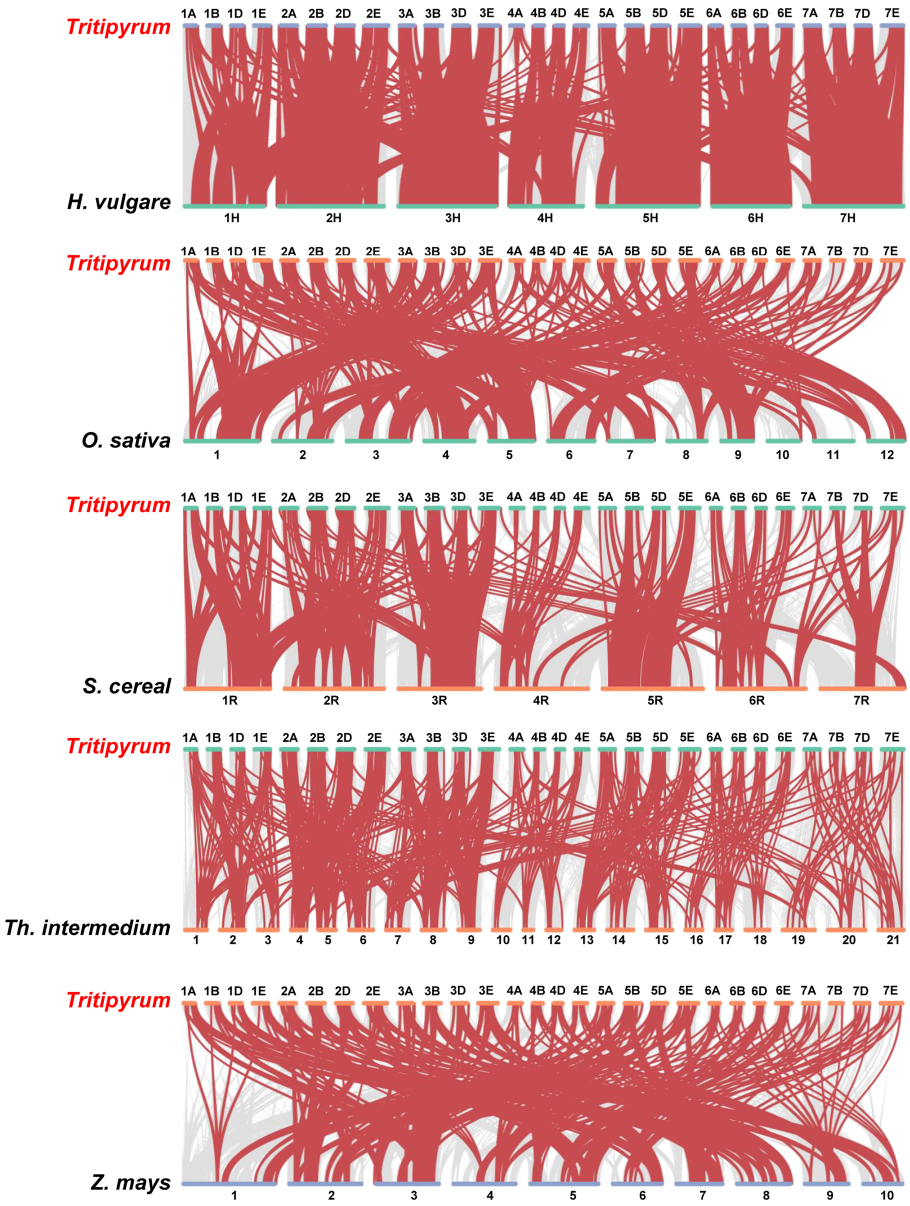


FIGURE 4
Synteny analyses between *Tritipyrum* and five representative plant species. Gray lines in the background indicate collinear blocks within *Tritipyrum* and other plant genomes, while red lines highlight syntenic WRKY gene pairs.

the 346 *TtWRKY* genes, 249 *TtWRKY* genes were expressed in all 11 samples tested, and 107 genes showed constitutive expression (TPM>1 in all samples). These 249 *TtWRKY* genes were mainly from Group IIc and Group III subfamilies, and the clustering analysis showed that the corresponding WRKY genes were not significantly correlated with subfamily types for salt stress and recovery treatments (Figure 5A). 87 *TtWRKY* genes were not expressed in all detected samples, which may be pseudogenes or not a response to salt stress and recovery treatment. GO annotation of 249 expressed genes showed that the biological process and molecular function of these genes was cellular process, metabolic process, response to stimulus, regulation of biological process, biological regulation, biosynthetic process, regulation of metabolic process, cellular metabolic process and transcription regulator activity, DNA-binding transcription factor activity, organic cyclic compound binding, heterocyclic compound binding, nucleic acid binding, DNA binding (Figures 5B, C).

To further verify the accuracy of transcriptome data, 54 *TtWRKY* members, whose mRNA levels were relatively high

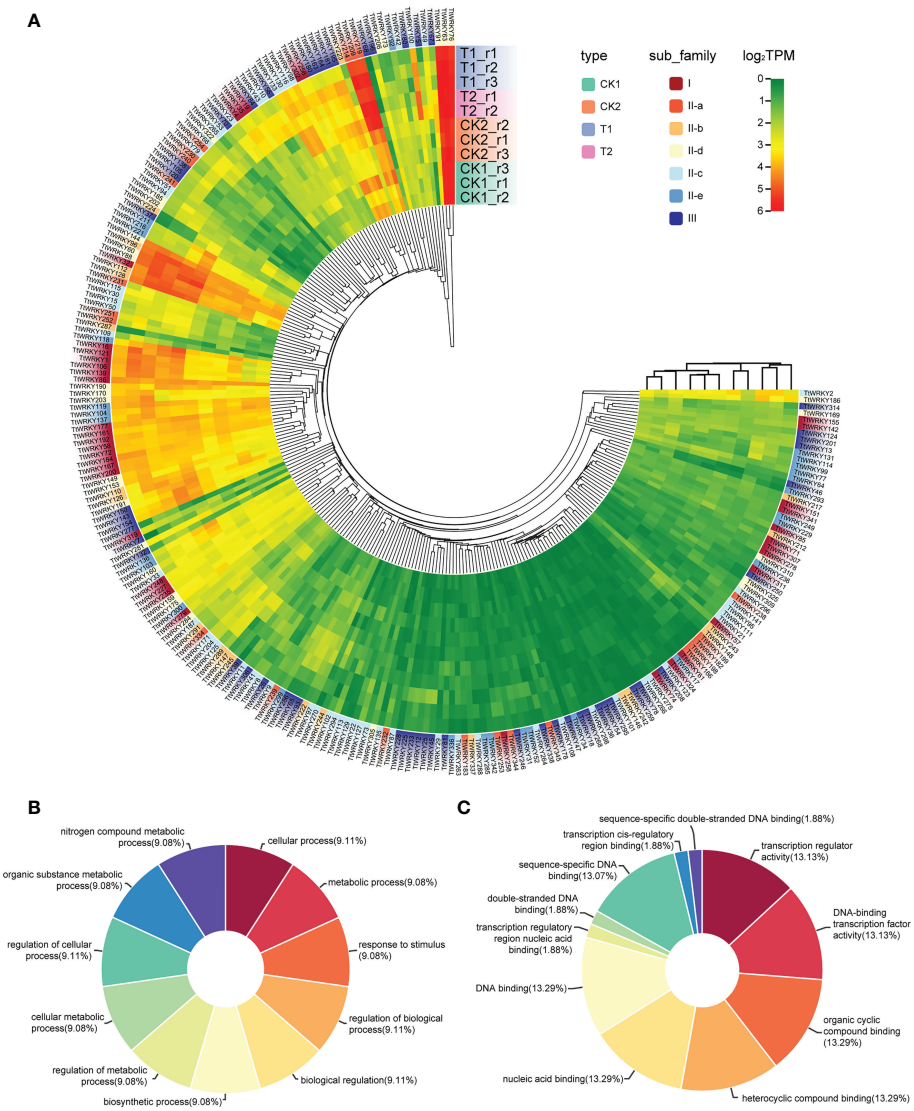


FIGURE 5
Expression patterns of *TtWRKY* genes under salt stresses and recovery treatments. **(A)** Hierachical clustering of expression profiles of 249 *TtWRKY* genes were expressed in 11 samples including salt stress and recovery treatment. **(B, C)** the BP **(B)** and MF **(C)** analysis of 249 expression genes.

(log_2FoldChange>1) across different salt stresses and recovery treatments, were carefully selected from 249 *Tritipyrum* WRKY genes. We designed specific primers for the fifty-four genes (Supplementary Table S2). Quantitative real-time PCR (qPCR) experiments were further performed to analyze their expression patterns in response to different salt stress and recovery treatments. Quantitative variations of the selected genes between qPCR and transcriptome were roughly similar (Figures 6A–C); gene expression trends were significantly similar (r^2 = 0.82) with those from the transcriptome data, indicating that our transcriptome results were reliable (Figure 6D).

TtWRKY256 cloning and sequence analysis

Earlier studies showed that overexpression of *AtWRKY25* (AT5G07100) and *AtWRKY33* (AT2G38470) genes in *A. thaliana* enhances plant salt tolerance by regulating the SOS pathway. *TtWRKY256* (Tel1E01T143800) gene of the *Tritipyrum* WRKY gene family and AT5G07100 and AT2G38470 are located in the same evolutionary tree branch (Figure S1) and are up-regulated in different salt stresses and recovery treatments (Figure 5). Using Tel1E01T143800-specific primers, a 1422 bp

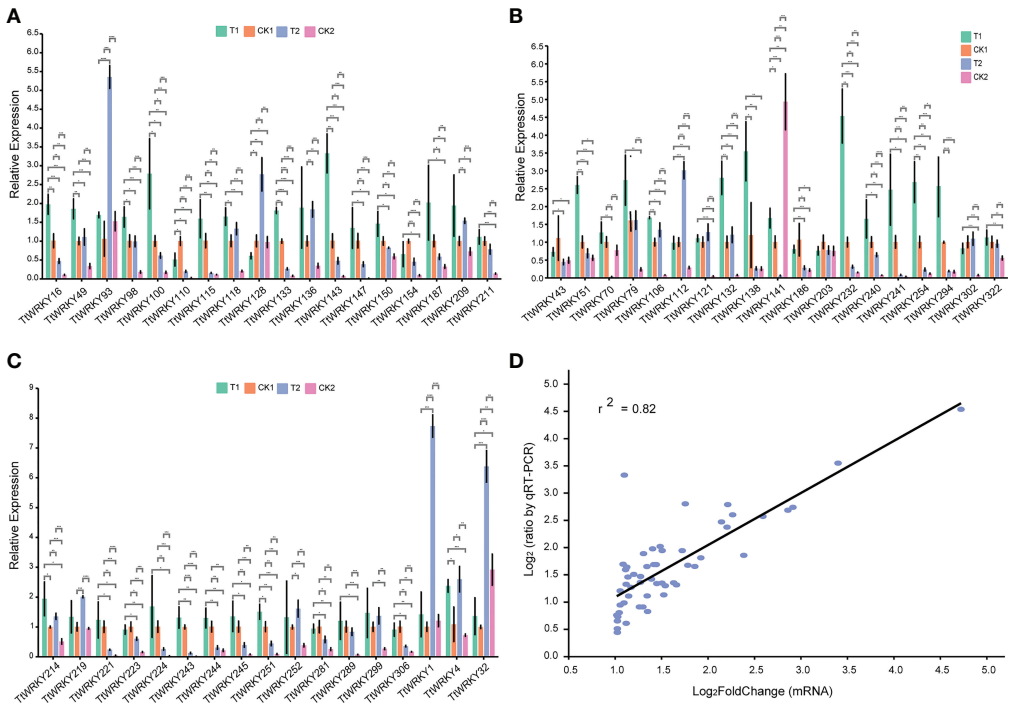


FIGURE 6
(A–C) Expression analysis of 54 WRKY genes in eleven samples by qPCR. Data were normalized to β-actin gene and vertical bars indicate standard deviation. **(D)** The relationships between qPCR and transcriptional of 54 up-regulated expression genes. Values are the log_2ratio (salt stress or recovery treatment/CK treatment) for genes. The determine coefficient (r^2) is indicated in the figure. All qPCR reactions were performed in three biological replicates. Asterisk, double and triple asterisks indicate significant differences ($p < 0.05$, 0.01 and 0.001, respectively) between groups.

cDNA fragment corresponding to Tel1E01T143800was amplified and cloned from *Tritipyrum* 'Y1805' by PCR (Figure S2) and named *TtWRKY256*. The *TtWRKY256* sequence had 99.41% identity to Tel1E01T143800, with only 7 bp nucleotides changes between them (Figure 7A). Therefore, *TtWRKY256* was similar to Tel1E01T143800 according to their cDNA sequences. A phylogenetic tree based on nucleic acid sequences of different species showed that *TtWRKY256* displayed maximum homology with *Th. elongatum* (Tel1E01T143800, Figure 7B). The genetic distance of *TtWRKY256* with *T.aestivum* was closer than *T.urartu*, *T.dicoccoides* and *A.tauschii*. The *Triticum* crops and *A.sativa* were clustered together, with *O.sativa*, *Z.mays*, *S.italica*, *B.distachyon* and *L.rigidum* being more distantly related, which is consistent with the distant affinities among the species (Figure 7B).

TtWRKY256 expression correlation analysis

To explore the biological processes associated with *TtWRKY256* expression in salt stress and recovery treatments, a Pearson correlation analysis was performed on *TtWRKY256* and other genes in transcriptome data. The results showed that 181 genes ($R > 0.9$) were positively correlated with the expression of *TtWRKY256* (Figure 8A). Biological process (BP) and KEGG analyses were performed on these 181 genes by Blast2GO software and the KEGG website (https://www.genome.jp/kegg/). The results are listed in Figures 8B, C, wherein 5 BPs for 173 genes were observed, including metabolic process, cellular process, response to stimulus, biological regulation, and regulation of biological process. As for KEGG analysis, amino acid metabolism, environmental information processing, carbohydrate metabolism, protein families: metabolism, organismal systems, and protein families: signaling and cellular processes were included. All the above results indicated that *TtWRKY256* could be associated with abiotic stress tolerance in plants.

Expression patterns of *TtWRKY256* and subcellular localization

To investigate the spatial and temporal expression pattern of *TtWRKY256*, qPCR analysis was used to examine the expression levels of *TtWRKY256* in roots under salt stress and recovery

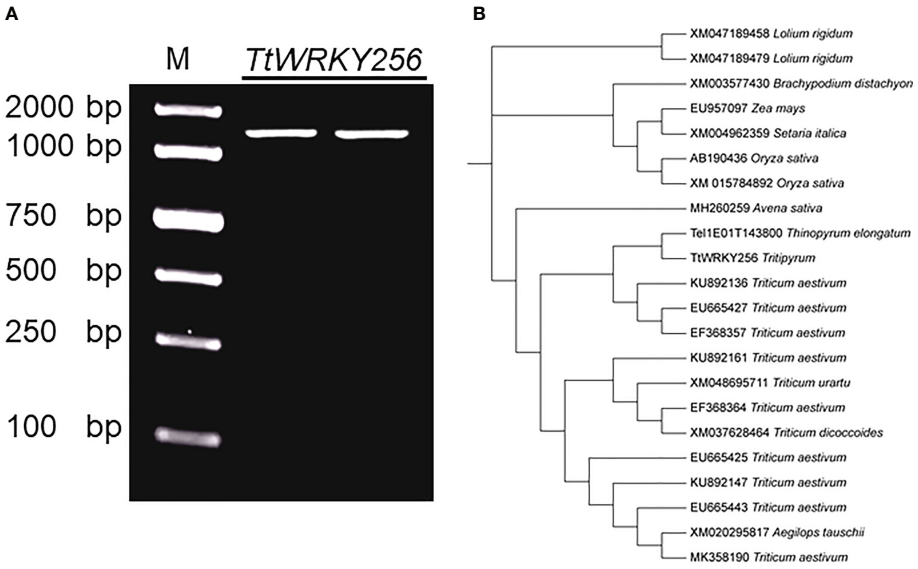


FIGURE 7
Clone and Phylogenetic relationships of *TtWRKY256*. **(A)** Amplification products using 'Y1805' cDNA; **(B)** Phylogenetic tree of WRKY nucleic acid sequences in various plant species. The Maximum Likelihood (ML) tree was generated using MEGA X with 1000 bootstrap replicates.

conditions and in diverse tissues. The *TtWRKY256* expression level was significantly (4.1-fold) higher than the control under salt stress and recovery in the roots (Figure 9A). These results were consistent with the above transcriptome data. In addition, the relative expression level of the *TtWRKY256* gene was the highest under salt stress in the leaves of Y1805, followed by stems, and then roots (Figure 9B). In short, all of the above results showed that the 'Y1805'-specific *TtWRKY256*

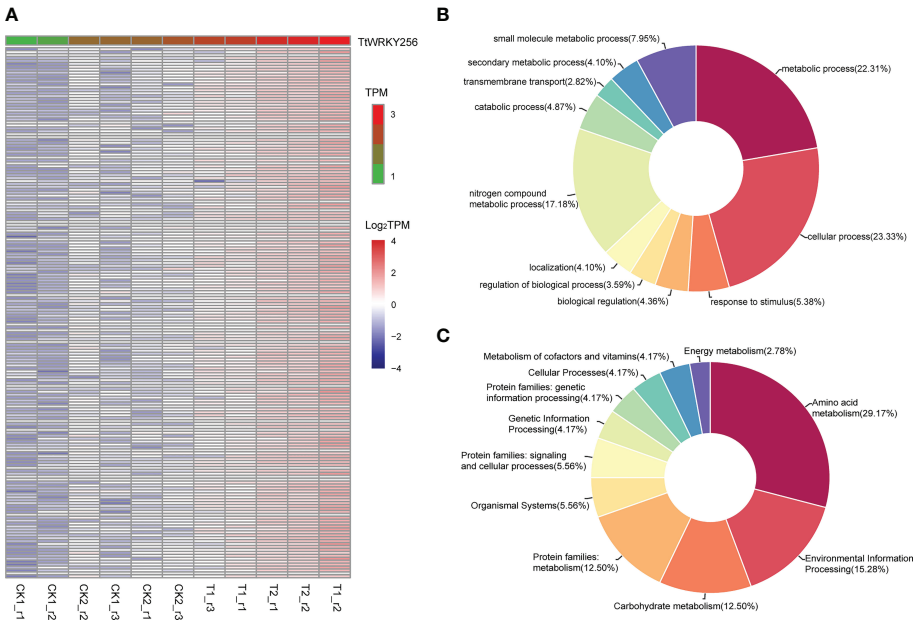


FIGURE 8
Expression correlation analysis of *TtWRKY256*. **(A)** One hundred eighty-one genes positively related (R > 0.9) with *TtWRKY256* expression; **(B, C)** the BP **(B)** and KEGG **(C)** analysis of 181 positively related to *TtWRKY256* expression genes.

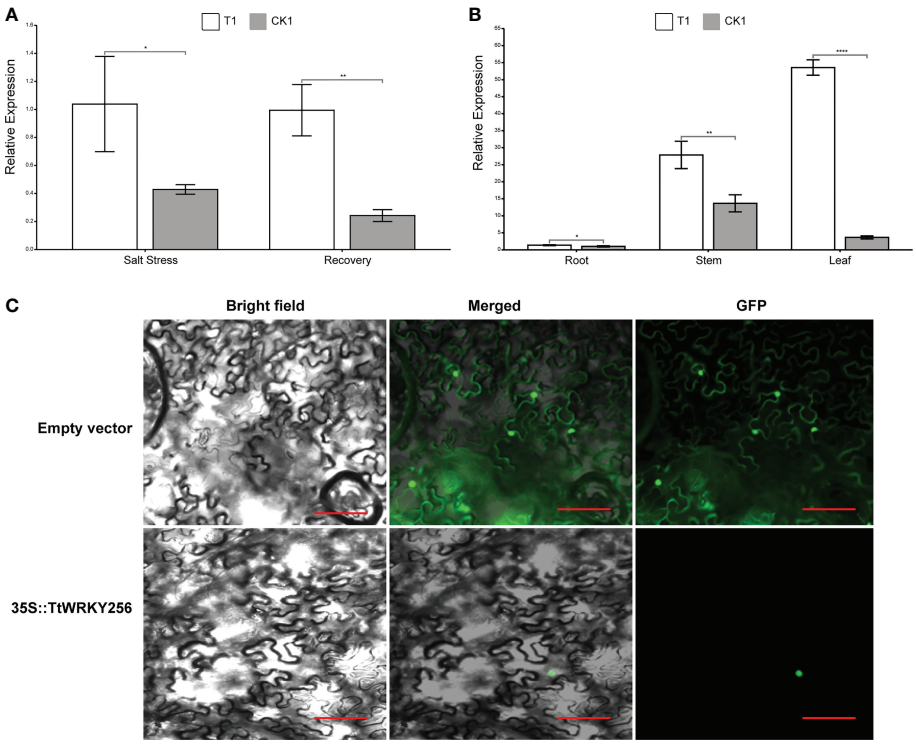


FIGURE 9
Expression and Subcellular localization of *TtWRKY256*. **(A)** Relative expression levels of *TtWRKY256* in roots under salt stress and recovery conditions; **(B)** Relative expression levels of *TtWRKY256* in roots, stems, and leaves under salt stress; **(C)** Subcellular localization of the TtWRKY256 protein in tobacco protoplasts, Scale bar = 20 mm. Asterisk, double and triple asterisks indicate significant differences ($p < 0.05$, 0.01 and 0.001, respectively) between groups.

gene was highly and sensitively expressed in the whole plant when salt stress was high. To determine the subcellular localization of *TtWRKY256*, a fusion protein transiently expressing 35S-*TtWRKY256*-GFP in tobacco epidermal cells was produced. It was found that the fluorescence emitted by the fusion protein was localized to the nucleus (Figure 9C). The results showed that *TtWRKY256* might contribute to transcription regulation or the protection of heredity substances and cellular components.

Discussion

In numerous plant species, the evolution, function, abiotic stress response, and other features of WRKY, one of the largest families of plant transcription factors, have been studied. In the present work, 346 members of the WRKY gene family were discovered in *Tritipyrum*. The 346 WRKY genes discovered in *Tritipyrum* can be categorized into three distinct groups. The proportion of WRKY genes represented by each group varied; group II contained the highest proportion of WRKY genes. Meanwhile, the genetic distances of IIs and IIIs were similar, which might be due to the group IIb originated from group III by the duplication of WRKY genes (Brand et al., 2013); the evolutionary patterns of this WRKY transcription factor have also been validated in wild rice (Jiang et al., 2017a). MEME analysis of WRKY protein sequences uncovered distinct group specificities. Group IIc WRKY proteins, for instance, are highly conserved and contain few additional motifs. Duplication of genes is essential for species evolution, genome amplification, and gene family evolution (Lynch and Conery, 2000). Whole genome duplication, tandem duplication, and segmental duplication are the three primary kinds of gene duplication (Yu et al., 2005). Genes that are duplicated may have distinct expressions. It seems that some evolutionary events such as duplication and polyploidy in plant have been extended the gene family members (Rezaee et al., 2020; Faraji et al., 2021). On the other hand, some changes in gene structure including point mutations in coding DNA sequence regions and regulatory site of duplicated members have affected the function of new members (Faraji et al., 2020; Heidari et al., 2021). The synteny of the WRKY genomes of six species, including *Tritipyrum*, *H. vulgare*, *O. sativa*, *S. cereal*, *Th. intermedium*, and *Z. mays*, was studied. The most syntenic linkages were discovered between

Tritipyrum and *Th. intermedium*. This data implies that these taxa share a tight evolutionary relationship, compatible with traditional Gramineae classifications. *Tritipyrum* and *S. cereal*, which belong to separate genera, were determined to have the fewest number of syntenic links. In the five sets of syntenic relationships, 1-17 WRKY genes were shared by at least three species, which may shed light on the evolution of WRKY genes across species.

Salinity is a significant environmental danger to crop production because the excessive concentration of salt in the soil has a devastating impact on plant performance by disrupting cellular metabolism. These negative consequences of increasing salinity are mostly caused by osmotic stress and the accumulation of harmful ions in plant cells. Salinity is a major environmental threat for crop production because high concentration of salt in the soil severely affects plant performance by disturbing cellular metabolism. Such adverse effects of increased salinity occur mainly due to the osmotic stress and continuous accumulation of toxic ions within the plant cells (Munns and Tester, 2008). Genes encoding TFs are differentially expressed in plants as part of their complex stress response systems. Several transcription factors, including bHLH, WRKY, MYC, NAC, MYB, and ERF/AP2, have been linked to salt tolerance pathways (Okushima et al., 2007; Golldack et al., 2011; Lan Thi Hoang et al., 2017; Meraj et al., 2020). Many WRKY genes have been identified in Arabidopsis, rice, maize, and wheat, respectively, in response to various adversity stresses. After overexpression of the *ZmWRKY106* gene, the drought tolerance of *A. thaliana* was improved. Under drought conditions (Wang et al., 2018). *AtWRKY53* can accelerate the metabolism of starch in the guard cells and reduce H_2O_2 levels, hence promoting stomatal movement (Sun and Yu, 2015). *TaWRKY44* has multiple abiotic stress tolerance in transgenic tobacco, including drought, salt, and osmotic stress (Wang et al., 2015). Twelve GmWRKY genes appear differentially expressed in soybean under salt stress (Song et al., 2016). Plant WRKY genes with generally high expression could activate transcription of downstream target genes and thus regulate plant growth and development (Xu et al., 2016). Tissue-specific expression of WRKY genes may affect the growth and developmental processes of target tissues/organs by regulating transcriptional processes (Tang et al., 2014). In this study, 249 *TtWRKY* genes were detected in *Tritipyrum* root tissues that significantly responded to the induction of salt stress, of which 107 *TtWRKY* genes had constitutive expression characteristics, and these *TtWRKY*s may be involved in the regulation of cellular basal life processes. Fifty-four of these *TtWRKY* genes were further selected to examine their role in salt stress response. The results of expression profiling and qPCR analysis showed that these 54 *TtWRKYs* may positively regulate salt stress response in *Tritipyrum* root tissues. Transgenic experiments will be used to learn more about the exact biological functions of these *TtWRKYs*, and their use in genetic engineering to improve crop stress resistance and other agronomic traits will also be looked into.

Similar to the previous investigation of *T. aestivum*, which identified 12 TaWRKY genes as candidate genes for salt stress response, the majority of *TtWRKY* genes responded to salinity stress (Ning et al., 2017). Overexpression of *AtWRKY33* has also been demonstrated to increase *A. thaliana* tolerance to salt stress (Bao et al., 2018). *AtWRKY33* is controlled not only by salt but also by oxidative stress (Jiang and Deyholos, 2009; Birkenbihl et al., 2012). In addition, genes regulated by *AtWRKY33* are associated with ROS detoxification processes, indicating that WRKY TFs are essential regulators for stress adaptation (Jiang and Deyholos, 2009). The *TtWRKY256* gene in the WRKY gene family of *Tritipyrum* and *AtWRKY33* in *A. thaliana* are located in the same evolutionary branch, and the *TtWRKY256* gene is up-regulated in different salt stresses and recovery treatments. As a result, *TtWRKY256* was selected for additional functional research in this experiment. Here, the relative expression level of the *TtWRKY256* gene was the highest in the leaves of *Tritipyrum* under salt stress, followed by stems, and then roots. It was found that the root system was damaged directly and seriously in a high salt solution. In addition, the *TtWRKY256* expression level in the whole plant was significantly higher than in the control under salt stress and recovery. These results were consistent with our transcriptome data and previous reports (Ning et al., 2017). Because of this, *TtWRKY256* was highly and sensitively expressed in the whole plant to help the plant deal with salt stress. GO and KEGG analysis of highly related genes of *TtWRKY256* demonstrated that the highly related genes are associated with in metabolic process, cellular process, response to stimulus, biological regulation, and regulation of biological process, which may contribute to the study of the salt tolerance mechanism of *TtWRKY256*.

Soil salinization is a growing problem and has become a major factor limiting seed germination, seedling growth and crop yield. *Th. elongatum* is a closely related species of *T. aestivum* that can grow in salt concentrations similar to those of seawater. *Tritipyrum*, obtained by intergeneric hybridization between *T. aestivum* and *Th. elongatum*, is an important bridge material for introducing salt tolerance genes of *Th. elongatum* into *T. aestivum* (Baker et al., 2020; McKenna et al., 2020). Improving salt tolerance in plants mainly induces the activation of stress-responsive genes, whose expression products are involved in the repair of various aspects of primary and secondary stresses induced by salt stress. In contrast to single functional genes, transcription factor can regulate a set of downstream target genes, which in turn regulate physiological and biochemical processes in response to salt stress (Jiang et al., 2017b; Lan Thi Hoang et al., 2017). In this study, bioinformatics such as phylogenetic analysis, motif analysis and correlation analysis were used to conduct a comprehensive analysis of the

WRKY family in *Tritipyrum*. RNA-seq was used to look for WRKY transcription factors in *Tritipyrum* salt stress response as part of a transcriptomic analysis of the plant's response to salt stress. The *TtWRKY256* gene was cloned and identified using the Arabidopsis salt stress response-related gene *AtWRKY33* as a reference gene, and the analysis of gene expression levels and subcellular localization of the *TtWRKY256* gene under salt stress and recovery were completed. The response of plants to salt stress is very complex, with multiple genes involved in the regulatory network and multiple pathways acting together. The functions of the *TtWRKY* transcription factors that were screened in this paper have not been transgenically characterized. It needs to be found out if these transcription factors interact with each other, if they interact with other proteins, and if they have downstream target genes. With the discovery and wide use of salt tolerance-related transcription factor candidate genes and the continuous improvement of the understanding of the salt tolerance mechanism involved in transcription factors, genetic engineering will make it easier to grow crops that can handle salt.

Conclusions

In this study, a thorough examination of the WRKY gene family in *Tritipyrum* was conducted. 346 full-length WRKY genes were described and categorized further into three primary categories, with extremely similar motif compositions within the same groups and subgroups. Synteny analysis and phylogenetic comparison of WRKY genes from a variety of plant species yielded useful insights into the evolutionary properties of WRKY genes in *Tritipyrum*. Fifty-four *TtWRKY* genes play an important role in salt stress response in *Tritipyrum*, as evidenced by their expression patterns in different tissues and in response to salt stress and recovery treatments. In addition, *TtWRKY256* may be a potential target gene for enhancing wheat's salt tolerance *via* biotechnology or molecular breeding. These data provide a great resource for gaining a deeper comprehension of the biological functions of particular WRKY genes in *Tritipyrum*.

Data availability statement

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

Author contributions

KL planned and designed the research and analysed the data. KL and XL wrote the manuscript. FH, SC, and GZ studied gene expression by qPCR. FH identified the *Tritipyrum* WRKY gene family and analysed gene structure. YW and SZ studied chromosome distribution and gene duplication. YY analysed the evolutionary relationship of WRKY genes in several different species. MR supervised the research. YY and MR revised the manuscript. All authors read and approved the final manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Research Fund from the Science and Technology Department of Guizhou Province, China (ZK2022-315, 2109-4246, 2109-1073, and 2017-5788), the National Natural Science Foundation of China (32160474, 32001433, 32160456 and 31660390), and the Key R & D Program of Shandong Province (Major Science and Technology Innovation Project) (2021LZGC009), Cooperation project from Jinan Academy of Agricultural Sciences (YY201902).

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1042078/full#supplementary-material

References

Baker, L., Grewal, S., Yang, C.-y., Hubbart-Edwards, S., Scholefield, D., Ashling, S., et al. (2020). Exploiting the genome of thinopyrum elongatum to expand the gene pool of hexaploid wheat. *Theor. Appl. Genet.* 133 (7), 2213–2226.

Bao, W., Wang, X., Chen, M., Chai, T., and Wang, H. (2018). A WRKY transcription factor, PcWRKY33, from polygonum cuspidatum reduces salt tolerance in transgenic arabidopsis thaliana. *Plant Cell Rep.* 37 (7), 1033–1048.

Birkenbihl, R. P., Diezel, C., and Somssich, I. E. (2012). Arabidopsis WRKY33 is a key transcriptional regulator of hormonal and metabolic responses toward botrytis cinerea infection. *Plant Physiol.* 159 (1), 266–285.

Brand, L. H., Fischer, N. M., Harter, K., Kohlbacher, O., and Wanke, D. (2013). Elucidating the evolutionary conserved DNA-binding specificities of WRKY transcription factors by molecular dynamics and *in vitro* binding assays. *Nucleic Acids Res.* 41 (21), 9764–9778.

Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12 (1), 59–60.

Chen, F., Hu, Y., Vannozzi, A., Wu, K., Cai, H., Qin, Y., et al. (2017). The WRKY transcription factor family in model plants and crops. *Crit. Rev. Plant Sci.* 36 (5-6), 311–335.

Chen, L., Zhang, L., Li, D., Wang, F., and Yu, D. (2013). WRKY8 transcription factor functions in the TMV-cg defense response by mediating both abscisic acid and ethylene signaling in arabidopsis. *Proc. Natl. Acad. Sci.* 110 (21), E1963–E1971.

Dai, W., Wang, M., Gong, X., and Liu, J. H. (2018). The transcription factor fc WRKY 40 of fortunella crassifolia functions positively in salt tolerance through modulation of ion homeostasis and proline biosynthesis by directly regulating SOS 2 and P5 CS 1 homologs. *New Phytol.* 219 (3), 972–989.

Duan, M.-R., Nan, J., Liang, Y.-H., Mao, P., Lu, L., Li, L., et al. (2007). DNA Binding mechanism revealed by high resolution crystal structure of arabidopsis thaliana WRKY1 protein. *Nucleic Acids Res.* 35 (4), 1145–1154.

Eulgem, T., Rushton, P. J., Robatzek, S., and Somssich, I. E. (2000). The WRKY superfamily of plant transcription factors. *Trends Plant Sci.* 5 (5), 199–206.

Faraji, S., Filiz, E., Kazemitabar, S. K., Vannozzi, A., Palumbo, F., Barcaccia, G., et al. (2020). The AP2/ERF gene family in triticum durum: Genome-wide identification and expression analysis under drought and salinity stresses. *Genes* 11 (12), 1464.

Faraji, S., Heidari, P., Amouei, H., Filiz, E., and Poczai, P. (2021). Investigation and computational analysis of the sulfotransferase (SOT) gene family in potato (Solanum tuberosum): Insights into sulfur adjustment for proper development and stimuli responses. *Plants* 10 (12), 2597.

Golldack, D., Lüking, I., and Yang, O. (2011). Plant tolerance to drought and salinity: stress regulating transcription factors and their functional significance in the cellular transcriptional network. *Plant Cell Rep.* 30 (8), 1383–1391.

Heidari, P., Faraji, S., and Poczai, P. (2021). Magnesium transporter gene family: Genome-wide identification and characterization in theobroma cacao, corchorus capsularis, and gossypium hirsutum of family malvaceae. *Agronomy* 11 (8), 1651.

Hou, H., Jia, H., Yan, Q., and Wang, X. (2018). Overexpression of a SBP-box gene (VpSBP16) from Chinese wild vitis species in arabidopsis improves salinity and drought stress tolerance. *Int. J. Mol. Sci.* 19 (4), 940.

Jiang, Y., and Deyholos, M. K. (2009). Functional characterization of arabidopsis NaCl-inducible WRKY25 and WRKY33 transcription factors in abiotic stresses. *Plant Mol. Biol.* 69 (1), 91–105.

Jiang, J., Ma, S., Ye, N., Jiang, M., Cao, J., and Zhang, J. (2017b). WRKY transcription factors in plant responses to stresses. *J. Integr. Plant Biol.* 59 (2), 86–101.

Jiang, C., Shen, Q. J., Wang, B., He, B., Xiao, S., Chen, L., et al. (2017a). Transcriptome analysis of WRKY gene family in oryza officinalis wall ex watt and WRKY genes involved in responses to xanthomonas oryzae pv. oryzae stress. *PloS One* 12 (11), e0188742.

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

Lan Thi Hoang, X., Du Nhi, N. H., Binh Anh Thu, N., Phuong Thao, N., and Phan Tran, L.-S. (2017). Transcription factors and their roles in signal transduction in plants under abiotic stresses. *Curr. Genomics* 18 (6), 483–497.

Letunic, I., Doerks, T., and Bork, P. (2012). SMART 7: Recent updates to the protein domain annotation resource. *Nucleic Acids Res.* 40 (D1), D302–D305.

Liang, W., Ma, X., Wan, P., and Liu, L. (2018). Plant salt-tolerance mechanism: A review. *Biochem. Biophys. Res. Commun.* 495 (1), 286–291.

Liang, Q.-y., Wu, Y.-h., Wang, K., Bai, Z.-y., Liu, Q.-l., Pan, Y.-z., et al. (2017). Chrysanthemum WRKY gene DgWRKY5 enhances tolerance to salt stress in transgenic chrysanthemum. *Sci. Rep.* 7 (1), 1–10.

Liao, Y., Smyth, G. K., and Shi, W. (2014). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30 (7), 923–930.

Lynch, M., and Conery, J. S. (2000). The evolutionary fate and consequences of duplicate genes. *science* 290 (5494), 1151–1155.

Maeo, K., Hayashi, S., Kojima-Suzuki, H., MoRIKAMI, A., and NAKAMURA, K. (2001). Role of conserved residues of the WRKY domain in the DNA-binding of tobacco WRKY family proteins. *Bioscience biotechnology Biochem.* 65 (11), 2428–2436.

Mayer, K. F. X., Rogers, J., Dolezel, J., Pozniak, C., Eversole, K., Feuillet, C., et al. (2014). A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. *Science* 345 (6194). doi: 10.1126/science.1251788

McKenna, T. P., Koziol, L., Bever, J. D., Crews, T. E., and Sikes, B. A. (2020). Abiotic and biotic context dependency of perennial crop yield. *PloS One* 15 (6), e0234546.

Meraj, T. A., Fu, J., Raza, M. A., Zhu, C., Shen, Q., Xu, D., et al. (2020). Transcriptional factors regulate plant stress responses through mediating secondary metabolism. *Genes* 11 (4), 346.

Munns, R., and Tester, M. (2008). Mechanisms of salinity tolerance. *Annu. Rev. Plant Biol.* 59, 651–681.

Ning, P., Liu, C., Kang, J., and Lv, J. (2017). Genome-wide analysis of WRKY transcription factors in wheat (Triticum aestivum l.) and differential expression under water deficit condition. *PeerJ* 5, e3232.

Okushima, Y., Fukaki, H., Onoda, M., Theologis, A., and Tasaka, M. (2007). ARF7 and ARF19 regulate lateral root formation *via* direct activation of LBD/ASL genes in arabidopsis. *Plant Cell* 19 (1), 118–130.

Rajappa, S., Krishnamurthy, P., and Kumar, P. P. (2020). Regulation of AtKUP2 expression by bHLH and WRKY transcription factors helps to confer increased salt tolerance to arabidopsis thaliana plants. *Front. Plant Sci.* 111311.

Rezaee, S., Ahmadizadeh, M., and Heidari, P. (2020). Genome-wide characterization, expression profiling, and post-transcriptional study of GASA gene family. *Gene Rep.* 20, 100795.

Song, Y., Li, J., Sui, Y., Han, G., Zhang, Y., Guo, S., et al. (2020). The sweet sorghum SbWRKY50 is negatively involved in salt response by regulating ion homeostasis. *Plant Mol. Biol.* 102 (6), 603–614.

Song, H., Wang, P., Hou, L., Zhao, S., Zhao, C., Xia, H., et al. (2016). Global analysis of WRKY genes and their response to dehydration and salt stress in soybean. *Front. Plant Sci.* 7, 9.

Sun, Y., and Yu, D. (2015). Activated expression of AtWRKY53 negatively regulates drought tolerance by mediating stomatal movement. *Plant Cell Rep.* 34 (8), 1295–1306.

Tang, J., Wang, F., Hou, X.-L., Wang, Z., and Huang, Z.-N. (2014). Genome-wide fractionation and identification of WRKY transcription factors in Chinese cabbage (Brassica rapa ssp. pekinensis) reveals collinearity and their expression patterns under abiotic and biotic stresses. *Plant Mol. Biol. Rep.* 32 (4), 781–795.

Ülker, B., and Somssich, I. E. (2004). WRKY transcription factors: from DNA binding towards biological function. *Curr. Opin. Plant Biol.* 7 (5), 491–498.

Van Zelm, E., Zhang, Y., and Testerink, C. (2020). Salt tolerance mechanisms of plants. *Annu. Rev. Plant Biol.* 71, 403–433.

Wang, C.-T., Ru, J.-N., Liu, Y.-W., Li, M., Zhao, D., Yang, J.-F., et al. (2018). Maize WRKY transcription factor ZmWRKY106 confers drought and heat tolerance in transgenic plants. *Int. J. Mol. Sci.* 19 (10), 3046.

Wang, H., Sun, S., Ge, W., Zhao, L., Hou, B., Wang, K., et al. (2020). Horizontal gene transfer of Fhb7 from fungus underlies fusarium head blight resistance in wheat. *Science* 368 (6493).

Wang, X., Zeng, J., Li, Y., Rong, X., Sun, J., Sun, T., et al. (2015). Expression of TaWRKY44, a wheat WRKY gene, in transgenic tobacco confers multiple abiotic stress tolerances. *Front. Plant Sci.* 6, 615.

Xu, H., Watanabe, K. A., Zhang, L., and Shen, Q. J. (2016). WRKY transcription factor genes in wild rice oryza nivara. *DNA Res.* 23 (4), 311–323.

Yang, P., Chen, C., Wang, Z., Fan, B., and Chen, Z. (1999). A pathogen-and salicylic acid-induced WRKY DNA-binding activity recognizes the elicitor response element of the tobacco class I chitinase gene promoter. *Plant J.* 18 (2), 141–149.

Yang, Y., and Guo, Y. (2018). Elucidating the molecular mechanisms mediating plant salt-stress responses. *New Phytol.* 217 (2), 523–539.

Yan, H., Jia, H., Chen, X., Hao, L., An, H., and Guo, X. (2014). The cotton WRKY transcription factor GhWRKY17 functions in drought and salt stress in transgenic nicotiana benthamiana through ABA signaling and the modulation of reactive oxygen species production. *Plant Cell Physiol.* 55 (12), 2060–2076.

Yu, J., Wang, J., Lin, W., Li, S., Li, H., Zhou, J., et al. (2005). The genomes of oryza sativa: A history of duplications. *PloS Biol.* 3 (2), e38.

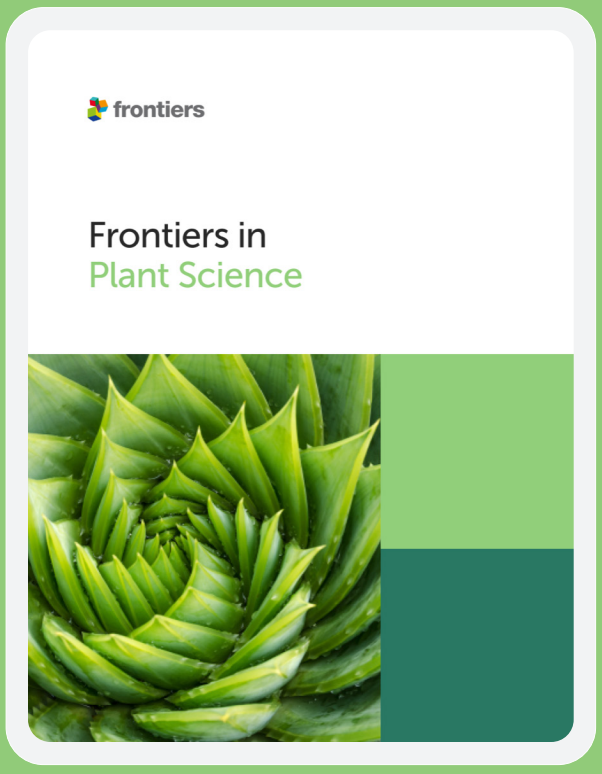
Frontiers in
Plant Science

Cultivates the science of plant biology and its applications

The most cited plant science journal, which advances our understanding of plant biology for sustainable food security, functional ecosystems and human health.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact