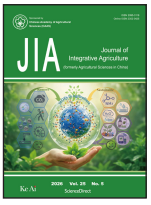
Journal of Integrative Agriculture 2026, 25(5): 1813–1821

Available online at ScienceDirect

Journal of Integrative Agriculture



RESEARCH ARTICLE

Identification and fine mapping of a major QTL for grain protein content, *qGPC4D*, using wheat–*Aegilops tauschii* introgression lines

Yijun Wang*, Jinhao Han*, Tenglong Zhang, Mengjia Sun, Hongyu Ren, Cunyao Bo, Yuqing Diao, Xin Ma, Hongwei Wang, Xiaoqian Wang#

National Key Laboratory of Wheat Improvement/College of Agronomy, Shandong Agricultural University, Tai'an 271018, China

Highlights

- A major and stable quantitative trait locus (QTL) for grain protein content, *qGPC4D*, was identified from *Aegilops tauschii* and fine-mapped to a 9.88 Mb interval.
- High-throughput kompetitive allele-specific PCR (KASP) markers were developed for *qGPC4D*, enabling efficient marker-assisted selection in breeding.

Abstract

Wheat is a staple cereal crop that is crucial for food security and human health. Improving wheat quality has become an essential task for allowing breeders to meet escalating market demand. In this study, a set of wheat–*Aegilops tauschii* introgression lines was developed from a cross between the high-yielding wheat variety Jimai 22 and *Ae. tauschii* Y215. A high-density genetic map containing 2,727 single nucleotide polymorphisms (SNPs) was constructed using a 55K SNP array, and it was used to conduct a quantitative trait loci (QTLs) analysis for grain quality-related traits. Eight QTLs were identified for grain protein content (GPC), starch content, and wet gluten content in two environments. Among them, a major and environmentally stable QTL for GPC, *qGPC4D*, was identified, with favorable alleles contributed by *Ae. tauschii* Y215. Subsequently, *qGPC4D* was narrowed down to a 9.88 Mb physical interval through further fine mapping utilizing the introgression lines. In addition, three linked SNP of *qGPC4D* were converted into high-throughput kompetitive allele-specific PCR (KASP) markers and validated in the introgression population. These findings offer promising candidate genes, elite introgression lines, and KASP markers for high-quality wheat breeding.

Keywords: wheat, *Aegilops tauschii*, protein content, QTL, KASP

1. Introduction

As one of the most important cereals globally, wheat (*Triticum aestivum* L.) plays a vital role in improving world food security by contributing approximately 20% of the dietary calories for human consumption (Brenchley *et al.* 2012). Compared to cereal grains, wheat contributes half of the protein and 40% of the fat originating from all cereal grain sources (Uthayakumaran *et al.* 2010). The improvement of wheat grain quality must be addressed to confront the substantial problems caused by climate change and increasingly diversified market demand (Juliana *et al.* 2019; Li *et al.* 2025). However, wheat diversity is relatively narrow due to genetic bottlenecks caused by domestication and polyploidization (He *et al.* 2019). Wild relatives with considerable genetic variability remain valuable resources for genetically improving agronomic traits (Jiang *et al.* 2023; Song *et al.* 2023).

Aegilops tauschii Coss. (2n=2x=DD), a diploid D genome-progenitor of hexaploid wheat (2n=6x=AABBDD), is an excellent reservoir of genetic diversity that is underexploited (Marcussen *et al.* 2014). Morphologically, *Ae. tauschii* traditionally comprises two subspecies: *Ae. tauschii* ssp. *tauschii* (with moniliform spikes) and *Ae. tauschii* ssp. *strangulate* (with cylindrical spikes). The general consensus is that only a few *Ae. tauschii* ssp. *strangulata* accessions

Received 12 April 2024; Received in revised form 11 June 2024; Accepted 13 June 2024; Available online 19 July 2024
#Correspondence Xiaoqian Wang, E-mail: xqwang1988@126.com
* These authors contributed equally to this study.

© 2026 CAAS. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/). Peer review under responsibility of Editorial Board of *Journal of Integrative Agriculture*.
doi: 10.1016/j.jia.2024.07.029

were involved in the origin of wheat (Yu *et al.* 2014), leading to limited diversity in the wheat D-genome (Zhou *et al.* 2021). The low genetic diversity of the wheat D genome has become a bottleneck for the improvement of bread wheat (Mirzaghaderi *et al.* 2019). Cultivating and utilizing desirable genes in *Ae*. *tauschii* is an effective way to broaden the genetic base of wheat (Rasheed *et al.* 2018). Several genes/QTLs have been identified in *Ae*. *tauschii* and successfully applied in wheat breeding. A leaf rust resistance gene, *Lr21*, and a stem rust resistance gene, *Sr33*, both introgressed from *Ae*. *tauschii* into bread wheat, were cloned by Huang *et al.* (2003) and Sambasivam *et al.* (2013), respectively. The stripe rust resistance gene derived from *Ae*. *tauschii*, *Yr28*, is located on chromosome 4DS (Singh *et al.* 2000). The powdery mildew resistance genes derived from *Ae*. *tauschii*, *Pm2*, *Pm19*, *Pm34*, *Pm35*, and *Pm58*, were mapped to chromosomes 5DS, 7D, 5DL, 5DL, and 2DS, respectively (Lutz *et al.* 1995; Miranda *et al.* 2006; Miranda *et al.* 2007; Wiersma *et al.* 2017). Leaf rust resistance genes, *Lr22*, *Lr32*, *Lr39*, and *Lr42* are located on chromosomes 2DS, 3D, 2DS, and 1DS, respectively, and they have been transferred to bread wheat from *Ae*. *tauschii* (Rowland *et al.* 1974; Kerber 1987; Raupp *et al.* 2001; Gill *et al.* 2019). Using F_2 and $F_{2:3}$ populations derived from a common allohexaploid wheat line TAA10 and a synthetic allohexaploid wheat line XX329, three environmentally stable QTL regions associated with grain size and shape were mapped on chromosomes 2DS, 2DL and 7DS, with favorable alleles contributed by *Ae*. *tauschii* (Yan *et al.* 2017). Using recombinant inbred lines (RILs) derived from a cross between synthetic hexaploid wheat SHW-L1 and hexaploid wheat Chuanmai 32, four environment-independent QTLs for agronomic traits were identified on chromosome 2D (Yu *et al.* 2014). Synthetic allohexaploid wheat Chuanmai 42, derived from Wheat and Maize Improvement Center (CIMMYT) synthetic wheat (pedigree: Syn769/SW3243//Chuan6415), shows excellent yield potential, strong stripe rust resistance, wide adaptation and good quality. The high-yielding gene in Chuanmai 42 was mapped to 4DL using two RILs (Li *et al.* 2011). Therefore, *Ae*. *tauschii* is considered to have a promising gene pool for the improvement of common wheat.

Two main approaches have been proposed to introduce *Ae*. *tauschii* germplasm into the hexaploid gene pool (Cox *et al.* 2017): "synthetic hexaploid" and "direct hybridization". For the "synthetic hexaploid" approach, *Ae*. *tauschii* is crossed with a tetraploid wheat (*T. turgidum*) to produce a triploid ABD hybrid and the chromosome number is subsequently doubled to produce hexaploid seeds. For the "direct hybridization" approach, *T*. *aestivum* plants are pollinated by *Ae*. *tauschii* and the resulting embryos are rescued after pollination for 12□15 days. A subsequent backcross with *T*. *aestivum* is conducted to obtain stable 42-chromosome AABBDD progeny. Synthetic hexaploid has been widely used as a bridge to transfer genes involved in biotic resistance (Ghaffary *et al.* 2012), grain yield (Lutz *et al.* 1995; Hao *et al.* 2019), preharvest sprouting (Imtiaz *et al.* 2008), grain quality (Kunert *et al.* 2007) and abiotic tolerance (Ogbonnaya *et al.* 2013). As an improvement over the "synthetic hexaploid" approach, the "direct hybridization" approach can introgress desired D genome regions accompanied by complete conservation of the recurrent parent's A and B genomes (Cox *et al.* 2017; Gaurav *et al.* 2021). The "direct hybridization" approach has been widely utilized to improve disease and pest resistance, e.g., Hessian fly resistance (Gill *et al.* 1991), greenbug resistance (Gill *et al.* 1991), soilborne mosaic virus resistance (Gill *et al.* 1991; Cox *et al.* 1994), wheat spindle streak mosaic virus resistance (Cox *et al.* 1994), leaf rust resistance, stem rust resistance (Olson *et al.* 2013), and powdery mildew resistance (Cox *et al.* 1994). However, few studies have reported improving wheat quality by the "direct hybridization" approach.

In the present study, a set of wheat–*Ae*. *tauschii* introgression lines (ILs) was developed from a cross between the common wheat Jimai 22 and *Ae*. *tauschii* Y215. These ILs were genotyped using a 55K SNP array and phenotyped for grain quality-related traits in two environments. The objectives were to: 1) detect favorable QTLs/genes from *Ae*. *tauschii*, 2) identify candidate genes for major and stable QTLs through fine mapping and expression analysis, and 3) develop kompetitive allele-specific PCR (KASP) markers for marker-assisted selection to improve wheat quality.

2. Materials and methods

2.1. Plant materials

The diploid *Ae*. *tauschii* ssp. *tauschii* accession Y215 was crossed with the high-yielding winter wheat cultivar Jimai 22. The resulting hybrid F_1 embryos were rescued after pollination for 12–15 days. Subsequently, three sequential backcrosses with the recurrent parent Jimai 22 were conducted to obtain the BC_3F_1 population. Several self-pollination cycles were used to construct a BC_3F_5 introgression population, designated as BC3-ILs, with a population size of 170 individuals.

2.2. Field trials

The ILs and *Ae*. *tauschii* accessions were planted in two locations in the Shandong Province, China, namely, Dezhou (DZ, 37.38°N, 116.39°E) and Tai'an (TA, 36.17°N, 117.17°E), from 2020 to 2021. For the ILs, each accession was planted in a two-row plot with 50 seeds per row at a row spacing of 25 cm and a row length of 2.0 m with two replicates for each accession in both environments. Standard practices of the local farmers were followed for field management.

2.3. Trait measurements

At the maturity stage, six middle uniform plants for each accession were bulk harvested and air-dried in drying houses for three months. Subsequently, the grain protein content (GPC), starch content (SC), and wet gluten content (WGC) of the grain seeds were measured using near-infrared spectroscopy DA7200 (Perteninstruments 7200 type, Huddinge, Sweden). Three replicates were performed for each line, and the average value was used for subsequent analyses.

2.4. Genotyping

The BC3-ILs and their parents, Jimai 22 and Y215, were

genotyped using a Wheat55K SNP array designed by the Chinese Academy of Agricultural Sciences. Heterozygous loci were treated as missing data and SNPs with >10% missing data were removed. The SNP genotype data for *Ae. tauschii* accessions were generated by resequencing at a sequencing depth of >10×. Polymorphic SNPs within specific QTL intervals were used for association mapping.

2.5. Statistical analysis and trait correlations

The best linear unbiased estimate (BLUE) values of the two replicates in each environment were calculated using a mixed linear model with genotypes as fixed factors and replicates as random factors. The distribution of phenotypes was plotted using the ggplot2 package in R software. Phenotypic correlations were calculated using the "rcorr" function implemented in the R package Hmisc (Harrell 2019).

2.6. QTL analysis and fine mapping of major QTLs

The following QTL mapping procedures were administered. First, the loci in the population with the same genotype as Jimai 22 were designated as A, and the loci with the same genotype as Y215 were designated as B. Then, consecutive markers with identical genotypes in the population were combined into a single bin marker. Next, the "aov" function in R software was used to analyze the correlations between individual bin markers and traits, with a significance threshold of $-\log_{10}(P)\geq 5.0$. The additive effect of a bin marker was (B–A)/2. The physical interval determined by a two-unit reduction in the *P*-value of the peak marker was designated as the confidence interval for the QTL.

The 196 ILs were used to conduct the fine mapping of the major QTLs. Markers located within the QTL interval were selected for the fine mapping of the major QTLs. The ILs were categorized based on the genotypes present within the specified interval. The group that had no introgressed segment within the specified interval was used to conduct *t*-tests with each of the other groups to fine-map each QTL.

2.7. Expression analysis of putative candidate genes

Transcriptomic data of five tissues from different growth stages of the bread wheat variety Chinese Spring were obtained from the expVIP platform7 (http://www.wheat-expression.com), and used to identify the differential expression characteristics of putative candidate genes located within each target QTL (Borrill *et al.* 2016). Transcripts per million (TPM) was used to evaluate the expression levels of putative candidate genes.

2.8. Development of KASP markers

The SNPs derived from the resequencing of *Ae. tauschii* Y215 were compared with those obtained from the resequencing of Jimai 22, as reported by Yang *et al.* (2020). Polymorphic SNPs located within the *qGPC4D* interval were selected to develop the KASP markers. The primers were synthesized by LGC, Biosearch Technologies (Hoddesdon, UK). KASP genotyping was performed in the molecular breeding laboratory of the Beijing Vegetable Research Center using the LGC SNP-line platform (LGC, https://www.lgcgroup.com/), following the procedures described by Zhang *et al.* (2023).

3. Results

3.1. Phenotypic variations and trait correlations

The three grain quality-related traits exhibited normal distributions and showed large consecutive variations in both DZ and TA. The ILs exhibited higher values for GPC and WGC, but lower SC in TA than in DZ (Fig. 1). Pairwise phenotypic correlations among the three measured traits are shown in Fig. 2, and all the traits exhibited strong correlations ($P<0.01$). GPC was strongly positively correlated with WGC but negatively correlated with SC. Strong positive correlations were detected between the two environments for each trait.

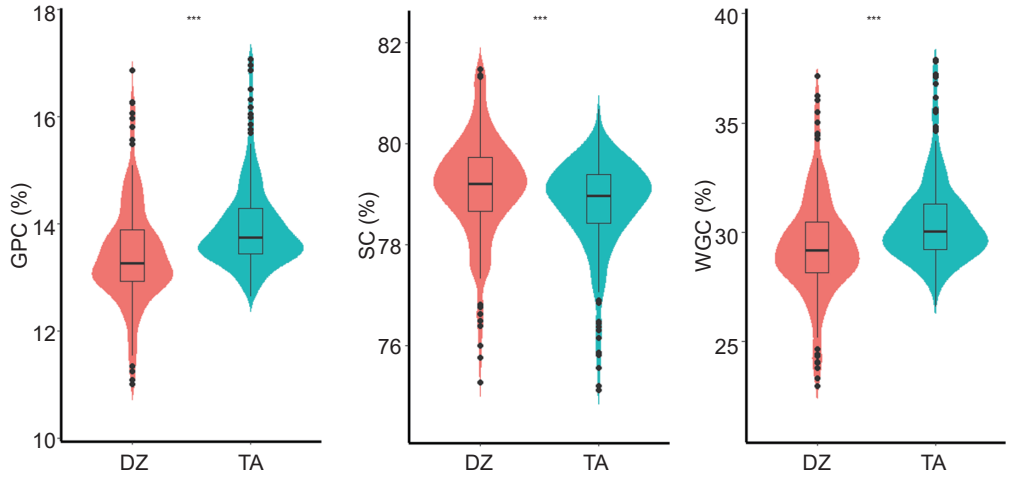


Fig. 1 Phenotype distribution of the introgression lines (ILs) population in Dezhou (DZ) and Tai'an (TA), Shandong Province, China. GPC, grain protein content; SC, starch content; WGC, wet gluten content. *** means the phenotypic difference between DZ and TA was significant at 0.001 level.

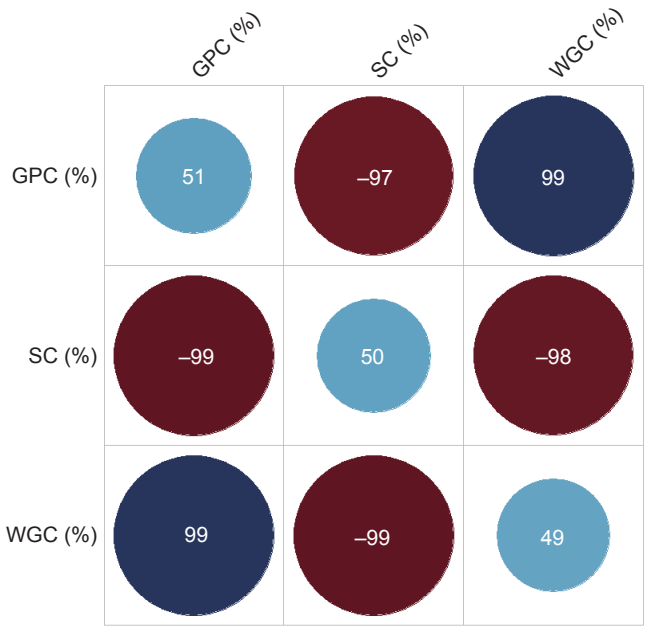


Fig. 2 Phenotypic correlations between the three evaluated traits in Dezhou (DZ) (lower triangle) and Tai'an (TA) (upper triangle), Shandong Province, China. The values are correlation coefficients (*r*) multiplied by 100. The areas and colors of the circles show the absolute value of the corresponding *r*. Blue and red colors indicate positive and negative correlations, respectively. The values without glyphs indicate insignificance at 0.01. GPC, grain protein content; SC, starch content; WGC, wet gluten content.

3.2. Basic statistics for the markers

In this study, 2,727 high-quality polymorphic SNPs were identified. The number of markers per chromosome ranged from 198 on chromosome 4D to 568 on chromosome 2D (Table 1; Fig. 3). The average coverage rate across all chromosomes was 94.94%, with the lowest coverage of 95.25% for chromosome 6D and the highest coverage of 99.19% for chromosome 3D. The average marker density was 1.56 Mb, varying between 0.99 Mb for chromosome 1D to 2.55 Mb for chromosome 4D. Overall, the 2,727 markers across the genome were categorized into 484 bins, with an average bin size of 9.16 Mb. The average introgression frequency per chromosome was 11.1%, with the highest introgression frequency of 18.3% for chromosome 1D and the lowest of 4.6% for chromosome 3D (Table 1). The size of the introgressed genome of each line is shown in Fig. 4. The average size of the introgressed segments across all ILs was 274.8 Mb. In 33% of individuals, the size of the introgressed segments was less than 50 Mb (Fig. 4).

3.3. QTL mapping

In total, eight QTLs associated with grain quality-related traits were detected, with three QTLs identified in both the TA and DZ environments (Table 2; Fig. 5). For GPC, two QTLs were identified on chromosomes 4D and 6D. Of them, *qGPC4D* was detected in both environments, explaining 15.2 and 13.1% of the phenotypic variance in DZ and TA, respectively. The *qGPC6D* QTL was detected only in DZ and accounted for 11.3% of the phenotypic variance. For SC, three QTLs were identified on chromosomes 2D, 4D and 6D. Of them, *qSC4D* was identified in both environments, explaining 14.6 and 13.3% of the phenotypic variance in DZ and TA, respectively. The *qSC2D* and *qSC6D* QTLs were detected in DZ, explaining 12.8 and 16.2% of the phenotypic variance, respectively. For WGC, *qWGC4D* was detected in both environments and explained 11.2 and 11.6% of the phenotypic variance in DZ and TA, respectively. The *qWGC2D* and *qWGC6D* QTLs were detected in DZ and explained 11.0 and 10.9% of the phenotypic variance, respectively.

Three QTL clusters associated with grain quality-related traits were identified. In Cluster 1, which comprised *qGPC4D*, *qSC4D*, and *qWGC4D*, the alleles that increased GPC and WGC while reducing SC were contributed by *Ae. tauschii*. Cluster 2, comprising *qGPC6D*, *qSC6D*, and *qWGC6D*, was detected on chromosome 6D. The alleles responsible for increasing GPC and WGC and reducing SC originated from *Ae. tauschii*. In Cluster 3, which included *qSC2D* and *qWGC2D*, the alleles that enhanced SC and reduced WGC were derived from *Ae. tauschii*.

3.4. Fine mapping of *qGPC4D*

The environmentally stable and major QTL *qGPC4D* was further fine mapped using ILs and five polymorphic markers within the target region (Fig. 6). Genotype 1 exhibited significant differences from genotypes 2, 3, and 5, but no significant difference was observed from genotype 4. Consequently, the interval was narrowed down to the region between 7.20 Mb (AX-110005953) and 17.08 Mb (AX-89398511), corresponding to a 9.88 Mb physical interval on chromosome 4D of Chinese Spring. Within the corresponding physical interval of 7.20 to 17.08 Mb on

Table 1 Distribution of 2,727 single nucleotide polymorphism (SNP) markers on the seven chromosomes

Chromosome	Marker number	Genome size (Mb)	Start (Mb)	End (Mb)	Coverage (%)	Marker density (Mb)	Number of bins	Size of bins (Mb)	Introgression rate (%)
1D	498	498.6	0.08	492.89	98.84	0.99	87	5.66	18.3
2D	568	656.5	4.50	651.04	98.48	1.14	87	7.43	9.9
3D	383	619.6	0.87	615.48	99.19	1.60	33	18.62	4.6
4D	198	518.3	3.30	508.50	97.47	2.55	48	10.53	9.0
5D	353	570.0	3.51	564.75	98.46	1.59	85	6.60	10.9
6D	332	495.4	0.90	472.77	95.25	1.42	70	6.74	15.7
7D	395	642.9	3.79	633.25	97.91	1.59	74	8.51	9.3
Total	2,727	4,001.4			97.94	1.56	484	9.16	11.1

the Chinese Spring genome assembly in WheatOmics 1.0 (https://202.194.139.32), a total of 236 annotated genes were identified.

3.5. Development of KASP markers for *qGPC4D*

Based on the resequencing data of common wheat Jimai 22 and *Ae. tauschii*, three polymorphic and locus-specific markers within the fine mapping interval of *qGPC4D* were selected and successfully converted into KASP markers (Fig. 7; Appendix A).

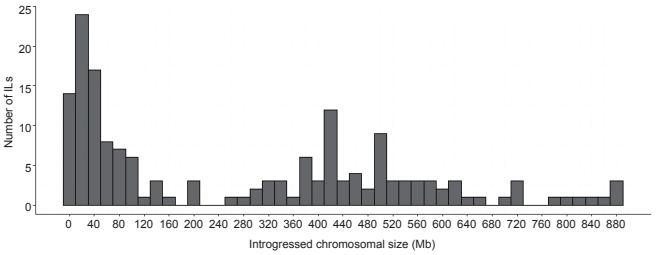


Fig. 4 Size of the donor-derived introgressed segment in the introgression lines (ILs).

4. Discussion

4.1. Trait performance and correlations

Proteins are the primary factors responsible for the typical viscoelastic behavior of wheat flour dough (Shevkani *et al.* 2024). As the most abundant storage carbohydrate in the endosperm of wheat grains, starch plays a vital role in determining grain texture (Shang *et al.* 2021). Both protein and starch play crucial roles in wheat processing and the determination of end-use quality (Muqaddasi *et al.* 2020; Guo *et al.* 2023). However, GPC was strongly negatively correlated with SC, which was confirmed by the contrasting additive effects of the QTLs for GPC and SC. Previous studies have demonstrated that SC is positively correlated with grain yield but negatively correlated with GPC, indicating that achieving a balance is crucial for obtaining the desired protein level and yield (Guo *et al.* 2023). There were strong correlations between each pair of traits in the two environments. However, a notable difference was observed between the GPC in the DZ and TA locations, indicating that GPC exhibited environmental sensitivity (Muqaddasi *et al.* 2020; Liu *et al.* 2024).

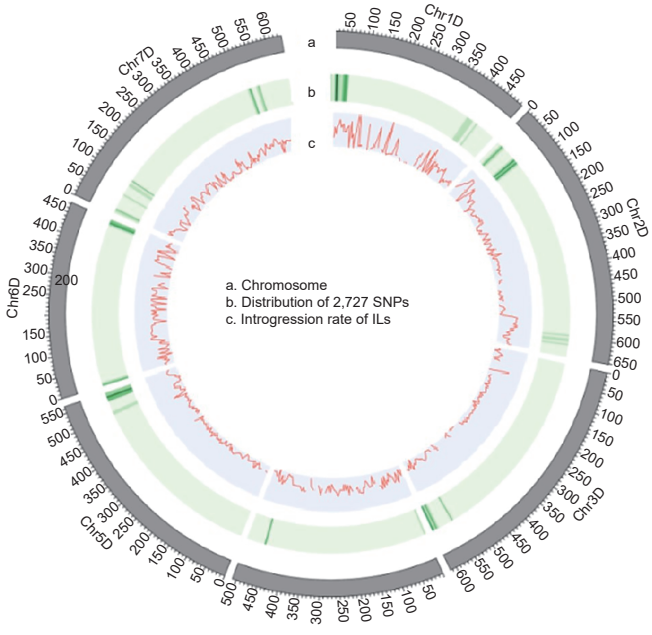


Fig. 3 Frequency of the donor allele along individual chromosomes in the introgression lines (ILs).

4.2. Comparison with previous studies

Numerous studies have used traditional biparental linkage mapping and genome-wide association studies (GWAS) to investigate grain quality-related traits (Zheng *et al.* 2013; Jiang *et al.* 2021; Zhang G *et al.* 2021; Pu *et al.* 2022; Li *et al.* 2023; Hao *et al.* 2024). Jiang *et al.* (2021) performed association mapping for GPC in a natural population comprising 103 local cultivated varieties, and identified a QTL controlling GPC on chromosome 2D. Su *et al.* (2020) found a QTL for GPC on chromosome 2D using an advanced BC_2F_6 population derived from a cross of hexaploid wheat Zhoumai 18 and *Ae. tauschii* T093. Mahjourimajd *et al.* (2016) found a QTL for GPC on chromosome 2D using a sub-population of

Table 2 QTL mapping results for wheat grain quality-related traits in Dezhou (DZ) and Tai'an (TA), Shandong Province, China

Trait[1)]	QTL	Env	Chr	Peak marker	Physical position (bp)	$-\log_{10}(P)$	Left physical position	Right physical position	Effect	PVE (%)[2)]
GPC	*qGPC4D*	DZ	4D	AX-89398511	17,083,872	5.65	11,870,112	31,398,343	0.54	15.2
GPC	*qGPC4D*	TA	4D	AX-89398511	17,083,872	5.01	11,870,112	31,398,343	0.44	13.1
GPC	*qGPC6D*	DZ	6D	AX-110416062	306,323,757	4.30	33,143,556	389,517,107	0.27	11.3
SC	*qSC2D*	DZ	2D	AX-89728114	26,344,368	4.78	10,283,122	30,493,658	0.38	12.8
SC	*qSC4D*	DZ	4D	AX-89398511	17,083,872	5.43	11,870,112	31,398,343	–0.57	14.6
SC	*qSC4D*	TA	4D	AX-89398511	17,083,872	5.08	11,870,112	31,398,343	–0.53	13.3
SC	*qSC6D*	DZ	6D	AX-110416062	306,323,757	6.07	291,244,822	389,517,107	–0.35	16.2
WGC	*qWGC2D*	DZ	2D	AX-89728114	26,344,368	4.15	10,283,122	30,493,658	–0.79	11.0
WGC	*qWGC4D*	DZ	4D	AX-89398511	17,083,872	4.23	11,870,112	55,045,649	1.13	11.2
WGC	*qWGC4D*	TA	4D	AX-89398511	17,083,872	4.47	11,870,112	55,045,649	1.04	11.6
WGC	*qWGC6D*	DZ	6D	AX-110416062	306,323,757	4.17	33,143,556	389,517,107	0.64	10.9

[1)] GPC, grain protein content; SC, starch content; WGC, wet gluten content.
[2)] PVE, phenotypic variance explained.

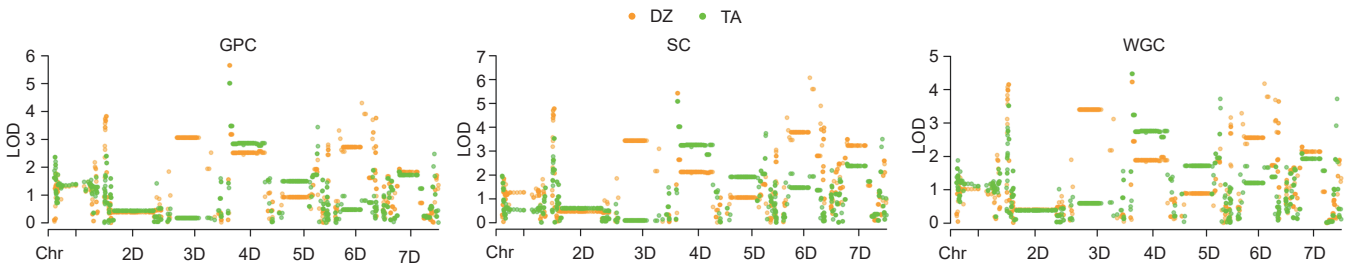
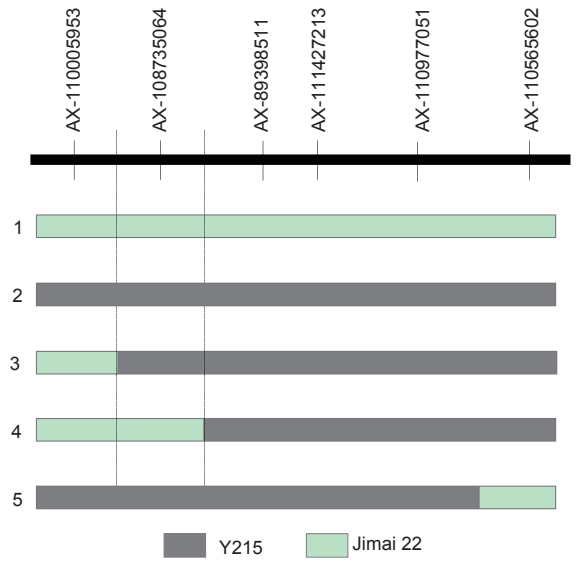


Fig. 5 QTL mapping results for the introgression lines (ILs) in Dezhou (DZ) and Tai'an (TA), Shandong Province, China. GPC, grain protein content; SC, starch content; WGC, wet gluten content.



Genotype types	No. of lines	Environments	GPC (%)	*t*-test	Significance level
1	143	20TA	13.68±0.55		
		20DZ	13.22±0.80		
2	8	20TA	17.02+0.08	1.46E–14	***
		20DZ	16.47±0.57	8.50E–08	***
3	7	20TA	15.39±0.74	2.68E–10	***
		20DZ	15.89±0.36	7.10E–06	***
4	5	20TA	14.30±0.27	0.112	
		20DZ	14.48±0.12	0.030	
5	7	20TA	16.25±0.69	9.10E–43	***
		20DZ	14.88±0.69	2.45E–06	***

Fig. 6 Fine mapping of *qGPC4D*. TA, Tai'an; DZ, Dezhou, Shandong Province, China. ***, $P<0.001$.

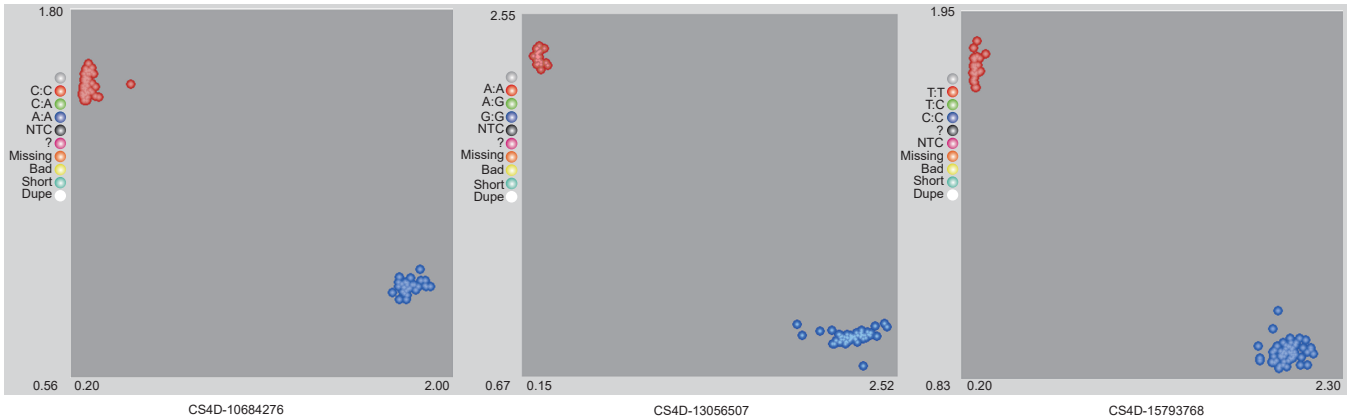


Fig. 7 Genotyping results for the introgression lines (ILs) using three developed kompetitive allele-specific PCR (KASP) markers.

doubled-haploid (DH) lines derived from a cross between two Australian varieties. These QTLs were located in the same regions as *qWGC2D* and *qSC2D* in this study. Kumar *et al.* (2019) identified a QTL for GPC on chromosome 6D using a RIL population developed from a cross of an elite HRSW and a non-adapted wheat genotype. Zhang G *et al.* (2021) found a QTL that controls GPC, *QFPC.ksu-6DL*, using a DH population derived from a cross between Tiger and 'Danby'. These QTLs on chromosome 6D were co-located with *qGPC6D*, *qSC6D*, and *qWGC6D* in this study. To the best of our knowledge, no QTLs for GPC have been detected in the same region as *qGPC4D*, indicating that *qGPC4D* is a novel

locus controlling GPC with the favorable alleles from *Ae. tauschii*.

4.3. Candidate gene analysis for *qGPC4D*

Transcriptome data downloaded from expVIP were used to identify more credible candidate genes. Genes with TPM>5 in the grain development stages, including 14 or 30 days after anthesis (DAA), were identified as the most likely candidate genes (Appendices B and C). Ultimately, 11 potential candidate genes were identified, and they could be categorized into two distinct expression patterns. In Class I, the genes were highly expressed at both the 14 and 30 DAA stages. Candidate genes in Class II exhibited high expression, specifically at 30 DAA.

Globulin, a salt-soluble protein, is an important non-gluten protein in wheat (Zhu *et al.* 2022) that accounts for 6–10% of the embryo protein (Zhang *et al.* 2022). Previous studies have shown that globulin enhances grain quality by forming a more rigid protein network structure (Zhang L *et al.* 2021; Marcone *et al.* 1998). In maize, the major globulin component GLB I is encoded by a single gene *Globulin-1* (*Glbl*) and the three most common *Glbl* alleles have been designated L, I, and S for large, intermediate, and small proteins, respectively (Belanger *et al.* 1991). Loit *et al.* (2009) identified three new globulin gene sequences, *Glo-3A*, *Glo-3B* and *Glo-3C*, from hexaploid wheat by screening a *T. aestivum* BAC genomic library with cDNA from a diabetes-associated globulin, suggesting the existence of a diverse group of globulins in wheat. Wang *et al.* (2017) purified proteins encoded by the globulin-1 S allele. In the current study, *TraesCS4D02G016100*, which encodes globulin 1, was highly expressed at both the 14 and 30 DAA stages, with TMP values of 8.8 and 12.5, respectively. Zhao *et al.* (2022) cloned and characterized a novel C2H2 zinc finger protein that regulates amylose content and amylopectin chain-length distribution in rice, thereby affecting rice eating and cooking quality. The candidate gene *TraesCS4D02G038700* identified in this study encodes a zinc finger family protein that might affect GPC by regulating the SC in wheat grains. Wang *et al.* (2020) identified a glycine- and proline-rich protein, *OsGPRP3*, which is highly expressed in seeds at 10 days after pollination and regulates grain size and GPC in rice. The candidate gene *TraesCS4D02G027100* identified in this study encodes a glycine-rich family protein. Lv *et al.* (2021) found that transgenic wheat grains with an RNAi-mediated knockdown of lipoxygenase-encoding genes exhibited substantial reductions in starch content and dough stability. In the current study, *TraesCS4D02G035200* was found to encode a lipoxygenase protein. Functional verification of these candidate genes for *qGPC4D* through CRISPR/Cas9 editing and overexpression may further elucidate their effects on wheat grain quality.

4.4. Implications for wheat breeding

The mining and introgression of favorable genes from *Ae. tauschii* are conductive to wheat improvement in breeding programs (Ma *et al.* 2023). Previous studies have demonstrated successful improvements in wheat accessions for curl mite resistance, PHS resistance, grain size, and grain weight using a direct hybridization approach (Rudd *et al.* 2014; Gaurav *et al.* 2021; Zhou *et al.* 2021; Ma *et al.* 2023). In the current study, a set of BC3-ILs was constructed, with an average introgression frequency of 11.1% per chromosome. Half of the ILs had introgressed segments smaller than 100 Mb (Fig. 3), indicating that the quality of the backcrossed parent could be improved despite a relatively low genetic load. Consequently, the efficiency of genetic breeding could be substantially enhanced. Furthermore, QTL mapping and subsequent fine mapping of the major QTLs using ILs are effective methods for identifying candidate genes (Cox *et al.* 2017). In the current study, the major and stable QTL *qGPC4D* was fine mapped to a 9.88 Mb physical interval on chromosome 4D of Chinese Spring, and it is a novel QTL. By combining the transcriptome information of the genes, the 11 most likely candidate genes with TPM>5 were screened, facilitating subsequent fine mapping and their application in breeding. Therefore, the direct hybridization approach should be adopted to harness the beneficial genes/QTLs from wild relatives of wheat (Zhou *et al.* 2021; Bo *et al.* 2022; Ma *et al.* 2023). In addition, the KASP assay is a promising technology for high-throughput SNP genotyping, which provides an excellent opportunity for the effective implementation of marker-assisted breeding (Saini *et al.* 2020). In the present study, we developed three KASP markers specific to the *qGPC4D* interval, which will enhance its utilization in high-quality wheat breeding. Genetic bottlenecks resulting from domestication and polyploidization have influenced wheat diversity (He *et al.* 2019). Introgressing favorable genes/QTLs from *Ae. tauschii* into the D genome of common wheat is an effective way to increase the genetic diversity of common wheat (Zhou *et al.* 2021). ILs carrying the favorable alleles of *qGPC4D* from *Ae. tauschii* can be utilized in high-quality wheat breeding.

5. Conclusion

A set of BC3-ILs was developed by crossing the high-yielding winter variety Jimai 22 and *Ae. tauschii* accession Y215. A high-density genetic linkage map containing 2,727 SNPs was successfully constructed. QTL mapping identified eight QTLs associated with grain quality-related traits. Of them, the major and stable QTL *qGPC4D* carried favorable alleles contributed by *Ae. tauschii* accession Y215. Further fine mapping and candidate gene analysis were conducted to identify the candidate genes. High-throughput KASP markers for *qGPC4D* were successfully developed and validated.

Acknowledgements

This work was supported by the Shandong Province Agricultural Fine Seeds Project, China (2022LZGC005), the National Natural Science Foundation of China (31901491 and 32030081), and the Young and Innovative Team Plan of Shandong Higher Education Institutions, China (2023KJ336).

Declaration of competing interest

The authors declare that they have no conflict of interest.

Declaration of generative AI and AI-assisted technologies in the writing process

The authors declare that they did not use AI in the preparation and writing of this manuscript.

Appendices associated with this paper are available at https://doi.org/10.1016/j.jia.2024.07.029

References

Belanger F C, Kriz A L. 1991. Molecular basis for allelic polymorphism of the maize *Globulin-1* gene. *Genetics*, **129**, 863–872.

Bo C, Fan Z, Ma X, Li A, Wang H, Kong L, Wang X. 2022. Identification and introgression of a novel HMW-GS gene from *Aegilops tauschii*. *Agronomy*, **12**, 2709.

Borrill P, Ramirez-Gonzalez R, Uauy C. 2016. expVIP: A customizable RNA-seq data analysis and visualization platform *Plant Physiology*, **170**, 2172–2186.

Brenchley R, Spannagl M, Pfeifer M, Barker G L, D'amore R, Allen A M, Mckenzie N, Kramer M, Kerhornou A, Bolser D. 2012. Analysis of the bread wheat genome using whole-genome shotgun sequencing. *Nature*, **491**, 705.

Cox T S, Sorrells M, Bergstrom G, Sears R, Gill B, Walsh E, Leath S, Murphy J. 1994. Registration of KS92WGRC21 and KS92WGRC22 hard red winter wheat germplasms resistant to wheat spindle-streak mosaic virus, wheat soilborne mosaic virus, and powdery mildew. *Crop Science*, **34**, 546.

Cox T S, Wu J, Wang S, Cai J, Zhong Q, Fu B. 2017. Comparing two approaches for introgression of germplasm from *Aegilops tauschii* into common wheat. *The Crop Journal*, **5**, 355–362.

Gaurav K, Arora S, Silva P, Sánchez-Martín J, Horsnell R, Gao L, Brar G S, Widrig V, John Raupp W, Singh N, Wu S, Kale S M, Chinoy C, Nicholson P, Quiroz-Chávez J, Simmonds J, Hayta S, Smedley M A, Harwood W, Pearce S, *et al.* 2021. Population genomic analysis of *Aegilops tauschii* identifies targets for bread wheat improvement. *Nature Biotechnology*, **40**, 422–431.

Ghaffary S M T, Faris J D, Friesen T L, Visser R G, Van Der Lee T A, Robert O, Kema G H. 2012. New broad-spectrum resistance to septoria tritici blotch derived from synthetic hexaploid wheat. *Theoretical and Applied Genetics*, **124**, 125–142.

Gill B, Wilson D, Raupp W, Hatchett J, Harvey T, Cox T, Sears R. 1991. Registration of KS89WGRC4 hard red winter wheat germplasm with resistance to Hessian fly, greenbug, and soil-borne mosaic virus. *Crop Science*, **31**, 246.

Gill H S, Li C, Sidhu J S, Liu W, Sehgal S K. 2019. Fine mapping of the wheat leaf rust resistance gene *Lr42*. *International Journal of Molecular Sciences*, **20**, doi: 10.3390/ijms20102445.

Guo Y, Wang G, Guo X, Chi S, Yu H, Jin K, Huang H, Wang D, Wu C, Tian J, Chen J, Bao Y, Zhang W, Deng Z. 2023. Genetic dissection of protein and starch during wheat grain development using QTL mapping and GWAS. *Frontiers in Plant Science*, **14**, 1189887.

Hao M, Zhang L Q, Zhao L B, Dai S F, Li A L, Yang W Y, Xie D E, Li Q C, Ning S Z, Yan Z H. 2019. A breeding strategy targeting the secondary gene pool of bread wheat: Introgression from a synthetic hexaploid wheat. *Theoretical and Applied Genetics*, **132**, 2285–2294.

Hao Y C, Kong F M, Wang L L, Zhao Y, Li M Y, Che N X, Li S, Wang M, Hao M, Zhang X, Zhao Y. 2024. Genome-wide association study of grain micronutrient concentrations in bread wheat. *Journal of Integrative Agriculture*, **23**, 1468–1480.

Harrell F E. 2019. Hmisc: harrell miscellaneous, R package version 4.2-0. https://CRAN.R-project.org/package=Hmisc

He F, Pasam R, Shi F, Kant S, Keeble-Gagnere G, Kay P, Forrest K, Fritz A, Hucl P, Wiebe K, Knox R, Cuthbert R, Pozniak C, Akhunova A, Morrell P L, Davies J P, Webb S R, Spangenberg G, Hayes B, Daetwyler H, *et al.* 2019. Exome sequencing highlights the role of wild-relative introgression in shaping the adaptive landscape of the wheat genome. *Nature Genetics*, **51**, 896–904.

Huang L, Brooks S A, Li W, Fellers J P, Trick H N, Gill B S. 2003. Map-based cloning of leaf rust resistance gene *Lr21* from the large and polyploid genome of bread wheat. *Genetics*, **164**, 655–664.

Imtiaz M, Ogbonnaya F C, Oman J, van Ginkel M. 2008. Characterization of quantitative trait loci controlling genetic variation for preharvest sprouting in synthetic backcross-derived wheat lines. *Genetics*, **178**, 1725–1736.

Jiang P, Zhang P, Wu L, He Y, Li C, Ma H, Zhang X. 2021. Linkage and association mapping and kompetitive allele-specific PCR marker development for improving grain protein content in wheat. *Theoretical and Applied Genetics*, **134**, 3563–3575.

Jiang Y, Wang D L, Hao M, Zhang J, Liu D C. 2023. Development and characterization of wheat–*Aegilops kotschyi* $1U^k$(1A) substitution line with positive dough quality parameters. *Journal of Integrative Agriculture*, **22**, 999–1008.

Juliana P, Poland J, Huerta-Espino J, Shrestha S, Crossa J, Crespo-Herrera L, Toledo F H, Govindan V, Mondal S, Kumar U, Bhavani S, Singh P K, Randhawa M S, He X, Guzman C, Dreisigacker S, Rouse M N, Jin Y, Pérez-Rodríguez P, Montesinos-López O A, *et al.* 2019. Improving grain yield, stress resilience and quality of bread wheat using large-scale genomics. *Nature Genetics*, **51**, 1530–1539.

Kerber E R. 1987. Resistance to leaf rust in hexaploid wheat: *Lr32* a third gene derived from *Triticum tauschii*. *Crop Science*, **27**, 204–206.

Kumar A, Mantovani E E, Simsek S, Jain S, Elias E M, Mergoum M. 2019. Genome wide genetic dissection of wheat quality and yield related traits and their relationship with grain shape and size traits in an elite×non-adapted bread wheat cross. *PLoS ONE*, **14**, e0221826.

Kunert A, Naz A A, Dedeck O, Pillen K, Léon J. 2007. AB-QTL analysis in winter wheat: I. Synthetic hexaploid wheat (*T. turgidum* ssp. *dicoccoides*×*T. tauschii*) as a source of favourable alleles for milling and baking quality traits. *Theoretical and Applied Genetics*, **115**, 683–695.

Li J, Wei H, Hu X, Li C, Tang Y, Liu D, Yang W. 2011. Identification of a high-yield introgression locus in Chuanmai 42 inherited from synthetic hexaploid wheat. *Acta Agronomica Sinica*, **37**, 255–261. (in Chinese)

Li L L, Mao Z Q, Wang P, Cai J, Zhou Q, Zhong Y X, Jiang D, Wang X. 2025. Drought priming enhances wheat grain starch and protein quality under drought stress during grain filling. *Journal of Integrative Agriculture*, **24**, 2888–2901.

Li N, Miao Y, Ma J, Zhang P, Chen T, Liu Y, Che Z, Shahinnia F, Yang D. 2023. Consensus genomic regions for grain quality traits in wheat revealed by Meta-QTL analysis and in silico transcriptome integration. *The Plant Genome*, **16**, e20336.

Liu Z H, Lai X J, Chen Y J, Zhao P, Wang X M, Ji W Q, Xu S B. 2024. Selection and application of four QTLs of grain protein content in modern wheat cultivars. *Journal of Integrative Agriculture*, **23**, 2557–2570.

Loit E, Melnyk C W, Macfarlane A J, Scott F W, Altosaar I. 2009. Identification of three wheat globulin genes by screening a *Triticum aestivum* BAC genomic library with cDNA from a diabetes-associated globulin. *BMC Plant Biology*, **9**, 93.

Lutz J, Hsam S L K, Limpert E, 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.

Lv G, Tian Q, Zhang F, Chen J, Niaz M, Liu C, Hu H, Sun C, Chen F. 2021. Reduced expression of lipoxygenase genes improves flour processing quality in soft wheat. *Journal of Experimental Botany*, **72**, 6247–6259.

Ma F F, Li R Z, Guo G H, Nie F, Zhu L L, Liu W J, Lyu L, Bai S L, Zhao X P, Li Z, Zhang D L, Li H, Li S P, Zhou Y, Song C P. 2023. Introgression of QTL from *Aegilops tauschii* enhances yield-related traits in common wheat. *The Crop Journal*, **11**, 1521–1532.

Mahjourimajd S, Taylor J, Rengel Z, Khabaz-Saberi H, Kuchel H, Okamoto M, Langridge P. 2016. The genetic control of grain protein content under variable nitrogen supply in an Australian wheat mapping population. *PLoS ONE*, **11**, e0159371.

Marcone M F, Kakuda Y, Yada R Y. 1998. Salt-soluble seed globulins of various dicotyledonous and monocotyledonous plants — I. Isolation/purification and characterization. *Food Chemistry*, **62**, 27–47.

Marcussen T, Sandve S R, Heier L, Spannagl M, Pfeifer M, Jakobsen K S, Wulff B, Steuernagel B, Mayer K, Olsen O A. 2014. Ancient hybridizations among the ancestral genomes of bread wheat. *Science*, **345**, 1250092.

Miranda L M, Murphy J P, Marshall D, Cowger C, Leath S. 2007. Chromosomal location of *Pm35*, a novel *Aegilops tauschii* derived powdery mildew resistance gene introgressed into common wheat

(*Triticum aestivum* L.). *Theoretical and Applied Genetics*, **114**, 1451–1456.
Miranda L M, Murphy J P, Marshall D, Leath S. 2006. *Pm34*: A new powdery mildew resistance gene transferred from *Aegilops tauschii* Coss. to common wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, **113**, 1497.
Mirzaghaderi G, Mason A S. 2019. Broadening the bread wheat D genome. *Theoretical and Applied Genetics*, **132**, 1295–1307.
Muqaddasi Q H, Brassac J, Ebmeyer E, Kollers S, Korzun V, Argillier O, Stiewe G, Plieske J, Ganal M W, Röder M S. 2020. Prospects of GWAS and predictive breeding for European winter wheat's grain protein content, grain starch content, and grain hardness. *Scientific Reports*, **10**, 12541.
Ogbonnaya F C, Abdalla O, Mujeeb-Kazi A, Kazi A G, Xu S S, Gosman N, Lagudah E S, Bonnett D, Sorrells M E, Tsujimoto H. 2013. Synthetic hexaploids: harnessing species of the primary gene pool for wheat improvement. *Plant Breeding Reviews*, **37**, 35–122.
Olson E L, Rouse M N, Pumphrey M O, Bowden R L, Gill B S, Poland J A. 2013. Introgression of stem rust resistance genes *SrTA10187* and *SrTA10171* from *Aegilops tauschii* to wheat. *Theoretical & Applied Genetics*, **126**, 2477–2484.
Pu Z E, Ye X L, Li Y, Shi B X, Guo Z, Dai S F, Ma J, Liu Z H, Jiang Y F, Li W, Jiang Q T, Chen G Y, Wei Y M, Zheng Y L. 2022. Identification and validation of novel loci associated with wheat quality through a genome-wide association study. *Journal of Integrative Agriculture*, **21**, 3131–3147.
Rasheed A, Ogbonnaya F C, Lagudah E, Appels R, He Z. 2018. The goat grass genome's role in wheat improvement. *Nature Plants*, **4**, 56–58.
Raupp W J, Sukhwinder-Singh, Brown-Guedira G L, Gill B S. 2001. Cytogenetic and molecular mapping of the leaf rust resistance gene *Lr39* in wheat. *Theoretical and Applied Genetics*, **102**, 347–352.
Rowland G G, Kerber E R. 1974. Telocentric mapping in hexaploid wheat of genes for leaf rust resistance and other characters derived from *Aegilops squarrosa*. *Canadian Journal of Genetics & Cytology*, **16**, 137–144.
Rudd J C, Devkota R N, Baker J A, Peterson G L, Lazar M D, Bean B, Worrall D, Baughman T, Marshall D, Sutton R, Rooney L W, Nelson L R, Fritz A K, Weng Y, Morgan G D, Seabourn B W. 2014. 'TAM 112' wheat, resistant to greenbug and wheat curl mite and adapted to the dryland production system in the southern high plains. *Journal of Plant Registrations*, **8**, 291–297.
Saini D K, Devi P, Kaushik P. 2020. Advances in genomic interventions for wheat biofortification: A review. *Agronomy*, **10**, 62.
Sambasivam P, John M, Michael A, Urmil B, Xiaojing W, Li H, Karin D, Mingcheng L, Xiuying K, Harbans B. 2013. The gene *Sr33*, an ortholog of barley *Mla* genes, encodes resistance to wheat stem rust race Ug99. *Science*, **341**, 786–788.
Shang J, Li L, Liu C, Hong J, Liu M, Zhao B, Zheng X. 2021. Relationships of flour characteristics with isolated starch properties in different Chinese wheat varieties. *Journal of Cereal Science*, **99**, 103210.
Shevkani K, Katyal M, Singh N. 2024. A comparative review of protein and starch characteristics and end-use quality of soft and hard wheat. *Food Chemistry Advances*, **4**, 100613.
Singh R P, Nelson J C, E. S M. 2000. Mapping *Yr28* and other genes for resistance to stripe rust in wheat. *Crop Science*, **40**, 1148–1155.
Song Z P, Zuo Y Y, Xiang Q, Li W J, Li J, Liu G, Dai S F, Yan Z H. 2023. Investigation of *Aegilops umbellulata* for stripe rust resistance, heading date, and the contents of iron, zinc, and gluten protein. *Journal of Integrative Agriculture*, **22**, 1258–1265.
Su Y, Liao P, Song D, Huang S, He J, Gao X, Li S. 2020. Application of *Aegilops tauschii–Triticum aestivum* recombinant inbred lines for grain protein content quantitative trait loci detection and wheat improvement. *Canadian Journal of Plant Science*, **100**, 425–434.
Uthayakumaran S, Wrigley C W. 2010. Wheat: Characteristics and quality requirements. In: Wrigley C W, Batey I L, eds., *Cereal Grains: Assessing and Managing Quality*. Woodhead Publishing/ CRC Press, Cambridge, UK. pp. 59–111.
Wang X, Yan X, Tian X, Zhang Z, Wu W, Shang J, Ouyang J, Yao W, Li S. 2020. Glycine- and proline-rich protein OsGPRP3 regulates grain size and quality in rice. *Journal of Agricultural and Food Chemistry*, **68**, 7581–7590.
Wang Z, Li Y, Yang Y, Liu X, Qin H, Dong Z, Zheng S, Zhang K, Wang D. 2017. New insight into the function of wheat glutenin proteins as investigated with two series of genetic mutants. *Scientific Reports*, **7**, 3428.
Wiersma A T, Pulman J A, Brown L K, Cowger C, Olson E L. 2017. Identification of *Pm58* from *Aegilops tauschii*. *Theoretical and Applied Genetics*, **130**, 1123–1133.
Yan L, Liang F, Xu H, Zhang X, Zhai H, Sun Q, Ni Z. 2017. Identification of QTL for grain size and shape on the D genome of natural and synthetic allohexaploid wheats with near-identical AABB genomes. *Frontiers in Plant Science*, **8**, 1705.
Yang Z, Wang Z, Hu Z, Xin M, Yao Y, Peng H, You M, Su Z, Guo W. 2020. Comparative analysis of the genomic sequences between commercial wheat varieties Jimai 22 and Liangxing 99. *Acta Agronomica Sinica*, **46**, 1870–1883. (in Chinese)
Yu M, Chen G Y, Zhang L Q, Liu Y X, Liu D C, Wang J R, Pu Z E, Zhang L, Lan X J, Wei Y M, Liu C J, Zheng Y L. 2014. QTL mapping for important agronomic traits in synthetic hexaploid wheat derived from *Aegiliops tauschii* ssp. *tauschii*. *Journal of Integrative Agriculture*, **13**, 1835–1844.
Zhang G, Chen R Y, Shao M, Bai G, Seabourn B W. 2021. Genetic analysis of end-use quality traits in wheat. *Crop Science*, **61**, 1709–1723.
Zhang J, Ren J, Yang J, Fu S, Zhang X, Xia C, Zhao H, Yang K, Wen C. 2023. Evaluation of SNP fingerprinting for variety identification of tomato by DUS testing. *Agriculture Communications*, **1**, 100006.
Zhang L, Guan E, Yang Y, Liu Y, Zhang T, Bian K. 2021. Impact of wheat globulin addition on dough rheological properties and quality of cooked noodles. *Food Chemistry*, **362**, 130170.
Zhang L, Guan E, Zhang K, Zhang T, Bian K. 2022. The aggregation characteristics of wheat globulin induced by heating and edible salts and its effects on noodle processing quality. *LWT*, **154**, 112803.
Zhao G, Xie S, Zong S, Wang T, Mao C, Shi J, Li J. 2022. Mutation of *TL1*, encoding a novel C2H2 zinc finger protein, improves grains eating and cooking quality in rice. *Theoretical and Applied Genetics*, **135**, 3531–3543.
Zheng F F, Deng Z Y, Shi C L, Zhang X Y, Tian J C. 2013. QTL mapping for dough mixing characteristics in a recombinant inbred population derived from a waxy × strong gluten wheat (*Triticum aestivum* L.). *Journal of Integrative Agriculture*, **12**, 951–961.
Zhou Y, Bai S, Li H, Sun G, Zhang D, Ma F, Zhao X, Nie F, Li J, Chen L, Lv L, Zhu L, Fan R, Ge Y, Shaheen A, Guo G, Zhang Z, Ma J, Liang H, Qiu X, *et al.* 2021. Introgressing the *Aegilops tauschii* genome into wheat as a basis for cereal improvement. *Nature Plants*, **7**, 774–786.
Zhu C, Wang C, Zhou J, Wang Y, Chen Q, Fu L. 2022. Purification and identification of globulin-1 S allele as a novel allergen with N-glycans in wheat (*Triticum aestivum*). *Food Chemistry*, **390**, 133189.

Executive Editor-in-Chief Xueyong Zhang
Managing Editor Ning Wang