



Genome-Wide Association Study Reveals the Genetic Basis of Five Quality Traits in Chinese Wheat

Shuiyuan Hao^{1,2}, Hongyao Lou³, Haiwei Wang⁴, Jinghong Shi⁴, Dan Liu⁵, Baogerile⁶, Jianguang Tao⁷, Sanming Miao⁸, Qunce Pei⁸, Liangliang Yu⁷, Min Wu⁹, Ming Gao⁴, Naihu Zhao⁴, Jinchao Dong⁴, Mingshan You^{1*} and Mingming Xin^{1*}

¹ College of Agronomy, China Agricultural University, Beijing, China, ² Safety Production and Early Warning Control Laboratory of Green Agricultural Products in Hetao Region, Hetao College, Bayannur, China, ³ Institute of Hybrid Wheat, Beijing Academy of Agriculture Forestry Sciences, Beijing, China, ⁴ Department of Agriculture, Hetao College, Bayannur, China, ⁵ Department of Medicine, Hetao College, Bayannur, China, ⁶ Department of Library, Hetao College, Bayannur, China, ⁷ Bayannur City Meteorological Bureau, Bayannur, China, ⁸ Bureau of Agriculture and Animal Husbandry of Linhe District of Bayannur, Bayannur, China, ⁹ Bureau of Agriculture and Animal Husbandry of Urat Middle Banner of Bayannur, Bayannur, China

OPEN ACCESS

Edited by:

Guoyou Ye,
International Rice Research Institute
(IRRI), Philippines

Reviewed by:

Huihui Li,
Institute of Crop Sciences (CAAS),
China

Ernesto Igarua,
Spanish National Research Council
(CSIC), Spain

*Correspondence:

Mingshan You
msyou67@cau.edu.cn
Mingming Xin
1766094979@qq.com

Specialty section:

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

Received: 14 December 2021

Accepted: 14 February 2022

Published: 03 March 2022

Citation:

Hao S, Lou H, Wang H, Shi J,
Liu D, Baogerile, Tao J, Miao S, Pei Q,
Yu L, Wu M, Gao M, Zhao N, Dong J,
You M and Xin M (2022)
Genome-Wide Association Study
Reveals the Genetic Basis of Five
Quality Traits in Chinese Wheat.
Front. Plant Sci. 13:835306.
doi: 10.3389/fpls.2022.835306

Bread wheat is a highly adaptable food crop grown extensively around the world and its quality genetic improvement has received wide attention. In this study, the genetic loci associated with five quality traits including protein content (PC), gluten content (GC), baking value (BV), grain hardness (HA), and sedimentation value (SV) in a population of 253 Chinese wheat grown in Inner Mongolia were investigated through genome wide association mapping. A total of 103 QTL containing 556 SNPs were significantly related to the five quality traits based on the phenotypic data collected from three environments and BLUP data. Of these QTL, 32 QTL were continuously detected under at least two experiments. Some QTL such as *qBV3D.2/qHA3D.2* on 3D, *qPC5A.3/qGC5A* on 5A, *qBV5D/qHA5D* on 5D, *qBV6B.2/qHA6B.3* on 6B, and *qBV6D/qHA6D.1* on 6D were associated with multiple traits. In addition, distribution of favorable alleles of the stable QTL in the association panel and their effects on five quality traits were validated. Analysis of existing transcriptome data revealed that 34 genes were specifically highly expressed in grains during reproductive growth stages. The functions of these genes will be characterized in future experiments. This study provides novel insights into the genetic basis of quality traits in wheat.

Keywords: wheat, quality trait, GWAS, QTL, favorable alleles

INTRODUCTION

Bread wheat (*Triticum aestivum* L.), which is a highly adaptable food crop cultivated globally, supplies approximately 20% of the total calories consumed by humans (Shiferaw et al., 2013; Shewry and Hey, 2015). As a staple food for about one-third of the global population, wheat is an important source of protein, vitamins, dietary fiber, and minerals. Because of its unique gluten properties, wheat can be used to produce various foods (Kumar et al., 2018). The improvement of living standards for many people worldwide and the development of the flour processing industry have

forced breeders to focus on generating new high-quality wheat varieties (Curtis and Halford, 2014; Gao et al., 2021).

Wheat quality is a comprehensive trait that is influenced by the individual or combined effects of diverse traits, including grain protein content, gluten protein quality, grain hardness, flour color, starch quality, and nutrient element content; these traits are controlled by multiple quantitative trait loci (QTL) and are very sensitive to agro-environmental factors (Wang et al., 2018). Exploring the genetic basis of quality traits will accelerate the breeding of high-quality wheat varieties. Despite the complexity and limited genetic diversity of the wheat genome, there has been substantial molecular genetic research into wheat quality traits, which has resulted in the identification of some genes/QTL associated with quality traits through reverse genetics and traditional mapping methods (Uauy et al., 2006; He et al., 2007, 2009; Kumar et al., 2009; Crawford et al., 2011; Sun et al., 2016; Boehm et al., 2018; Jin et al., 2018; Goel et al., 2019; Semagn et al., 2021).

The traditional QTL mapping method usually involves the hybrid population of two parents and the detection of target trait loci via linkage mapping (Bettgenhaeuser and Krattinger, 2019). The limitation of this method is that relatively few recombination events occur during artificial hybridization, which leads to low resolution (Rabbi et al., 2021). Fine mapping requires the investment of considerable time, labor, and resources to develop many recombinant populations. Additionally, the evaluation of wheat quality traits necessitates the large-scale production of seeds to be ground into flour, which is one of challenges to traditional QTL mapping. With the rapid development of genome sequencing technology and high-density SNP arrays for genotyping (e.g., 9, 35, 55, 90, 660, and 820), the genome-wide association study (GWAS) has become an increasingly popular and promising alternative to linkage mapping for elucidating the genetic basis of complex traits, especially quality traits (Yu et al., 2006; Zhu et al., 2008; Ji et al., 2021). This method enables the use of natural populations without interbreeding to produce progeny, making it convenient for phenotype and genotype identifications. Thus, GWAS has been widely used for studying multiple traits in various plant species, such as *Arabidopsis* (Meijon et al., 2014; Proietti et al., 2018), rice (Yano et al., 2016; Tang et al., 2019), maize (Wang et al., 2016; Luo et al., 2019), and sorghum (Xie et al., 2019), as well as for investigating wheat agronomic traits (Liu et al., 2017; Sun et al., 2017) and disease resistance (Wu et al., 2021). The results of these studies provide useful information for identifying the genetic loci affecting wheat quality traits via GWAS.

In China, which is the largest wheat producer after the European Union, there are 10 wheat-growing regions (Sun et al., 2017). The wheat cultivated in the Hetao region of Inner Mongolia in the northern spring wheat region is well-known for its excellent quality, which is the result of the local climatic conditions (e.g., large day-to-night temperature changes, long-day conditions, and a long frost-free period). Improving wheat quality through molecular breeding will enhance the competitiveness of Hetao wheat in the international market. Furthermore, to the best of our knowledge, there have been no genetic studies on the Hetao wheat accessions conducted

to identify and characterize the QTL associated with quality-related traits. Therefore, the present study was undertaken to identify genes/QTL related to wheat quality traits across different environments in Inner Mongolia using an association mapping approach combined with a high-density SNP marker assay.

MATERIALS AND METHODS

Plant Materials and Experimental Design

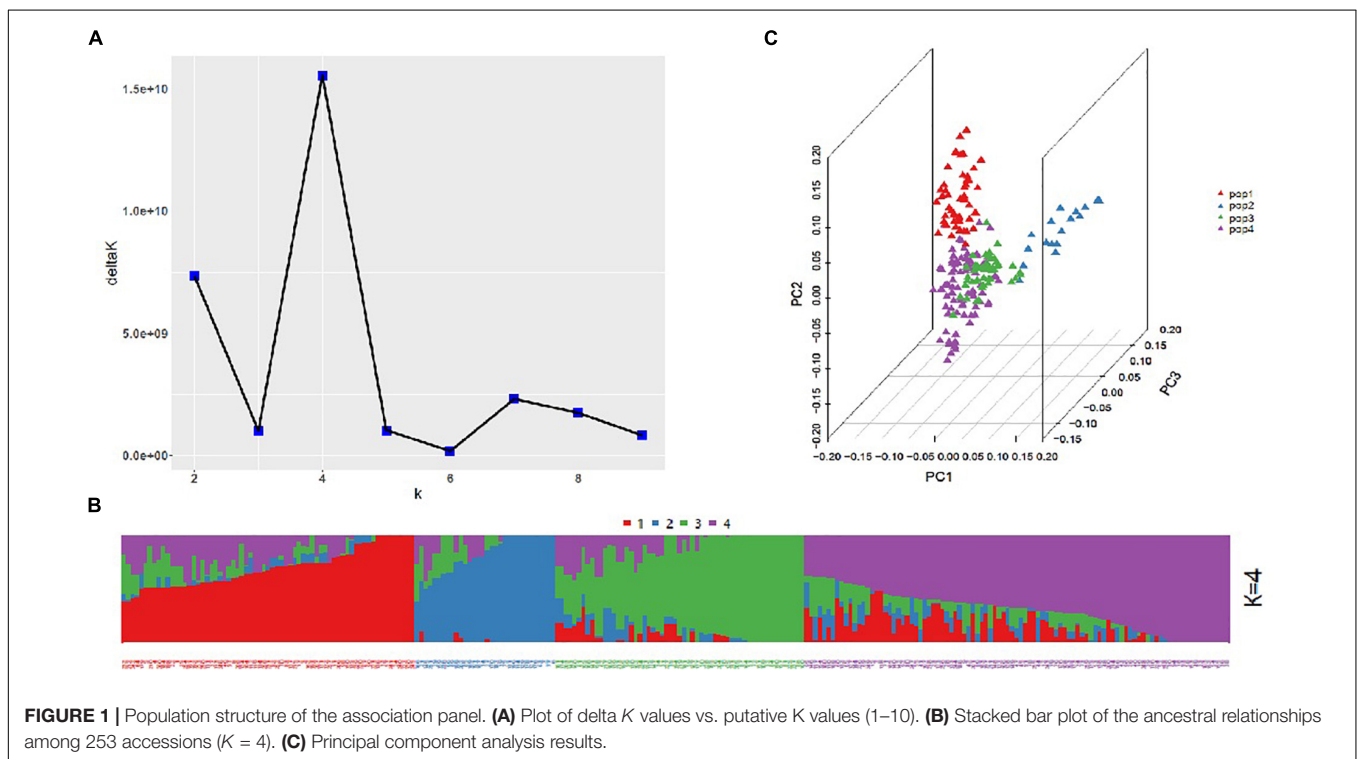
The association panel used in this study consisted of 253 spring and semi-winter common wheat accessions collected from various wheat-producing regions in China, but mostly from the Inner Mongolia Autonomous Region and Sichuan province. These accessions were grown in Hangjihouqi (HJHQ; 40.85°N, 107.12°E), Linhe (LH; 40.73°N, 107.37°E), and Wulateqianqi (WLT; 40.73°N, 108.64°E) in the Inner Mongolia Autonomous Region during the 2019–2020 cropping season. Field trials were conducted using a randomized complete block design with three replications. Each accession was grown in a single row (2 m long) with 25 cm separating rows. Fertilizers were applied according to the following schedule: (1) 35 kg of diammonium phosphate and 10 kg of potassium nitrate were used as seed fertilizers (2) 40 kg of urea and 25 kg of compound fertilizer (15:15:15) were applied at the first and second watering, respectively. Other cultivation and management measures refer to local agronomic practices.

Phenotypic Data Analysis

After the seeds harvested from the plants were dried and cleaned, the following five quality-related traits were analyzed by near-infrared spectroscopy using an Infraneo Senior instrument (Chopin Technologies, France): protein content (PC), gluten content (GC), baking value (BV), sedimentation value (SV), and grain hardness (HA). The near-infrared calibration model of wheat was developed and calibrated by the Application Department of Chopin Technologies, France, using the partial least squares method. Each sample was examined twice and the average value was used for further analyses. The R software package “pastecs” was used for the statistical analysis of phenotypic data and the analysis of variance for all five traits in different environments. The formula for the best linear unbiased estimate (BLUP) of phenotype is: $y = Xb + Zu + e$, where y , b , u , and e represent the observed phenotype, fixed effect vector, random effect vector, and residual, respectively, and X and Z represent incidence matrices. The R package “lme4” was used to estimate BLUP values for each line, considering environmental effects as fixed and genotype as random: $y \sim (1 | rep\% \text{ in}\% \text{ env}) + (1 | env) + (1 | lines) + (1 | env: lines)$, where $rep\% \text{ in}\% \text{ env}$ represents replications were nested within the environments. The broad-sense heritability (h^2) of each trait was calculated as follows: $h^2 = \sigma_g^2 / (\sigma_g^2 + \sigma_{ge}^2/n + \sigma_{er}^2/nr)$, where σ_g^2 , σ_{ge}^2 , and σ_{er}^2 represent the variance components for genotype, genotype \times environment, and error, respectively, and n and r are the number of environments and replications, respectively. Correlations between traits were analyzed using the R package “corrplot.”

TABLE 1 | SNP marker distribution in the A, B, D, and whole genomes.

Chr	No. of SNPs				Map length (Mb)				SNP density (Mb/SNP)			
	A	B	D	Total	A	B	D	Total	A	B	D	Total
1	1753	1729	1233	4715	593.25	688.70	495.13	1777.08	0.34	0.40	0.40	0.38
2	2011	1646	1280	4937	778.74	798.75	650.88	2228.37	0.39	0.49	0.51	0.45
3	1479	1761	1035	4275	749.87	829.73	615.46	2195.06	0.51	0.47	0.59	0.51
4	1138	1940	597	3675	743.32	673.47	509.18	1925.97	0.65	0.35	0.85	0.52
5	1877	1865	1116	4858	709.19	712.40	564.75	1986.34	0.38	0.38	0.51	0.41
6	1078	1563	1061	3702	616.94	720.49	473.52	1810.95	0.57	0.46	0.45	0.49
7	1543	1638	1449	4630	733.94	747.32	634.28	2115.54	0.48	0.46	0.44	0.46
All	10879	12142	7771	30792	4925.25	5170.86	3943.20	14039.31	0.45	0.43	0.51	0.46



SNP Genotyping

A wheat 55 K SNP array comprising 53,063 markers was used to genotype the association panel by China Golden Marker (Beijing) Biotech Co., Ltd¹. The SNP data were analyzed using the Illumina GenomeStudio genotyping software. After filtering out the SNPs with a minor allele frequency (MAF) < 5% and a missing rate > 10%, the remaining 30,792 high-quality SNPs were included in the subsequent analyses.

Population Structure, Principal Component, and Linkage Disequilibrium Analyses

The genetic structure of the natural population was assessed using two programs. First, the ADMIXTURE software

(Alexander et al., 2009) was used to estimate the population structure of the panel according to polymorphic SNPs. The K value (number of hypothetical groups) ranged from 1 to 10, and the highest delta K value was set as the optimal value as previously described (Evanno et al., 2005). Second, the GCTA software (Yang et al., 2011) was used to perform a principal component analysis for the association panel. The R packages “Plot3D” and “Scatterplot3d” were used to prepare the figures presenting the results of the principal component analysis. The kinship matrix (K) was calculated using the Centered IBS method of the TASSEL software (Bradbury et al., 2007). LD analysis was performed using the TASSEL software. LD was estimated by pairwise comparisons among filtered SNP markers using squared allele frequency correlations (r^2). The parameters for calculating r^2 are to set “Select LD Type” to sliding window and “LD Window Size” to 1000. The LD decay distance was determined as the distance at which r^2 decreased to half of the maximum value.

¹<http://www.cgmb.com.cn>

TABLE 2 | Descriptive statistics, broad-sense heritability, and analysis of variance for the five quality traits in the association panel.

Trait	Environment	Mean	SD	CV%	Min	Max	Range	h^2	F-values from ANOVA	
									Genotype	Environment
PC	HJHQ	16.51	0.71	4.31	14.51	19.56	5.05	76.09	3.86***	14.85***
	LH	16.38	0.77	4.71	14.34	19.27	4.93			
	WLTQQ	16.49	0.88	5.29	14.39	19.68	5.29			
	BLUP	16.51	0.47	2.89	14.98	18.05	3.07			
GC	HJHQ	40.65	2.48	6.09	35.74	49.89	14.15	73.13	3.54***	16.05***
	LH	41.40	2.57	6.21	34.54	50.50	15.96			
	WLTQQ	41.54	2.81	6.78	35.12	51.53	16.41			
	BLUP	41.19	1.51	3.67	36.93	45.96	9.03			
BV	HJHQ	395.87	27.68	6.99	301.49	517.87	216.38	73.98	3.26***	201.84***
	LH	364.18	31.32	8.60	294.20	472.32	178.12			
	WLTQQ	401.36	31.37	7.82	318.45	503.83	185.38			
	BLUP	387.30	18.18	4.69	330.11	450.55	120.44			
SV	HJHQ	41.66	7.14	17.14	21.84	63.41	41.57	76.10	3.51***	347.91***
	LH	33.10	6.44	19.47	21.13	54.61	33.48			
	WLTQQ	44.36	7.05	15.92	28.44	68.78	40.34			
	BLUP	39.63	4.26	10.76	28.86	54.73	25.87			
HA	HJHQ	67.96	1.17	1.74	64.47	72.32	7.85	76.96	3.98***	168.22***
	LH	66.98	1.21	1.80	64.38	72.17	7.79			
	WLTQQ	68.34	1.32	1.93	65.37	73.95	8.58			
	BLUP	67.76	0.76	1.12	65.77	71.15	5.38			

***Represents significance at 0.001 level.

Genome-Wide Association Analysis

A GWAS was conducted using the rMVP 1.0 software (Yin et al., 2021). More specifically, the General Linear Model (GLM), Mixed Linear Model (MLM), and Fixed and random model Circulating Probability Unification (FarmCPU) were used, of which MLM considered the population structure and relative kinship simultaneously. Marker–trait associations were identified on the basis of the polymorphic SNPs as well as the phenotypic data for the five wheat quality traits in different environments and the BLUP values. Considering the differences in the efficiency of GWAS for determining the marker–trait associations for different traits, two threshold P values were calculated according to the number of markers ($P_1 = 1/n$ for BV, GC, and PC; $P_2 = 0.05/n$ for HA and SV; $n =$ total number of SNPs with $MAF > 0.05$). The R package “man” was used to prepare quantile–quantile plots and Manhattan plots.

Putative Candidate Gene Analysis

To identify potential candidate genes, the IWGSC database² was used to retrieve high-confidence genes located within 4 Mb of the LD region flanking significant SNPs included in stable QTL revealed by GWAS. Gene annotations and the homologs of specific wheat genes in rice and Arabidopsis were investigated using the Triticeae-Gene Tribe website³ (Chen et al., 2020). The reference genomes of the following species were used in this study: *T. aestivum* (IWGSC RefSeq v1.1), *Oryza sativa japonica* group (IRGSP 1.0), and Arabidopsis (TAIR10). The

²<http://www.wheatgenome.org/>

³<http://wheat.cau.edu.cn/TGT/>

publicly available wheat transcriptome database was used to investigate the spatiotemporal expression signature of candidate genes, and transcripts per kilobase million (TPM) values were used to represent the expression levels of candidate genes. TPM values for candidate genes in five tissues (root, leaf, shoot, spike, and grain) and two growth stage (vegetative and reproductive) obtained from the wheat expression Database⁴ (Borrill et al., 2016; Ramírez-González et al., 2018).

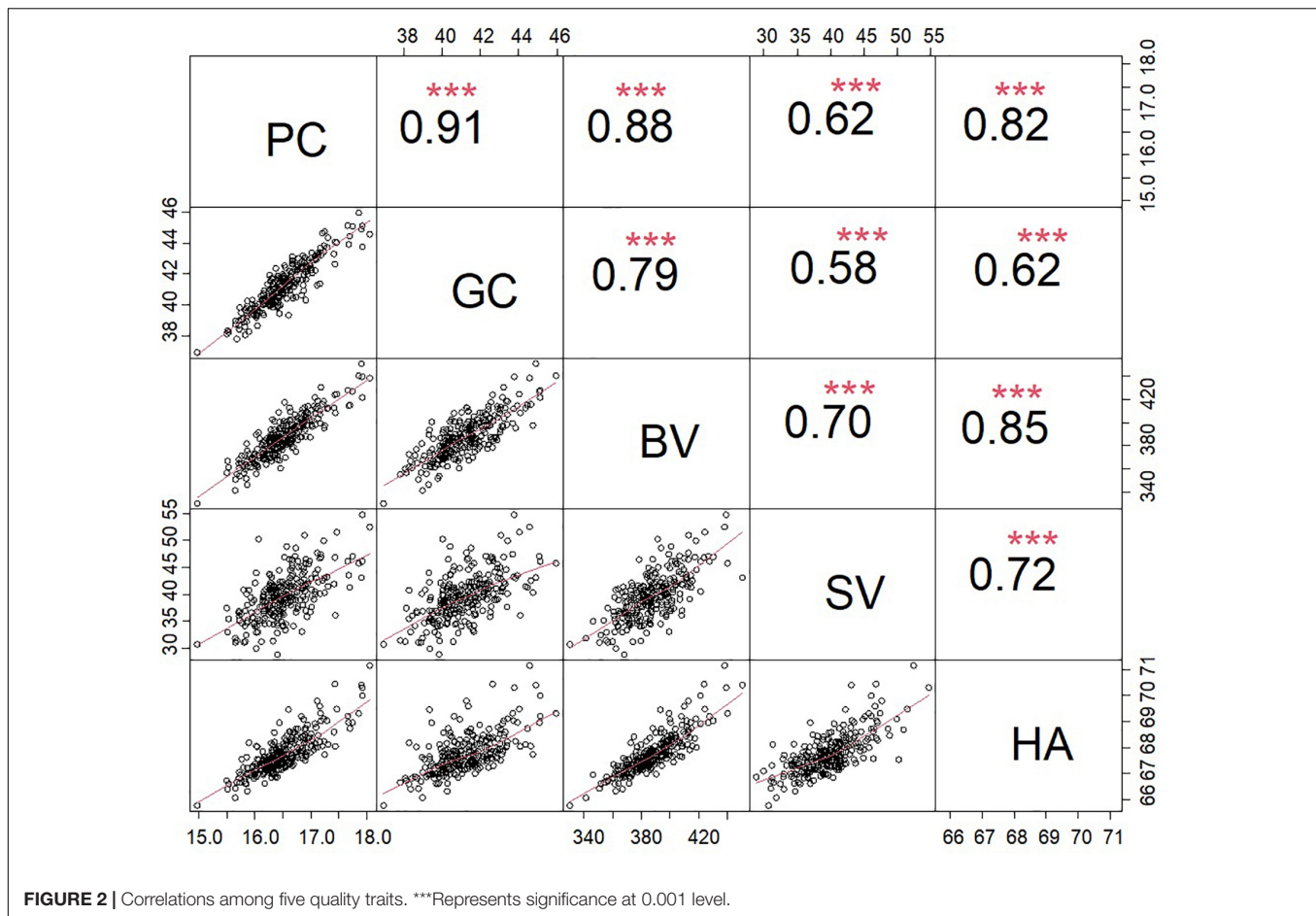
RESULTS

Molecular Markers, Population Structure, Principal Component Analysis, and LD

A total of 30,729 high-quality SNP markers were identified across all 21 wheat chromosomes (Table 1 and Supplementary Figure 1), of which 10,879, 12,142, and 7,771 markers were located in the A, B, and D subgenomes, respectively. The total chromosome length spanned by these SNPs was 14039.31 Mb, with 6D and 3B having the shortest (473.52 Mb) and longest (829.73 Mb) chromosome length, respectively. The distance between SNPs ranged from 0.34 Mb on chromosome 1A to 0.65 Mb on chromosome 4D, with a genome-wide average of 0.46 Mb.

The highest delta K value representing genetic clusters revealed by ADMIXTURE was calculated as 4 when 30,729

⁴<http://www.wheat-expression.com>



polymorphic markers were used; hence, the 253 accessions were divided into four subgroups (pop1, pop2, pop3, and pop4) (Figure 1), which was consistent with the result of principal components analysis based on the first three components contribution. Specifically, pop1, pop2, pop3, and pop4 comprised 67, 32, 57, and 97 accessions, respectively. The LD decay patterns varied by subgenomes. The LD decay distance was largest (approximately 7.6 Mb) and smallest (approximately 3.5 Mb) in the B and D subgenomes, respectively. The average LD decay distance for the whole genome was approximately 4 Mb (Supplementary Figure 2).

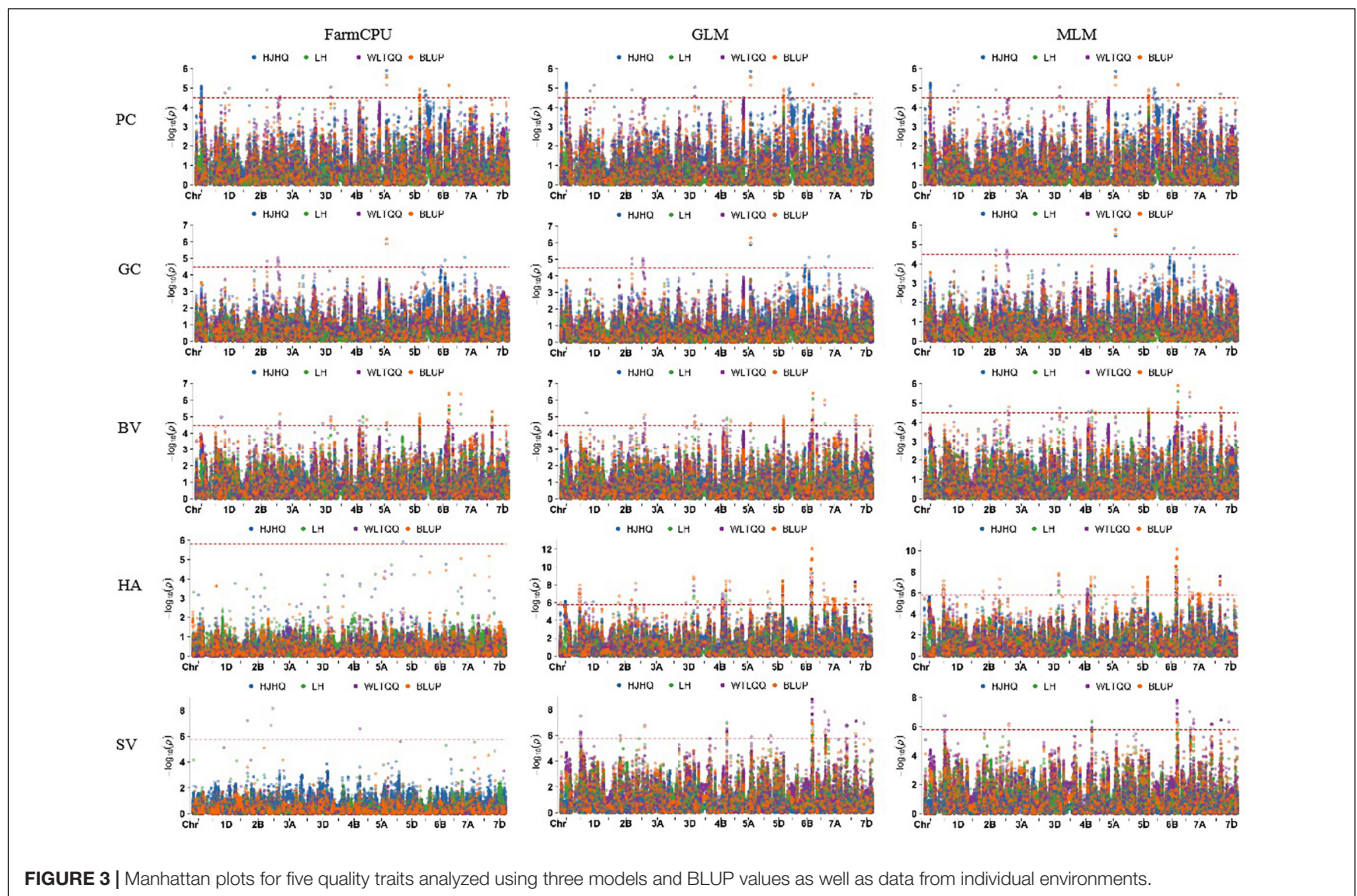
Phenotypic Variations and Correlations

The analysis of the association panel revealed considerable variations in all traits evaluated across all environments (Supplementary Table 1). Descriptive statistics for these traits are presented in Table 2. The variations ranged from 14.34 to 19.68% ($16.51\% \pm 0.47\%$) for PC, 34.54% to 51.53% ($41.49\% \pm 1.51\%$) for GC, 294.20 to 517.87 (387.30 ± 18.18) for BV, 21.13 to 68.78 mm (39.63 ± 4.26 mm) for SV, and 64.38 to 73.95 (67.76 ± 0.76) for HA. The combined variance analysis of the five traits detected significant differences between the genotypes of the association populations, the environments, and the genotype \times environment interactions. The heritability (h^2) of the five traits ranged from

73.13% to 76.96% (Table 2). Significant positive correlations were detected among the five traits according to the phenotypic BLUP values, with the correlation coefficient ranging from 0.80 (between PC and GC) to 0.91 (between GC and SV) (Figure 2).

Loci Significantly Associated With Five Quality Traits

Using 30,792 polymorphic SNPs (missing rate $\leq 10\%$ and MAF $\geq 5\%$), a GWAS involving the GLM, MLM, and FarmCPU models was performed to analyze five quality traits. A total of 556 SNPs significantly associated with the five traits were detected on the basis of the phenotypic data for the different environments and the BLUP values (Supplementary Table 2, Figure 3, and Supplementary Figure 3). These SNPs were distributed over all wheat chromosomes, with the exception of chromosome 3A. Additionally, the SNPs were relatively concentrated in some chromosomal regions (e.g., the long arm of chromosomes 1A, 4B, 2D, 5D, and 6B). Among the examined traits, HA was associated with the most SNPs (260), which was followed by PC (108), SV (83), BV (70), and GC (35). On the basis of the average LD decay distance (about 4 Mb) in the genomes of the 253 accessions, significant SNPs within a 4-Mb region were combined as a single QTL. Thus, 103 QTL were identified for the five quality traits (Supplementary Table 3). The number of QTL per trait ranged



from 15 for GC to 37 for HA. The 103 significant QTL were almost evenly distributed among the three subgenomes, with 32 on the A subgenome, 37 on the B subgenome, and 34 on the D subgenome. But they were unevenly distributed on different chromosomes, ranging from 1 (2D and 6B) to 11 (2A and 5B), indicating that the power of QTL identification was affected by marker density. Of these QTL, 32 were identified simultaneously in at least one environment and on the basis of BLUP values, including 2 for each for PC and GC, 8 for BV, 15 for HA and 5 for SV (Table 3).

Interestingly, several QTL were associated with multiple quality traits. For example, the QTL on chromosome 5A (*qPC5A.3/qGC5A*; 479.28–479.32 Mb) were associated with PC and GC, whereas the QTL on chromosomes 3D (*qBV3D.2/qHA3D.2*; 571.36–572.83 Mb), 5D (*qBV5D/qHA5D*; 547.50–549.49 Mb), and 6D (*qBV6D/qHA6D.1*; 437.62 Mb) were associated with BV and HA. Additionally, two QTL on chromosome 6B (*qBV6B.1/qHA6B.2*; 613.35–632.44 Mb and *qBV6B.2/qHA6B.3*; 653.06–662.45 Mb) related to BV were also associated with HA and SV.

Because stability across different environments and models is an important parameter for evaluating the reliability of QTL, during our haplotype analysis, we focused on the QTL on chromosomes 3D, 5A, 5D, 6B, and 6D that were detected in at least two environments by all three models and had pleiotropic effects on multiple traits. We divided 252 accessions into two

haplotypes according to the allele types of the three SNPs in the common QTL for PC and GC on chromosome 5A. Hap1 had a higher average PC (16.88%) and GC (42.33%) than Hap2 (16.47% for PC and 41.06% for GC). The 252 accessions were also divided into two haplotypes according to the allele types of the SNPs in the five common QTL for BV and HA on chromosomes 3D, 5D, 6B, and 6D. There were significant differences in the phenotypic values between the haplotypes (Table 4 and Figure 4).

Relationship Between the Quality Traits and the Number of Favorable Alleles

In this study, marker alleles with positive additive effects leading to increased phenotypic values were considered to be favorable alleles. Allelic effects were simulated for all five examined quality traits using the most significant SNPs in each QTL interval. The relationship between the effects of the favorable and unfavorable alleles was determined by analyzing the BLUP values for the phenotypes. The number of favorable alleles in individual accessions ranged from 0 to 16, 1 to 15, 0 to 16, 1 to 35, and 0 to 16 for PC, GC, BV, HA, and SV, respectively (Figure 5). The accessions were grouped according to the number of favorable alleles for each trait. A significant linear relationship was observed between the traits and the number of favorable alleles. In other words, increases in the number of favorable alleles were associated with increases in the phenotypic values

TABLE 3 | Details regarding the 32 stable QTL identified simultaneously in at least one environment and on the basis of BLUP values.

Trait	QTL	No. SNPs	Chr	Region (Mb)	Environments (Model)
PC	<i>qPC1A</i>	68	1A	301.16–318.19	HJHQ (FarmCPU/GLM/MLM), BLUP (GLM)
	<i>qPC5A.3</i>	3	5A	479.28–479.32	HJHQ (FarmCPU/GLM/MLM), BLUP (FarmCPU/GLM/MLM)
GC	<i>qGC2B.2</i>	2	2B	723.96–731.60	WLTQQ (FarmCPU/GLM/MLM), BLUP (FarmCPU)
	<i>qGC5A</i>	3	5A	479.28–479.32	HJHQ (FarmCPU/GLM/MLM), BLUP (FarmCPU/GLM/MLM)
BV	<i>qBV2D.2</i>	2	2D	481.59–489.72	WLTQQ (FarmCPU/GLM), BLUP (FarmCPU/GLM/MLM)
	<i>qBV3D.2</i>	3	3D	571.36–572.83	WLTQQ (FarmCPU/GLM), BLUP (FarmCPU/GLM)
	<i>qBV5D</i>	25	5D	547.50–549.49	LH (FarmCPU/GLM/MLM), BLUP (FarmCPU/GLM/MLM)
	<i>qBV6B.1</i>	12	6B	613.35–632.44	LH (FarmCPU/GLM/MLM), BLUP (FarmCPU/GLM)
	<i>qBV6B.2</i>	7	6B	658.73–662.45	LH (FarmCPU/GLM/MLM), WLTQQ (FarmCPU/GLM/MLM), BLUP (FarmCPU/GLM/MLM)
	<i>qBV6D</i>	1	6D	437.62	LH (FarmCPU/GLM), WLTQQ (FarmCPU/GLM/MLM), BLUP (FarmCPU/GLM/MLM)
	<i>qBV7B.1</i>	1	7B	605.96	WLTQQ (GLM/MLM), BLUP (FarmCPU/GLM)
	<i>qBV7B.2</i>	7	7B	634.30–635.02	LH (FarmCPU/GLM), WLTQQ (FarmCPU/GLM), BLUP (FarmCPU/GLM/MLM)
HA	<i>qHA1A.1</i>	1	1A	54.5	WLTQQ (GLM), BLUP (GLM)
	<i>qHA1B.1</i>	8	1B	336.20–343.25	WLTQQ (GLM/MLM), BLUP (GLM/MLM)
	<i>qHA2B.2</i>	1	2B	105.48	WLTQQ (GLM), BLUP (GLM/MLM)
	<i>qHA2D.1</i>	1	2D	34.23	WLTQQ (GLM/MLM), BLUP (GLM/MLM)
	<i>qHA3D.2</i>	3	3D	571.36–572.83	LH (FarmCPU/GLM), WLTQQ (FarmCPU/GLM/MLM), BLUP (FarmCPU/GLM/MLM)
	<i>qHA4B.1</i>	2	4B	465.74–465.90	WLTQQ (GLM), BLUP (GLM)
	<i>qHA4B.2</i>	29	4B	477.23–481.80	WLTQQ (GLM/MLM), BLUP (GLM)
	<i>qHA4B.3</i>	12	4B	658.66–659.16	LH (GLM/MLM), WLTQQ (GLM), BLUP (GLM/MLM)
	<i>qHA4D</i>	1	4D	47.32	LH (GLM/MLM), WLTQQ (GLM/MLM), BLUP (GLM/MLM)
	<i>qHA5D</i>	39	5D	547.50–550.44	LH (GLM/MLM), WLTQQ (GLM/MLM), BLUP (GLM/MLM)
	<i>qHA6B.2</i>	23	6B	613.35–632.44	HJHQ (GLM/MLM), LH (GLM/MLM), WLTQQ (GLM/MLM), BLUP (GLM/MLM)
	<i>qHA6B.3</i>	19	6B	653.06–662.45	HJHQ (GLM/MLM), LH (GLM/MLM), WLTQQ (GLM/MLM), BLUP (GLM/MLM)
	<i>qHA6D.1</i>	1	6D	437.62	WLTQQ (GLM/MLM), BLUP (GLM/MLM)
	<i>qHA7A.1</i>	3	7A	280.11–284.15	WLTQQ (GLM), BLUP (GLM/MLM)
<i>qHA7B.2</i>	7	7B	605.96–635.02	WLTQQ (GLM/MLM), BLUP (GLM/MLM)	
SV	<i>qSV2D.3</i>	1	2D	489.72	WLTQQ (GLM/MLM), BLUP (FarmCPU/GLM)
	<i>qSV4B</i>	13	4B	649.58–659.16	LH (GLM/MLM), WLTQQ (FarmCPU/GLM/MLM), BLUP (GLM)
	<i>qSV6B.1</i>	23	6B	613.35–632.44	LH (GLM), WLTQQ (GLM/MLM), BLUP (GLM/MLM)
	<i>qSV6D.1</i>	1	6D	437.62	WLTQQ (GLM/MLM), BLUP (GLM)
	<i>qSV6D.2</i>	11	6D	470.09–470.96	LH (GLM), WLTQQ (FarmCPU/GLM)

for the traits. More specifically, as the number of favorable alleles increased, the values for PC, GC, BV, HA, and SV increased by 1.05%, 2.99%, 41.00, 3.60, and 10.85 ml, respectively (**Figure 5**).

Identification and Functional Characterization of Candidate Genes

In this study, the 4-Mb sequences flanking the significant SNPs in each stable QTL associated with five quality traits were regarded as potential candidate gene regions. A total of 4,176 putative candidate genes were identified by screening the annotated genes in the “Chinese Spring” reference genome database (IWGSC RefSeq v1.1) (**Supplementary Table 4**). To further speculate on the causal genes responsible for the formation of wheat quality, we analyzed the expression characteristics of these genes in different developmental stages and tissues using RNA-seq data available in an online wheat gene expression database (see text footnote 4). It was found that 34 genes (16 of those were associated with at least two traits) were specifically and highly expressed in grains at the reproductive growth stage, from 4 for PC to 23 for HA (**Supplementary Table 5** and **Supplementary**

Figure 4). These genes were predicted to encode proteins with various functions, such as aquaporin (*TraesCS1A02G165400*), vacuolar-processing enzyme (*TraesCS2D02G382400*, *TraesCS2D02G383400*, *TraesCS2D02G383500*), cytochrome P450 (*TraesCS5D02G535400*), zinc finger CCCH domain-containing protein (*TraesCS2B02G138100*), phosphoglycerate kinase (*TraesCS2D02G081700*), divalent metal cation transporter (*TraesCS4B02G381200*), tropinone reductase-like protein (*TraesCS5D02G530100*), but also include some proteins with unknown functions (**Supplementary Table 5**). Given that wheat quality is closely related to grain development and its composition, these specifically expressed genes are most likely to be key causal genes for wheat quality traits, and their functions were worth further investigation in the future.

DISCUSSION

Wheat has been cultivated in China for more than 4,000 years, and it is now grown in 10 major agro-ecological regions (Hao et al., 2008; Sun et al., 2017). High-quality spring wheat is

TABLE 4 | Analysis of the haplotype effects of the stable QTL on chromosomes 3D, 5A, 5D, 6B, and 6D.

Trait	QTL	Haplotype	Allele	No. accession	BLUP
PC	<i>qPC5A.3</i>	Hap1-PC	GCC	24	16.88
		Hap2-PC	TGT	228	16.47
GC	<i>qGC5A</i>	Hap1-GC	GCC	24	42.33
		Hap2-GC	TGT	228	41.06
BV	<i>qBV3D.2</i>	Hap1-BV	CCC	47	394.96
		Hap2-BV	ATT	200	385.37
	<i>qBV5D</i>	Hap1-BV	CCG...CTT	55	393.84
		Hap2-BV	AGA...TGG	197	385.27
	<i>qBV6B.1</i>	Hap1-BV	TAC...CAA	28	397.5
		Hap2-BV	CGT...TGG	225	385.96
	<i>qBV6B.2</i>	Hap1-BV	GCTAATA	21	402.87
		Hap2-BV	TTCGCCG	232	385.83
	<i>qBV6D</i>	Hap1-BV	T	54	395.79
		Hap2-BV	C	198	384.78
HA	<i>qHA3D.2</i>	Hap1-HA	CCC	47	68.31
		Hap2-HA	ATT	200	67.63
	<i>qHA5D</i>	Hap1-HA	CCG...CTT	55	68.22
		Hap2-HA	AGA...TGG	197	67.62
	<i>qHA6B.2</i>	Hap1-HA	TAC...CAA	28	68.57
		Hap2-HA	CGT...TGG	225	67.66
	<i>qHA6B.3</i>	Hap1-HA	GCTAATA	21	68.81
		Hap2-HA	TTCGCCG	232	67.66
	<i>qHA6D.1</i>	Hap1-HA	T	54	68.25
		Hap2-HA	C	198	67.62

cultivated in the Hetao Irrigation District in Inner Mongolia because of the superior regional environmental conditions (e.g., light and heat) as well as the availability of water sources. However, the genetic diversity of the wheat in this region is low, resulting in varietal degradation and poor stress resistance. Clarifying the genetic basis of the quality traits in Hetao wheat varieties is important for ensuring the efficient and sustainable development and competitiveness of the Hetao wheat industry through the breeding of improved varieties. In this study, we evaluated five quality traits and the population structure in a panel comprising 253 wheat accessions. Additionally, we conducted a GWAS using a wheat 55 K SNP array to analyze wheat grown in multiple environments in Inner Mongolia.

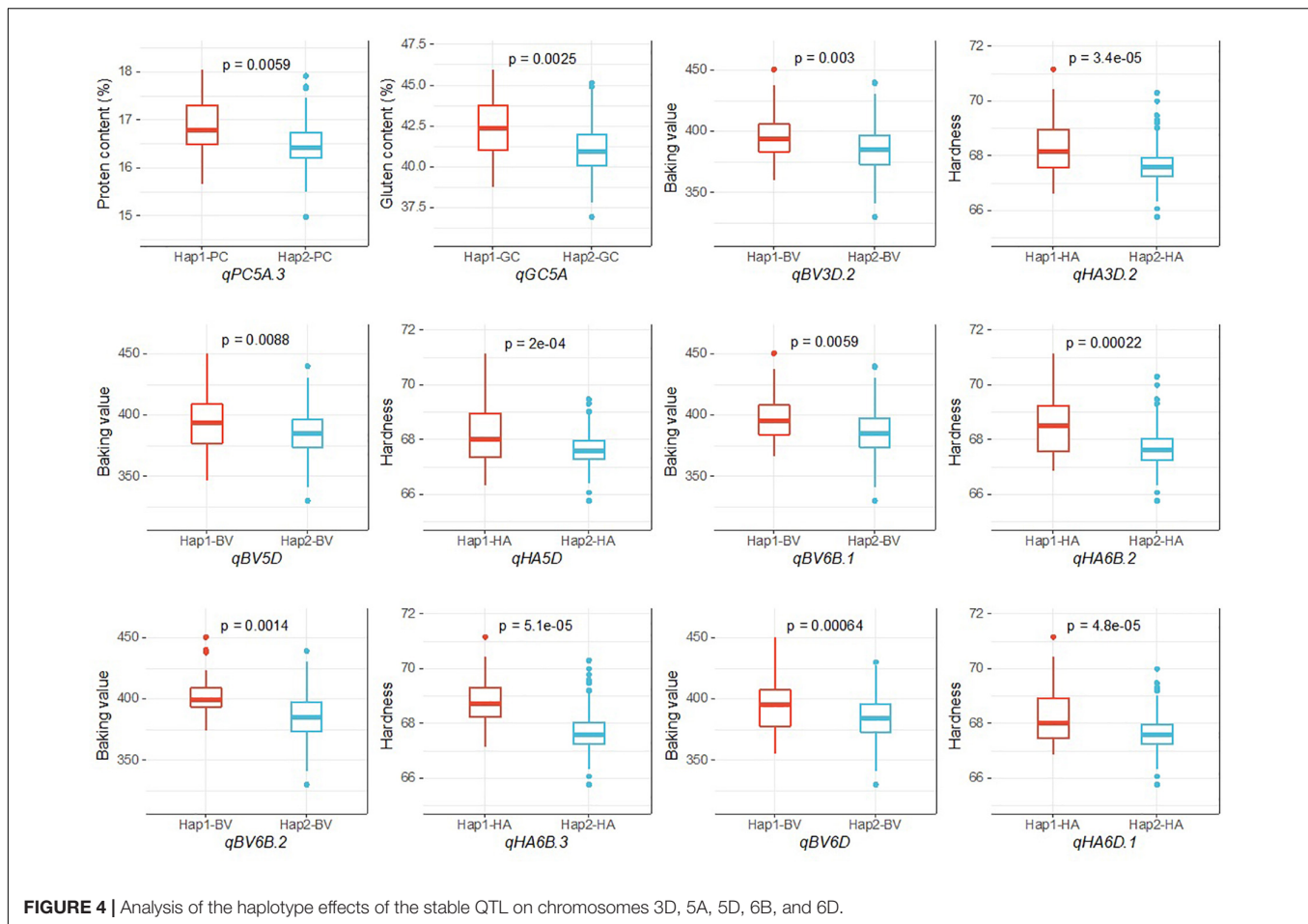
Wheat quality is a complex and comprehensive trait influenced by individual traits and their interrelationships and interactions. In the current study, strong correlations between individual traits were revealed by the BLUP data (Figure 2), which are consistent with those of other recent studies (Gao et al., 2021; Lou et al., 2021). Additionally, significant variations in individual traits were observed in the 253 accessions included in this study. However, few materials (Longmai 15–5049) performed well in multiple traits at the same time. In this regard, there is an urgent need to screen or introduce more excellent parents, which is crucial for the continuous cultivation of high-quality wheat with excellent comprehensive traits.

The LD within a population forms the basis of GWAS, but it is often influenced by mutations, genomic rearrangements,

genetic draft, population stratification, and natural selection (Al-Maskri et al., 2012). Population stratification is considered to be the major factor affecting the accuracy of GWAS results (Cardon and Palmer, 2003). In the current study, the ADMIXTURE and principal component analysis produced the same results regarding the population structure, with the 253 accessions divided into four subgroups (Figure 1). Because wheat is a self-pollinating species with a very large genome, its LD decay distance is larger than that of other plants, including maize, rice, and *Arabidopsis* (Andersen et al., 2007; Kim et al., 2007; Mather et al., 2007; Al-Maskri et al., 2012). Moreover, LD decay can vary among the mapping populations of the same species. In this study, LD decay was greatest in the D genome, whereas Ain et al. (2015) and Sun et al. (2017) reported that LD decay was greatest in the A and B genomes, respectively.

LD decay distance is generally used to determine candidate intervals for the target QTL/genes. The LD decay distance shown in this study are similar to those reported in previous studies (Jin et al., 2020; Wu et al., 2021). It is important to note, however, this result should be interpreted with caution. Due to the large variation in LD between chromosomes and within chromosomes, and the diversity of methods for determining LD decay distance thresholds, the LD decay distance used to specify candidate intervals is only an approximation, and also that candidate genes in some cases may have not been captured within the identified intervals.

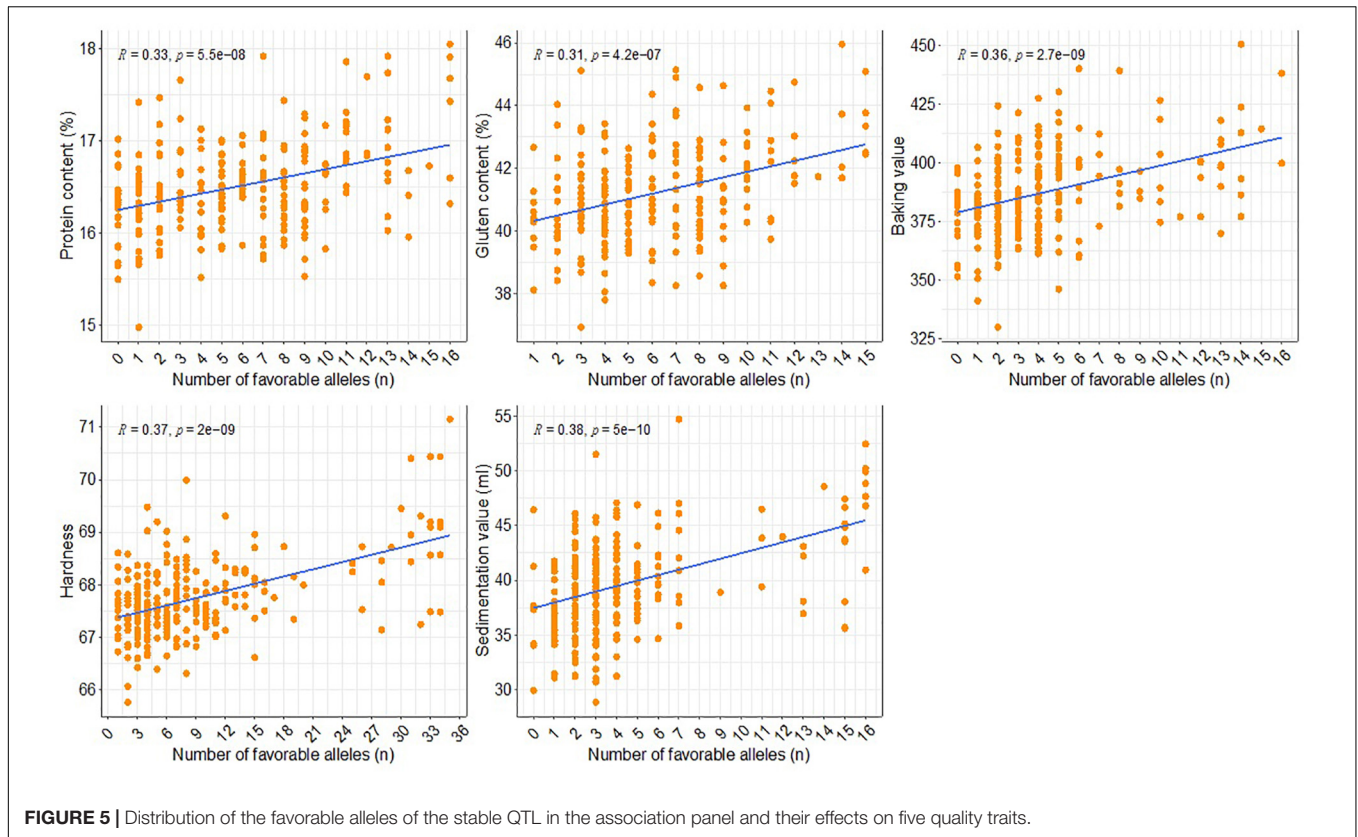
In our GWAS, 556 significant SNPs in 103 QTL were associated with the five examined quality traits. Because genetic analyses performed to identify the genomic regions corresponding to quantitative traits are often influenced by the genotype \times environment interaction, stability across diverse environments is an important parameter for assessing the value and reliability of a particular locus. We focused on 32 QTL identified in at least two environments. These QTL are located on 12 of the 21 chromosomes in wheat (Table 3). We compared our results with those of previous investigations. For PC and GC, a common QTL (*qPC5A.3/qGC5A*) was detected on chromosome 5A; several QTL on this chromosome have been identified (Sun et al., 2008, 2016; Tsilo et al., 2010; Echeverry-Solarte et al., 2015; Tiwari et al., 2015; Li et al., 2016). Another two QTL (*QGpc.WY-1A.1/QWgc.WY-1A.1* and *QGpc.WY-5A.1/QWgc.WY-5A.2*) associated with PC and GC were detected by Li et al. (2011). These results reflect the high correlation between PC and GC. Li et al. (2011) also identified two QTL (*QWgc.WJ-2B.1a* and *QWgc.WJ-2B.1b*) for GC on chromosome 2B. In the present study, a predicted GC QTL (*qGC2B.2*) was detected between the SNPs *Affx-111472438* and *Affx-109061076* on chromosome 2B. However, as QTL mapping provides only limited resolution, further experiments were required to determine whether the genome regions on 2B associated with GC in the two studies were the same. Regarding BV, relatively few QTL/genes were revealed in earlier studies involving map-based cloning or GWAS. Therefore, the eight stable QTL detected in this study may provide useful insights for future investigations of the genetic basis of BV. For HA, in addition to the *Pina/Pinb* gene on chromosome 5D, dozens of QTL related to this trait were mapped in previous



studies (Arbelbide and Bernardo, 2006; Ma et al., 2010; Heo and Sherman, 2013; Boehm et al., 2018; Sun et al., 2018; Kumar et al., 2019). Several of these QTL are within one LD of a QTL (*qHA1A.1*, *qHA1B.1*, *qHA5D*, and *qHA7A.1*) revealed in the current study. Notably, two pleiotropic loci (*qHA6B.2/qHA6B.3*) affecting HA and BV were detected on chromosome 6B in all tested environments, indicating they may be major loci. Moreover, we detected five stable QTL for SV on four chromosomes (2D, 4B, 6B, and 6D). To the best of our knowledge, there are no published reports describing these QTL. Furthermore, the QTL (*qSV6B.1*) on chromosome 6B contains up to 23 significant SNPs. Hence, it should be more thoroughly examined in future studies to identify potential causal genes.

There is substantial interest in identifying the candidate genes underlying specific traits relevant for plant breeding. To date, researchers have identified many QTL related to wheat quality traits, but there has been little research aimed at identifying putative candidate genes in these genomic regions, with a few exceptions (e.g., Uauy et al., 2006; Chen et al., 2019). The PC is significantly higher in tetraploid wheat cultivars containing *GPC-B1* than in cultivars containing mutant or missing alleles. *Glu-1*, *Glu-3* and *Gli* genes encoding high molecular weight glutenin subunit, low molecular weight glutenin subunit and gliadin, respectively, were confirmed to be directly involved in

the formation of grain proteins and had a great impact on GPC and WGC (Plessis et al., 2013; Wang et al., 2018). *Pina* and *Pinb* gene, encoding the friabilin components puroindoline a and b, are widely known to be major determinants of grain texture in common wheat (Giroux and Morris, 1998; Wang et al., 2018). In this study, candidate genes were predicted in the genomic regions surrounding 32 major and stable QTL (Table 3 and Supplementary Table 4). Our analysis of existing transcriptome data of wheat indicated that 34 genes were specifically highly expressed in grains during reproductive growth (Supplementary Table 5 and Supplementary Figure 4). For example, *TraesCS2D02G382400*, *TraesCS2D02G383400*, and *TraesCS2D02G383500*, encoding vacuolar processing enzymes, were identified as candidate genes for BV and SV. In rice, *Glup3/OsVPE1* (vacuole-processing enzyme) belongs to the cysteine protease family, which processes gluten precursors into acidic and basic subunits in protein storage vacuole (PSV), and thus plays a key role in gluten maturation (Wang et al., 2009; Kumamaru et al., 2010; Ueda et al., 2010). *TraesCS2B02G138100*, which is related to HA, was predicted to encode a Zinc finger CCCH domain-containing protein. In rice, the down-regulated expression of *OsGZF1* (Cys3His1, a zinc finger gene) by RNAi reportedly leads to increased grain nitrogen concentration in transgenic plants (Chen et al., 2014). However, as molecular



mechanism of wheat quality formation is quite complex, the role of genes encoding unknown functional proteins and the actual functions of presumed candidate genes should be carefully verified in the future.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

MY and MX conceived and designed the research. SH, HL, and HW performed the experiments. JS, Baogerile, JT, SM, QP, LY, MW, and MG analyzed the data. SH, HL, NZ, and JD performed the statistical data analysis and visualizations. SH, HL, MY, and MX wrote the original draft. MY, MX, and DL reviewed and edited the manuscript. All authors read and approved the last version.

FUNDING

This work was supported by Science and Technology Project of Autonomous Region: Research and Development of Quality

Genetic Key Technology of High Quality Spring Wheat in Hetao Irrigation area, and Laboratory of Universities of Autonomous Region: Hetao Green Agricultural Product Safety Production and Early Warning Control Laboratory

ACKNOWLEDGMENTS

We thank Liwen Bianji (Edanz) (www.liwenbianji.cn/) for editing the English text of a draft of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.835306/full#supplementary-material>

Supplementary Figure 1 | Chromosomal distribution of the SNPs used for the GWAS.

Supplementary Figure 2 | Average LD decay distance in the genomes of 253 wheat accessions.

Supplementary Figure 3 | Quantile–quantile plots for five quality traits analyzed using three models and BLUP values as well as data from individual environments.

Supplementary Figure 4 | Expression heatmap of 34 key candidate genes based on transcriptome data.

Supplementary Table 1 | Phenotypic datas of five quality traits in 253 accessions for three environments and BLUP values.

Supplementary Table 2 | Significant SNP for five quality traits in different environments identified by GWAS.

Supplementary Table 3 | A total of 103 QTL composed of significant snp detected by GWAS.

Supplementary Table 4 | Annotation genes within the 4 Mb region flanking of the significant SNPs included in stable QTL for five quality traits.

Supplementary Table 5 | Candidate genes specifically and highly expressed in grain at reproductive growth stages based on available transcriptome data.

REFERENCES

- Ain, Q., Rasheed, A., Anwar, A., Mahmood, T., Imtiaz, M., Mahmood, T., et al. (2015). Genome-wide association for grain yield under rainfed conditions in historical wheat cultivars from Pakistan. *Front. Plant Sci.* 6, 743. doi: 10.3389/fpls.2015.00743
- Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. doi: 10.1101/gr.094052.109
- Al-Maskri, A. Y., Sajjad, M., and Khan, S. H. (2012). Association mapping: a step forward to discovering new alleles for crop improvement. *Int. J. Agric. Biol.* 14, 153–160.
- Andersen, J. R., Zein, I., Wenzel, G., Krutzfeldt, B., Eder, J., Ouzunova, M., et al. (2007). High levels of linkage disequilibrium and associations with forage quality at a phenylalanine ammonia-lyase locus in European maize (*Zea mays* L.) inbreds. *Theor. Appl. Genet.* 114, 307–319. doi: 10.1007/s00122-006-0434-8
- Arbelvide, M., and Bernardo, R. (2006). Mixed-model QTL mapping for kernel hardness and dough strength in bread wheat. *Theor. Appl. Genet.* 112, 885–890. doi: 10.1007/s00122-005-0190-1
- Bettgenhauser, J., and Krattinger, S. G. (2019). Rapid gene cloning in cereals. *Theor. Appl. Genet.* 132, 699–711. doi: 10.1007/s00122-018-3210-7
- Boehm, J. D., Ibba, M. I., Kiszonas, A. M., See, D. R., Skinner, D. Z., and Morris, C. F. (2018). Genetic analysis of kernel texture (grain hardness) in a hard red spring wheat (*Triticum aestivum* L.) bi-parental population. *J. Cereal Sci.* 79, 57–65.
- Borrill, P., Ramirez-Gonzalez, R., and Uauy, C. (2016). expVIP: a customizable RNA-seq data analysis and visualization platform. *Plant Physiol.* 170, 2172–2186. doi: 10.1104/pp.15.01667
- Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308
- Cardon, L. R., and Palmer, L. J. (2003). Population stratification and spurious allelic association. *Lancet* 361, 598–604. doi: 10.1016/S0140-6736(03)12520-2
- Chen, J., Zhang, F., Zhao, C., Lv, G., Sun, C., Pan, Y., et al. (2019). Genome-wide association study of six quality traits reveals the association of the TaRPP13L1 gene with flour colour in Chinese bread wheat. *Plant Biotechnol. J.* 17, 2106–2122. doi: 10.1111/pbi.13126
- Chen, Y., Song, W., Xie, X., Wang, Z., Guan, P., Peng, H., et al. (2020). A collinearity-incorporating homology inference strategy for connecting emerging assemblies in the Triticeae tribe as a pilot practice in the plant pangenomic era. *Mol. Plant* 13, 1694–1708. doi: 10.1016/j.molp.2020.09.019
- Chen, Y., Sun, A., Wang, M., Zhu, Z., and Ouwerkerk, P. B. (2014). Functions of the CCCH type zinc finger protein OsGZF1 in regulation of the seed storage protein GluB-1 from rice. *Plant Mol. Biol.* 84, 621–634. doi: 10.1007/s11103-013-0158-5
- Crawford, A. C., Stefanova, K., Lambe, W., McLean, R., Wilson, R., Barclay, I., et al. (2011). Functional relationships of phytoene synthase 1 alleles on chromosome 7A controlling flour colour variation in selected Australian wheat genotypes. *Theor. Appl. Genet.* 123, 95–108. doi: 10.1007/s00122-011-1569-9
- Curtis, T., and Halford, N. G. (2014). Food security: the challenge of increasing wheat yield and the importance of not compromising food safety. *Ann. Appl. Biol.* 164, 354–372. doi: 10.1111/aab.12108
- Echeverry-Solarte, M., Kumar, A., Kianian, S., Simsek, S., Alamri, M. S., Mantovani, E. E., et al. (2015). New QTL alleles for quality-related traits in spring wheat revealed by RIL population derived from supernumerary x non-supernumerary spikelet genotypes. *Theor. Appl. Genet.* 128, 893–912. doi: 10.1007/s00122-015-2478-0
- Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x
- Gao, L., Meng, C. S., Yi, T. F., Xu, K., Cao, H. W., Zhang, S. H., et al. (2021). Genome-wide association study reveals the genetic basis of yield- and quality-related traits in wheat. *BMC Plant Biol.* 21:144. doi: 10.1186/s12870-021-02925-7
- Giroux, M. J., and Morris, C. F. (1998). Wheat grain hardness results from highly conserved mutations in the friabilin components puroindoline a and b. *Proc. Natl. Acad. Sci. U.S.A.* 95, 6262–6266. doi: 10.1073/pnas.95.11.6262
- Goel, S., Singh, K., Singh, B., Grewal, S., Dwivedi, N., Alqarawi, A. A., et al. (2019). Analysis of genetic control and QTL mapping of essential wheat grain quality traits in a recombinant inbred population. *PLoS One* 14:e0200669. doi: 10.1371/journal.pone.0200669
- Hao, C., Dong, Y., Wang, L., You, G., Zhang, H., Ge, H., et al. (2008). Genetic diversity and construction of core collection in Chinese wheat genetic resources. *Sci. Bull.* 53, 1518–1526.
- He, X. Y., He, Z. H., Ma, W., Appels, R., and Xia, X. C. (2009). Allelic variants of phytoene synthase 1 (Psy1) genes in Chinese and CIMMYT wheat cultivars and development of functional markers for flour colour. *Mol. Breed.* 23, 553–563.
- He, X. Y., He, Z. H., Zhang, L. P., Sun, D. J., Morris, C. F., Fuerst, E. P., et al. (2007). Allelic variation of polyphenol oxidase (PPO) genes located on chromosomes 2A and 2D and development of functional markers for the PPO genes in common wheat. *Theor. Appl. Genet.* 115, 47–58. doi: 10.1007/s00122-007-0539-8
- Heo, H., and Sherman, J. (2013). Identification of QTL for grain protein content and grain hardness from winter wheat for genetic improvement of spring wheat. *Plant Breed. Biotechnol.* 1, 347–353.
- Ji, M., Fang, W., Li, W., Zhao, Y., Guo, Y., Wang, W., et al. (2021). Genome wide association study of the whiteness and colour related traits of flour and dough sheets in common wheat. *Sci. Rep.* 11:8790. doi: 10.1038/s41598-021-88241-4
- Jin, J., Duan, S., Qi, Y., Yan, S., Li, W., Li, B., et al. (2020). Identification of a novel genomic region associated with resistance to *Fusarium* crown rot in wheat. *Theor. Appl. Genet.* 133, 2063–2073. doi: 10.1007/s00122-020-03577-1
- Jin, X., Feng, B., Xu, Z., Fan, X., Liu, J., Liu, Q., et al. (2018). TaAAP6-3B, a regulator of grain protein content selected during wheat improvement. *BMC Plant Biol.* 18:71. doi: 10.1186/s12870-018-1280-y
- Kim, S., Plagnol, V., Hu, T. T., Toomajian, C., Clark, R. M., Ossowski, S., et al. (2007). Recombination and linkage disequilibrium in *Arabidopsis thaliana*. *Nat. Genet.* 39, 1151–1155. doi: 10.1038/ng2115
- Kumamaru, T., Uemura, Y., Inoue, Y., Takemoto, Y., Siddiqui, S. U., Ogawa, M., et al. (2010). Vacuolar processing enzyme plays an essential role in the crystalline structure of glutelin in rice seed. *Plant Cell Physiol.* 51, 38–46. doi: 10.1093/pcp/pcp165
- Kumar, A., Jain, S., Elias, E. M., Ibrahim, M., and Sharma, L. K. (2018). “An overview of QTL identification and marker-assisted selection for grain protein content in wheat,” in *Eco-Friendly Agro-Biological Techniques for Enhancing Crop Productivity*, eds R. S. Sengar and A. Singh (Singapore: Springer), 245–274. doi: 10.1007/978-981-10-6934-5_11
- Kumar, A., Kumar, J., Singh, R., Garg, T., Chhuneja, P., Balyan, H. S., et al. (2009). QTL analysis for grain colour and pre-harvest sprouting in bread wheat. *Plant Sci.* 177, 114–122.
- Kumar, A., Mantovani, E. E., Simsek, S., Jain, S., Elias, E. M., and 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 x non-adapted bread wheat cross. *PLoS One* 14:e0221826. doi: 10.1371/journal.pone.0221826
- Li, C., Bai, G., Chao, S., Carver, B., and Wang, Z. (2016). Single nucleotide polymorphisms linked to quantitative trait loci for grain quality traits in wheat. *Crop J.* 4, 1–11. doi: 10.1016/j.cj.2015.10.002
- Li, J., Cui, F., Ding, A.-m., Zhao, C.-h., Wang, X.-q., Wang, L., et al. (2011). QTL detection of seven quality traits in wheat using two related recombinant inbred line populations. *Euphytica* 183, 207–226.

- Liu, Y., Lin, Y., Gao, S., Li, Z., Ma, J., Deng, M., et al. (2017). A genome-wide association study of 23 agronomic traits in Chinese wheat landraces. *Plant J.* 91, 861–873. doi: 10.1111/tpj.13614
- Lou, H., Zhang, R., Liu, Y., Guo, D., Zhai, S., Chen, A., et al. (2021). Genome-wide association study of six quality-related traits in common wheat (*Triticum aestivum* L.) under two sowing conditions. *Theor. Appl. Genet.* 134, 399–418. doi: 10.1007/s00122-020-03704-y
- Luo, X., Wang, B., Gao, S., Zhang, F., Terzaghi, W., and Dai, M. (2019). Genome-wide association study dissects the genetic bases of salt tolerance in maize seedlings. *J. Integr. Plant Biol.* 61, 658–674. doi: 10.1111/jipb.12797
- Ma, D., Yan, J., He, Z., Wu, L., and Xia, X. (2010). Characterization of a cell wall invertase gene TaCwi-A1 on common wheat chromosome 2A and development of functional markers. *Mol. Breed.* 29, 43–52.
- Mather, K. A., Caicedo, A. L., Polato, N. R., Olsen, K. M., McCouch, S., and Purugganan, M. D. (2007). The extent of linkage disequilibrium in rice (*Oryza sativa* L.). *Genetics* 177, 2223–2232. doi: 10.1534/genetics.107.079616
- Meijon, M., Satbhai, S. B., Tsuchimatsu, T., and Busch, W. (2014). Genome-wide association study using cellular traits identifies a new regulator of root development in *Arabidopsis*. *Nat. Genet.* 46, 77–81. doi: 10.1038/ng.2824
- Plessis, A., Ravel, C., Bordes, J., Balfourier, F., and Martre, P. (2013). Association study of wheat grain protein composition reveals that gliadin and glutenin composition are trans-regulated by different chromosome regions. *J. Exp. Bot.* 64, 3627–3644. doi: 10.1093/jxb/ert188
- Proietti, S., Caarls, L., Coolen, S., Van Pelt, J. A., Van Wees, S. C. M., and Pieterse, C. M. J. (2018). Genome-wide association study reveals novel players in defense hormone crosstalk in *Arabidopsis*. *Plant Cell Environ.* 41, 2342–2356. doi: 10.1111/pce.13357
- Rabbi, S., Kumar, A., Mohajeri Naraghi, S., Simsek, S., Sapkota, S., Solanki, S., et al. (2021). Genome-wide association mapping for yield and related traits under drought stressed and non-stressed environments in wheat. *Front. Genet.* 12:649988. doi: 10.3389/fgene.2021.649988
- Ramírez-González, R., Borrill, P., Lang, D., Harrington, S. A., Brinton, J., Venturini, L., et al. (2018). The transcriptional landscape of hexaploid wheat across tissues and cultivars. *Science* 361:eaar6089. doi: 10.1126/science.aar6089
- Semagn, K., Iqbal, M., Chen, H., Perez-Lara, E., Bemister, D. H., Xiang, R., et al. (2021). Physical mapping of QTL associated with agronomic and end-use quality traits in spring wheat under conventional and organic management systems. *Theor. Appl. Genet.* 134, 3699–3719. doi: 10.1007/s00122-021-03923-x
- 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
- Shiferaw, B., Smale, M., Braun, H.-J., Duveiller, E., Reynolds, M., and Muricho, G. (2013). Crops that feed the world 10. Past successes and future challenges to the role played by wheat in global food security. *Food Secur.* 5, 291–317. doi: 10.1007/s12571-013-0263-y
- 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
- Sun, H., Lü, J., Fan, Y., Zhao, Y., Kong, F., Li, R., et al. (2008). Quantitative trait loci (QTLs) for quality traits related to protein and starch in wheat. *Prog. Nat. Sci.* 18, 825–831. doi: 10.1016/j.pnsc.2007.12.013
- Sun, X., Liu, T., Ning, T., Liu, K., Duan, X., Wang, X., et al. (2018). Genetic dissection of wheat kernel hardness using conditional QTL mapping of kernel size and protein-related traits. *Plant Mol. Biol. Rep.* 36, 1–12. doi: 10.1007/s11105-017-1061-z
- Sun, X., Wu, K., Zhao, Y., Qian, Z., Kong, F., Guo, Y., et al. (2016). Molecular genetic analysis of grain protein content and flour whiteness degree using RILs in common wheat. *J. Genet.* 95, 317–324. doi: 10.1007/s12041-016-0639-9
- Tang, W., Ye, J., Yao, X., Zhao, P., Xuan, W., Tian, Y., et al. (2019). Genome-wide associated study identifies NAC42-activated nitrate transporter conferring high nitrogen use efficiency in rice. *Nat. Commun.* 10:5279. doi: 10.1038/s41467-019-13187-1
- Tiwari, C., Wallwork, H., Arun, B., Mishra, V. K., Velu, G., Stangoulis, J., et al. (2015). Molecular mapping of quantitative trait loci for zinc, iron and protein content in the grains of hexaploid wheat. *Euphytica* 207, 563–570. doi: 10.1371/journal.pone.0179851
- Tsilo, T. J., Hareland, G. A., Simsek, S., Chao, S., and Anderson, J. A. (2010). Genome mapping of kernel characteristics in hard red spring wheat breeding lines. *Theor. Appl. Genet.* 121, 717–730. doi: 10.1007/s00122-010-1343-4
- Uauy, C., Distelfeld, A., Fahima, T., Blechl, A., and Dubcovsky, J. (2006). A NAC Gene regulating senescence improves grain protein, zinc, and iron content in wheat. *Science* 314, 1298–1301. doi: 10.1126/science.1133649
- Ueda, Y., Satoh-Cruz, M., Matsusaka, H., Takemoto-Kuno, Y., Fukuda, M., Okita, T. W., et al. (2010). Gene-gene interactions between mutants that accumulate abnormally high amounts of proglutelin in rice seed. *Breed. Sci.* 60, 568–574. doi: 10.1270/jsbbs.60.568
- Wang, D., Zhang, K., Dong, L., Dong, Z., Li, Y., Hussain, A., et al. (2018). Molecular genetic and genomic analysis of wheat milling and end-use traits in China: progress and perspectives. *Crop J.* 6, 68–81. doi: 10.1016/j.cj.2017.10.001
- Wang, X., Wang, H., Liu, S., Ferjani, A., Li, J., Yan, J., et al. (2016). Genetic variation in ZmVPP1 contributes to drought tolerance in maize seedlings. *Nat. Genet.* 48, 1233–1241. doi: 10.1038/ng.3636
- Wang, Y., Zhu, S., Liu, S., Jiang, L., Chen, L., Ren, Y., et al. (2009). The vacuolar processing enzyme OsVPE1 is required for efficient glutelin processing in rice. *Plant J.* 58, 606–617. doi: 10.1111/j.1365-313X.2009.03801.x
- Wu, J., Yu, R., Wang, H., Zhou, C., Huang, S., Jiao, H., et al. (2021). A large-scale genomic association analysis identifies the candidate causal genes conferring stripe rust resistance under multiple field environments. *Plant Biotechnol. J.* 19, 177–191. doi: 10.1111/pbi.13452
- Xie, P., Shi, J., Tang, S., Chen, C., Khan, A., Zhang, F., et al. (2019). Control of Bird feeding behavior by Tannin1 through modulating the biosynthesis of polyphenols and fatty acid-derived volatiles in *Sorghum*. *Mol. Plant* 12, 1315–1324. doi: 10.1016/j.molp.2019.08.004
- Yang, J., Lee, S. H., Goddard, M. E., and Visscher, P. M. (2011). GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* 88, 76–82. doi: 10.1016/j.ajhg.2010.11.011
- Yano, K., Yamamoto, E., Aya, K., Takeuchi, H., Lo, P. C., Hu, L., et al. (2016). Genome-wide association study using whole-genome sequencing rapidly identifies new genes influencing agronomic traits in rice. *Nat. Genet.* 48, 927–934. doi: 10.1038/ng.3596
- Yin, L., Zhang, H., Tang, Z., Xu, J., Yin, D., Zhang, Z., et al. (2021). rMVP: a memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study. *Genomics Proteomics Bioinformatics* doi: 10.1016/j.gpb.2020.10.007
- Yu, J., Pressoir, G., Briggs, W. H., Vroh Bi, I., Yamasaki, M., Doebley, J. F., et al. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* 38, 203–208. doi: 10.1038/ng1702
- Zhu, C., Gore, M., Buckler, E. S., and Yu, J. (2008). Status and prospects of association mapping in plants. *Plant Genome* 1, 5–20.

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.

Copyright © 2022 Hao, Lou, Wang, Shi, Liu, Baogerile, Tao, Miao, Pei, Yu, Wu, Gao, Zhao, Dong, You and Xin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.