



Genome-Wide Association Study of Kernel Traits Using a 35K SNP Array in Bread Wheat (*Triticum aestivum* L.)

Peng Wang^{1,2}, Tian Tian^{1,2}, Jingfu Ma^{1,3}, Yuan Liu^{1,2}, Peipei Zhang¹, Tao Chen^{1,2}, Fahimeh Shahinnia⁴ and Delong Yang^{1,2*}

¹ State Key Laboratory of Aridland Crop Science, Lanzhou, China, ² College of Life Science and Technology, Gansu Agricultural University, Lanzhou, China, ³ College of Agronomy, Gansu Agricultural University, Lanzhou, China, ⁴ Bavarian State Research Center for Agriculture, Institute for Crop Science and Plant Breeding, Freising, Germany

OPEN ACCESS

Edited by:

Xue-Feng Ma,
Forage Genetics International,
United States

Reviewed by:

Vandana Jaiswal,
Institute of Himalayan Bioresource
Technology (CSIR), India
Jindong Liu,
Agricultural Genomics Institute
at Shenzhen (CAAS), China

*Correspondence:

Delong Yang
yangdl@gsau.edu.cn

Specialty section:

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

Received: 27 March 2022

Accepted: 09 May 2022

Published: 06 June 2022

Citation:

Wang P, Tian T, Ma J, Liu Y,
Zhang P, Chen T, Shahinnia F and
Yang D (2022) Genome-Wide
Association Study of Kernel Traits
Using a 35K SNP Array in Bread
Wheat (*Triticum aestivum* L.).
Front. Plant Sci. 13:905660.
doi: 10.3389/fpls.2022.905660

Kernel size and weight are crucial components of grain yield in wheat. Deciphering their genetic basis is essential for improving yield potential in wheat breeding. In this study, five kernel traits, including kernel length (KL), kernel width (KW), kernel diameter ratio (KDR), kernel perimeter (KP), and thousand-kernel weight (TKW), were evaluated in a panel consisting of 198 wheat accessions under six environments. Wheat accessions were genotyped using the 35K SNP iSelect chip array, resulting in a set of 13,228 polymorphic SNP markers that were used for genome-wide association study (GWAS). A total of 146 significant marker-trait associations (MTAs) were identified for five kernel traits on 21 chromosomes [$-\log_{10}(P) \geq 3$], which explained 5.91–15.02% of the phenotypic variation. Of these, 12 stable MTAs were identified in multiple environments, and six superior alleles showed positive effects on KL, KP, and KDR. Four potential candidate genes underlying the associated SNP markers were predicted for encoding ML protein, F-box protein, ethylene-responsive transcription factor, and 1,4- α -glucan branching enzyme. These genes were strongly expressed in grain development at different growth stages. The results will provide new insights into the genetic basis of kernel traits in wheat. The associated SNP markers and predicted candidate genes will facilitate marker-assisted selection in wheat breeding.

Keywords: bread wheat, kernel traits, SNP array, GWAS, candidate genes

INTRODUCTION

Bread wheat (*Triticum aestivum* L., $2n = 6 \times = 42$, BBAADD) is one of the most important cereals worldwide, providing calories and proteins for ~40% of the world population (Rajaram, 2001; Golan et al., 2015). Over the past 40 years, wheat production has increased significantly, but the rate of increase in annual wheat production has tended to decline. Current increases in wheat production are likely to be insufficient to meet the food needs of the global population by 2050 (Ray et al., 2012; Ray et al., 2013). Given the increasing mismatch between acreage and world population, increasing wheat yield has become particularly crucial (Li M. et al., 2015). Thousand kernel weight (TKW), as one of the three essential components of wheat yield, has relatively higher heritability than the number of spikes per plant and the number of grains per spike. Therefore, genetic improvement of TKW is a viable approach to increase wheat yield. Kernel length (KL), kernel width (KW), kernel diameter ratio (KDR), and kernel perimeter (KP) determine kernel size

and affect kernel weight and milling quality (Gegas et al., 2010; Ramya et al., 2010). Therefore, understanding the genetic basis of kernel traits is of great importance for improving wheat yield.

Most agronomic traits of wheat are complex quantitative traits controlled by multiple genes. Their genetic structure is difficult to decipher using conventional methods, so complementing them with molecular approaches is essential (Ayoub et al., 2002; Börner et al., 2002). Quantitative trait locus (QTL) mapping based on linkage analysis has been widely used to analyze the molecular genetic basis of complex quantitative traits. However, this method is only applicable to bi-parental populations and leads to low-resolution mapping of the QTLs and decreases the usefulness of detected QTLs in breeding practices. Alternatively, genome-wide association study (GWAS) provides an effective strategy to identify associations between genotypes and phenotypes. Compared with the former, GWAS can shorten breeding years because segregation population do not need to be established, broader genetic variations can be explored, and much higher accuracy in QTL detection can be achieved (Zhang P. et al., 2021). In recent years, with the development of molecular DNA markers and high-throughput genotyping technology, the single-nucleotide polymorphism (SNP) array has become a powerful tool for marker-assisted selection (MAS). GWAS using SNP chip arrays has been applied to a variety of crops such as rice (*Oryza sativa* L.) (Yang et al., 2014; Yano et al., 2016), maize (*Zea mays* L.) (Liu et al., 2016; Zhu et al., 2018), cotton (*Gossypium* spp.) (Su et al., 2018, 2020a), and other crops. GWAS is also an efficient approach for the genetic analysis of complex quantitative traits in wheat and has been reported for various traits, such as disease resistance (Riaz et al., 2017; Juliana et al., 2018; Mourad et al., 2018; Yang et al., 2019) and quality-related traits (Kristensen et al., 2018; Li et al., 2018; Liu et al., 2018; Chen et al., 2019). Pang et al. (2020) reported 90 marker-trait associations (MTAs) and identified eight putative candidate genes for TKW, KL, KW, and KDR using GWAS with 768 wheat cultivars. In addition, several studies have used the GWAS method to identify MTAs for wheat kernel traits and to determine the underlying candidate genes and function markers (Sun et al., 2017; Li et al., 2019; Alemu et al., 2020; Soumya et al., 2021). However, due to the complexity and large size of the wheat genome, the genetic basis of kernel traits is still poorly understood. In this study, we performed a GWAS for five kernel traits in six environments using the 35K SNP iSelect chip array. Our main objectives were to identify MTAs and search for useful SNP markers and potential candidate genes associated with kernel traits in wheat to be used for marker-assisted selection and genetic improvement of grain size and weight in wheat breeding.

MATERIALS AND METHODS

Plant Materials and Field Conditions

A diverse set of 198 bread wheat accessions was used for this study. Of these, 183 varieties were from eight provinces in China, including 114 from Gansu, 33 from Shanxi, 11 from Hebei, ten from Beijing, six from Shaanxi, six from Shangdong, two from Tianjin, and one from Henan. The remaining 15 cultivars were

from America (**Supplementary Table 1**). All cultivars were grown in six environments with different locations and years at Yuzhong farm station, Lanzhou, Gansu (35°51'N, 104°07'E; altitude 1900 m) in 2015–2016 (E1) and 2017–2018 (E3), and Tongwei farm station, Dingxi, Gansu (35°11'N, 105°19'E; altitude 1750 m) in 2015–2016 (E2), 2017–2018 (E4), 2018–2019 (E5), and 2019–2020 (E6). The two cultivation sites are characterized by a typical arid inland climate in Northwestern China, where the annual average temperature is about 7.0°C, the annual rainfall is less than 400 mm, with nearly 60% falling from July to September, but the annual evaporation capacity is more than 1,500 mm (Miao et al., 2022). A randomized complete block design was conducted with three replications, where the row length of each plot was 1 m and the row spacing was 20 cm, and 60 seeds were sown in each row. Local wheat cultivation practices were considered in field management.

Trait Measurement and Statistical Analysis

At grain maturity, five individual plants of each cultivar were randomly selected for threshing, and incomplete seeds were removed. Kernel length (KL), kernel width (KW), kernel diameter ratio (KDR), kernel perimeter (KP), and thousand kernel weight (TKW) were measured by image analysis using the SC-G software (Hangzhou Wanshen Detection Technology Co., Ltd., Hangzhou, China). All trait measurements were repeated three times.

The best linear unbiased prediction (BLUP) of kernel traits for wheat was calculated by using the R package “lme4.” The broad-sense heritability (h^2_B) was calculated by the following equation (Sukumaran et al., 2015): $h^2_B = \sigma^2_G / (\sigma^2_G + \sigma^2_{G \times E} / l + \sigma^2_e / rl)$, where σ^2_G is the genetic variance, $\sigma^2_{G \times E}$ is the genotype-environment variance, σ^2_e is the residual variance, l is the number of environments, and r is the number of replications. Descriptive statistics, ANOVA, and Pearson's correlation coefficient were calculated using IBM SPSS Statistics V.25.

Single-Nucleotide Polymorphism Genotyping and Population Structure Analysis

A panel of 198 wheat accessions was genotyped by the 35K SNP iSelect chip array (Beijing Zhongyujin Marking Co., Ltd., Beijing, China), which contained 35,143 SNP markers (Allen et al., 2017). The markers without clear physical position information on the chromosomes were removed. The genotyping data were filtered by removing markers with missing values > 20% and minor allele frequency (MAF) < 5%. As a result, 13,228 polymorphic SNP markers were used for subsequent analysis.

The ADMIXTURE V1.3 software was applied to infer population structure based on the filtered polymorphic SNP markers. The genotypes were clustered by assuming the number of groups with K value set as 2–10. The optimal number of clusters was determined according to the cross-validation error (CV error) rate, and the K -value with the minimum CV error rate corresponds to the optimal number of subpopulations

(Alexander et al., 2009). A neighbor-joining (NJ) phylogenetic tree of genotypes was constructed by the PowerMarker V3.25 software (Liu and Muse, 2005).

Genome-Wide Association Analysis

The GWAS analysis was conducted by the TASSEL V5.0 software using the mixed linear model (Q + K, MLM), where population structure (Q) and kinship (K) matrices were applied to avoid spurious associations caused by population structure and control false positives simultaneously. To declare the significant MTAs, P -value after Bonferroni correction was considered with high stringency. Therefore, $P \leq 0.001$ [$-\log_{10}(P) \geq 3$] was regarded as the threshold using a liberal approach to reduce ignoring any significantly associated SNP marker for the kernel traits (Kumar et al., 2020). Manhattan plots and Quantile-Quantile (Q-Q) plots were drawn using the R package “CMplot.”¹

Identification of Candidate Genes

After the significant MTAs were identified in multiple environments, the flanking sequence spanning 1 kb upstream and downstream of the significant SNP position was used to query against the Chinese Spring wheat reference genome IWGSC RefSeq v1.1.² Subsequently, JBrowse was used to examine candidate genes that overlapped with the flanking region of the SNPs (Kumar et al., 2020). Functional annotations

of candidate genes were extracted from WheatGmap³ (Zhang L. et al., 2021). To investigate the function of candidate genes, expression analysis was performed based on the publicly available expression data of a drought-tolerant cultivar Jinmai 47, obtained from roots, stems, leaves, developing spikes, spikes at the anthesis stage, and developing grains at 5, 10, and 20 days after anthesis (DAA). Total RNA was extracted using the E.Z.N.A.® Plant RNA Kit (Omega Bio-Tek, Norcross, GA, United States). Three biological replications were used for quantitative real-time PCR (qRT-PCR) analysis in each treatment. *TaActin* was used as an endogenous reference, and the $2^{-\Delta\Delta C(t)}$ method was used to determine relative gene expression levels. Primer sequences used for qRT-PCR are listed in **Supplementary Table 2**.

RESULTS

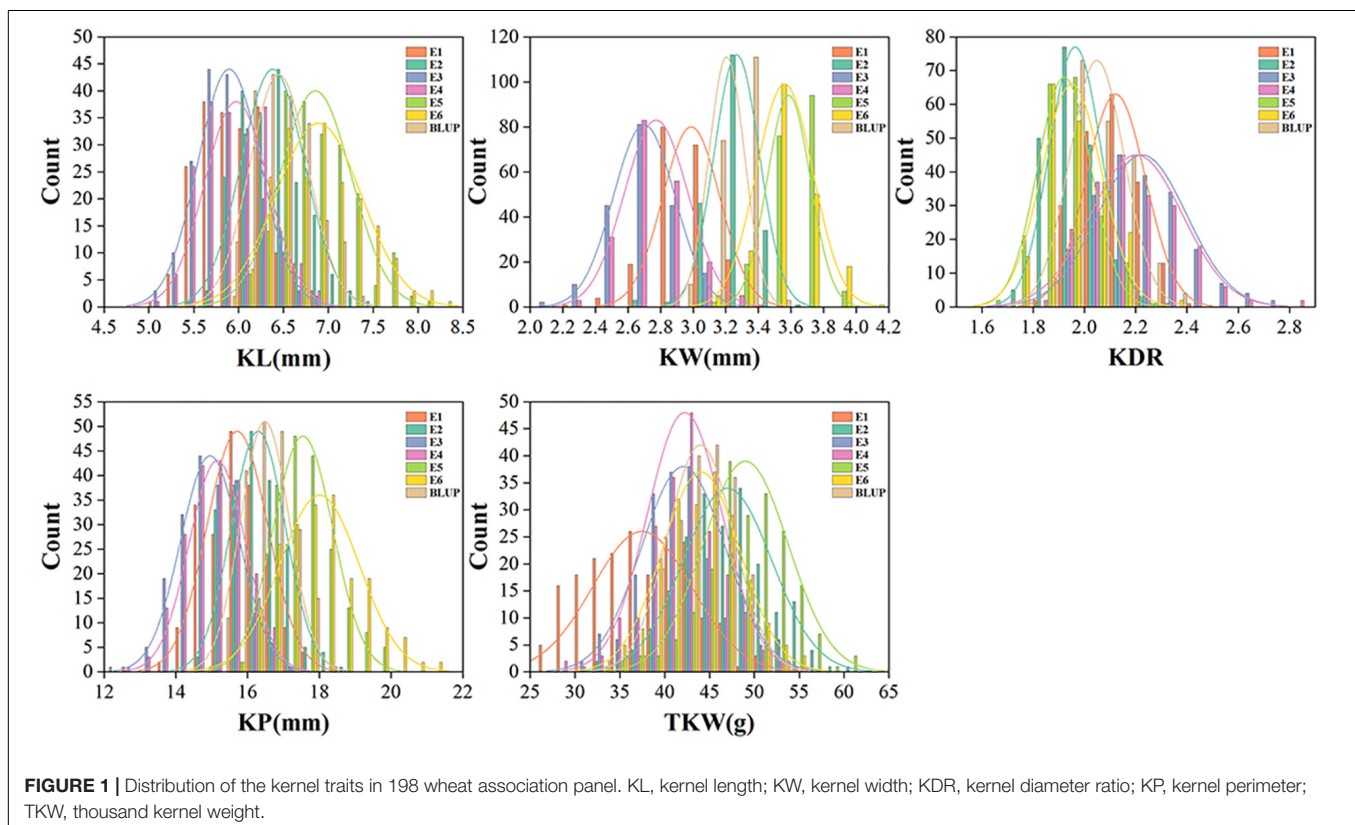
Phenotypic Evaluation

Phenotypic performance of kernel traits showed wide variation and normal distribution in different environments (**Figure 1**). The coefficients of variation for each trait in different environments ranged from 3.44 to 14.79%, indicating broad phenotypic variation among genotypes. The broad-sense heritabilities of KL, KW, KDR, KP, and TKW were moderate to high, ranging from 0.60 to 0.85. In comparison, KW had the

¹<https://github.com/YinLiLin/R-CMplot>

²<http://www.wheatgenome.org/>

³<https://www.wheatgmap.org/>



lowest h^2_B value (0.60), whereas KL had the highest h^2_B value (0.85) (Table 1).

Pearson's correlation coefficients based on BLUP for the kernel traits are shown in Figure 2. All traits were significantly correlated with each other ($P < 0.001$). KL, KW, KDR, and KP showed positive correlations with TKW ($r = 0.24$ – 0.83). KW showed a negative correlation with KDR ($r = -0.37$), whereas the other traits were positively correlated with each other. The highest correlation was between KL and KP ($r = 0.98$).

Marker Distribution and Population Structure

A total of 35,143 SNPs were genotyped for 198 wheat accessions. After filtering, 13,228 SNPs were polymorphic and were used

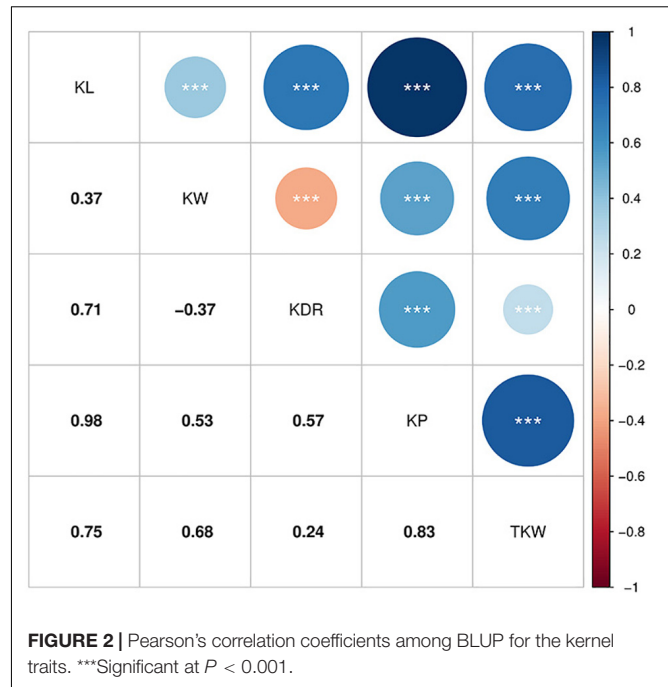
TABLE 1 | Descriptive statistics and heritability (h^2_B) of five kernel traits.

Trait	Environment	Min	Max	Mean	SD	CV ^a (%)	F-test ^b	h^2_B ^c
KL (mm)	E1	5.22	7.36	6.28	0.38	5.99	34.77**	0.85
	E2	5.47	7.48	6.38	0.36	5.63	61.53**	
	E3	5.02	7.00	5.90	0.37	6.23	19.95**	
	E4	5.14	7.11	5.97	0.36	6.06	35.39**	
	E5	6.06	8.06	6.86	0.40	5.84	219.60**	
	E6	5.92	8.36	6.89	0.48	6.95	340.49**	
	BLUP	5.80	7.33	6.45	0.33	5.05	231.37**	
KW (mm)	E1	2.29	3.47	2.99	0.18	6.18	20.03**	0.60
	E2	2.64	3.63	3.27	0.15	4.44	22.37**	
	E3	2.01	3.29	2.70	0.21	7.63	10.53**	
	E4	2.23	3.42	2.77	0.21	7.40	21.13**	
	E5	3.14	3.87	3.59	0.13	3.73	72.20**	
	E6	3.00	4.04	3.56	0.18	4.92	134.71**	
	BLUP	2.84	3.44	3.21	0.11	3.44	58.96**	
KDR	E1	1.85	2.49	2.12	0.12	5.54	59.42**	0.80
	E2	1.72	2.35	1.96	0.10	5.19	75.06**	
	E3	1.87	2.94	2.23	0.18	8.04	32.97**	
	E4	1.83	2.92	2.20	0.19	8.46	25.41**	
	E5	1.69	2.23	1.92	0.11	5.47	208.48**	
	E6	1.69	2.35	1.95	0.12	6.04	220.81**	
	BLUP	1.84	2.35	2.05	0.11	5.22	153.76**	
KP (mm)	E1	12.32	17.95	15.71	0.87	5.54	35.93**	0.79
	E2	14.76	18.56	16.30	0.77	4.70	54.31**	
	E3	11.78	17.30	14.96	0.90	5.99	10.93**	
	E4	12.94	17.39	15.16	0.83	5.47	19.17**	
	E5	15.75	19.97	17.54	0.84	4.78	153.79**	
	E6	15.53	21.27	18.01	1.11	6.17	197.88**	
	BLUP	14.88	18.26	16.48	0.68	4.12	133.44**	
TKW (g)	E1	26.80	51.85	37.49	5.55	14.79	258.23**	0.80
	E2	32.95	62.19	47.01	4.97	10.57	144.18**	
	E3	29.87	56.12	42.09	4.70	11.17	170.71**	
	E4	28.72	53.51	42.24	4.24	10.03	103.50**	
	E5	34.55	61.72	49.02	4.78	9.76	128.47**	
	E6	32.40	55.71	44.05	4.38	9.94	168.24**	
	BLUP	32.88	54.14	44.00	3.70	8.40	515.26**	

^aCV, coefficient of variation.

^b** $P \leq 0.01$.

^c h^2_B , the broad sense heritability.



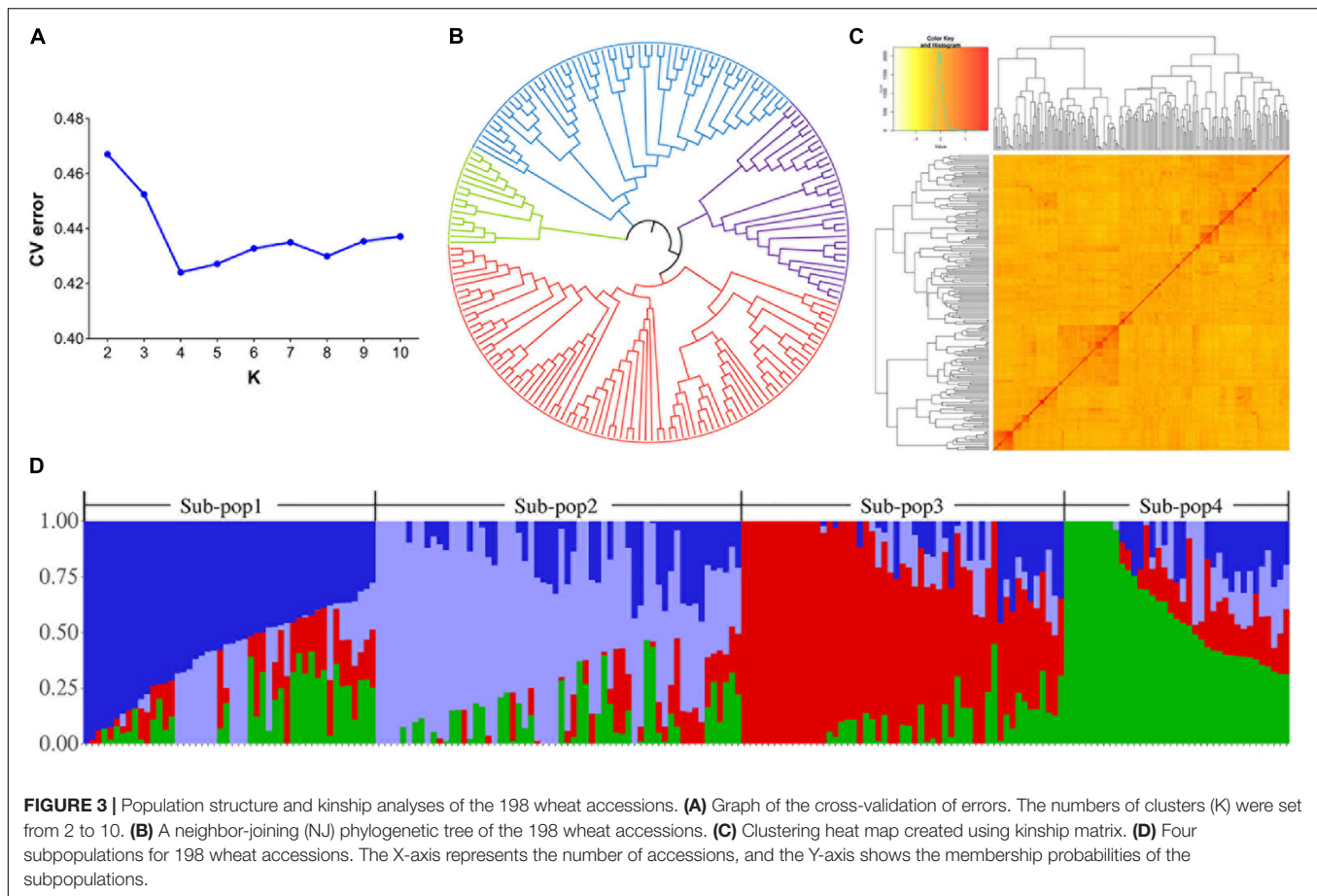
for subsequent analysis (Supplementary Figure 1). Of these, 4,704, 5,408, and 3,116 SNPs were observed in the A, B, and D sub-genomes, respectively. On average, 1,890 SNP markers were distributed per homeologous group. The smallest number of SNPs was found in the homeologous group IV (1,070 SNPs), and the largest number of SNPs was found in the homeologous group II (2,519 SNPs). The distribution of SNP markers varied greatly across 21 chromosomes, with the fewest SNPs on chromosome 4D (180 SNPs) and the most on chromosome 2B (980 SNPs).

The inferring population structure showed a minimum CV error rate value at $K = 4$, indicating that the association population was structured into four subpopulations. Each of them contained 48, 60, 44, and 46 accessions, respectively (Figures 3A,D and Supplementary Table 1). Genetic clustering with the kinship matrix and the NJ phylogenetic tree also showed that the association population could be clustered into four subpopulations (Figures 3B,C). Thus, it was suitable to divide the association population into four subpopulations.

Marker-Trait Associations

A total of 146 significant MTAs ($P \leq 0.001$) for the kernel traits were identified on all 21 chromosomes, explaining phenotypic variation (R^2) ranging from 5.91 to 15.02% (Supplementary Table 3 and Figure 4). The overall rate of false-positive association results was properly controlled, as shown by quantile-quantile plots (Supplementary Figure 2). Of these, 22, 43, 36, 23, and 22 MTAs were identified for KL, KW, KDR, KP, and TKW, respectively, in all environments (Supplementary Figure 3).

In total, twelve MTAs were repeatedly observed in multiple environments, with R^2 ranging from 5.91 to 15.02%. Therefore, these MTAs were considered stable loci. Of these, seven MTAs were repeatedly observed in two environments, four MTAs



were repeatedly observed in three environments, and one MTA was repeatedly observed in all environments except E1 and E6. In addition, three SNP markers involved in six MTAs were associated with multiple kernel traits. Two markers, *AX-94400331* and *AX-94509671*, were both significantly associated with KL and KP, and *AX-94393836* was significantly associated with KW and TKW (Table 2).

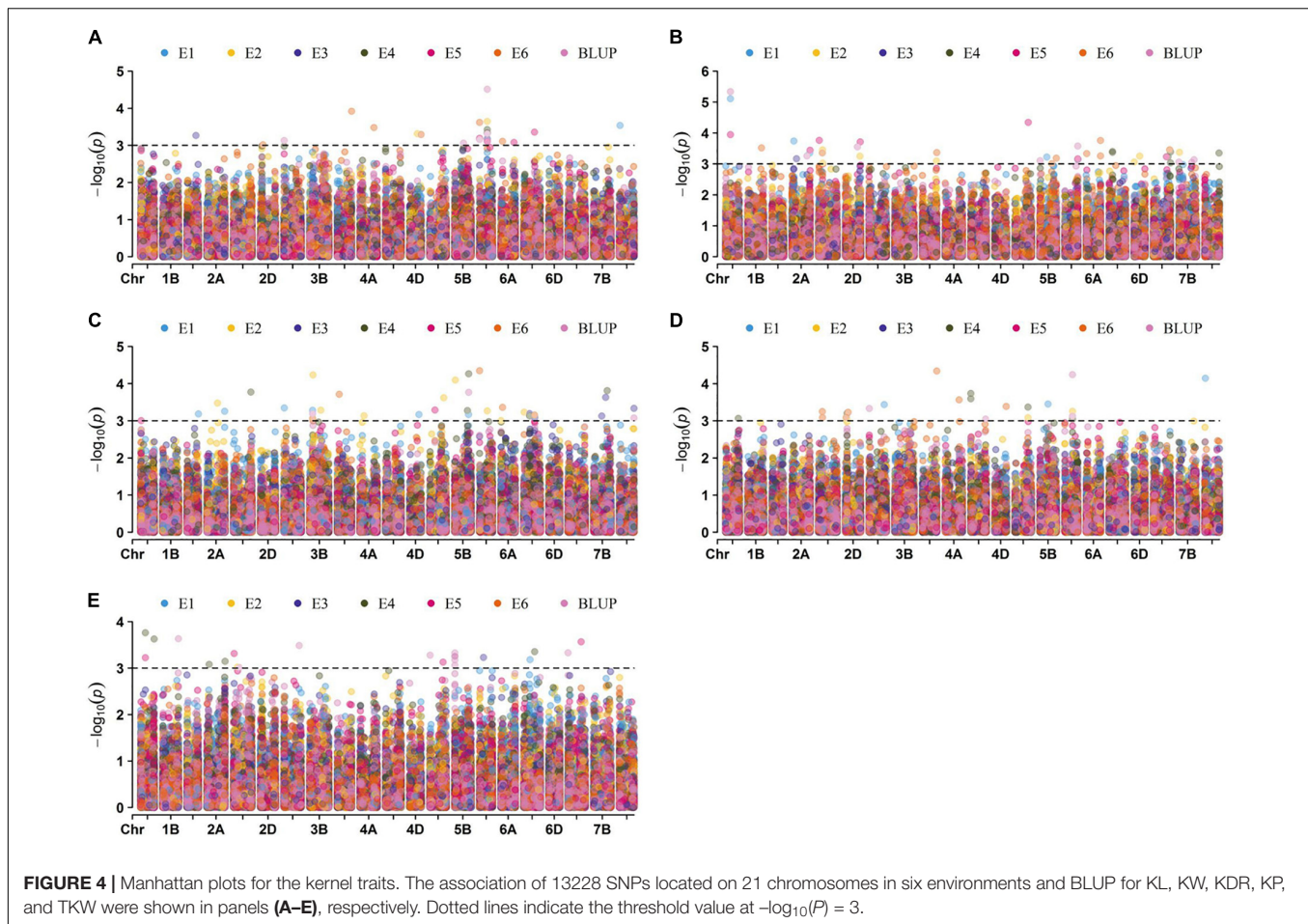
Allelic Effects of Associated Markers on Kernel Traits

The allelic effect was estimated from the MTAs repeatedly identified in multiple environments. Significant differences between SNP alleles were found for six MTAs except *AX-95629937* (Figure 5). The alleles of *AX-94400331* and *AX-94509671* showed positive additive effects, increasing both KL and KP. For *AX-94400331*, the allelic effect on mean KL between the superior allele TT and the inferior allele CC was 0.33 mm, and the corresponding effect on mean KP was 0.70 mm. For *AX-94509671*, the allelic effect on mean KL between the superior allele CC and the inferior allele AA was 0.32 mm, and the corresponding effect on mean KP was 0.64 mm. In addition, the mean KDR of the superior alleles (GG and AA) was slightly higher than that of the inferior alleles (AA and GG) for both markers *AX-94409249* and *AX-95629937*.

The average values for phenotypes of wheat accessions carrying different numbers of superior alleles (0–2) are shown in Figure 6. Under different environments, the average KL with two superior alleles was 6.65 mm (6.24–7.20 mm). In contrast, the average KL with one superior allele and neither superior allele was 6.61 mm (6.08–7.40 mm) and 6.32 mm (5.86–6.84 mm), respectively. For KDR, the mean score with two superior alleles was 2.18 (2.03–2.45), and the mean scores with one superior allele and neither superior allele were 2.09 (1.94–2.26) and 2.03 (1.90–2.18), respectively. In addition, the average KP with different superior alleles showed the same result as KL, indicating that the phenotypic values for these kernel traits are higher when they contain more superior alleles.

Candidate Genes for Kernel Traits

By using the flanking sequence alignment of significant SNP markers from 12 stable MTAs, nine candidate genes for the kernel traits were identified, controlling diverse functional groups of proteins (Table 3). Functional annotations showed that four KDR-associated candidate genes, namely, *TraesCS3B02G139500* on chromosome 3B, *TraesCS5B02G428200* on chromosome 5B, *TraesCS6B02G323800LC* on chromosome 6B, and *TraesCS7D02G535400* on chromosome 7D. Those genes encoded a rotundifolia-like protein, an MD-2-related lipid



recognition domain-containing protein, an F-box protein, and a 1,4- α -glucan branching enzyme, respectively. On chromosome 5D, two candidate genes, namely, *TraesCS5D02G011900* and *TraesCS5D02G657200LC*, correlated with KL and KW, and they encoded hydroxysteroid dehydrogenase and chromosomal replication initiator protein DnaA, respectively. Of the other three candidate genes, which were involved in two different traits, the gene *TraesCS5D02G223200* on chromosome 5D for KL and KP encoded serine protease HtrA-like and the gene *TraesCS5D02G223700* on chromosome 5D for KL and KP encoded ethylene-responsive transcription factor. *TraesCS1A02G135300* on chromosome 1A for KW and TKW encoded a protein of the superfamily of nucleoside triphosphate hydrolases with P-loop. The expression levels of the candidate genes were examined in various tissues, including roots, stems, leaves, developing spikes, spikes at anthesis, and developing grains at 5, 10, and 20 DAA. Among the nine candidate genes identified, the CDS sequences of *TraesCS5D02G657200LC* and *TraesCS5B02G428200* were short, and it was difficult to design suitable specific primers. Therefore, we only presented the qRT-PCR results for the other seven candidate genes (Figure 7). The relative expression of each gene was different at different growth stages. Among them, both *TraesCS5D02G223700* and *TraesCS7D02G535400* were strongly expressed in the developing

grains at different growth stages, with the highest expression level shown at 10 DAA. Both *TraesCS5B02G428200* and *TraesCS6B02G323800LC* were strongly expressed in developing grains at 20 DAA. This suggests that the above four genes may play a key role in regulating grain development.

DISCUSSION

Phenotypic Variation and Population Structure

We used image analysis technology for the phenotyping of kernel traits, which drastically reduces artificial errors and ensures the accuracy of measurement results. The phenotypic distribution of the kernel traits showed normal distribution of the traits, indicating that these traits are quantitative traits and suitable for GWAS analysis. Large phenotypic variation was observed in all target traits, of which the largest variation was observed in TKW with coefficients of variation ranging from 8.40 to 14.79% in different environments (Figure 1). In addition, TKW showed a higher h^2_B of 0.80 (Table 1), which is consistent with previous studies (Li Q. et al., 2015; Su et al., 2018). The h^2_B of other kernel traits was moderate to high, ranging from 0.60 to 0.85. This indicated that the traits were largely controlled

TABLE 2 | Significant marker-trait associations for kernel traits under multiple environments.

Trait	SNP marker	Allele ^a	Chr ^b	Position (bp)	Environment	P-value	R ^{2c} (%)
KL	AX-95248961	A/ <u>T</u>	5D	6450092	E5	6.41E-04	7.91
					E6	2.40E-04	8.91
					BLUP	6.88E-04	7.80
	AX-94400331	C/ <u>T</u>	5D	330891375	E4	3.77E-04	8.76
					BLUP	4.81E-04	8.44
					E2	2.25E-04	8.99
	AX-94509671	A/ <u>C</u>	5D	331402475	E3	6.67E-04	7.69
					E4	4.57E-04	8.20
					E5	7.87E-04	7.57
BLUP					3.08E-05	11.23	
E1					7.78E-06	14.68	
E5					1.13E-04	10.62	
KW	AX-94393836	<u>C</u> /T	1A	208219190	BLUP	4.59E-06	15.02
					E3	7.14E-04	6.14
					E5	6.86E-04	6.36
	AX-94438072	<u>C</u> /T	5D	561626204	BLUP	2.63E-04	7.35
					E3	7.14E-04	6.14
					E5	6.86E-04	6.36
KDR	AX-94409249	A/ <u>G</u>	3B	126000659	E1	5.21E-04	8.13
					E6	8.73E-04	7.81
					BLUP	6.46E-04	8.19
	AX-95629937	A/ <u>G</u>	5B	604022658	E4	5.44E-05	11.90
					BLUP	1.72E-04	10.38
					E3	6.42E-04	6.32
	AX-94711022	A/ <u>G</u>	6B	252383444	E4	7.36E-04	6.06
					E3	4.63E-04	6.58
					BLUP	8.40E-04	5.91
AX-94578940	A/ <u>C</u>	7D	627323708	E3	4.63E-04	6.58	
				E3	4.63E-04	6.58	
				BLUP	8.40E-04	5.91	
KP	AX-94400331	C/ <u>T</u>	5D	330891375	E4	9.09E-04	7.75
					BLUP	7.48E-04	7.89
					E2	5.55E-04	7.97
	AX-94509671	A/ <u>C</u>	5D	331402475	BLUP	5.72E-05	10.49
					E4	1.73E-04	10.66
					E5	5.95E-04	8.68
TKW	AX-94393836	<u>C</u> /T	1A	208219190	E4	1.73E-04	10.66
					E5	5.95E-04	8.68
					E5	5.95E-04	8.68

^aSuperior alleles are underlined.

^bChr, chromosome.

^cPhenotypic variance explained by the MTAs.

by genetic factors. Correlation analysis indicated that KL, KW, KDR, and KP contributed to TKW (Figure 2), suggesting that kernel size traits should be targeted for improving kernel weight and yield in wheat.

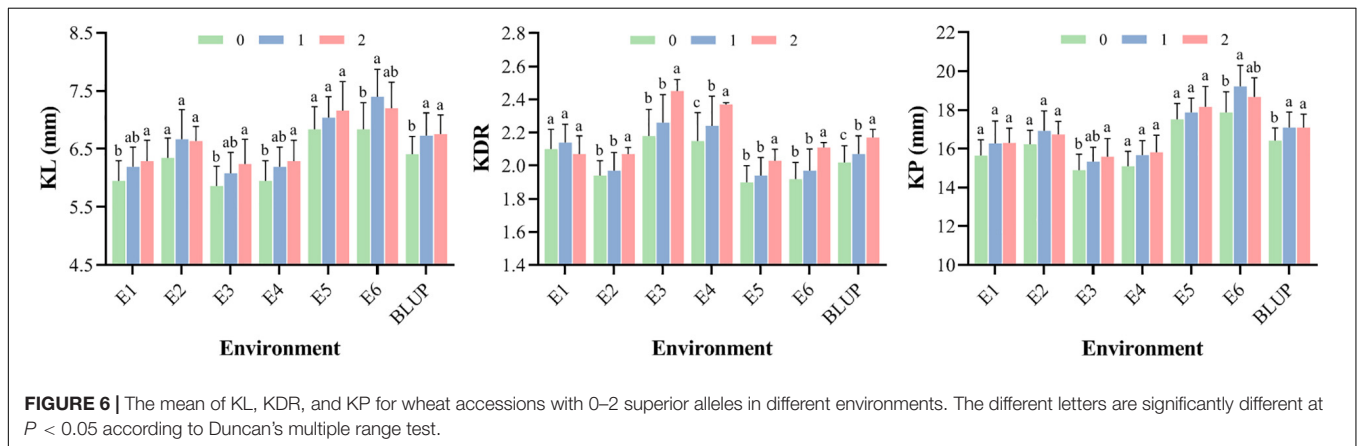
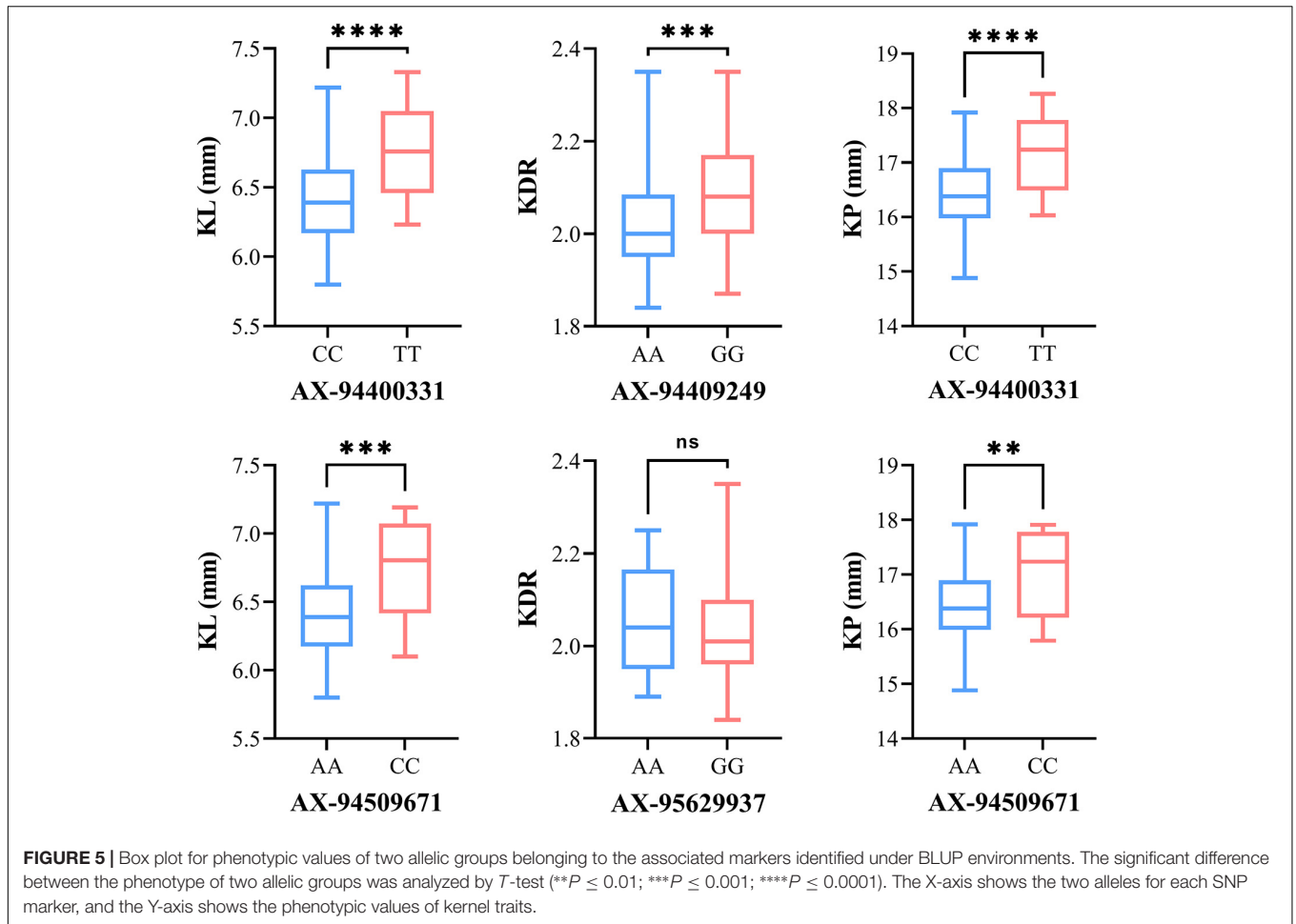
Since the rich germplasm resources contain substantial genetic variations and superior alleles (Walsh, 1981), we collected a diverse set of bread wheat accessions from eight different ecological regions in China and America that provides the basis for identifying MTAs. To reduce spurious associations between markers and traits, it was necessary to evaluate the genetic structure of the population before performing GWAS. The presence of four subpopulations in this associated population was confirmed by population structure and diversity analyses (Figure 3).

Stable Marker-Trait Associations

To ensure the accuracy of the association results, we used the MLM model with population structure and kinship. This approach has been widely used in previous studies and was able

to efficiently control false positives (Ain et al., 2015; Zuo et al., 2019; Alemu et al., 2020). As a result, a total of 146 significant MTAs were identified for the five kernel traits (Supplementary Table 2). Of these, 12 MTAs were repeatedly detected in multiple environments (Table 2), which were considered stable MTAs. Accordingly, nine SNP markers were involved in these MTAs. Three SNP markers of them were simultaneously correlated with multiple traits. For example, two SNP markers, namely, AX-94400331 and AX-94509671 were significantly associated with KL and KP, and another marker, namely, AX-94393836 was significantly associated with KW and TKW.

Some studies have performed QTL mapping for wheat kernel traits using different molecular markers (Ramya et al., 2010; Okamoto et al., 2013; Cui et al., 2014). In this study, 12 stable MTAs were identified and compared with earlier studies while taking into account the correlations among traits, of which five MTAs were co-localized with QTLs for kernel traits previously reported. The SNP marker AX-94393836 associated with KW and TKW on chromosome 1A was mapped within the confidence



interval (*wmc24~wmc278*) of a TKW QTL (*QTkw.ncl-1A.1*) reported by Ramya et al. (2010) and co-localized with a QTL (*QTkw-1A.1*) for TKW reported by Cui et al. (2014). Both AX-95629937 and AX-94578940 that were associated with KDR co-localized with *QTkw-5B.1* on chromosome 5B for KL and *QTkw-7D.2* on chromosome 7D for TKW, respectively, as reported by Cui et al. (2014). In addition, the KL locus AX-95248961 on chromosome 5D was near to a QTL (*QTkw-7D.2*) for hundred

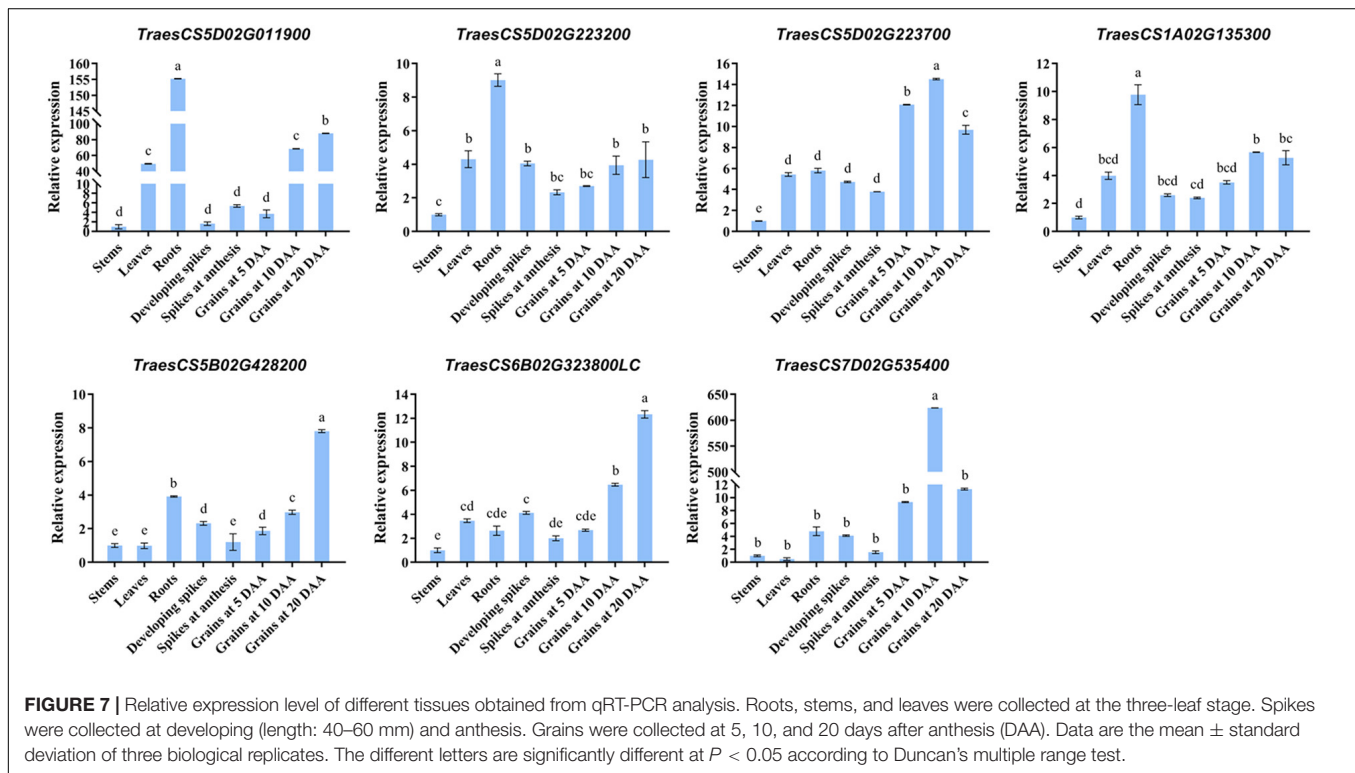
grain weight reported in Okamoto et al. (2013). Notably, the remaining seven MTAs were not reported previously and are likely novel loci for kernel traits.

Discovery of Superior Alleles

In this study, we identified 12 stable MTAs in multiple environments based on the GWAS results. The superior alleles identified from these MTAs were of greater importance.

TABLE 3 | Candidate genes underlying the significant markers associated with kernel traits.

Trait	SNP marker	Chr	Position (bp)	Gene position (bp)	Gene ID	Functional annotation
KL	AX-95248961	5D	6450092	6444885–6450213	<i>TraesCS5D02G011900</i>	Hydroxysteroid dehydrogenase, putative
	AX-94400331	5D	330891375	330889652–330899414	<i>TraesCS5D02G223200</i>	Serine protease HtrA-like
	AX-94509671	5D	331402475	331402389–331406349	<i>TraesCS5D02G223700</i>	Ethylene-responsive transcription factor
KW	AX-94393836	1A	208219190	208208304–208219356	<i>TraesCS1A02G135300</i>	P-loop containing nucleoside triphosphate hydrolases superfamily protein
	AX-94438072	5D	561626204	561625578–561630640	<i>TraesCS5D02G657200LC</i>	Chromosomal replication initiator protein DnaA
KDR	AX-94409249	3B	126000659	125995818–125996030	<i>TraesCS3B02G139500</i>	Rotundifolia-like protein
	AX-95629937	5B	604022658	604021148–604023499	<i>TraesCS5B02G428200</i>	MD-2-related lipid recognition domain-containing protein/ML domain-containing protein
	AX-94711022	6B	252383444	252385913–252388254	<i>TraesCS6B02G323800LC</i>	F-box protein (DUF295)
KP	AX-94578940	7D	627323708	627322122–627328145	<i>TraesCS7D02G535400</i>	1,4- α -glucan branching enzyme
	AX-94400331	5D	330891375	330889652–330899414	<i>TraesCS5D02G223200</i>	Serine protease HtrA-like
	AX-94509671	5D	331402475	331402389–331406349	<i>TraesCS5D02G223700</i>	Ethylene-responsive transcription factor
TKW	AX-94393836	1A	208219190	208208304–208219356	<i>TraesCS1A02G135300</i>	P-loop containing nucleoside triphosphate hydrolases superfamily protein



Therefore, we compared the average BLUP value of phenotype regulated by superior and inferior alleles based on six MTAs, and all superior alleles showed positive effects on KL, KP, and KDR. In particular, the superior allele TT for AX-94400331 and CC for AX-94509671 simultaneously showed highly positive effects on KL and KP (Figure 6).

Exploring superior alleles is very valuable for wheat breeding programs. Although the contribution of a single marker in influencing phenotypic variation might be small, the combination of superior alleles from different markers can have much larger effects in a single variety (Li et al., 2020). In our study, phenotype scores for KL, KDR, and KP were

found to be positively correlated with the number of superior alleles, suggesting that pyramiding superior alleles improved wheat kernel traits performance. With the development of sequencing technology, allelic pyramiding becomes more feasible and powerful.

Potential Candidate Genes for the Kernel Traits

In wheat, several genes have been identified that are related to kernel size and weight. For example, *TaCYP78A3*, which encodes cytochrome P450 CYP78A3, has a significant effect on kernel size by affecting seed coat cell number, and its silencing can

cause an 11% decrease in wheat kernel size (Ma et al., 2015). In addition, *TaGW2* encodes E3 RING ligase, which affects thousand kernel weight by influencing kernel length and width (Simmonds et al., 2016; Zhang et al., 2018). In this study, a total of nine candidate genes were functionally annotated based on the Chinese Spring wheat reference genome (IWGSC RefSeq v1.1) (Table 3). Using qRT-PCR analysis, four promising genes were strongly expressed in developing grain at different growth stages (Figure 7), which were considered as potential candidate genes to control kernel traits.

We predicted four potential candidate genes for kernel traits. The marker AX-95629937 was associated with KDR in E4 and BLUP, with a potential candidate gene *TraesCS5B02G428200* on chromosome 5B encoding the MD-2-related lipid recognition/ML domain-containing protein. The MD-2-related lipid recognition domain is defined as an ML protein (Inohara and Nuñez, 2002). Hakenjos et al. (2013) identified ML3 as a NEDD8- and ubiquitin-modified protein in *Arabidopsis* (*Arabidopsis thaliana*), which was also conjugated to ubiquitin pathway with a critical role in controlling seed size in plants (Li and Li, 2014).

The gene *TraesCS6B02G323800LC* on chromosome 6B underlying the marker AX-94711022 was associated with KDR in E3 and E4 environments. The gene encoded the F-box protein, which is involved in diverse hormone signal transduction and cellular processes (Kuroda et al., 2002). Li et al. (2011) found that *LARGER PANICLE* (*LP*), which encodes an F-box protein and is involved in the regulation of cytokinin levels, can improve rice yield by regulating panicle structure.

The marker AX-94509671 was associated with KL in all environments except E1 and E6 and with KP in E2 and BLUP. A potential candidate gene *TraesCS5D02G223700* encoding ethylene-responsive transcription factor (ERF) was found near the physical map position of AX-94509671 on chromosome 5D. ERF is a plant-specific transcription factor involved in plant growth and development processes (Jin and Liu, 2008). Xu et al. (2016) reported that *OsERF* interacts with a transcription factor (*OsNF-YB1*) specifically expressed in the aleurone layer of the endosperm. This process affected kernel filling and endosperm development, regulating rice kernel size and weight. Li et al. (2017) found that the transcription factor *ZmEREB94* regulates several starch-synthetic genes, further affecting endosperm development.

The marker AX-94578940 associated with KDR in the E3 environment and BLUP corresponded to a candidate gene *TraesCS7D02G535400* (7D) encoding the 1,4- α -glucan branching enzyme (GBE). GBE is the key enzyme that catalyzes the formation of α -1,6-linked branch in starch and belongs to the glycoside hydrolase 13 (GH13) family (Xia et al., 2021). In tapioca and maize, GBE treatment led to molecular restructuring of starch, resulting in a change in the amylose and amylopectin content of starch. This suggests that GBE is important for starch modification (Li et al., 2016a,b; Ren et al., 2018). We speculated that the above four genes might be very important in influencing kernel size and weight in wheat.

CONCLUSION

In this study, GWAS for kernel traits was performed on a diverse panel of 198 bread wheat accessions, genotyped with the 35K SNP iSelect chip array. As a result, 146 significant MTAs were identified for five kernel traits. Of these, 12 stable significant MTAs were identified in different environments. Six superior alleles of nine major SNP markers showed positive effects on KL, KP, and KDR. Four potential candidate genes were predicted to be highly expressed in the developing grain at different growth stages. The results provide not only new insights into the genetic basis of kernel traits in wheat but also diagnostic makers that can be used for marker-assisted selection in wheat breeding.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

PW and DY conceived the study. PW, TT, and JM performed the phenotypic data measurement and data analysis. YL, PZ, TC, and FS provided scientific feedback and revised the content. PW wrote the first draft of the manuscript. DY and FS revised and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

FUNDING

This study was financially supported by the Research Program Sponsored by the State Key Laboratory of Aridland Crop Science, China (GHSJ 2020-Z4); the Key Research and Development Program of Gansu Province, China (21YF5NA089); and the National Natural Science Foundation of China (31760385).

ACKNOWLEDGMENTS

We thank Prof. Ruilian Jing (Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing, China) and Prof. Guohong Zhang (Institute of Dryland Agriculture, Gansu Academy of Agricultural Science, Lanzhou, China) for providing the wheat accessions.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2022.905660/full#supplementary-material>

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.* 6:743–757. doi: 10.3389/fpls.2015.00743
- Alemu, A., Feyissa, T., Tuberosa, R., Maccaferri, M., Sciara, G., Letta, T., et al. (2020). Genome-wide association mapping for grain shape and color traits in Ethiopian durum wheat (*Triticum turgidum* ssp. durum). *Crop J.* 8, 757–768. doi: 10.1016/j.cj.2020.01.001
- Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. doi: 10.1161/01.ATV.0000137190.63214.c5
- Allen, A. M., Winfield, M. O., Burrridge, A. J., Downie, R. C., Benbow, H. R., Barker, G. L. A., et al. (2017). Characterization of a wheat breeders' array suitable for high-throughput SNP genotyping of global accessions of hexaploid bread wheat (*Triticum aestivum*). *Plant Biotechnol. J.* 15, 390–401. doi: 10.1111/pbi.12635
- Ayoub, M., Symons, S., Edney, M., and Mather, D. (2002). QTLs affecting kernel size and shape in a two-rowed by six-rowed barley cross. *Theor. Appl. Genet.* 105, 237–247. doi: 10.1007/s00122-002-0941-1
- Börner, A., Schumann, E., Fürste, A., Cöster, H., Leithold, B., Röder, M., et al. (2002). Mapping of quantitative trait loci determining agronomic important characters in hexaploid wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 105, 921–936. doi: 10.1007/s00122-002-0994-1
- 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.* 11, 2106–2122. doi: 10.1111/pbi.13126
- Cui, F., Zhao, C., Ding, A., Li, J., Wang, L., Li, X., et al. (2014). Construction of an integrative linkage map and QTL mapping of grain yield-related traits using three related wheat RIL populations. *Theor. Appl. Genet.* 127, 659–675. doi: 10.1007/s00122-013-2249-8
- Gegas, V. C., Nazari, A., Griffiths, S., Simmonds, J., Fish, L., Orford, S., et al. (2010). A genetic framework for grain size and shape variation in wheat. *Plant Cell.* 22, 1046–1056. doi: 10.1105/tpc.110.074153
- Golan, G., Oksenberg, A., and Peleg, Z. (2015). Genetic evidence for differential selection of grain and embryo weight during wheat evolution under domestication. *J. Exp. Bot.* 66, 5703–5711. doi: 10.1093/jxb/erv249
- Hakenjos, J. P., Bejai, S., Ranftl, Q., Behringer, C., Vlot, A. C., Absmanner, B., et al. (2013). ML3 is a NEDD8- and ubiquitin-modified protein. *Plant Physiol.* 163, 135–149. doi: 10.1104/pp.113.221341
- Inohara, N., and Nuñez, G. (2002). ML-a conserved domain involved in innate immunity and lipid metabolism. *Trends Biochem. Sci.* 27, 219–221. doi: 10.1016/S0968-0004(02)02084-4
- Jin, L., and Liu, J. (2008). Molecular cloning, expression profile and promoter analysis of a novel ethylene responsive transcription factor gene GhERF4 from cotton (*Gossypium hirsutum*). *Plant Physiol. Biochem.* 46, 46–53. doi: 10.1016/j.plaphy.2007.10.004
- Juliana, P., Singh, R. P., Singh, P. K., Poland, J. A., Bergstrom, C. C., Huerta-Espino, J., et al. (2018). Genome-wide association mapping for resistance to leaf rust, stripe rust and tan spot in wheat reveals potential candidate genes. *Theor. Appl. Genet.* 131, 1405–1422. doi: 10.1007/s00122-018-3086-6
- Kristensen, P. S., Jahoor, A., Andersen, J. R., Cericola, F., Orabi, J., Janss, L. L., et al. (2018). Genome-wide association studies and comparison of models and cross-validation strategies for genomic prediction of quality traits in advanced winter wheat breeding lines. *Front. Plant Sci.* 9:69–83. doi: 10.3389/fpls.2018.00069
- Kumar, D., Kumar, A., Chhokar, V., Gangwar, O. P., Bhardwaj, S. C., Sivasamy, M., et al. (2020). Genome-wide association studies in diverse spring wheat panel for stripe, stem, and leaf rust resistance. *Front. Plant Sci.* 11:748–775. doi: 10.3389/fpls.2020.00748
- Kuroda, H., Takahashi, N., Shimada, H., Seki, M., Shinozaki, K., and Matsui, M. (2002). Classification and expression analysis of Arabidopsis F-box-containing protein genes. *Plant Cell Physiol.* 43, 1073–1085. doi: 10.1093/pcp/pcf151
- Li, F., Wen, W., Liu, J., Zhang, Y., Cao, S., He, Z., et al. (2019). Genetic architecture of grain yield in bread wheat based on genome-wide association studies. *BMC Plant Biol.* 19:168–186. doi: 10.1186/s12870-019-1781-3
- Li, H., Xiao, Q., Zhang, C., Du, J., Li, X., Huang, H., et al. (2017). Identification and characterization of transcription factor ZmEREB94 involved in starch synthesis in maize. *J. Plant Physiol.* 216, 11–16. doi: 10.1016/j.jplph.2017.04.016
- Li, J., Liu, J., Wen, W., Zhang, P., Wan, Y., Xia, X., et al. (2018). Genome-wide association mapping of vitamins B1 and B2 in common wheat. *Crop J.* 6, 263–270. doi: 10.1016/j.cj.2017.08.002
- Li, M., Tang, D., Wang, K., Wu, X., Lu, L., Yu, H., et al. (2011). Mutations in the F-box gene LARGER PANICLE improve the panicle architecture and enhance the grain yield in rice. *Plant Biotechnol. J.* 9, 1002–1013. doi: 10.1111/j.1467-7652.2011.00610.x
- Li, M., Wang, Z., Liang, Z., Shen, W., Sun, F., Xi, Y., et al. (2015). Quantitative trait loci analysis for kernel-related characteristics in common wheat (*Triticum aestivum* L.). *Crop Sci.* 55, 1–9. doi: 10.2135/cropsci2014.09.0616
- Li, N., and Li, Y. (2014). Ubiquitin-mediated control of seed size in plants. *Front. Plant Sci.* 5:332–346. doi: 10.3389/fpls.2014.00332
- Li, Q., Zhang, Y., Liu, T., Wang, F., Liu, K., Chen, J., et al. (2015). Genetic analysis of kernel weight and kernel size in wheat (*Triticum aestivum* L.) using unconditional and conditional QTL mapping. *Mol. Breed.* 35, 194–208. doi: 10.1007/s11032-015-0384-4
- Li, W., Li, C., Gu, Z., Qiu, Y., Cheng, L., Hong, Y., et al. (2016a). Retrogradation behavior of corn starch treated with 1,4- α -glucan branching enzyme. *Food Chem.* 203, 308–313. doi: 10.1016/j.foodchem.2016.02.059
- Li, W., Li, C., Gu, Z., Qiu, Y., Cheng, L., Hong, Y., et al. (2016b). Relationship between structure and retrogradation properties of corn starch treated with 1,4- α -glucan branching enzyme. *Food Hydrocoll.* 52, 868–875. doi: 10.1016/j.foodhyd.2015.09.009
- Li, X., Xu, X., Liu, W., Li, X., Yang, X., Ru, Z., et al. (2020). Dissection of superior alleles for yield-related traits and their distribution in important cultivars of wheat by association mapping. *Front. Plant Sci.* 11:175–188. doi: 10.3389/fpls.2020.00175
- Liu, J., Feng, B., Xu, Z., Fan, X., Jiang, F., Jin, X., et al. (2018). A genome-wide association study of wheat yield and quality-related traits in southwest China. *Mol. Breed.* 38, 1–10. doi: 10.1007/s11032-017-0759-9
- Liu, K., and Muse, S. V. (2005). Power marker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 21, 2128–2129. doi: 10.1093/bioinformatics/bti282
- Liu, N., Xue, Y., Guo, Z., Li, W., and Tang, J. (2016). Genome-wide association study identifies candidate genes for starch content regulation in maize kernels. *Front. Plant Sci.* 7:1046–1053. doi: 10.3389/fpls.2016.01046
- Ma, M., Wang, Q., Li, Z., Cheng, H., Li, Z., Liu, X., et al. (2015). Expression of TaCYP78A3, a gene encoding cytochrome P450 CYP78A3 protein in wheat (*Triticum aestivum* L.), affects seed size. *Plant J.* 83, 312–325. doi: 10.1111/tpj.12896
- Miao, Y., Jing, F., Ma, J., Liu, Y., Zhang, P., Chen, T., et al. (2022). Major genomic regions for wheat grain weight as revealed by QTL linkage mapping and meta-analysis. *Front. Plant Sci.* 13:802310. doi: 10.3389/fpls.2022.802310
- Mourad, A. M. I., Sallam, A., Belamkar, V., Wegulo, S., Bowden, R., Jin, Y., et al. (2018). Genome-wide association study for identification and validation of novel SNP markers for Sr6 stem rust resistance gene in bread wheat. *Front. Plant Sci.* 9:00380. doi: 10.3389/fpls.2018.00380
- Okamoto, Y., Nguyen, A. T., Yoshioka, M., Iehisa, J. C. M., and Takumi, S. (2013). Identification of quantitative trait loci controlling grain size and shape in the D genome of synthetic hexaploid wheat lines. *Breed. Sci.* 63, 423–429. doi: 10.1270/jsbbs.63.423
- Pang, Y., Liu, C., Wang, D., Amand, P. S., Bernardo, A., Li, W., et al. (2020). High-resolution genome-wide association study identifies genomic regions and candidate genes for important agronomic traits in wheat. *Mol. Plant* 13, 1311–1327. doi: 10.1016/j.molp.2020.07.008
- Rajaram, S. (2001). Prospects and promise of wheat breeding in the 21st century. *Euphytica* 119, 3–15. doi: 10.1023/A:1017538304429
- Ramya, P., Chaubal, A., Kulkarni, K., Gupta, L., Kadoo, N., Dhaliwal, H. S., et al. (2010). QTL mapping of 1000-kernel weight, kernel length, and kernel width in bread wheat (*Triticum aestivum* L.). *J. Appl. Genet.* 51, 421–429. doi: 10.1007/BF03208872
- Ray, D. K., Mueller, N. D., West, P. C., and Foley, J. A. (2013). Yield trends are insufficient to double global crop production by 2050. *PLoS One* 8:e66428. doi: 10.1371/journal.pone.0066428

- Ray, D. K., Ramankutty, N., Mueller, N. D., West, P. C., and Foley, J. A. (2012). Recent patterns of crop yield growth and stagnation. *Nat. Commun.* 3, 1293–1299. doi: 10.1038/ncomms2296
- Ren, J., Li, C., Gu, Z., Cheng, L., Hong, Y., and Li, Z. (2018). Digestion rate of tapioca starch was lowered through molecular rearrangement catalyzed by 1,4- α -glucan branching enzyme. *Food Hydrocoll.* 84, 117–124. doi: 10.1016/j.foodhyd.2018.06.005
- Riaz, A., Athiyannan, N., Periyannan, S. K., Afanasenko, O., Mitrofanova, O. P., Platz, G. J., et al. (2017). Unlocking new alleles for leaf rust resistance in the Vavilov wheat collection. *Theor. Appl. Genet.* 131, 127–144. doi: 10.1007/s00122-017-2990-5
- Simmonds, J., Scott, P., Brinton, J., Mestre, T. C., Bush, M., Blanco, A. D., et al. (2016). A splice acceptor site mutation in TaGW2-A1 increases thousand grain weight in tetraploid and hexaploid wheat through wider and longer grains. *Theor. Appl. Genet.* 129, 1099–1112. doi: 10.1007/s00122-016-2686-2
- Soumya, P. R., Burridge, A. J., Singh, N., Batra, R., Pandey, R., Kalia, S., et al. (2021). Population structure and genome-wide association studies in bread wheat for phosphorus efficiency traits using 35 K wheat breeder's affymetrix array. *Sci. Rep.* 11, 7601–7617. doi: 10.1038/s41598-021-87182-2
- Su, J., Wang, C., Ma, Q., Zhang, A., Shi, C., Liu, J., et al. (2020a). An RTM-GWAS procedure reveals the QTL alleles and candidate genes for three yield-related traits in upland cotton. *BMC Plant Biol.* 20:416–430. doi: 10.1186/s12870-020-02613-y
- Su, J., Wang, C., Yang, D., Shi, C., Zhang, A., Ma, Q., et al. (2020b). Decryption of favourable haplotypes and potential candidate genes for five fibre quality properties using a relatively novel genome-wide association study procedure in upland cotton. *Ind. Crop Prod.* 158:113004. doi: 10.1016/j.indcrop.2020.113004
- Su, Q., Zhang, X., Zhang, W., Zhang, N., Song, L., Liu, L., et al. (2018). QTL detection for kernel size and weight in bread wheat (*Triticum aestivum* L.) using a high-density SNP and SSR-based linkage map. *Front. Plant Sci.* 9:1484–1496. doi: 10.3389/fpls.2018.01484
- Sukumaran, S., Dreisigacker, S., Lopes, M., Chavez, P., and Reynolds, M. (2015). Genome-wide association study for grain yield and related traits in an elite spring wheat population grown in temperate irrigated environments. *Theor. Appl. Genet.* 128, 353–363. doi: 10.1007/s00122-014-2435-3
- 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
- Walsh, J. (1981). Genetic vulnerability down on the farm. *Science* 214, 161–164. doi: 10.1126/science.214.4517.161
- Xia, C., Qiao, Y., Dong, C., Chen, X., Wang, Y., Li, D., et al. (2021). Enzymatic properties of an efficient glucan branching enzyme and its potential application in starch modification. *Protein Exp. Purif.* 178:105779. doi: 10.1016/j.pep.2020.105779
- Xu, J., Zhang, X., and Xue, H. (2016). Rice aleurone layer specific OsNF-YB1 regulates grain filling and endosperm development by interacting with an ERF transcriptionfactor. *J. Exp. Bot.* 67, 6399–6411. doi: 10.1093/jxb/erw409
- Yang, W., Guo, Z., Huang, C., Duan, L., Chen, G., Jiang, N., et al. (2014). Combining high-throughput phenotyping and genome-wide association studies to reveal natural genetic variation in rice. *Nat. Commun.* 5, 5087–5095. doi: 10.1038/ncomms6087
- Yang, X., Pan, Y., Singh, P. K., He, X., Ren, Y., Zhao, L., et al. (2019). Investigation and genome-wide association study for *Fusarium crown rot* resistance in Chinese common wheat. *BMC Plant Biol.* 19:153–166. doi: 10.1186/s12870-019-1758-2
- Yano, K., Yamamoto, E., Aya, K., Takeuchi, H., Lo, P., 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
- Zhang, L., Dong, C., Chen, Z., Gui, L., Chen, C., Li, D., et al. (2021). WheatGmap: a comprehensive platform for wheat gene mapping and genomic studies. *Mol. Plant* 14, 187–190. doi: 10.1016/j.molp.2020.11.018
- Zhang, P., Yan, X., Gebrewahid, T. W., Zhou, Y., Yang, E., Xia, X., et al. (2021). Genome-wide association mapping of leaf rust and stripe rust resistance in wheat accessions using the 90K SNP array. *Theor. Appl. Genet.* 134, 1233–1251. doi: 10.1007/s00122-021-03769-3
- Zhang, Y., Li, D., Zhang, D., Zhao, X., Cao, X., Dong, L., et al. (2018). Analysis of the functions of TaGW2 homoeologs in wheat grain weight and protein content traits. *Plant J.* 94, 857–866. doi: 10.1111/tpj.13903
- Zhu, X., Shao, X., Pei, Y., Guo, X., Li, J., Song, X., et al. (2018). Genetic diversity and genome-wide association study of major ear quantitative traits using high-density SNPs in maize. *Front. Plant Sci.* 9:966–981. doi: 10.3389/fpls.2018.00966
- Zuo, J., Lin, C., Cao, H., Chen, F., Liu, Y., and Liu, J. (2019). Genome-wide association study and quantitative trait loci mapping of seed dormancy in common wheat (*Triticum aestivum* L.). *Planta* 250, 187–198. doi: 10.1007/s00425-019-03164-9

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 Wang, Tian, Ma, Liu, Zhang, Chen, Shahinnia and Yang. 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.