Journal of Integrative Agriculture 2026, 25(7): 2714–2722

Available online at ScienceDirect

Journal of Integrative Agriculture



RESEARCH ARTICLE

Genetic dissection and validation of a stable QTL for grain roundness on chromosome 5A in bread wheat (*Triticum aestivum* L.)

Jiajie He[1, 2], Zhibin Xu[1], Bo Feng[1], Qiang Zhou[1], Xiaofeng Liu[3], Guangsi Ji[4], Shaodan Guo[1], Xiaoli Fan[1], Tao Wang[1#]

[1] *Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu 610213, China*
[2] *University of Chinese Academy of Sciences, Beijing 101408, China*
[3] *Sichuan Academy of Agricultural Sciences, Chengdu 610066, China*
[4] *School of Life Science and Engineering, Henan University of Urban Construction, Pingdingshan 467044, China*

Highlights

- A stable QTL for grain roundness, *QGr.cib-5A*, was consistently detected across all environments.
- The effect of *QGr.cib-5A* on grain roundness was stable in various genetic backgrounds.
- Rounder grain shape has been relatively widely selected during wheat breeding.

Abstract

Wheat grain morphology is an important breeding target due to its impact on yield and end-use properties. However, the genetic basis of grain roundness, which is a major determinant of grain morphology, remains largely unexplored. In this study, F_2 and recombinant inbred line (RIL) populations from a Zhongkemai 138 (ZKM138)×Chinese Spring (CS) cross were employed to analyze the genetic basis of grain shape variation. Kompetitive allele specific PCR (KASP) markers were developed according to single nucleotide polymorphisms (SNPs) from bulked segregant exome sequencing (BSE-Seq) of F_2 and Wheat 55K SNP array data online. Those SNPs were then used to construct two genetic maps of the F_2 and RIL populations, spanning 148.89 cM (30 KASP markers) and 129.82 cM (25 KASP markers), respectively. Using the traditional QTL mapping method based on these two maps, a stable quantitative trait locus (QTL) for grain roundness (GR), *QGr.cib-5A*, could be repeatedly detected in the interval of 444.8–455.5 Mb on chromosome 5A. Further conditional QTL mapping analysis revealed that grain width was the major contributor to GR at this locus. In addition, the use of two tightly linked markers, *5A4-15* and *55k-31*, showed a 96.27% transmissibility of ZKM138-derived alleles in 134 ZKM138 derivatives, along with a 7.38% increase in GR and a 65.19% distribution among worldwide varieties. Finally, *TraesCS5A02G236400*, which may encode a hydroxyproline-rich glycoprotein family protein, was deduced to be the candidate gene. Collectively, these results can facilitate wheat grain shape improvement and enhance the market value of wheat.

Keywords: wheat, grain roundness, QTL mapping, candidate gene

1. Introduction

Wheat (*Triticum aestivum* L.) has been subjected to natural selection and human-designed domestication with a greater emphasis on nourishing the world's ever-growing population. During its evolutionary trajectory, the wheat grain has experienced a dramatic transformation in size and shape from a slim grain to a plump grain (Gegas *et al.* 2010). Grain size is characterized by grain length (GL), width (GW), and thickness (GT) while grain shape is described by the roundness (GR, the width to length ratio), and these two key factors determine grain weight and end-use suitability (Dholakia *et al.* 2003). Therefore, decomposition of the complex morphology to dissect the genetic basis of grain morphological traits is crucial for improving wheat shape and market value.

Wheat grain size and shape are complex traits inherited quantitatively (Campbell *et al.* 1999; Varshney *et al.* 2000;

Received 8 July 2024; Received in revised form 4 September 2024; Accepted 24 September 2024; Available online 13 November 2024
Jiajie He, E-mail: hejj@cib.ac.cn; #Correspondence Tao Wang, Tel: +86-28-82890308, E-mail: wangtao@cib.ac.cn

© 2026 CAAS. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/). Peer review under responsibility of Editorial Board of *Journal of Integrative Agriculture*.
doi: 10.1016/j.jia.2024.11.030

Ammiraju *et al*. 2001). Selecting larger grains, often accompanied by an increase in thousand-grain weight (TGW) compared with grain shape, has invariably been the breeding hotspot. With the completion of wheat reference genome sequencing (https://wheat-urgi.versailles.inrae.fr/Seq-Repository/Assemblies) and the availability of more high-throughput genotyping SNP arrays, a wide range of QTLs linked to grain size and weight have been detected across all the wheat chromosomes, and some of the related genes have been cloned in the post-genomics period (Williams *et al.* 2013; Cui *et al.* 2014; Zhang G *et al.* 2015; Mohler *et al.* 2016; Xu *et al.* 2017; Sukumaran *et al.* 2018; Cao *et al.* 2019; Zhao *et al.* 2021; Song *et al*. 2022; Lyu *et al.* 2024). For example, *TaGW2*, the equivalent to rice *OsGW2*, enhances TGW because of the broadened GW (Yang *et al.* 2012; Geng *et al.* 2017). *TaGS-D1*, an orthologous gene of *OsGS3*, is related to GL and TGW (Zhang *et al.* 2014). *TaGASR7-A1*, homologous with *OsGASR7*, controls GW through GL (Dong *et al.* 2014; Zhang D *et al.* 2015). However, the contribution of grain shape characteristics, like grain roundness, to TGW and final yield remains unclear. For example, Cheng *et al.* (2017) found that GR and GW are negatively correlated. Li *et al.* (2019) observed the lowest correlation coefficients between TGW and GR compared with the grain size components, like GW and GL. However, Su *et al.* (2018) found that two QTLs for TGW were entirely contributed by GW and grain diameter ratio (GDR) through conditional QTL analysis, indicating that the possible effects of grain shape on TGW remain unclear. This discrepancy may be associated with the specific genetic backgrounds used in those studies, and an in-depth exploration of the underlying relationship between grain shape and yield is needed.

However, grain shape clearly has an immediate association with wheat quality traits, such as milling performance and test weight. For example, fine flour quality partly lies in a uniform grain shape (Zhang *et al.* 2008). Studies have found that larger sorghum (*Sorghum bicolor* (L.) Moench) grains sustained less relative mass loss and presented better milling performance (Lee *et al.* 2002), and the same is true of wheat. Rounder or plumper grains make it easier to grind off the outer layers of bran and result in a larger amount of high-quality flour yield under the same conditions (Mares *et al.* 1986). In addition, Sahand's predictive model also demonstrated that circular grains produced more flour than elongated grains, independent of grain size (Assadzadeh *et al.* 2022). Pomeranz and Afework (1984) found that the grain grinding time and flour particle size increased with reduced grain size, and that tendency was more obvious in shriveled grains than in plump grains. Given the positive correlation between grain shape and milling performance, any related genetic improvements would be welcomed by the market.

In this study, a traditional QTL analysis of GL, GW and GR was conducted with an F_2 population and a recombinant inbred line (RIL) population from the cross of Zhongkemai 138 (ZKM138) and Chinese Spring (CS). In addition, a conditional analysis was performed to identify the major contributor to GR in this genetic background. Finally, a series of natural populations was used to verify the genetic effects of the target QTL. The results of this study can help us to gain deeper insights into the underlying basis of grain shape variation.

2. Materials and methods

2.1. Plant materials

An F_2 population derived from a ZKM138×CS cross and a subsequent F_6 RILs population (1C) advanced by the single-seed descent method were used in this study. ZKM138, a cultivar selected from Chuanmai 42 (CM42)×Chuanyu 16 (CY16), offers many of the crop output benefits with larger grains than CS, which was derived from Sichuan white landrace Chengduguangtou (CDGT). The larger grain is one of the characteristics that make these populations suitable for studying the phenotypic plasticity and adaption during evolution. Aside from the great differences in morphological traits between the two parents, CS boasts a nearly fully annotated and ordered genome sequence, which enables more precise, reliable mapping results and candidate gene prediction. Another group of 154 RILs derived from ZKM138×CM44 (BC-RIL) and a set of 134 ZKM138-derivatives were used to validate and assess the impact of the target QTL according to their phenotypes. In addition, 117 Chinese cultivars and 158 overseas varieties gathered from 15 countries were tested to determine the prevalence of the detected locus based on the mostly linked KASP markers.

2.2. Phenotypic evaluation

A total of 494 F_2 plants were planted at Shifang (103°51′E, 30°43′N) in 2018–2019, and 184 RIL lines were planted in Shuangliu (103°52′E, 30°34′N) and Shifang in 2022–2023 and 2021–2023, respectively. A completely randomized block design was used for all the trials in each environment. The seeds were sown in 1 m single-row plots with 10 seeds per row and a 0.25 m spacing interval between adjacent rows. Fertilizer (N, 25%; P_2O_5, 10%; K_2O, 10%) was applied during the sowing period at a rate of 450 kg ha^{-1}. Field management and disease control were carried out as per local procedures.

At the seedling stage, a 10 cm fragment of the tender leaves of F_2 and RIL plants was cut from each corresponding accession for extracting DNA. After maturity, the main spike of each F_2 plant and five plants per RIL line were collected separately and dried to a constant weight. The grain-related traits TGW, GL, GW, GR, grain length to width ratio (LWR), grain area (GA), grain perimeter (GP), and grain diameter (GD) were recorded by SC-G software (Wanshen Detection Technology Co., Ltd., Hangzhou, China). The mean values of related traits for each RIL were calculated as the representatives of that specific line, and data analysis was completed with Microsoft Office Excel 2019.

2.3. Data analysis and QTL mapping

The statistical analysis and correlation coefficients among traits were evaluated with SPSS 27 (IBM Corp., Armonk, NY, USA). Estimation of the broad-sense heritability (H^2) of each trait followed the method proposed by Smith *et al.* (1998). The best linear unbiased estimation (BLUE) was calculated using the QTL IciMapping V4.2 software. The genetic map

was generated by JoinMap 4.1 (https://www.kyazma.nl/index.php/JoinMap/) with the Kosambi algorithm, which converts the recombinant fractions to centiMorgans (cM). Inclusive composite interval mapping with the additive tool (ICIM-ADD) in IciMapping V4.2 was performed to detect the QTLs (Meng *et al.* 2015). Individual environmental QTL values were measured using the bi-parental populations (BIP) module with walking step=1.00 cM, PIN=0.001, and logarithm of odds (LOD) scores≥2.5. Each QTL captured more than once was considered a stable QTL.

2.4. BSE-Seq analysis

According to the phenotypic data sets (Appendix A), the genomic DNA of 49 F_2 lines with extremely high and low grain sizes together with their parents was extracted and bulked in equal ratios to construct four DNA pools, i.e., a large-size pool, a small-size pool, a ZKM138 pool, and a CS pool. Genomic DNA was extracted from individual plants for all accessions using the CTAB method. The concentration and purity were tested by NanoDrop 2000 (Thermo Scientific, Waltham, MA, USA). Bulked segregant exome sequencing (BSE-Seq) analysis of the four DNA pools was conducted by Tcuni Biotechnology Co., Ltd. (Chengdu, China). The 4th power of the original ED (ED^4) was used as the correlation value to eliminate background noise. The regions above that threshold were likely to be the linked regions.

2.5. Development of KASP markers and genotyping

With the aims of QTL detection and further narrowing down the flanking intervals of the target locus, the BSE-Seq-detected and array-based polymorphic SNPs that separated ZKM138 from CS on chromosome 5A were transferred to detectable kompetitive allele specific PCR (KASP) primers (http://www.polymarker.info) against the backdrop of the Chinese Spring reference genome sequence IWGSC RefSeq v1.0. Each KASP marker system contained two allele-specific forward primers and a common reverse primer. The allele-specific primers were tailed with the FAM signal sequence (5′-GAAGGTGACCAAGTTCATGCT-3′) or the HEX signal sequence (5′-GAAGGTCGGAGTCAACGG ATT-3′). The common primer was specific to the SNP flanking sequences. The KASP reaction system included 5 μL 2× KASP Master Mix, 0.4 μL primer mixture (0.2 μL forward primers and 0.2 μL common primer), 2 μL DNA template (50–200 ng μL^{-1}), and 3 μL ddH_2O. PCR was performed as follows: a hot start at 95°C for 10 min, 10 touchdown cycles (94°C for 20 s, touchdown at 61°C with reduction by 0.6°C per cycle for 40 s), followed by 30–34 additional cycles of 95°C for 20 s and 55°C for 40 s. The genotyping data were processed using the Quantstudio3 Real-Time PCR System (Thermo Fisher Scientific, USA), and genotypes with homozygous alleles from ZKM138 were designated as “ZKM138” while those from CS were as designated at “CS”.

2.6. Conditional QTL analysis for GR

The conditional analysis method proposed by Zhu (1995) allows for the exclusion of the contribution of a causal trait from the variation of a certain trait. The remaining variation of the resultant trait is defined as conditional variation, which describes the links between a complex trait and its causal components at the QTL level (Cui *et al.* 2011b). Therefore, a conditional QTL analysis was conducted in this study. The term (T1|T2) obtained from QGAStation2.0 (Chen *et al.* 2012) means the value of Trait 1 conditioned on Trait 2, so ‘GR|GL’ and ‘GR|GW’ here refer to the values of GR excluding the impacts of GL and GW, respectively. QTL IciMapping v4.2 software was then used to detect the conditional loci with the inclusion of conditional phenotypic parameters. For the new mapping results, if a QTL was detected only in the unconditional QTL analysis, that locus was considered to make a large contribution to the corresponding trait; whereas a QTL was considered to be unassociated with the corresponding trait if it was detected in both the conditional and unconditional QTL analyses.

2.7. Prediction of candidate genes for *QGr.cib-5A*

The physical positions and sequences of flanking markers of the QTL were obtained by blasting against the Chinese Spring genome (IWGSC RefSeq v1.0). Genes predicted with high confidence in the candidate region were extracted from the JBrowse website (https://urgi.versailles.inra.fr/jbrowseiwgsc). The functional descriptions of the genes were retrieved from UniProt (https://www.uniprot.org). The gene expression data in different tissues were obtained from expVIP (http://www.wheat-expression.com) and further visualized using TBtools software (v.1.098) (Chen *et al.* 2020). Orthologs of the candidate genes in rice (*Oryza sativa*) were retrieved from Ensembl Plants (https://plants.ensembl.org/biomart/).

3. Results

3.1. Phenotypic evaluation and correlation analysis

A phenotypic comparison of the parents is shown in Fig. 1-A. Clear differences between ZKM138 and CS were found in TGW, GL and GW, with ZKM138 showing better performance in most cases. For F_2 individuals, transgressive segregation beyond the boundary of both parents was observed (Table 1), with GL ranging from 4.50 to 7.29 mm, GW from 2.16 to 3.29 mm, GR from 0.39 to 0.59, and TGW from 19.19 to 59.50 g. These broad and significant variations indicated that various alleles were dispersed between the two parents.

For the RIL population (Appendix B), the morphometric parameters followed continuous distributions, with GL ranging from 5.07 to 6.64 mm, GW from 2.34 to 3.00 mm, GR from 0.40 to 0.59, and TGW from 23.49 to 48.97 g. In addition, the absolute values of skewness and kurtosis among these traits were mostly lower than 1. The broad sense heritability (H^2) of each trait was greater than 0.7 across environments, demonstrating that the genetic factors behind them mattered more than the environment. These results indicated that this ZKM138×CS RIL population was a favorable mapping population for QTL analysis.

Pearson’s correlation coefficient analysis (Fig. 1-B) was conducted among the morphometric traits. The results showed that GW and GR (r=0.19) had significant and

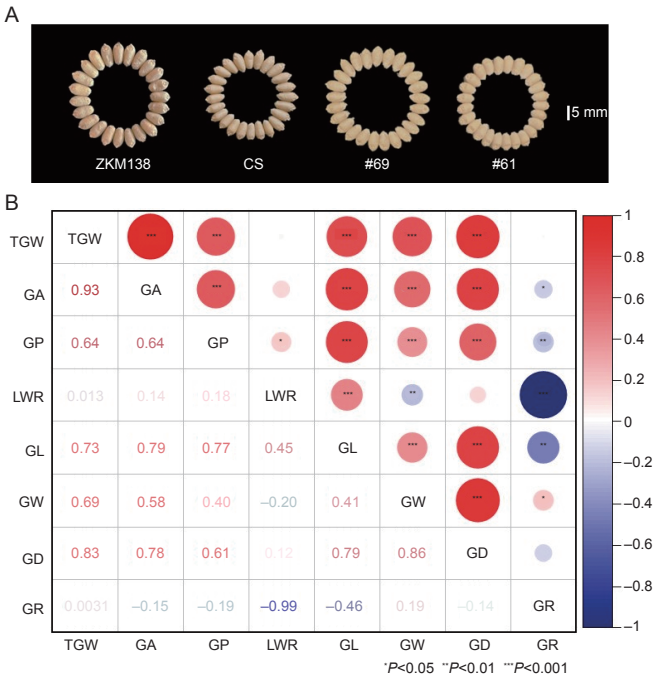


Fig. 1 Performance of the measured grain related traits in the 1C-recombinant inbred lines (RILs) population. A, grain morphology of the parents Zhongkemai 138 (ZKM138) and Chinese Spring (CS), and two representative lines of the 1C-RILs population. The white line represents a scale of 5 mm. B, correlation analysis between grain-related traits in the 1C-RILs population based on the best linear unbiased prediction (BLUE) dataset. TGW, thousand-grain weight; GA, grain area; GP, grain perimeter; LWR, grain length to width ratio; GL, grain length; GW, grain width; GD, grain diameter; GR, grain roundness. *, ** and *** indicate significances at $P<0.05$, $P<0.01$ and $P<0.001$, respectively.

Table 1 Phenotypic variation of grain-related traits for the F_2 population in the environment of 2018–2019 in Shuangliu, Sichuan Province, China

Trait[1)]	Min	Max	Mean	SD	CV (%)	Skewness	Kurtosis
TGW (g)	19.19	59.50	40.23	6.92	0.17	–0.45	–0.01
GR	0.39	0.59	0.49	0.04	0.07	–0.25	–0.37
GW (mm)	2.18	3.29	2.83	0.21	0.08	–0.51	–0.14
GL (mm)	4.50	7.29	5.78	0.31	0.05	–0.03	1.99
LWR	1.73	2.58	2.07	0.16	0.08	0.58	0.00
GD (mm)	3.29	4.61	4.06	0.23	0.06	–0.61	0.36
GA (mm^2)	8.60	23.98	13.22	1.52	0.11	0.39	4.76
GP (mm)	11.89	18.30	14.75	0.74	0.05	–0.12	1.71

[1)] TGW, thousand-grain weight; GR, grain roundness; GW, grain width; GL, grain length; LWR, grain length to width ratio; GD, grain diameter; GA, grain area; GP, grain perimeter.

positive correlations ($P<0.05$), while GL and GR ($r=0.46$) had significant and negative correlations ($P<0.01$). There were no significant correlations between GR and TGW. For GW and GL, the correlation coefficient of TGW-GL ($r=0.73$) was almost the same as TGW-GW ($r=0.69$), indicating that they were equivalent contributors to TGW in the 1C-RILs.

3.2. BSE-Seq analysis

Bulked segregant exome capture sequencing (BSE-Seq) is an efficient high-throughput sequencing method for identifying causal mutations or candidate regions within the large and redundant wheat genome (Dong *et al.* 2020). The BSE-Seq results (Appendices C and D) showed that compared with the Chinese Spring reference genome RefSeq v1.0, many SNPs (6,743, 84.91%) and InDels (1,198, 15.09%) were detected from the four pools, with allelic frequency (AF) values of >0.3 or <0.7. The SNPs featured an uneven distribution across the chromosomes, in which chromosome 5A had the second densest variation after chromosome 1A. The ED algorithm (Cutoff=Median+3×SD) indicated that the potential regions on chromosome 5A could be effective for gene mapping (Fig. 2). SNPs relevant to grain shape in the candidate segment could be instrumental in QTL mapping, so they were converted to KASP markers to determine the genotypes of the populations.

3.3. Linkage map construction and QTL mapping

In this present study, seven pairs of KASP markers (Appendix E) were successfully designed based on the polymorphic SNPs between parents stemming from the BSE-Seq results. With access to the Wheat 55K SNP array data online, another 23 KASP markers (Appendix F) were converted. For the F_2 population, the genotyping results processed by JoinMap's Kosambi function established a genetic map consisting of 30 KASP markers with a total length of 148.89 cM (Appendix F). The BSE-Seq results were then validated in the RILs using the same markers. Together with all the markers available, a genetic map spanning 129.82 cM and consisting of 25 KASP markers with an average distance of 5.19 cM per marker was generated (Appendix F). The two mapping populations were used simultaneously to map a QTL pertinent to GR assisted by the BIP module of IciMapping v4.2 (Table 2; Appendix G). When more than one QTL overlapped in the confidence interval, those QTLs were considered as the same QTL. As shown in Table 2, four QTLs were detected in F_2 for GA, GW, GL, and GR. The LOD thresholds of those QTLs carrying 2.46, 2.89, 4.35, and 3.14% of the phenotypic variation explained were 2.67, 3.09, 4.57, and 3.41, respectively, with an additive effect range from –0.08 to 0.32. The genetic distance of

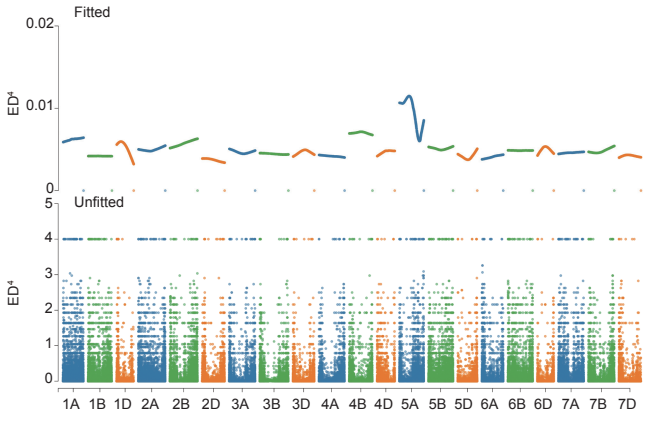


Fig. 2 The Euclidean distance (ED) at each SNP using a Loess (locally weighted scatterplot smoothing) fitting curve with the cutoff of 0.5. ED^4, the 4th power of the original ED.

Table 2 Grain-related traits detected in the 1C-F_2 population[1)]

Trait[2)]	QTL	Interval (cM)	Flanking markers	Physical position (bp)	LOD	PVE (%)	Add
GA	*QGa.cib-5A*	0–7.5	*55k-5–55k-24*	6,252,892–9,502,338	2.67	2.46	0.32
GW	*QGw.cib-5A*	7.5–20.5	*55k-24–55k-10*	9,502,338–18,712,331	3.09	2.89	0.05
GL	*QGl.cib-5A*	48.5–51.5	*5A4-15–55k-31*	445,243,555–445,436,340	4.57	4.35	–0.08
GR	*QGr.cib-5A*	60.5–72.5	*55k-41–55k-55*	458,898,875–485,845,842	3.41	3.14	0.01

[1)] Physical position, from IWGSC RefSeq v1.0.; LOD, logarithm of odds; PVE, phenotypic variation explained; Add, additive effect, where positive and negative values indicate that alleles are from Zhongkemai 138 (ZKM138) and Chinese Spring (CS), respectively.
[2)] GA, grain area; GW, grain width; GL, grain length; GR, grain roundness.

the QTL for GR was 0.85 cM, corresponding physically to 445.2–445.4 Mb and flanked by KASP markers *55k-31* and *5A4-15*. For the RILs, GR was repeatedly identified in all environments and the BLUE dataset (Appendix G), a sign of the reliability and stability of the detected QTL (hereafter *QGr.cib-5A*). *QGr.cib-5A* was flanked by markers *55k-30* and *55k-39* with a physical interval of around 10.65 Mb between 444.8 and 455.5 Mb (Fig. 3). The LOD value of the QTL was from 2.51–3.60, explaining 6.16–8.73% of the phenotypic variation. The agreement of the mapping results for the F_2 population indicates it was an identical QTL sharing the common interval in the RILs, and also that the positive locus which contributed a rounder grain originated from ZKM138.

3.4. Conditional QTL analysis for GR

GR is a comprehensive trait composed of GL and GW. Of the conditional QTLs, when GR was independent of GW, GR|GW failed to be detected with an LOD threshold of 2.5. However, the LOD value of GR|GL was close to that of the target region and its additive effects were higher than in the unconditioned case, suggesting GL played an inhibitory role in round shape determination (Appendix H). GW is therefore primarily responsible for GR in the target region, as inferred from the above observations. One likely reason for another emergence of GR|GL across the genetic map is that the locus controlling GR was undetectable unless the impact of GL was excluded. Hence, GR is more likely a comprehensive trait affected by various moderate or minor QTLs in the ZKM138×CS background.

3.5. Predicting candidate genes of *QGr.cib-5A*

According to differential SNPs between the parents in the 10.65 Mb physical interval against the IWGSC RefSeq v1.0 genome, seven high-confidence genes (Appendix I) were filtered out from all the genes in the candidate region. Furthermore, based on the IWGSC RefSeq v2.1 genome, *QGr.cib-5A* was mapped to 445.1–456.1 Mb, so it identified the same candidate genes as in IWGSC RefSeq v1.0. This indicated that these genes have stable positions and sequences, and exhibit high consistency and reliability in different versions of genome assembly and annotation. The expression pattern analysis showed that *TraesCS5A02G236400* was favorably expressed in spike and grain tissues (Appendix J). Gene annotation and orthologous analysis measured against the sequences of rice (Appendix I), combined with the previous expression pattern, suggested that *TraesCS5A02G236400* may be associated with the development of grain and hence affect grain shape.

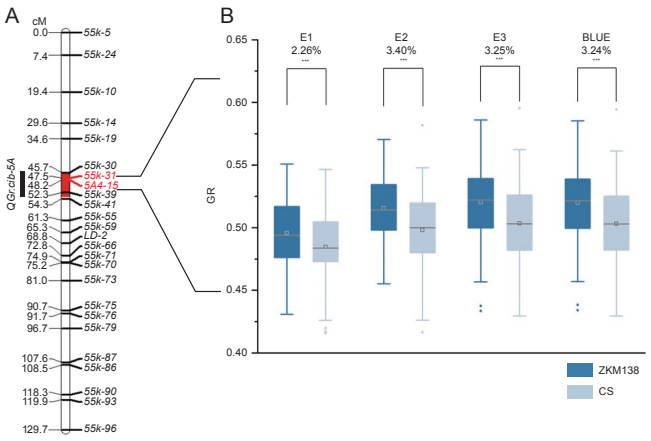


Fig. 3 Genetic map of the quantitative trait locus (QTL) *QGr.cib-5A*, and estimation its effect on grain roundness (GR) in the 1C-recombinant inbred lines (RILs). A, genetic map of *QGr.cib-5A*. B, its effects on GR in the 1C-RILs. ZKM138 and CS represent lines with alleles from Zhongkemai 138 (ZKM138) and Chinese Spring (CS), respectively. E1 and E2, 2021–2022 and 2022–2023 in Shifang, Sichuan Province, China, respectively; E3, 2022–2023 in Shuangliu, Sichuan Province, China; BLUE, best linear unbiased prediction. *** represents significance at $P<0.001$.

3.6. Verification of *QGr.cib-5A* in similar genetic backgrounds

ZKM138 is one of the core backbone parents in our lab in terms of the breeding process since it has bred multiple varieties. For example, another RIL population (ZKM138×CM44, BC-RIL) was constructed from the common parent ZKM138. Since the GR increasing allele of *QGr.cib-5A* was inherited from ZKM138, this RIL population was used to evaluate its impact on GR under a similar genetic background. Since *55k-31* can separate ZKM138 from CM44, it was used to scan the 154 BC-RIL lines. As expected, the lines of two homologous genotypes exhibited notable differences in GR ($P<0.01$) and GW ($P<0.05$), with the ZKM138-derived genotype higher in both parameters (Fig. 4-A). These results confirmed the reality of *QGr.cib-5A* and its influence on grain shape.

Furthermore, the 134 ZKM138-derivatives were employed in aggregate to assess the impact of *QGr.cib-5A* as well as the degree of its transmissibility. Based on the best discriminative marker couple *5A4-15* and *55k-31*, 96.27% of these lines with genotyping results of “ZKM138” could be traced back to ZKM138, demonstrating that this locus could be better used for a small breeding program

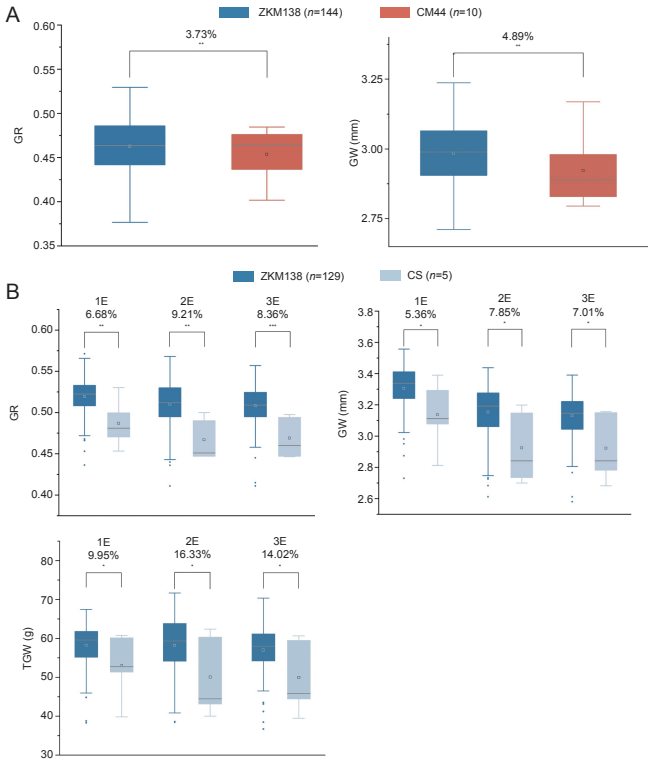


Fig. 4 Effects of *QGr.cib-5A* on grain roundness (GR) and related traits under similar genetic backgrounds. A, effects of *QGr.cib-5A* on GR and grain width (GW) in the BC-recombinant inbred lines (RILs). B, effects of *QGr.cib-5A* on GR, GW and thousand-grain weight (TGW) in the Zhongkemai 138 (ZKM138) derivatives. ZKM138, CM44 and CS indicate the lines with alleles from ZKM138, Chuanmai 44 (CM44) and Chinese Spring (CS), respectively. *, ** and *** represent significances at $P<0.05$, $P<0.01$ and $P<0.001$, respectively.

which included ZKM138 (Fig. 4-B). Meanwhile, the ZKM138 group exhibited an average 7.38% increase in GR ($P<0.05$) compared to the other recessive one, which confirmed the positive effect of the ZKM138-derived allele. Notably, the rounder-shaped group also had statistically broader GW and higher TGW, indicating that the ZKM138-derived allele might also make grains larger and heavier.

3.7. Geographic patterns of *QGr.cib-5A*

Outside the scope of the ZKM138 breeding programs, 117 Chinese cultivars covering most parts of the wheat-producing regions were acquired to determine the effect and geographical spread of the same allele throughout China (Appendix J). Notably, *QGr.cib-5A* also exerted a clear enhancement of roundness even in those complicated genetic backgrounds (Appendix K). The mean frequency of this locus was 76.07%, with the Southwestern Autumn-Sown Spring Wheat Zone having the highest value. To assess its global extent, 158 germplasm varieties from 15 other countries were used to navigate its dispersion pathway (Appendix L). The samples with genotypes consistent with ZKM138 covered all four continents, making up 65.19% of the genotyped lines, among which the frequency of this locus fluctuated from 1.08% (Africa) to 41.77% (Asia). According to these results, this ZKM138-derived superior allele has been subjected to moderate artificial selection during wheat breeding and may potentially contribute to the rounder morphology of existing varieties.

4. Discussion

4.1. Comparing *QGr.cib-5A* to previous reports

The present study investigated wheat grain morphology using an F_2 population and a RIL population from a ZKM138×CS cross. A stable QTL for grain roundness, *QGr.cib-5A*, was identified on chromosome 5A with a 10.65 Mb physical interval between 444.8 and 455.5 Mb. Its mapped position was similar to that of *QKdr-WJ-5A-1.4* (~455.4 Mb), a major and stable QTL for KDR (GL/GW ratio) that accounted for more than 10% of the PVE with LOD values above 3.0 against a Weimai 8×Jimai 20 (WJ) background (Cui *et al.* 2011a). The higher LOD and PVE values of *QKdr-WJ-5A-1.4* in WJ-RIL indicated that the genetic effects of the QTL for GR were dependent on specific genetic backgrounds. However, Cui *et al.* (2011a) did not further explore this QTL. Given the significant differences in grain traits between ZKM138 and CS (Fig. 1-A), a complex gene network might account for the variations in grain shape. Though affected by numerous genes, *QGr.cib-5A* was also the leading and stable QTL for GR in the ZKM138×CS background. Thus, this research can serve as a supplement for validating its genetic effects and for determining the pattern of its presence.

4.2. *QGr.cib-5A* shows robust stability in various genetic backgrounds

The reality of *QGr.cib-5A* and its stable effects on GR were validated in the BC-RILs using the *55k-31* KASP marker. Prominent differences were observed between two genotypic groups in terms of GW ($P<0.05$) and GR ($P<0.01$), respectively (Fig. 4-A). In addition, a pair of markers, *5A4-15* and *55k-31*, could split the ZKM138 derivatives into two groups, where the ZKM138 and CS group exhibited statistically significant differences in the aspects of TGW, GW, GR (Fig. 4-B; Appendix M). The discriminative power was also manifested in the natural accessions with genetic backgrounds far from that of ZKM138. The GR of the ZKM138 lines was higher than the CS lines, with an average increase up to 2.32% (Appendix K). These results suggested that the superior allele from ZKM138 could exert stable effects on GR in various genetic backgrounds and proved that the KASP markers used here were tightly linked to GR. Therefore, the use of these two KASP markers can accelerate the marker-assisted selection (MAS) programs aimed at improving wheat grain shape, thereby enhancing its market value.

4.3. *QGr.cib-5A* might have been indirectly selected during wheat breeding

Wheat grain shape does not seem to have been an artificial selection hotpot, unlike in other cereals such as rice, where selection for grain size and shape was linked together in the domestication process due their direct impacts on yield

and quality (Zhao *et al.* 2022; Yang *et al.* 2024). However, the physical appearance of wheat has transformed from the long and thin grain of ancient species to the much broader and shorter modern grain of commercial cultivars (Kumar *et al.* 2016). This was likely due to the reduction in genetic diversity during wheat evolution (Gegas *et al.* 2010). The same also holds true for CS and ZKM138. CS is thought to be a selection from the CDGT landrace, which is a primitive variety characterized by a more abundant pool of genes that regulate shape (Liu *et al.* 2018). While it is a released variety primarily planted in Chengdu, ZKM138 has been exposed to continuous selection, so it has more superior allelic variations controlling grain shape that have effectively spread to the southwest and most wheat zones in China. Regarding the especially high presence of *QGr.cib-5A* in southwestern China, one of ZKM138's parents (CM42) might have been the driver because it also carries this allele (Appendix J). Moreover, due to the direct effects of *QGr.cib-5A* on GR and its potential effects on GW and TGW, most breeders worldwide also selected this allelic variation with the same goal of cultivating larger, heavier and more uniform grains. As a result, more than half of the varieties tested in this study carry this favorable allele. Accordingly, the allele located in *QGr.cib-5A* was relatively strongly selected by breeders into modern lines both in China and around the world.

4.4. The round shape could be an alternative breeding target

In recent decades, researchers have amassed abundant evidence that grain weight gain is due more to grain size than shape traits (Wang *et al.* 2019; Liu *et al.* 2022). Few studies have focused on grain roundness. However, as a key indicator of grain shape, grain roundness also influences crop yield and quality. Larger and plumper grains contain more stored nutrient substances than smaller grains, thereby enhancing seed vigor and early growth as well as reducing the risk of environmental threats and achieving higher yield potentials (Su *et al.* 2011; Paolo *et al.* 2021). Moreover, grain shape is a key component of breeding programs due to its connection with market preferences and industry requirements, such as milling performance (i.e., flour quality and yield) (Kumar *et al.* 2016). For a given grain volume, optimizing the grain shape toward a rounder form could boost flour yield and quality to some extent. In the meantime, GR is a high-heritability trait (Appendix B) and a suitable breeding target. Therefore, the findings of this study will deepen our understanding of the genetic basis for grain roundness in wheat and provide an alternative perspective for round-shaped variety breeding programs to achieve greater food production.

4.5. Potential candidate gene(s) for *QGr.cib-5A*

Based on the flanking KASP markers in the *QGr.cib-5A* segment and differential SNPs within the parents, seven high-confidence genes were preliminarily selected from the IWGSC RefSeq v1.0 genome (Appendix I). From the expression pattern, gene annotation, and orthologue analyses, *TraesCS5A02G236400* emerged the most likely candidate gene associated with the regulation of grain development (Appendix J). One missense variant (G to A) occurred in the exon of the candidate gene, resulting in a Ser to Phe transition. According to IWGSC RefSeq v2.1, *TraesCS5A02G236400* encodes a hydroxyproline-rich glycoprotein family protein (HRGP), the same as its orthologous gene in rice, *Os09g0454200*. HRGP functions as a structural plant cell wall protein that participates in various aspects of growth and development (Kieliszewski *et al.* 2010; Liu *et al.* 2020). In *Arabidopsis thaliana*, one of its family members, leucine-rich repeat extensin 1 (LRX1) is a vital part of cell wall formation (Ringli 2010). Therefore, the fairly high expression of *TraesCS5A02G236400* in wheat reproductive tissues suggests its role in the formation of grain shape.

5. Conclusion

In this study, assisted by two linkage maps of F_2 and RIL populations from a ZKM138×CS cross, a QTL named *QGr.cib-5A* was consistently captured in all environments and the BLUE dataset. The gain in GW could be the main explanation for the rounder shape in these populations. The locus originating from ZKM138 has undergone successful selection during the small breeding programs which included ZKM138, as demonstrated by the dominant presence in its derivatives along with a notable increase in grain roundness. The geographic pattern analysis revealed that the superior allele was also relatively highly selected both in China and around the world. Hence, the detected *QGr.cib-5A* and its linked markers might be useful for improving the shape and increasing the market value of wheat grains, especially in the current situation where food security is the core of cereal research worldwide.

Acknowledgements

This work was supported by the National Key Research and Development Program of China (2023YFD1200400), the Key Project of Wheat Breeding in Sichuan Province, China (2021YFYZ0002) and the Science and Technology Co-operation Programme of Chengdu-Chinese Academy of Sciences. We sincerely thank Profs. Zhiyong Liu and Miaomiao Li in the Institute of Genetics and Developmental Biology of the Chinese Academy of Sciences for providing us with the DNA of global cultivars used in this study. We thank the Triticeae Multi-omics Center (http://202.194.139.32/) for providing the integrated platform of tools and genomic data that brought convenience to our work, and Tcuni Biotechnology Co., Ltd. (http://www.tcuni.com/) for the exome sequencing. Also, we thank the anonymous referees for their critical reading and revisions of the manuscript.

Declaration of competing interest

The authors declare that they have no conflict of interest.

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

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

Appendices associated with this paper are available at https://doi.org/10.1016/j.jia.2024.11.030

References

Ammiraju J S S, Dholakia B B, Santra D K, Singh H, Lagu M D, Tamhankar S A, Dhaliwal H S, Rao V S, Gupta V S, Ranjekar P K. 2001. Identification of inter simple sequence repeat (ISSR) markers associated with seed size in wheat. *Theoretical and Applied Genetics*, **102**, 726–732.

Assadzadeh S, Walker C, McDonald L, Panozzo J. 2022. Prediction of milling yield in wheat with the use of spectral, colour, shape, and morphological features. *Biosystems Engineering*, **214**, 28–41.

Campbell K G, Bergman C J, Gualberto D G, Anderson J A, Giroux M J, Hareland G, Fulcher R G, Sorrells M E, Finney P L. 1999. Quantitative trait loci associated with kernel traits in a soft × hard wheat cross. *Crop Science*, **39**, 1184–1195.

Cao P, Liang X, Zhao H, Feng B, Xu E, Wang L, Hu Y. 2019. Identification of the quantitative trait loci controlling spike-related traits in hexaploid wheat (*Triticum aestivum* L.). *Planta*, **250**, 1967–1981.

Chen C, Chen H, Zhang Y, Thomas H R, Frank M H, He Y, Xia R. 2020. TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Molecular Plant*, **13**, 1194–1202.

Chen G, Zhu Z, Zhang F, Zhu J. 2012. Quantitative genetic analysis station for the genetic analysis of complex traits. *Chinese Science Bulletin*, **57**, 2721–2726.

Cheng R, Kong Z, Zhang L, Xie Q, Jia H, Yu D, Huang Y, Ma Z. 2017. Mapping QTLs controlling kernel dimensions in a wheat inter-varietal RIL mapping population. *Theoretical and Applied Genetics*, **130**, 1405–1414.

Cui F, Ding A, Li J, Zhao C, Li X, Feng D, Wang X, Wang L, Gao J, Wang H. 2011a. Wheat kernel dimensions: How do they contribute to kernel weight at an individual QTL level? *Journal of Genetics*, **90**, 409–425.

Cui F, Li J, Ding A, Zhao C, Wang L, Wang X, Li S, Bao Y, Li X, Feng D, Kong L, Wang H. 2011b. Conditional QTL mapping for plant height with respect to the length of the spike and internode in two mapping populations of wheat. *Theoretical and Applied Genetics*, **122**, 1517–1536.

Cui F, Zhao C, Ding A, Li J, Wang L, Li X, Bao Y, Li J, Wang H. 2014. Construction of an integrative linkage map and QTL mapping of grain yield-related traits using three related wheat RIL populations. *Theoretical and Applied Genetics*, **127**, 659–675.

Dholakia B B, Ammiraju J S S, Singh H, Lagu M D, Röder M S, Rao V S, Dhaliwal H S, Ranjekar P K, Gupta V S, Weber W E. 2003. Molecular marker analysis of kernel size and shape in bread wheat. *Plant Breeding*, **122**, 392–395.

Dong C, Zhang L, Chen Z, Xia C, Gu Y, Wang J, Li D, Xie Z, Zhang Q, Zhang X, Gui L, Liu X, Kong X. 2020. Combining a new exome capture panel with an effective varBScore algorithm accelerates BSA-based gene cloning in wheat. *Frontiers in Plant Science*, **11**, 1249.

Dong L, Wang F, Liu T, Dong Z, Li A, Jing R, Mao L, Li Y, Liu X, Zhang K, Wang D. 2014. Natural variation of *TaGASR7-A1* affects grain length in common wheat under multiple cultivation conditions. *Molecular Breeding*, **34**, 937–947.

Gegas V C, Nazari A, Griffiths S, Simmonds J, Fish L, Orford S, Sayers L, Doonan J H, Snape J W. 2010. A genetic framework for grain size and shape variation in wheat. *The Plant Cell*, **22**, 1046–1056.

Geng J, Li L, Lv Q, Zhao Y, Liu Y, Zhang L, Li X. 2017. *TaGW2-6A* allelic variation contributes to grain size possibly by regulating the expression of cytokinins and starch-related genes in wheat. *Planta*, **246**, 1153–1163.

Kieliszewski M J, Lamport D T A, Tan L, Cannon M C. 2010. Hydroxyproline-rich glycoproteins: Form and function. In: *Annual Plant Reviews*. John Wiley & Sons, USA. pp. 321–342.

Kumar A, Mantovani E E, Seetan R, Soltani A, Echeverry-Solarte M, Jain S, Simsek S, Doehlert D, Alamri M S, Elias E M, Kianian S F, Mergoum M. 2016. Dissection of genetic factors underlying wheat kernel shape and size in an elite × nonadapted cross using a high density SNP linkage map. *The Plant Genome*, **9**, doi: 10.3835/plantgenome2015.09.0081.

Lee W J, Pedersen J F, Shelton D R. 2002. Relationship of sorghum kernel size to physiochemical, milling, pasting, and cooking properties. *Food Research International*, **35**, 643–649.

Li F, Wen W, Liu J, Zhang Y, Cao S, He Z, Rasheed A, Jin H, Zhang C, Yan J, Zhang P, Wan Y, Xia X. 2019. Genetic architecture of grain yield in bread wheat based on genome-wide association studies. *BMC Plant Biology*, **19**, 168.

Liu D, Zhang L, Hao M, Ning S, Yuan Z, Dai S, Huang L, Wu B, Yan Z, Lan X, Zheng Y. 2018. Wheat breeding in the hometown of Chinese Spring. *The Crop Journal*, **6**, 82–90.

Liu X, McKenna S, Welch L R, Showalter A M. 2020. Bioinformatic identification of plant hydroxyproline-rich glycoproteins. *Methods in Molecular Biology*, **2149**, 463–481.

Liu X, Xu Z, Feng B, Zhou Q, Ji G, Guo S, Liao S, Lin D, Fan X, Wang T. 2022. Quantitative trait loci identification and breeding value estimation of grain weight-related traits based on a new wheat 50K single nucleotide polymorphism array-derived genetic map. *Frontiers in Plant Science*, **13**, 96743.

Lyu J, Wang D, Sun N, Yang F, Li X, Mu J, Zhou R, Zheng G, Yang X, Zhang C, Han C, Xia G M, Li G, Fan M, Xiao J, Bai M Y. 2024. The *TaSnRK1-TabHLH489* module integrates brassinosteroid and sugar signalling to regulate the grain length in bread wheat. *Plant Biotechnology Journal*, **22**, 1989–2006.

Mares D, Moss H, Ellison F. 1986. Effects of grain shape and size on milling yields in wheat. II. Experimental studies. *Australian Journal of Agricultural Research*, **37**, 4.

Meng L, Li H, Zhang L, Wang J. 2015. QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *The Crop Journal*, **3**, 269–283.

Mohler V, Albrecht T, Castell A, Diethelm M, Schweizer G, Hartl L. 2016. Considering causal genes in the genetic dissection of kernel traits in common wheat. *Journal of Applied Genetics*, **57**, 467–476.

Paolo D, Orozco-Arroyo G, Rotasperti L, Masiero S, Colombo L, de Folter S, Ambrose B A, Caporali E, Ezquer I, Mizzotti C. 2021. Genetic interaction of *SEEDSTICK*, *GORDITA* and *AUXIN RESPONSE FACTOR 2* during seed development. *Genes*, **12**, 1189.

Pomeranz Y, Afework S. 1984. The effects of kernel size in plump and shrunken kernels, and of sprouting, on kernel hardness in wheat. *Journal of Cereal Science*, **2**, 119–126.

Ringli C. 2010. The hydroxyproline-rich glycoprotein domain of the *Arabidopsis* LRX1 requires Tyr for function but not for insolubilization in the cell wall. *The Plant Journal*, **63**, 662–669.

Smith S E, Kuehl R O, Ray I M, Hui R, Soleri D. 1998. Evaluation of simple methods for estimating broad-sense heritability in stands of randomly planted genotypes. *Crop Science*, **38**, 1125–1129.

Song J, Xu D, Dong Y, Li F, Bian Y, Li L, Luo X, Fei S, Li L, Zhao C, Zhang Y, Xia X, Ni Z, He Z, Cao S. 2022. Fine mapping and characterization of a major QTL for grain weight on wheat chromosome arm 5DL. *Theoretical and Applied Genetics*, **135**, 3237–3246.

Su Q, Zhang X, Zhang W, Zhang N, Song L, Liu L, Xue X, Liu G, Liu J, Meng D, Zhi L, Ji J, Zhao X, Yang C, Tong Y, Liu Z, Li J. 2018. QTL detection for kernel size and weight in bread wheat (*Triticum aestivum* L.) using a high-density SNP and SSR-based linkage map. *Frontiers in Plant Science*, **9**, 1484.

Su Z, Hao C, Wang L, Dong Y, Zhang X. 2011. Identification and development of a functional marker of *TaGW2* associated with grain weight in bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics*, **122**, 211–223.

Sukumaran S, Lopes M, Dreisigacker S, Reynolds M. 2018. Genetic analysis of multi-environmental spring wheat trials identifies genomic regions for locus-specific trade-offs for grain weight and grain number. *Theoretical and Applied Genetics*, **131**, 985–998.

Varshney R K, Prasad M, Roy J K, Harjit-Singh N K, Dhaliwal H S, Balyan H S, Gupta P K. 2000. Identification of eight chromosomes and a microsatellite marker on 1AS associated with QTL for grain weight in bread wheat. *Theoretical and*

Applied Genetics, **100**, 1290–1294.
Wang X, Dong L, Hu J, Pang Y, Hu L, Xiao G, Ma X, Kong X, Jia J, Wang H, Kong L. 2019. Dissecting genetic loci affecting grain morphological traits to improve grain weight *via* nested association mapping. *Theoretical and Applied Genetics*, **132**, 3115–3128.
Williams K, Munkvold J, Sorrells M. 2013. Comparison of digital image analysis using elliptic Fourier descriptors and major dimensions to phenotype seed shape in hexaploid wheat (*Triticum aestivum* L.). *Euphytica*, **190**, 99–116.
Xu Y F, Li S S, Li L H, Ma F F, Fu X Y, Shi Z L, Xu H X, Ma P T, An D G. 2017. QTL mapping for yield and photosynthetic related traits under different water regimes in wheat. *Molecular Breeding*, **37**, 34.
Yang X, Lu J, Shi W J, Chen Y H, Yu J W, Chen S H, Zhao D S, Huang L C, Fan X L, Zhang C Q, Zhang L, Liu Q Q, Li Q F. 2024. *RGA1* regulates grain size, rice quality and seed germination in the small and round grain mutant *srg5*. *BMC Plant Biology*, **24**, 1–13.
Yang Z, Bai Z, Li X, Wang P, Wu Q, Yang L, Li L, Li X. 2012. SNP identification and allelic-specific PCR markers development for *TaGW2*, a gene linked to wheat kernel weight. *Theoretical and Applied Genetics*, **125**, 1057–1068.
Zhang D, Wang B, Zhao J, Zhao X, Zhang L, Liu D, Dong L, Wang D, Mao L, Li A. 2015. Divergence in homoeolog expression of the grain length-associated gene *GASR7* during wheat allohexaploidization. *The Crop Journal*, **3**, 1–9.
Zhang G, Wang Y, Guo Y, Zhao Y, Kong F, Li S. 2015. Characterization and mapping of QTLs on chromosome 2D for grain size and yield traits using a mutant line induced by EMS in wheat. *The Crop Journal*, **3**, 135–144.
Zhang Y, Liu J, Xia X, He Z. 2014. *TaGS-D1*, an ortholog of rice *OsGS3*, is associated with grain weight and grain length in common wheat. *Molecular Breeding*, **34**, 1097–1107.
Zhang Y, Wang J, Zhou X, Yu L, Liu X. 2008. Influence of impurity types and content on the bulk density of wheat. *Journal of Henan University of Technology* (Natural Science Edition), **1**, 7–11. (in Chinese)
Zhao D, Yang L, Liu D, Zeng J, Cao S, Xia X, Yan J, Song X, He Z, Zhang Y. 2021. Fine mapping and validation of a major QTL for grain weight on chromosome 5B in bread wheat. *Theoretical and Applied Genetics*, **134**, 3731–3741.
Zhao D, Zhang C, Li Q, Liu Q. 2022. Genetic control of grain appearance quality in rice. *Biotechnology Advances*, **60**, 108014.
Zhu J. 1995. Analysis of conditional genetic effects and variance components in developmental genetics. *Genetics*, **141**, 1633–1639.

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