

Linkage and Association Mapping Identifies Loci and Develops Kompetitive Allele Specific PCR Markers for Improving Wheat Grain Protein Content

Peng Jiang

Jiangsu Academy of Agricultural Sciences <https://orcid.org/0000-0001-7840-8362>

Peng Zhang

Jiangsu Academy of Agricultural Sciences

Lei Wu

Jiangsu Academy of Agricultural Sciences

Yi He

Jiangsu Academy of Agricultural Sciences

Chang Li

Jiangsu Academy of Agricultural Sciences

Hongxiang Ma

Jiangsu Academy of Agricultural Sciences

Xu Zhang (✉ xuzhang@jaas.ac.cn)

Jiangsu Academy of Agricultural Sciences

Research Article

Keywords: wheat, linkage mapping, large-scale association mapping, kompetitive allele specific PCR, favorable allele, protein

Posted Date: April 2nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-365411/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at Theoretical and Applied Genetics on August 10th, 2021. See the published version at <https://doi.org/10.1007/s00122-021-03913-z>.

Abstract

Wheat grain protein content (GPC) is an important quality indicator. The GPC of wheat grown in the middle and lower reaches of the Yangtze River is often low. Marker-assisted selection (MAS) is an effective tool for improving quantitative traits; however, except *Gpc-B1*, most markers have not been effectively applied in GPC improvement, although many related loci have been identified. Linkage analysis using a recombinant inbred line population from the cross of core parents of Ningmai 9 and Yangmai 158 and association mapping using the local cultivated varieties were performed and nine candidate intervals were identified. The appropriate kompetitive allele specific PCR (KASP) markers associated with GPC were successfully developed and applied in 1163 F_4 breeding lines. Three markers, *Kgpc-2B*, *Kgpc-2D*, and *Kgpc-4A*, were validated to be significantly related to GPC by large-scale association mapping, and they were combined to achieve the highest efficiency to enhance GPC. We applied these markers in 164 F_6 breeding lines and obtained 15 lines with high GPC, indicating their high selective efficiency. Further, strategies for gene exploration in the three significant intervals were proposed. These results were expected to provide a novel route for improving GPC in wheat quality breeding.

Key Message

Linkage and association mapping identified 9 candidate intervals for wheat GPC, and large-scale association mapping based on 9 corresponding KASP markers and 1163 F_4 breeding lines revealed 3 significant markers.

Introduction

Wheat is one of the most widely grown food crops and provides 20% of human food energy requirements and 25% of proteins consumed on a daily basis, worldwide (Wheat Initiative, www.wheatinitiative.org/). Grain protein content (GPC) of wheat is also related to the end-use quality. Therefore, improving GPC is an important breeding objective for programs that aim to increase the nutritional value and quality of wheat.

GPC is a quantitative indicator of an important multi-gene trait (Balyan et al. 2013), and the quantitative trait loci (QTL) associated with GPC have been reported on all chromosomes (Fatiukha et al. 2020; Groos et al. 2003; Kunert et al. 2007; Ma et al. 2012). Echeverry-Solarte et al. (2015) used 163 recombinant inbred lines (RILs) derived from two widely divergent parents and detected 11 QTL on 9 chromosomes, and two of these QTL had major effects, explaining up to 16.5% and 16.9% of phenotypic variation. Fatiukha et al. (2020) identified 12 GPC QTL in an RIL population derived from the elite durum cultivar Svevo and a wild emmer wheat accession Y12-3, and four major QTL with the highest LOD scores and relatively high and stable phenotypic variance explained (PVE) were found on chromosomes 4BS, 5AS, 6BS, and 7BL. The QTL on 6BS had similar physical position as that on the cloned QTL, *Gpc-B1* (Uauy et al. 2006). A worldwide bread wheat core collection, including 372 accessions, was genotyped using DArT, SSR, and single nucleotide polymorphism (SNP) markers on the whole wheat genome, and 14 stable markers across 15 chromosomes were found to be associated with GPC (Bordes et al. 2011). A genome-wide association study was conducted using 189 spring bread wheat genotypes from the International Center for Agricultural Research in the Dry Areas, and 37 significant GPC SNPs were identified, some of which also exhibited an effect on gluten content (Suliman et al. 2021). However, the mapping results were variable with different mapping populations, methods, and environments, and most studies mainly focused on the identification of related loci, but the consequent validation and application were seldom reported except for *Gpc-B1*.

SNP markers have emerged as powerful tools for many genetic applications due to their low assay cost, high genomic abundance, locus specificity, and codominant inheritance. Kompetitive allele specific PCR (KASP) assay provides a rapid and cost-effective way to precisely determine SNP genotypes that perfectly fit the marker-assisted selection (MAS) on a large-scale (Semagn et al. 2014). Rasheed et al. (2016) developed and validated 70 functional KASP markers that are associated with wheat adaptability, grain yield, quality, and biotic and abiotic stress resistances, revealing that the KASP assays were 45 times faster compared with gel-based PCR markers. Lin et al. (2021) identified a major QTL related to grain number per spikelet on chromosome 2D and generated a closely linked KASP marker to validate its effects in different genetic backgrounds. Zhang et al. (2021) identified 22, 22, and 23 loci associated with seedling resistance to leaf rust, adult plant resistance to leaf rust, and adult plant resistance to stripe rust, using GWAS, respectively, and they converted 12 associated SNPs into KASP markers and verified in bi-parental populations.

A combination of linkage mapping and GWAS analysis can provide richer and more accurate information for QTL detection. Dhakal et al. (2018) identified a QTL for wheat curl mite resistance on chromosome 6DS in the wheat cultivar TAM 112 using linkage and association analyses, thus demonstrating the effectiveness of this locus to reduce symptom severity. *QTrl.saw-2D.2* that controls total root length, was identified in the RILs from SY95-71×CH7034 using linkage mapping and was validated in another population comprising 215 diverse lines, which suggests an effective and reliable result (Zheng et al. 2019).

The middle and lower reaches of the Yangtze River is the largest soft wheat production area and the second-largest wheat belt in China (He et al. 2002). The quality traits of wheat, including GPC, have received more attention in recent years; however, limited progress has been made in

genetic improvement of GPC in this area (He et al. 2018; Zhao et al. 2020). In this study, we performed linkage and association mapping in the wheat grown in this area to identify the genetic loci related to GPC, and successfully developed and validated the corresponding KASP markers. On the basis of these results, we obtained some effective molecular markers which can be utilized in MAS for GPC in the middle and lower reaches of the Yangtze River, and are expected to provide assistance for the quality breeding.

Material And Methods

Materials

A RIL population derived from Ningmai 9 × Yangmai 158 (282 F₈ RILs) was used for linkage mapping. Ningmai 9 is a soft wheat variety, whereas Yangmai 158 is a hard wheat variety. The RILs and their parents were planted using an augmented design with single row plots, 60 seeds per row with 1.6 m row length, 0.25 m row spacing in the growing seasons of 2015–2016 (E1), 2016–2017 (E2), 2017–2018 (E3), and 2018–2019 (E4) at Nanjing, China. Here 30 RILs were replicated twice and randomly distributed across the population (Lan et al. 2015).

A natural population comprising 103 local varieties released in the period of 1972–2016 in the middle and lower reaches of the Yangtze River (Supplemental Table 1) was planted in a randomized complete block design experiment with the same plots as that of RIL population, and two replicates in the growing seasons of 2015–2016 (E1), 2016–2017 (E2), 2017–2018 (E3), 2018–2019 (E4), and 2019–2020 (E5) at Nanjing, China.

All mature grains were harvested for GPC testing, and the 1163 F₄ lines and 164 F₆ lines were harvested from the breeding population in 2019–2020 at Nanjing, China. The GPC was measured using a Perton DA7200 (Sweden) device, according to the AACC Method 39-10.

Genotyping

Total genomic DNA was isolated from young leaf tissues using the CTAB method (Porebski et al. 1997). The RILs and their parents were genotyped using Illumina 90 K SNP assay (Wang et al. 2014). The genetic map containing 2285 bin markers with a total length of 3002 cM has been published in a previous study (Jiang et al. 2020). The natural population was genotyped using Affymetrix 50 K assay (CapitalBio Technology, Beijing, China). All markers were blasted in IWGSC Reference Sequence v1.0 to obtain the physical positions (The International Wheat Genome Sequencing Consortium 2018).

Phenotypic analysis and QTL analysis

Preliminary statistical analysis and correlation analysis for phenotype were performed in Microsoft Excel 2016. The test for analysis of variance (ANOVA) was conducted using IBM SPSS Statistics 19.0. The broad sense heritability (h^2) was estimated according to He et al. (2016): $h^2 = \delta_g^2 / (\delta_g^2 + \delta_{g \times y}^2 / y + \delta_e^2 / ry)$. δ_g^2 was genetic variance, $\delta_{g \times y}^2$ was genotype × environment variance, δ_e^2 was error variance, and y was number of environments, r was number of replication in each environment.

QTL mapping was conducted using the BIP function in IciMapping 4.1, and the algorithm of inclusive composite interval mapping (ICIM) was selected (Li et al. 2007). The walking step for QTL detection was set as 0.1 cM, and the threshold of LOD scores was set as 2.5.

Association mapping

Quality control for the genotype data from Affymetrix 50 K assay was performed in TASSEL software V5.2.13 (Bradbury et al. 2007), and the SNP markers with a minor allele frequency ≤ 5% and missing data >10% were removed.

The population structure was evaluated using STRUCTURE software 2.3.4 (Hubisz et al. 2009). The number of subsets (k) was set from 1 to 8 with five iterations each. Markov Chain Monte Carlo was used with a burn-in period of 10,000 repeats and 100,000 repeats after burn-in. The number of subgroups in the population (ΔK) was evaluated according to Evanno et al. (2005). The kinship matrix was calculated using the VanRaden method (VanRaden 2008).

GWAS was performed using R software through the GAPIT package (Lipka et al. 2012), and Bayesian-information and linkage-disequilibrium iteratively nested keyway method was selected (Huang et al. 2018). Principle components (PCs) calculated using TASSEL software V5.2.13 were used as covariates to reduce false positives. The threshold of the significant P values was set at 10^{-3} .

Development, validation, and application of KASP markers

KASP primers which contained two allele-specific forward primers and one common reverse primer were designed based on the sequences of identified SNPs using PolyMarker (Ramirez-Gonzalez et al. 2015) and synthesized by Shanghai Sangon Biotech Co., Ltd. (Supplemental Table S2). The KASP assays were performed in 384-well PCR plates with 2.5 μL KASP 2X Reaction Mix, 0.07 μL assay mix (12 μL each allele-specific

forward primer [100 μ M], 30 μ L reverse primer [100 μ M], and 46 μ L TRIS [10 mM, pH 8.3]), and 2.43 μ L 30 ng μ L⁻¹ genomic DNA. The PCR program was as follows: 94°C for 15 min; 10 touchdown cycles of 94°C for 20 s, 61°C–55°C for 60 s (decreasing by 0.6°C per cycle); and 26 cycles of 94°C for 20 s and 55°C for 60 s. The PCR reactions were carried out in Hydrocycler¹⁶ water instrument for PCR (LGC Genomics, Beverly, USA), and the PCR Fluorescence detection was performed using a PHERAstar (BMG LABTECH, Germany) microplate reader. KlusterCaller software (LGC Genomics, Beverly, USA) was used to analyze the data.

All the F₄ lines were screened with the developed KASP markers, and a large-scale association mapping was performed to identify significant markers by TASSEL software V5.2.13 based on general linear model (GLM). Then we used the significant KASP markers to screen the F₆ lines to test the select effect.

Exploration of candidate genes

To investigate candidate genes within the identified intervals, high confidence (HC) genes were extracted from the JBrowse website (https://urgi.versailles.inra.fr/jbrowseiwgsc/gmod_jbrowse/) according to IWGSC Reference Sequence v1.0 (The International Wheat Genome Sequencing Consortium 2018). The gene expression pattern was determined by blasting the predicted gene-id with WheatEXP (<https://wheat.pw.usda.gov/WheatExp/>) (Ramirez-Gonzalez et al. 2018).

For sequencing the candidate genes, multiple pairs of primers were designed to amplify the whole gene using Primer 6.0, and the sequencing services were provided by Shanghai Sangon Biotech Co., Ltd.

Results

Phenotype analysis

The description of genotype for the RIL and natural populations is shown in Table 1. In the RIL population, the parent Ningmai 9 showed a lower GPC than Yangmai 158 in all environments, which was consistent with their production performance. The max–min and coefficient of variation (CV) were above 5%, except for E1. The GPC in the natural population also presented a great variation, with CV ranging from 6.28% to 10.40% across the environments. GPC showed a significant positive correlation among almost all the environments in RIL and natural populations (Table 2), and the heritability reached 0.56 and 0.74, respectively (Table 3). Both genotype and environment had significant influence on the GPC of the RIL and natural populations, and their interaction also had significant influence on GPC of the natural population, but not on that of the RIL population (Table 3).

QTL mapping

Using QTL mapping, we identified 17 QTL on 10 chromosomes, including 1A (1), 1B (1), 2B (1), 2D (1), 3A (2), 3B (3), 4A (1), 5A (2), 5B (3), and 7A (2), which explained 3.94%–9.21% of phenotypic variance (Table 4 and Fig. 1). The additive effect of eight QTLs, including *Qgpc-3B.1*, *Qgpc-3B.3*, *Qgpc-4A*, *Qgpc-5A.1*, *Qgpc-5A.2*, *Qgpc-5B.3*, *Qgpc-7A.1*, and *Qgpc-7A.2*, was from Ningmai 9 and that of the other 9 QTL was from Yangmai 158. Two of the QTLs, *Qgpc-2D* and *Qgpc-3B.2*, were identified in two environments, and the other QTL were identified only in single environment.

Association mapping

The natural population was genotyped using Affymetrix 50 K assay. After quality control, 36360 SNPs were retained for further analysis. The SNPs on the A, B, and D genomes were 11415, 13133, and 11202, respectively, and the position of other 610 SNPs was uncertain. The number of SNPs per chromosome ranged from 875 (chromosome 4D) to 2698 (chromosome 3B) (Supplemental Table S3). According to the population structure analysis, the population was divided into three distinct subgroups ($\Delta K = 3$) (Fig. 2A), and a similar result was obtained from the kinship matrix analysis (Fig. 2C). Approximately one fifth of the materials was assigned to different subgroups by the two programs because of the different statistical methods used (Supplemental Table S1).

Using association mapping, we obtained 17 chromosome intervals containing significant markers associated with GPC distributed on 14 chromosomes, including 1B (1), 1D (1), 2A (1), 2B (1), 2D (1), 3A (2), 3B (1), 3D(2), 4A (1), 4B (1), 5A (2), 6B (1), 7A (1), and 7D (1) (Table 5 and Fig. 3). We then divided the intervals into three types for further analysis: Type A, P values below 10^{-4} and in multiple environments; Type B, P values below 10^{-4} and in single environment; and Type C, P values below 10^{-3} and in multiple environments.

Development and validation of KASP markers for candidate intervals

To further utilize the related QTL and associated markers, we tried to convert them into friendly KASP marker. Firstly, nine of these intervals were selected out according to the following three criteria (Table 6): (1) QTL detected in multi-environments; (2) Type A associated intervals;

(3) repeatedly regions between QTL mapping and association mapping. Secondly, the loci with low homology in these intervals were chosen for marker development (Supplemental Table S2). At last, KASP genotyping was performed in the materials from RIL population and natural population to test the developed markers.

Association mapping was performed in a large breeding population (1163 F_4 lines) with the nine successfully developed KASP markers based on GLM (Fig. 4). Then, we compared the GPC of the lines with 1~9 GPC-increased alleles and GPC-decreased alleles, and the selected order of the markers was according to the P values from lowest to highest (Fig. 5). It was found that the difference between the lines with GPC-increased alleles and GPC-decreased alleles increased as the number of markers increased, and remained relatively stable as the three markers with lowest P values were used in selection. Therefore, the three markers of *Kgpc-2B*, *Kgpc-2D*, and *Kgpc-4A* with low P values ($<10^{-10}$) were applicable for GPC selection, and their combination presented more effective.

Application of significant KASP markers

Further, we used the three markers to test 164 F_6 lines, and 15 lines with GPC-increased alleles showed an average GPC of 14.85%, which was significantly higher than 13.15% for eight lines with GPC-decreased alleles (Fig. 6 and Supplemental Table S4), thus indicating a good selection effect.

Exploration of candidate genes in the significant intervals

The published IWGSC reference sequence provided detailed information for the identification of candidate genes, and 607, 42, and 235 high confidence (HC) genes were identified in the intervals of *Kgpc-2B*, *Kgpc-2D*, and *Kgpc-4A*, respectively (Supplemental Table S5). We also analyzed the homologous genes of *Gpc-B1*, and a homologous gene of *TraesCS4A01G242700* was found in the interval of *Kgpc-4A*. Further, *TraesCS4A01G242700* of the parents, Ningmai 9 and Yangmai 158, was sequenced (Supplemental Table S6), however, no sequence difference were found. Gene expression data of *TraesCS4A01G242700* was extracted from WheatEXP, indicating that *TraesCS4A01G242700* was expressed in the grain, leaf, spike, and roots, and it showed the highest expression level in roots (Supplemental Table S7). In addition, 23 of the HC genes in the interval of *Kgpc-2D* expressed in different tissues at reproductive stage, and 17 expressed in the grain were analyzed (Supplemental Table S8).

Discussion

MAS is an effective route for increasing wheat GPC in the middle and lower reaches of the Yangtze River

The middle and lower reaches of the Yangtze River is the second largest area for wheat production in China. It was considered as a region with low GPC with a warm and humid climate (He et al. 2002), and little attention was paid to the selection of high GPC in wheat breeding in this area (Zhang et al. 2013a). The quality traits of common parents in this area were not outstanding, and the GPC were mostly at a low level (Qi et al. 2012; Wen et al. 2020). However, high GPC received more attention in local markets due to a rise in demand in recent years (Guo et al. 2020).

In traditional breeding, a high GPC parent from other ecological areas was used to improve local varieties, and quality testing is usually performed after the agronomic selection in field, leading to some lines with excellent quality being abandoned in early breeding generations. In addition, many quality traits, including GPC, are sensitive to environment and show significant difference among years or locations (Conti et al. 2011; Li et al. 2009; Ma et al. 2012). Therefore, the evaluation in a single environment may be inaccurate. MAS is independent of the environment and has a great potential in wheat quality breeding. There have been many reports on the application of high-molecular weight glutenin subunit related to dough rheological properties (Gautam et al. 2020; Izadi-Darbandi and Yazdi-Samadi 2012; Zhang et al. 2013b). Many studies have shown that the heritability of GPC was relatively high (Conti et al. 2011; Fatiukha et al. 2020; Li et al. 2009), which was sufficient for genetic selection and consistent with our study. *Gpc-B1*, which was identified in a wild emmer wheat, was previously reported as being used for GPC selection (Carter et al. 2012; Joppa et al. 1997). However, it was absent in most Chinese varieties (Rasheed et al. 2016), indicating that it was difficult to be directly utilized in our practical breeding.

Varieties with low GPC also contained some GPC-increased alleles, and the lines with expected high GPC could be produced as polymerizing more GPC-increased alleles. In this study, neither the parents of the RIL population showed good performance in GPC. The additive effect of eight and nine QTL was contributed from Ningmai 9 and Yangmai 158, respectively, and some derived RILs, which polymerize their GPC-increased alleles showed high GPC, which is consistent with many previous studies (Fatiukha et al. 2020; Li et al. 2009). Therefore, it is more feasible to improve GPC for exploring and polymerizing the GPC-increased alleles in local varieties.

Comparison of the candidate intervals for GPC with previous studies

In this study, we identified 17 QTL and 17 associated intervals using linkage mapping and association mapping, respectively. Most of the loci had a similar or close physical position to those reported in previous studies, particularly the ones selected for KASP marker development. Li et

al. (2009) identified an interval on 1B related to GPC, grain hardness and wet gluten content, and the physical position was close to *Kgpc-1B*. The interval of *Xgwm210-Xwmc25* on 2B related to GPC, flour protein content, and some mixograph parameters covered *Kgpc-2B* (McCartney et al. 2006). Laperche et al. (2007) found a QTL controlling GPC and grain yield near *Ppd-D1* on 2D, and it was also close to *Kgpc-2D*. *Kgpc-3A* was physically close to *QGpc.caas-3A* as reported by Li et al. (2012), which also affected mixograph parameters. A GPC QTL on 3B was co-located with those controlling the quantity of *Glu-A1* and *Glu-A3* (Zhang et al. 2011) and was near *Kgpc-3B*. Wang et al. (2012) identified a QTL on 4A, which was close to *Kgpc-4A*. *QGpc-5A1* contributed by Jing 411 had a similar position to *Kgpc-5A* (Xu et al. 2012). Liu et al. (2017) identified an associated marker neighboring *Kgpc-7A*. For *Kgpc-7D*, no adjacent loci have been reported, whereas a QTL on flour protein content was nearby (Smith et al. 2011). Using high-density SNP assays, some QTL, such as *Kgpc-2D* and *Kgpc-5A*, were mapped to a small interval, which provided closer markers and facilitated gene exploration. Some adjacent loci were reportedly related to several quality traits and were worthy of further research and utilization. In addition, some loci, such as *Qgpc-5B.1* and *Qgpc-5B.2*, and the associated intervals on 3D, were first reported to enrich the genetic study of GPC.

Development of available KASP markers for GPC

For quantitative traits, polymerizing more loci may produce a better effect. However, as the number of unlinked loci increased, the breeding population needed to be large enough to obtain a target genotype. When the number of unlinked loci was three or four, the minimum population sizes were 293 and 1177, respectively. As the number of loci increased to five and six, the population sizes expanded rapidly to 4714 and 18861, respectively (Wang et al. 2007). In our practical breeding program, a large number of loci and population sizes can make selection impracticable due to the high labor and time cost involved; therefore, 3–4 loci may be acceptable.

In this study, we obtained three efficient molecular markers, and their combination could produce a good selection effect, which depended on the appropriate materials, a reliable mapping method, and strict validation. First, both the parents of the RIL population, Ningmai 9 and Yangmai 158, were the main varieties; parents in the middle and lower reaches of the Yangtze River, and approximately 80% of the approved varieties in this area were derived from them during the past decade (Jiang et al. 2021). Therefore, their genetic information is widely distributed in the present breeding lines. The natural population, including most commercial varieties that were used during 1972–2016, covered the main genetic information from the middle and lower reaches of the Yangtze River. Second, linkage mapping and association mapping are both powerful methods in marker–trait association analysis, and their combination has successfully identified many loci associated with important traits, such as wheat agronomic traits and disease resistance (Liu et al. 2018; Shi et al. 2020). The statistical power of linkage mapping and association mapping were significantly related to population size, and in this study, their statistic powers were estimated to be approximately 90% and 80%, respectively (Hu et al. 2014; Li et al. 2010). Third, substantial artificial selection along with genetic recombination occurred in the separated generations of F_2 – F_5 , and the trait stability in the breeding population increased constantly, whereas the genetic heterozygosity and richness degraded gradually along with the generations. F_4 lines showing relatively stable traits and rich genetic information were chosen for marker validation and produced broadly applicable markers. Moreover, the large population (>1000) provided rich variations and greatly improved the accuracy and precision of the genetic effect estimates and false discovery rate (Hu et al. 2014), thus producing a more reliable result. In addition, all the markers were based on KASP genotyping arrays, which substantially facilitated MAS for GPC.

In the application trial, 15 lines with GPC-increased alleles were selected from 164 F_6 lines, and they were used as candidate lines with high GPC. In practical breeding, to obtain more candidate lines, it was better to apply the markers in early generations, such as F_2 and F_3 . F_2 enrichment increased the frequency of selected alleles and was a considerable strategy in MAS (Bonnert et al. 2005).

Exploration of candidate genes for GPC in the significant intervals

We obtained three significant markers by genetic analysis, and gene exploration in the intervals would be the next topic of study. On one hand, gene exploration is the precondition of the development of functional markers, whereas on the other hand, it can help to clarify the mechanism underlying grain protein synthesis. The physical sizes and HC genes differed greatly, so the method used for further gene exploration may be different. The database of high-quality genome and transcriptome provided great support (Kuzay et al. 2019; Zheng et al. 2019). *Gpc-B1*, which encoded an NAC transcription factor, has been previously cloned (Uauy et al. 2006). We found a homologous gene, *TraesCS4A01G242700*, in the interval of *Kgpc-4A*, but both the sequencing and expression pattern of *TraesCS4A01G242700* objected to a possible role in GPC accumulation. Therefore, fine mapping work should be considered for too many genes included in the interval of *Kgpc-4A* and that of *Kgpc-2B*. Combining the transcriptome data, 17 candidate genes expressed in the grain were identified in the interval of *Kgpc-2D*; however, further genetic and functional studies are needed.

Conclusion

In this study, we identified 17 QTL and 17 associated intervals by linkage and association mapping, respectively, and 9 of them were selected for KASP marker development. Large-scale association mapping was conducted on the basis of nine KASP markers and 1163 F₄ breeding lines, and the markers *Kgpc-2B*, *Kgpc-2D*, and *Kgpc-4A* were proved to be applicable for GPC selection and their combination was found to be more effective. The three markers were then used to test 164 F₆ lines, and candidate lines with high GPC were successfully obtained. Further, the strategies for gene exploration in the three significant intervals were discussed. These results are expected to be useful for wheat quality breeding in the middle and lower reaches of the Yangtze River.

Declarations

Acknowledgments This work was financially supported by the National Key Research and Development Program of China (2016YFD0101802 and 2017YFD0100806), and Jiangsu Agricultural Science and Technology Innovation Fund (CX(18)1001).

Author contribution statement PJ, HXM and XZ conceived and designed the research. PZ, LW and YH conducted the investigation of phenotype. PJ and CL conducted molecular marker analysis. PJ and PZ analyzed the data and wrote the article. HXM and XZ revised the article. All the authors read and approved the final version for publication.

Compliance with ethical standards

Conflicts of interest The authors declare that they have no conflict of interest.

References

1. Balyan HS, Gupta PK, Kumar S, Dhariwal R, Jaiswal V, Tyagi S, Agarwal P, Gahlaut V, Kumari S (2013) Genetic improvement of grain protein content and other health-related constituents of wheat grain. *Plant Breeding* 132:446-457
2. Bonnett D, Rebetzke G, Spielmeier W (2005) Strategies for efficient implementation of molecular markers in wheat breeding. *Molecular Breeding* 15:75-85
3. Bordes J, Ravel C, Gouis J, Lapierre A, Charmet G, Balfourier F (2011) Use of a global wheat core collection for association analysis of flour and dough quality traits. *Journal of Cereal Science* 54:137-147
4. Bradbury P, Zhang Z, Kroon D, Casstevens T, Ramdoss Y, Buckler E (2007) TASSEL: Software for Association Mapping of Complex Traits in Diverse Samples. *Bioinformatics* 23:2633-2635
5. Carter AH, Santra DK, Kidwell KK (2012) Assessment of the effects of the Gpc-B1 allele on senescence rate, grain protein concentration and mineral content in hard red spring wheat (*Triticum aestivum* L.) from the Pacific Northwest Region of the USA. *Plant Breeding* 131:62-68
6. Conti V, Roncallo P, Beaufort V, Cervigni G, Miranda R, Jensen C, Echenique V (2011) Mapping of main and epistatic effect QTLs associated to grain protein and gluten strength using a RIL population of durum wheat. *Journal of applied genetics* 52:287-298
7. Dhakal S, Tan C-T, Anderson V, Yu H, Fuentealba MP, Rudd JC, Haley SD, Xue Q, Ibrahim AMH, Garza L, Devkota RN, Liu S (2018) Mapping and KASP marker development for wheat curl mite resistance in "TAM 112" wheat using linkage and association analysis. *Molecular Breeding* 38:119
8. Echeverry-Solarte M, Kumar A, Kianian S, Simsek S, Alamri M, Mantovani E, McClean P, Deckard E, Elias E, Schatz B, Xu S, Mergoum M (2015) New QTL alleles for quality-related traits in spring wheat revealed by RIL population derived from supernumerary × non-supernumerary spikelet genotypes. *Theoretical and applied genetics* 128:893-912
9. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular ecology* 14:2611-2620
10. Fatiukha A, Filler N, Lupo I, Lidzbarsky G, Klymiuk V, Korol A, Pozniak C, Fahima T, Krugman T (2020) Grain protein content and thousand kernel weight QTLs identified in a durum × wild emmer wheat mapping population tested in five environments. *Theoretical and Applied Genetics* 133:119-131
11. Gautam T, Dhillon GS, Saripalli G, Rakhi, Singh VP, Prasad P, Kaur S, Chhuneja P, Sharma PK, Balyan HS, Gupta PK (2020) Marker-assisted pyramiding of genes/QTL for grain quality and rust resistance in wheat (*Triticum aestivum* L.). *Molecular Breeding* 40:49
12. Groos C, Robert N, Bervas E, Charmet G (2003) Genetic analysis of grain protein-content, grain yield and thousand-kernel weight in bread wheat. *Theoretical and Applied Genetics* 106:1032-1040
13. Guo R, Huang X, Wen M, Chen C, Liu J, Shan Y, Qu C, Qu D (2020) Effect of the Nitrogen Topdressing and Plant Density on Grain Yield and Quality of Spring Strong Gluten Wheat Cultivar Zhenmai 12. *Journal of Nuclear Agricultural Sciences* 34:1834-1839
14. He X, Lillemo M, Shi J, Wu J, Bjørnstad Å, Belova T, Dreisigacker S, Duveiller E, Singh P (2016) QTL Characterization of *Fusarium* Head Blight Resistance in CIMMYT Bread Wheat Line Soru#1. *Plos One* 11:e0158052

15. He Z, Lin Z, Wang L, Xiao Z, Wan F, Zhuang Q (2002) Classification on Chinese wheat regions based on quality. *Scientia Agricultura Sinica* 35:359-364
16. He Z, Qiaosheng Z, Cheng S, Yu Z, Zhao Z, Liu X (2018) Wheat Production and Technology Improvement in China. *Agronomy Journal* 8:99-106
17. Hu W, Kan H, Wang W, Xu C (2014) Statistical Genetics Approach for Functional Difference Identification of Allelic Variations and Its Application. *Acta Agronomica Sinica* 40:72
18. Huang M, Liu X, Yao Z, Summers R, Zhang Z (2018) BLINK: A package for the next level of genome-wide association studies with both individuals and markers in the millions. *GigaScience* 8
19. Hubisz M, Falush D, Stephens M, Pritchard J (2009) Inferring weak population structure with the assistance of sample group information. *Molecular ecology resources* 9:1322-1332
20. Izadi-Darbandi ALI, Yazdi-Samadi B (2012) Marker-assisted selection of high molecular weight glutenin alleles related to bread-making quality in Iranian common wheat (*Triticum aestivum* L.). *Journal of Genetics* 91:193-198
21. Jiang P, Zhang X, Wu L, He Y, Zhang PP, Ma HX, Kong LR (2021) Genetic analysis for yield related traits of wheat (*Triticum aestivum* L.) based on a recombinant inbred line population from Ningmai 9 and Yangmai 158. *Acta Agronomica Sinica* 47:869-881
22. Jiang P, Zhang X, Wu L, He Y, Zhuang W, Cheng X, Ge W, Ma H, Kong L (2020) A novel QTL on chromosome 5AL of Yangmai 158 increases resistance to *Fusarium* head blight in wheat. *Plant Pathology* 69:249-258
23. Joppa L, Du C, Hart G, Hareland G (1997) Mapping Gene(s) for Grain Protein in Tetraploid Wheat (*Triticum turgidum* L.) Using a Population of Recombinant Inbred Chromosome Lines. *Crop Science* 37:1586-1589
24. Kunert A, Naz AA, Dedek O, Pillen K, Léon J (2007) AB-QTL analysis in winter wheat: I. Synthetic hexaploid wheat (*T. turgidum ssp. dicoccoides* × *T. tauschii*) as a source of favourable alleles for milling and baking quality traits. *Theoretical and Applied Genetics* 115:683-695
25. Kuzay S, Xu Y, Zhang J, Katz A, Pearce S, Su Z, Fraser M, Anderson JA, Brown-Guedira G, DeWitt N, Peters Haugrud A, Faris JD, Akhunov E, Bai G, Dubcovsky J (2019) Identification of a candidate gene for a QTL for spikelet number per spike on wheat chromosome arm 7AL by high-resolution genetic mapping. *Theoretical and Applied Genetics* 132:2689-2705
26. Lan C, Zhang Y, Herrera-Foessel SA, Basnet BR, Huerta-Espino J, Lagudah ES, Singh RP (2015) Identification and characterization of pleiotropic and co-located resistance loci to leaf rust and stripe rust in bread wheat cultivar Sujata. *Theoretical and Applied Genetics* 128:549-561
27. Laperche A, Brancourt-Hulmel M, Heumez E, Gardet O, Hanocq E, Devienne-Barret F, Le Gouis J (2007) Using genotype × nitrogen interaction variables to evaluate the QTL involved in wheat tolerance to nitrogen constraints. *Theoretical and Applied Genetics* 115:399-415
28. Li H, Ye G, Wang J (2007) A modified algorithm for the improvement of composite interval mapping. *Genetics* 175:361-374
29. Li H, Zhang L, Wang J (2010) Analysis and answers to frequently asked questions in quantitative trait locus mapping. *Acta Agronomica Sinica* 36:918-931
30. Li Y, Song Y, Zhou R, Branlard G, Jia J (2009) Detection of QTLs for bread-making quality in wheat using a recombinant inbred line population. *Plant Breeding* 128:235-243
31. Li Y, Zhou R, Wang J, Liao X, Branlard G, Jia J (2012) Novel and favorable QTL allele clusters for end-use quality revealed by introgression lines derived from synthetic wheat. *Molecular Breeding* 29:627-643
32. Lin Y, Jiang X, Hu H, Zhou K, Wang Q, Yu S, Yang X, Wang Z, Wu F, Liu S, Li C, Deng M, Ma J, Chen G, Wei Y, Zheng Y, Liu Y (2021) QTL mapping for grain number per spikelet in wheat using a high-density genetic map. *The Crop Journal*
33. Lipka A, Tian F, Wang Q, Peiffer J, Li M, Bradbury P, Gore M, Buckler E, Zhang Z (2012) GAPIT: Genome Association and Prediction Integrated Tool. *Bioinformatics* 28:2397-2399
34. Liu J, Feng B, Xu Z, Fan X, Jiang F, Jin X, Cao J, Wang F, Liu Q, Yang L, Wang T (2017) A genome-wide association study of wheat yield and quality-related traits in southwest China. *Molecular Breeding* 38:1
35. Liu K, Sun X, Ning T, Duan X, Wang Q, Liu T, An Y, Guan X, Tian J, Chen J (2018) Genetic dissection of wheat panicle traits using linkage analysis and a genome-wide association study. *Theoretical and Applied Genetics* 131:1073-1090
36. Ma J, Zhang C, Yan G, Liu C (2012) Identification of QTLs conferring agronomic and quality traits in hexaploid wheat. *Journal of Integrative Agriculture* 11:1399-1408
37. McCartney CA, Somers DJ, Lukow O, Ames N, Noll J, Cloutier S, Humphreys DG, McCallum BD (2006) QTL analysis of quality traits in the spring wheat cross RL4452 × 'AC Domain'. *Plant Breeding* 125:565-575
38. Porebski S, Bailey L, Baum B (1997) Modification of CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Molecular Biology Reporter* 15:8-15

39. Qi L, Hu X, Zhou G, Wang S, Li J, Lu W, Wu L, Lu M, Sun L, Yang X, Song J, Wang B (2012) Analysis of wheat protein quality in main wheat producing areas of China from 2004 to 2011. *Scientia Agricultura Sinica*:4242-4251
40. Ramirez-Gonzalez R, Borrill P, Lang D, Harrington S, Brinton J, Venturini L, Davey M, Jacobs J, Ex F, Pasha A, Khedikar Y, Robinson S, Cory A, Florio T, Concia L, Juery C, Schoonbeek H-j, Steuernagel B, Xiang D, Uauy C (2018) The transcriptional landscape of polyploid wheat. *Science* 361
41. Ramirez-Gonzalez R, Uauy C, Caccamo M (2015) PolyMarker: A fast polyploid primer design pipeline. *Bioinformatics* 31
42. Rasheed A, Wen W, Gao F, Zhai S, Jin H, Liu J, Guo Q, Zhang Y, Dreisigacker S, Xia X (2016) Development and validation of KASP assays for genes underpinning key economic traits in bread wheat. *Theoretical and Applied Genetics* 129:1-18
43. Semagn K, Babu R, Hearne S, Olsen M (2014) Single nucleotide polymorphism genotyping using Kompetitive Allele Specific PCR (KASP): overview of the technology and its application in crop improvement. *Molecular Breeding* 33:1-14
44. Shi C, Zheng Y, Geng J, Liu C, Pei H, Ren Y, Dong Z, Zhao L, Zhang N, Chen F (2020) Identification of herbicide resistance loci using a genome-wide association study and linkage mapping in Chinese common wheat. *The Crop Journal* 8:666-675
45. Smith N, Guttieri M, Souza E, Shoots J, Sorrells M, Sneller C (2011) Identification and Validation of QTL for Grain Quality Traits in a Cross of Soft Wheat Cultivars Pioneer Brand 25R26 and Foster. *Crop Science* 51:1424
46. Suliman S, Alemu A, Abdelmula AA, Badawi GH, Al-Abdallat A, Tadesse W (2021) Genome-wide association analysis uncovers stable QTLs for yield and quality traits of spring bread wheat (*Triticum aestivum*.L) across contrasting environments. *Plant Gene* 25:100269
47. Uauy C, Distelfeld A, Fahima T, Blechl A, Dubcovsky J (2006) A NAC Gene Regulating Senescence Improves Grain Protein, Zinc, and Iron Content in Wheat. *Science* 314:1298-1301
48. VanRaden P (2008) Efficient Methods to Compute Genomic Predictions. *Journal of dairy science* 91:4414-4423
49. Wang J, Chapman S, Bonnett D, Rebetzke G, Crouch J (2007) Application of Population Genetic Theory and Simulation Models to Efficiently Pyramid Multiple Genes via Marker-Assisted Selection. *Crop Science* 47:582-590
50. Wang L, Cui F, Wang J, Jun LI, Ding A, Zhao C, Li X, Feng D, Gao J, Wang H (2012) Conditional QTL mapping of protein content in wheat with respect to grain yield and its components. *Journal of Genetics* 91:303-312
51. Wang S, Wong D, Forrest K, Allen A, Chao S, Huang B, Maccaferri M, Salvi S, Milner S, Cattivelli L (2014) Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. *Plant Biotechnology Journal* 12:787-796
52. Wen M, Li D, Chen C, Guo R, Lu P, Qu C (2020) Wheat Varieties in Regional Test in the Middle and Lower Reaches of the Yangtze River:Quality Analysis. *Chinese Agricultural Science Bulletin* 36:10-15
53. Xu Y, An D, Liu D, Zhang A, Xu H, Li B (2012) Molecular mapping of QTLs for grain zinc, iron and protein concentration of wheat across two environments. *Field Crops Research* 138:57-62
54. Zhang P, Yan X, Gebrewahid T-W, Zhou Y, Yang E, Xia X, He Z, Li Z, Liu D (2021) Genome-wide association mapping of leaf rust and stripe rust resistance in wheat accessions using the 90K SNP array. *Theoretical and Applied Genetics*
55. Zhang X, Zhang B, Zhu D, Lv G (2013a) Concept on Wheat Quality Improvement in Middle and Lower Reaches of Yangtze River Valley. *Journal of Triticeae Crops* 33:840-844
56. Zhang Y, Shen X, Zhang W, Chen X, Yan J, Zhang Y, Wang D, Wang Z, Liu Y, Tian Y, Xia X, He Z (2013b) Marker-Assisted Selection of HMW-Glutelin 1Dx5+1Dy10 Gene and 1B/1R Translocation for Improving Industry Quality in Common Wheat: Marker-Assisted Selection of HMW-Glutelin 1Dx5+1Dy10 Gene and 1B/1R Translocation for Improving Industry Quality in Common Wheat. *Acta agronomica sinica* 38:1743-1751
57. Zhang Y, Tang J, Zhang Y, Yan J, Xiao Y, Zhang Y, Xia X, He Z (2011) QTL mapping for quantities of protein fractions in bread wheat (*Triticum aestivum* L.). *Theoretical and Applied Genetics* 122:971-987
58. Zhao L, Song W, Yang X, Song Q, Zhang C, Wenli X, Xiao Z (2020) Application of Marker-assisted Backcrossing Breeding in Quality Improvement of Strong-gluten Wheat. *Chinese Agricultural Science Bulletin* 36:8-11
59. Zheng X, Wen X, Qiao L, Zhao J, Xiaojun Z, Li X, Zhang S, Yang Z, Chang Z, Chen J, Zheng J (2019) A novel QTL QTrl.saw-2D.2 associated with the total root length identified by linkage and association analyses in wheat (*Triticum aestivum* L.). *Planta* 250:129-143

Tables

Table 1 Phenotypic analysis of grain protein content in RIL and natural populations

Population	Environment	Ningmai 9 (%)	Yangmai 158 (%)	Max (%)	Min (%)	Mean (%)	Stdev	CV (%)
RIL population	E1	12.79	14.4	15.26	11.61	13.30	0.63	4.76
	E2	11.62	13.88	16.85	10.23	13.75	1.14	8.26
	E3	12.17	14.25	14.71	9.23	12.19	1.00	8.19
	E4	11.7	12.49	16.74	11.03	14.05	0.84	6.00
Natural population	E1	-	-	15.13	9.31	12.27	1.28	10.40
	E2	-	-	16.62	10.64	13.30	1.37	10.27
	E3	-	-	15.42	11.31	13.35	0.84	6.28
	E4	-	-	16.46	11.44	13.75	1.18	8.56
	E5	-	-	15.24	10.41	13.15	1.03	7.82

CV, coefficient of variation; RIL, recombinant inbred lines; Stdev, standard deviation

Table 2 Correlation analysis of grain protein content in the RIL and natural populations

	E1	E2	E3	E4	E5
E1	-	0.182**	0.215**	0.135*	-
E2	0.295**	-	0.528**	0.329**	-
E3	0.465**	0.321**	-	0.033	-
E4	0.479**	0.502**	0.371**	-	-
E5	0.378**	0.259**	0.387**	0.352**	-

* and ** indicates significant differences at the 0.05 and 0.01 probability level, respectively.

RIL, recombinant inbred lines

Table 3 ANOVA of grain protein content by F value

Item	RIL population	Natural population
Genotype	2.21**	9.14**
Environment	233.26**	85.94**
Genotype * Environment	0.91	2.34**
Heritability	0.56	0.74

** indicates significant differences at the 0.01 probability level.

Table 4 QTL analysis of grain protein content

No.	QTL	Environment	Genetic position (cM)	Physical interval (Mb)	Interval	LOD	PVE(%)	AE
1	<i>Qgpc-1A</i>	E3	19.95-25.35	15.4-18.5	IAAV1469~Tdurum_contig61492_684	3.23	4.84	-0.97
2	<i>Qgpc-1B</i>	E1	90.55-91.85	673.1-673.9	Tdurum_contig93330_263~Excalibur_c11190_617	3.03	4.48	-0.13
3	<i>Qgpc-2B</i>	E4	25.85-27.85	12.8-17.4	BobWhite_c19945_341~RAC875_rep_c69222_996	3.04	4.90	-0.16
4	<i>Qgpc-2D</i>	E2/E4	1.55-6.25	64.9-69.6	BS00073229_51~BobWhite_c47086_539	4.58	4.77	-0.27
5	<i>Qgpc-3A</i>	E4	53.65-54.35	624.5-686.8	GENE-3346_167~RAC875_rep_c109433_782	3.15	5.01	-0.21
6	<i>Qgpc-3A</i>	E2	69.65-73.75	697.2-701.0	Kukri_c49927_151~Excalibur_c96921_206	4.96	5.67	-0.28
7	<i>Qgpc-3B.1</i>	E3	59.55-62.45	24.9-101.4	BS00070456_51~BS00032830_51	4.21	5.21	1.02
8	<i>Qgpc-3B.2</i>	E2/E3	68.15-71.75	39.9-133.9	BS00087825_51~BS00092492_51	6.93	6.91	-0.32
9	<i>Qgpc-3B.3</i>	E4	106.25-106.4	146.8-153.7	IAAV1079~Kukri_c65979_1410	5.00	7.48	0.33
10	<i>Qgpc-4A</i>	E1	13.55-15.75	544.8-570.3	wsnp_Ex_c24443_33688268~GENE-2947_175	3.00	4.50	0.13
11	<i>Qgpc-5A.1</i>	E2	16.45-17.15	390.3-398.2	RAC875_rep_c113386_131~wsnp_Ex_c14839_22961659	4.05	3.94	0.23
12	<i>Qgpc-5A.2</i>	E1	43.65-46.05	503.1-503.8	GENE-3314_78~IAAV5294	3.89	5.84	0.14
13	<i>Qgpc-5B.1</i>	E1	6.45-8.95	4.8-9.0	BS00065164_51~Kukri_c65921_274	3.75	5.57	-0.14
14	<i>Qgpc-5B.2</i>	E4	45.35-46.15	545.4-546.9	Tdurum_contig25513_195~wsnp_Ex_c8659_14515623	3.23	4.74	-0.16
15	<i>Qgpc-5B.3</i>	E4	129.25-129.75	682.6-697.7	Excalibur_c6346_266~IAAV8023	6.07	9.14	0.24
16	<i>Qgpc-7A.1</i>	E3	55.25-56.8	205.4-210.9	Kukri_c10197_186~Tdurum_contig4676_3592	4.65	5.77	1.07
17	<i>Qgpc-7A.2</i>	E2	0-2.55	670.7-679.9	Ku_c62742_888~wsnp_Ku_c42539_50247333	9.11	9.21	0.39

Positive effect and negative effect were contributed by Ningmai 9 and Yangmai 158, respectively.

PVE, phenotypic variance explained; AE, additive effect; QTL, quantitative trait loci, LOD, logarithm of the odds

Table 5 Association mapping for grain protein content

Type	Chromosome	Physical interval (Mb)	Peak marker	P value	Environment
A	1B	633.1-656.7	AX-109109358	1.99×10 ⁻⁵	E1/E2
	3A	682.2-727.3	AX-109840225	1.04×10 ⁻⁵	E1/E2/E5
	7D	232.6-280.4	AX-110444859	6.02×10 ⁻⁵	E4/E5
B	1D	460.7-471.4	AX-94759528	5.97×10 ⁻⁵	E1
	3A	1.9-9.9	AX-110464884	2.10×10 ⁻⁵	E4
	3B	0-4.4	AX-94806155	3.01×10 ⁻⁵	E4
	3D	0-5.5	AX-110403551	8.78×10 ⁻⁶	E4
	4A	543.2-543.3	AX-108860370	3.43×10 ⁻⁵	E4
	6B	6.6-7.8	AX-94768259	8.19×10 ⁻⁶	E5
	7A	688.9-694.1	AX-109879999	1.01×10 ⁻⁵	E3
	C	2A	763.3-770.0	AX-111461751	5.91×10 ⁻⁴
2B		31.4-49.4	AX-109372952	1.86×10 ⁻⁴	E2/E3/E5
2D		8.7-29.9	AX-94514671	2.47×10 ⁻⁴	E2/E3/E4/E5
3D		363.7-382.3	AX-94457474	2.31×10 ⁻⁴	E3/E5
4B		9.9-22.0	AX-110976982	1.19×10 ⁻⁴	E3/E4/E5
5A		2.8-14.3	AX-110963298	3.12×10 ⁻⁴	E4/E5
5A		504.8-512.1	AX-111524034	2.18×10 ⁻⁴	E2/E5

Table 6 Candidate markers for grain protein content based on QTL mapping and association mapping

Name	Marker	SNP	GPC-increased allele	GPC-decreased allele	Physical interval (Mb)	Method
<i>Kgpc-1B</i>	AX-111461175	A/G	G	A	633.1-656.7	A
<i>Kgpc-2B</i>	BobWhite_c19945_341	T/G	G	T	12.8-49.4	L+A
<i>Kgpc-2D</i>	BS00067757_51	A/G	A	G	64.9-69.6	L
<i>Kgpc-3A</i>	AX-89352667	T/C	C	T	682.2-727.3	L+A
<i>Kgpc-3B</i>	BS00087825_51	T/C	C	T	39.9-133.9	L
<i>Kgpc-4A</i>	GEN×102947_175	T/C	C	T	543.2-570.3	L+A
<i>Kgpc-5A</i>	Excalibur_rep_c111546_167	A/G	G	A	503.1-503.8	L+A
<i>Kgpc-7A</i>	AX-108794609	A/G	A	G	670.7-694.1	L+A
<i>Kgpc-7D</i>	AX-110278908	A/G	A	G	232.6-280.4	A

L: QTL mapping; A: association mapping

Figures

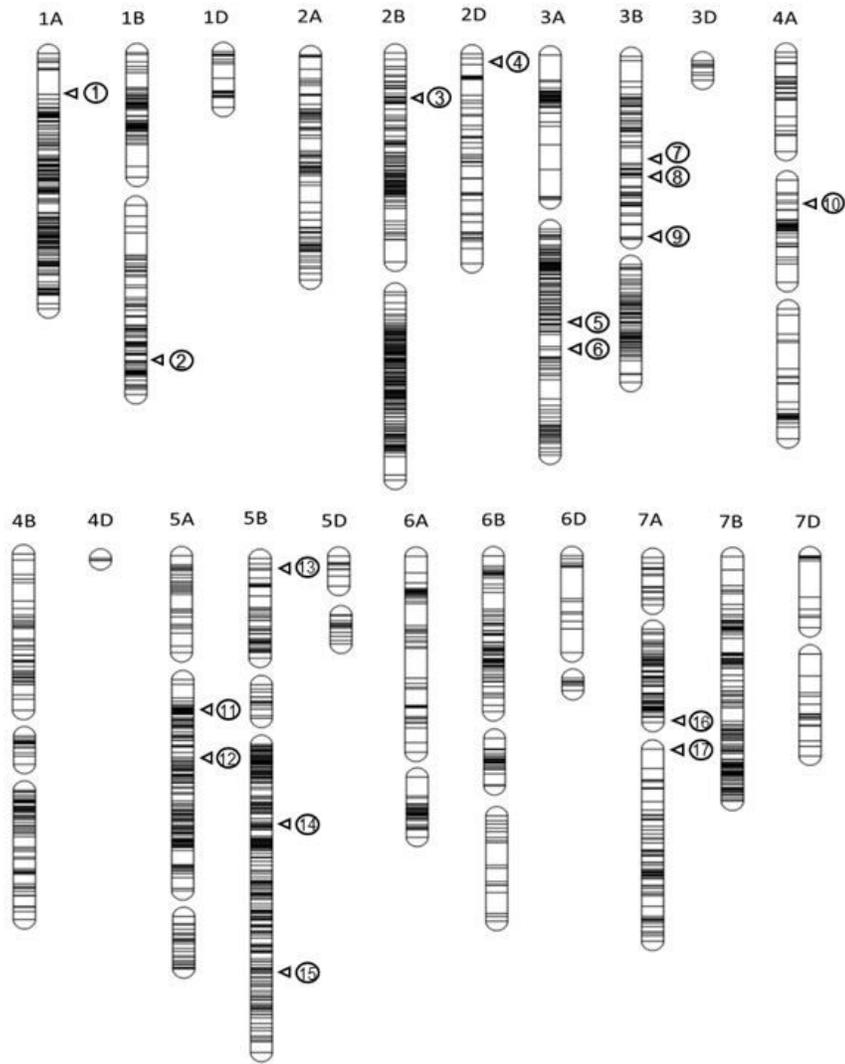


Figure 1

Quantitative trait loci (QTL) mapping for grain protein content. The numbers in the figure were consistent with the sequence numbers of QTL in Table 4.

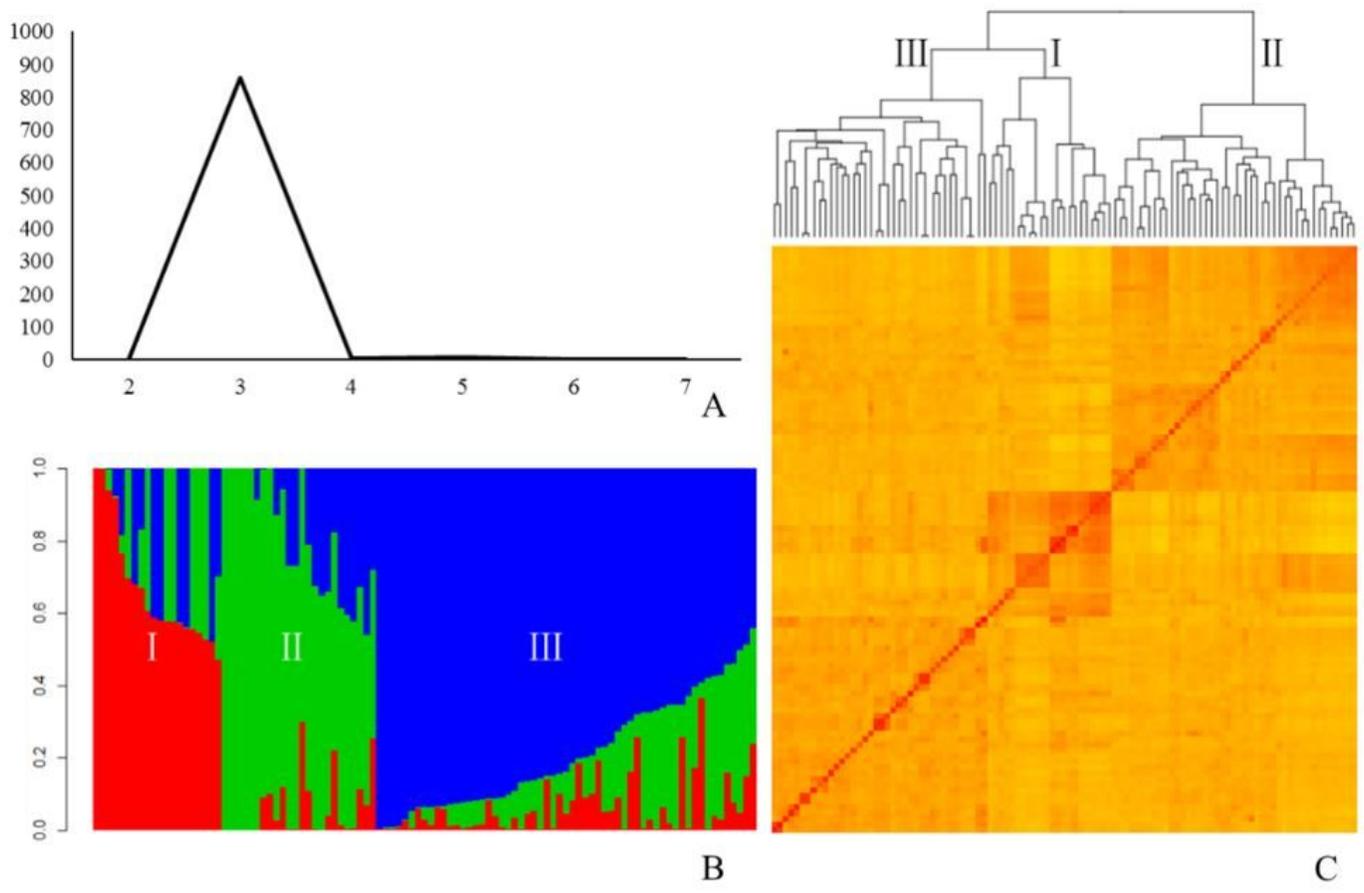


Figure 2

Population structure analysis of the entire population. A: Evaluation of the number of subgroups; B: Population structure of 103 wheat varieties; C: Kinship matrix of 103 wheat varieties.

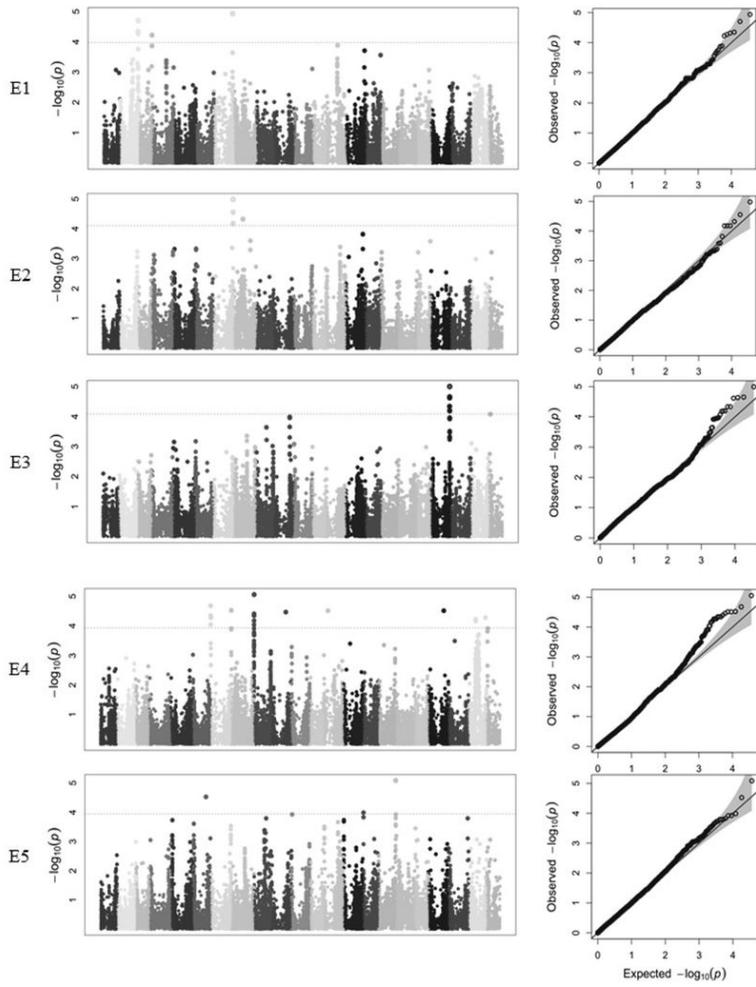


Figure 3

Associations of grain protein content in the natural population

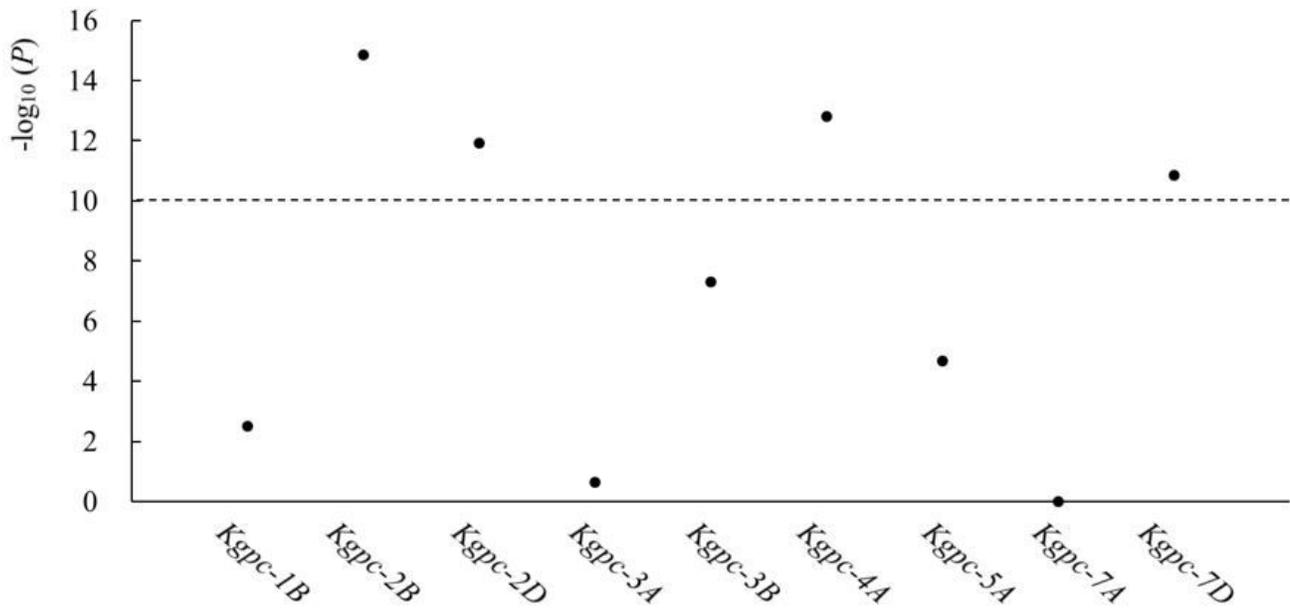


Figure 4

Significance of the candidate markers for grain protein content

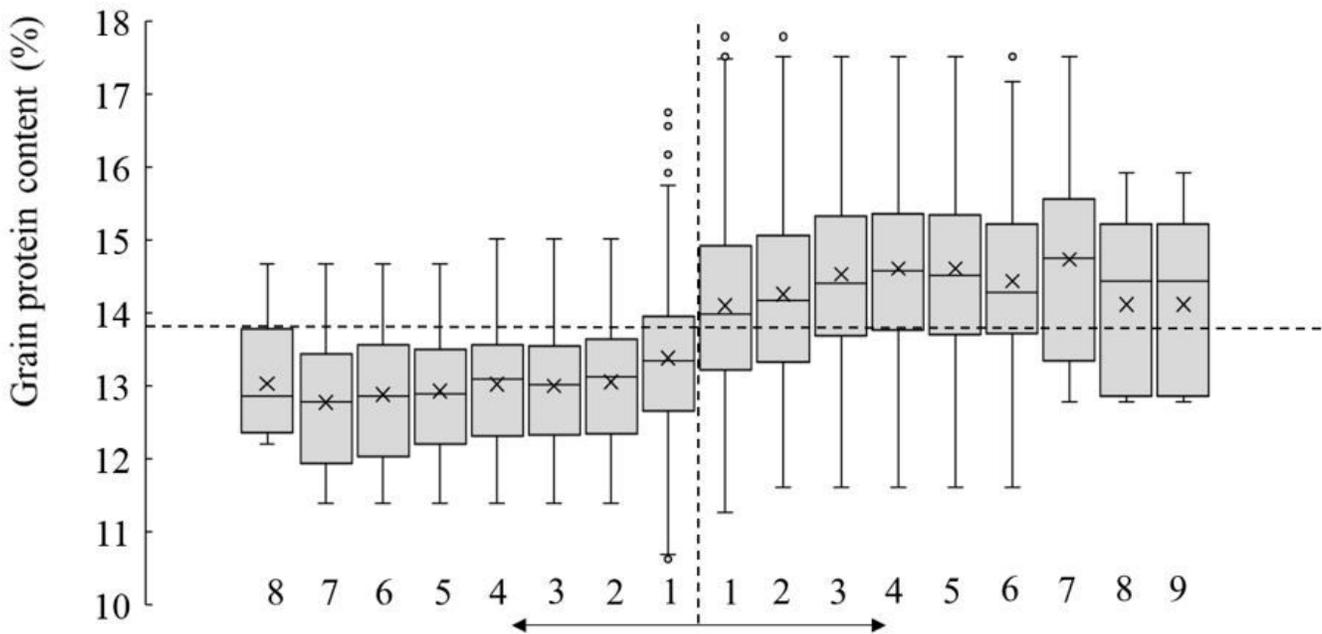


Figure 5

Comparison of grain protein content among the lines with different numbers of candidate markers. The numbers below the box indicated the numbers of candidate markers used for selection, and the select order of the markers were according to the P values from lowest to highest. The horizontal dashed line indicates the mean value of all lines; The horizontal line in the data box indicates the median; 'x' in the data box indicates the mean value.

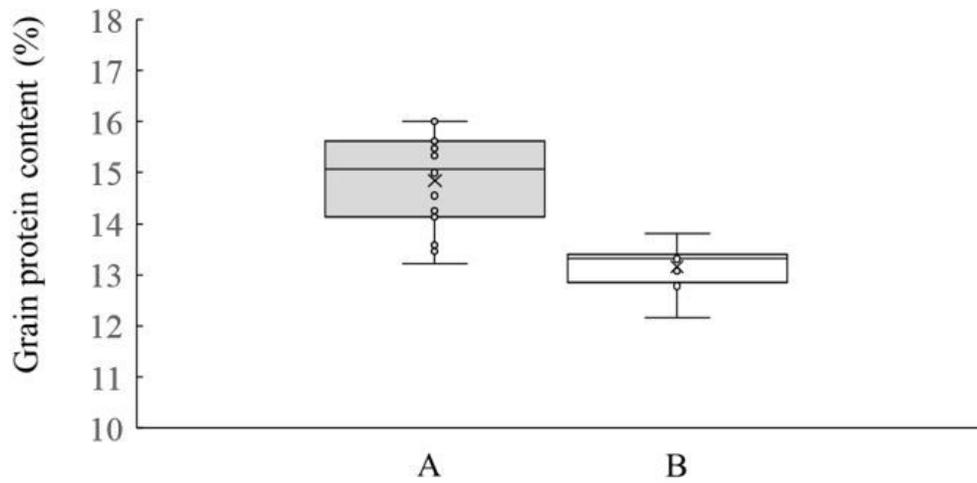


Figure 6

Comparison of grain protein content between the selected lines with the three markers A represented the lines with GPC-increased alleles; B represented the lines with GPC-decreased alleles. The horizontal line in the data box indicates the median; 'x' in the data box indicates the mean value.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementaltablex.xlsx](#)