Genetic dissection of grain iron concentration in hexaploid wheat (*Triticum aestivum* L.) using a genome-wide association analysis method (#68434)

First revision

Guidance from your Editor

Please submit by **28 Mar 2022** for the benefit of the authors .

Structure and Criteria
Please read the 'Structure and Criteria' page for general guidance.

Author notes
Have you read the author notes on the guidance page?

Raw data check
Review the raw data.

Image check
Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files
Download and review all files from the materials page.

1 Tracked changes manuscript(s)
1 Rebuttal letter(s)
3 Figure file(s)
4 Table file(s)
4 Raw data file(s)

Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- Impact and novelty not assessed. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- All underlying data have been provided; they are robust, statistically sound, & controlled.
- Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult. I suggest you have a colleague who is proficient in English and familiar with the subject matter review your manuscript, or contact a professional editing service.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Genetic dissection of grain iron concentration in hexaploid wheat (*Triticum aestivum* L.) using a genome-wide association analysis method

Jiansheng Wang [Equal first author, 1, 2, 3], **Xia Shi** [Equal first author, 2], **Zhengfu Zhou** [Corresp., 2], **Maomao Qin** [2], **Yahuan Wang** [2], **Wenxu Li** [2], **Pan Yang** [2], **Zhengqing Wu** [2], **Zhensheng Lei** [Corresp. 1, 2]

[1] College of Chemistry and Environment Engineering, Pingdingshan University, Pingdingshan, Henan Province, CHINA

[2] Wheat Research Institute, Henan Academy of Agricultural Sciences, Zhengzhou, Henan province, CHINA

[3] Henan Key Laboratory of Germplasm Innovation and Utilization of Eco-economic Woody Plant, Pingdingshan, Henan province, CHINA

Corresponding Authors: Zhengfu Zhou, Zhensheng Lei
Email address: zhouzf215@163.com, zhenshenglei@126.com

Iron (Fe) is an essential micronutrient of the body. Low concentrations of bioavailable Fe in staple food result in micronutrient malnutrition. Wheat (*Triticum aestivum* L.) is the most important global food crop and thus become an important source of iron for people. Breeding nutritious wheat with high grain –Fe has provides a method for alleviating malnutrition. Understanding the genetic basis of micronutrient concentration in wheat grains may provide useful information for breeding for high Fe varieties through marker-assisted selection (MAS). Hence, in the present study, a genome-wide association studies (GWAS) was conducted for grain Fe. An association panel of 207 accessions was genotyped using a 660K SNP array, and phenotyped for grain Fe content at three locations. The genotypic and phenotypic data obtained thus was utilized for GWAS. A total of 911 SNPs were significantly associated with grain Fe concentrations. These SNPs were distributed on all 21 wheat chromosomes, and each SNP explained from 5.79% to 25.31% of the phenotypic variations. Of particular note, the two significant SNPs (AX-108912427 and AX-94729264) that not only have a more significant effect on grain Fe concentration but have the reliability under the different environments. Furthermore, candidate genes potentially associated with grain Fe concentration were predicted, and 10 candidate genes were identified. These candidate genes were related to transport, translocation, remobilization, and accumulation of iron in wheat plants. These findings will not only help in better understanding the molecular basis of Fe accumulation in grains, but also provide elite wheat germplasms to develop Fe-rich wheat varieties through breeding.

Genetic Dissection of Grain Iron Concentration in Hexaploid Wheat (*Triticum aestivum* L.) Using a Genome-Wide Association Analysis Method

Jian Sheng Wang[1, 2, 3,#], Xia Shi[2#], Zheng Fu Zhou[2,*], Mao Mao Qin[2], Ya Huan Wang[2], Wen Xu Li[2], Pan Yang[2], Zheng Qing Wu[2], Zhen Sheng Lei[1,2,*]

[1]College of Chemistry and Environment Engineering, Pingdingshan University, Pingdingshan, 467000, Henan, China

[2]Wheat Research Institute, Henan Academy of Agricultural Sciences, Zhengzhou 450002, Henan, China

[3] Henan Key Laboratory of Germplasm Innovation and Utilization of Eco-economic Woody Plant, Pingdingshan, 467000, Henan, China.

[*]Corresponding authors: zhouzf215@163.com (Zhengfu Zhou); zhenshenglei@126.com (Zhensheng Lei).

[#] Contributed equally to this work.

ABSTRACT

Iron (Fe) is an essential micronutrient of the body. Low concentrations of bioavailable Fe in staple food result in micronutrient malnutrition. Wheat (*Triticum aestivum* L.) is the most important global food crop and thus become an important source of iron for people. Breeding nutritious wheat with high grain –Fe has provides a method for alleviating malnutrition. Understanding the genetic basis of micronutrient concentration in wheat grains may provide useful information for breeding for high Fe varieties through marker-assisted selection (MAS). Hence, in the present study, a genome-wide association studies (GWAS) was conducted for grain Fe. An association panel of 207 accessions was genotyped using a 660K SNP array, and phenotyped for grain Fe content at three locations. The genotypic and phenotypic data obtained thus was utilized for GWAS. A total of 911 SNPs were significantly associated with grain Fe concentrations. These SNPs were distributed on all 21 wheat chromosomes, and each SNP explained from 5.79% to 25.31% of the phenotypic variations. Of particular note, the two significant SNPs (AX-108912427 and AX-94729264) that not only have a more significant effect on grain Fe concentration but have the reliability under the different environments. Furthermore, candidate genes potentially associated with grain Fe concentration were predicted, and 10 candidate genes were identified. These candidate genes were related to transport, translocation, remobilization, and accumulation of iron in wheat plants. These findings will not only help in better understanding the molecular basis of Fe accumulation in grains, but also provide elite wheat germplasms to develop Fe-rich wheat varieties through breeding.

Abbreviations GWAS, Genome-Wide Association Study; SNP, Single-Nucleotide Polymorphism; LD, Linkage Disequilibrium; QTL, Quantitative Trait Loci; YY, Yuanyang; KF, Kaifeng; SQ, Shangqiu; Fe, Iron.

INTRODUCTION

Micronutrient malnutrition is caused by a lack of important micronutrients such as Fe. Due to its symptoms have few visible warning signs, so it is also called ‘hidden hunger’. This condition could impair the mental and physical development of humans and generate long-term effects on human health (*Alderman et al.*, *2006*). Microelement deficiencies are common in

developed and developing countries, and have become a major global health concern. Experts have estimated that one-third of the world population is at risk of Fe deficiency (*Alloway, 2009*). Women of child-bearing age and children are more prone to microelement deficiencies because they have greater micronutrient needs (*Grzeszczak et al., 2020*). Fe is a critical micronutrient, and has several important functions in the body such as its central role in transportation of blood oxygen, and reduced Fe intake can lead to impaired growth and behavioral problems (Welch & Graham, 1999). Therefore, adequate intake of essential minerals is important to eliminating 'hidden hunger'. Because the majority of the world's population depends on a few staple crops, such as wheat, rice, and maize, biofortification of these food crops seems to be a promising approach to address dietary mineral deficiencies (*Khush et al., 2012*). Several attempts have been made towards mineral improvement in plants, of which traditional breeding and genetic engineering techniques have been considered to be the most feasible and cost-effective approaches (*Bouis, 2003*).

Wheat (*Triticum aestivum* L.), one of the most important staple food crops in the world, contributes more than 50% of the diet and up to 60% of daily intakes of Fe and Zn in several developing countries (*Cakmak et al., 2010*). Because wheat has many advantages such as wide agronomic flexibility and ease of storage, billions of people depend on wheat for fulfilling their nutritional prerequisites. Hence, a sustainable way against mineral deficiencies is improvement of nutritional quality of wheat products through breeding approaches. Selection of wheat varieties with naturally high mineral content can easily facilitate this strategy. This requires a better understanding of the genetic basis of mineral element accumulation in wheat grains.

Inheritance of micronutrients is quantitative in nature. Linkage mapping and association analysis are useful methods in identifying QTLs for mineral elements. Linkage mapping generally involves specific populations such as recombinant inbred lines (RILs), doubled haploid (DH) lines, $F_{2:3}$ families, and backcross populations (BC_xF_y) to identify QTLs of target traits (*Groos et al., 2007; Wang et al., 2011*). However, these populations can only be used to identify QTLs for a limited number of traits. Furthermore, construction of these mapping populations is time-consuming and labor-intensive. Compared with linkage mapping, genome-wide association studies (GWAS) based on linkage disequilibrium (LD) capitalizes on historical recombination, and thus could identify more QTLs related to complex traits at a higher mapping resolution (Falconer & MacKay, 1996). Therefore, GWAS has become an important approach in QTL mapping for important and complex agronomic traits (Gao et al., *2016*). Various studies have been carried out to identify QTLs associated with wheat grain-Fe (*Arora et al., 2019; Alomari et al., 2018; Crespo-Herrera et al., 2017; Cu et al., 2020; Gorafi et al., 2016; Kumar et al., 2018; Peleg et al., 2009; Rathan et al., 2021; Roshanzamir et al., 2013; Shi et al., 2013; Srinivasa et al., 2014; Tiwari et al., 2016; Velu et al., 2017; Wang et al., 2021a; Wang et al., 2021b*).

Furthermore, *Uauy et al. (2006)* have reported a NAC gene (*NAM-B1*) that associated with increased grain Fe content.

In GWAS analysis, the low density of molecular markers could cause loss of linkages between markers and loci of target traits. High-density linkage maps are important to high-resolution QTL mapping and identification of candidate genes. Hence, plentiful molecular markers are imperative to construct wheat saturate genetic maps, and to significantly improve the efficiency of QTL mapping in GWAS. Compared with other types of molecular markers, single-nucleotide polymorphisms (SNPs) have many advantages such as the most abundant DNA sequence variation present in plant genomes, are virtually unlimited, evenly distributed along the genome, bi-allelic, and co-dominant (*Akpinar et al., 2017*), making SNPs ideal molecular markers in GWAS analysis. With the development of new sequencing technologies, methods of increasing the number of SNPs have been developed in wheat. Particularly, recently developed SNP gene chips have provided larger numbers of SNP markers. To date, SNP chips have been widely used in QTL analysis for important agronomic traits (*Gao et al., 2016; Cui et al., 2017*).

Elevation of essential mineral concentrations in grains is an effective strategy for improving the nutritional value of wheat to prevent micronutrient malnutrition. In the present study, we performed associative analysis with 207 wheat accessions from eight countries using 660K SNP chips to identify QTLs for the concentration of grain Fe. This study has two objectives. The first is to identify QTLs and candidate genes of essential microelements that may be useful for wheat biofortification. The second is to use the molecular markers significantly associated with major effect QTLs in marker-assisted wheat breeding, increasing the possibility of new elite wheat varieties with high grain Fe content.

MATERIALS & METHODS

Plant materials and field trials

The association panel used in the present study contained 207 wheat diverse accessions, comprising 194 accessions from the different wheat planting regions of China including Henan Province, Shannxi Province, Jiangsu Province, Shandong Province, Sichuan Province, Hebei Province, Shanxi Province, Beijing City, Anhui Province, Hubei Province, Guizhou Province, Yunnan Province, Ningxia Province, and Heilongjiang Province and 13 accessions from seven other countries, including Russia, France, Mexico, Japan, Australia, Bulgaria, and Romania. All of the accessions were grown at Yuanyang (YY, E 113°37′, N 35°12′), Kaifeng (KF, E 114°30′, N 34°80′) and Shangqiu (SQ, E 115°65′, N 34°45′) in Henan Province during the 2016 cropping seasons. Soil Fe content of the three experiment locations were measured. And the mean Fe

content of soil for Shangqiu, Yuanyang and Kaifeng was 26.7 mg/kg, 26.4 mg/kg and 26.1 mg/kg, respectively. Field trials were conducted in randomized complete blocks with three replicates at all locations. Each plot contained three 2-m rows spaced 20 cm apart. Agronomic management followed local practices. At wheat maturity, seeds were harvested separately for each accession under the different planted locations. In every location, one sample of seeds was collected for each replicated field plot and total three samples were obtained for each accession. Then the three samples were used to evaluate Fe concentration of each accession.

Determination of grain Fe concentration

Grains were harvested from each accession of the association population in the Kaifeng, Shangqiu and Yuanyang environments when they were matured. The following method was used for the analysis of grain Fe concentration (*Zarcinas et al., 1987*). First, grains were washed thoroughly with purified water three times to remove soil and dirt. Grains were dried in an oven at 70°C for 72 h and ground into a fine powder that could pass through a 1-mm screen. Then, 50 micrograms of powdered samples from each sample were microwave digested with 5mL nitric acid (HNO_3) and 2 mL hydrogen peroxide (H_2O_2) in polypropylene tubes using a microwave accelerate reaction system (CEM USA). Subsequently, Fe concentrations in the solutions were measured by a flame Atomic Absorption Spectrometer (AAS) (model 1100, Perkin-Elmer). Meanwhile, blank samples and standard samples were added each time for reference. All of the results represent the average of three replications.

Genotyping and quality control

The samples consisting of 207 wheat accessions were genotyped using the Affymetrix 660K SNP array comprising 630,517 SNPs and performed by Capital Bio-Corporation, Beijing, China (http://www.capitalbiotech.com/), following the manufacturer's protocol as described by *Akhunov et al.(2009*). To ensure the quality of genotyping data, sample call rate, SNP call rate, minor allele frequency (MAF), and Hardy-Weinberg equilibrium (HWE) were analyzed. In addition, accuracy was checked for SNP clustering, and manual adjustments were made for incorrectly clustered SNPs. The SNPs with a minor allele frequency (MAF) < 0.05 and missing data over 20% were excluded from further data analysis. The physical positions of SNP markers from 660K SNP arrays were obtained from the International Wheat Genome Sequencing Consortium website (IWGSC, http://www.wheatgenome.org/). Finally, a total of 224,706 high-quality SNP markers were used for GWAS analysis.

Genome-wide association analysis and haplotype analysis

In the present study, a total of 224,706 SNPs with a minor allele frequency ≥ 5% were used for the GWAS. Associations between genotypic and phenotypic data were analyzed using the kinship matrix by GAPIT package in R software (*Lipka et al., 2012*). The P value determining whether a SNP marker was associated with Fe concentration and R^2 were used to evaluate the magnitude of the MTA (significant marker-trait association) effects. GWAS was conducted for wheat grain Fe concentration at YY, KF, and SQ environments. Common haplotype patterns were assessed in Haploview version 4.2 and haplotype blocks were defined with the confidence interval method.

Candidate gene identification

In order to identify the candidate genes for SNP flanking regions, the flanking DNA sequences corresponding to the SNP markers significantly associated with Fe concentration were used in BLAST searches against the reported common wheat reference genome sequence in NCBI databases (https://blast.ncbi.nlm.nih.gov/Blast.cgi). The high confidence gene list of wheat was also obtained from the International Wheat Genome Sequence Consortium (IWGSC) website (https://wheat-urgi.versailles.inra.fr/) and used to identify possible candidate genes for each identified loci. The annotation of the candidate genes was accomplished with InterProScan (http://www.ebi.ac.uk/interpro/scan.html). The
transcript and the corresponding annotation of candidate genes were obtained from the website of IWGSC. For the loci that no candidates found in its mapping interval, the gene close to the peak SNP of the loci was assigned as the candidate.

RESULTS

Phenotypic variation for grain Fe concentration in wheat populations

Grain Fe concentration was tested for 207 genotypes in three environments (Supplemental Table S1). The range and mean of grain Fe concentration in the accessions are presented in Table 1 and Figure 1. Grain Fe concentration of accessions for the different environments are also depicted in Figure 1. Grain Fe concentration varied from 1.33–250.62 mg/kg at Kaifeng, 20.23–158.27 mg/kg at Shangqiu, and 28.15–143.78 mg/kg at Yuanyang. Mean Fe concentration was 99.57 mg/kg at Kaifeng, 74.00 mg/kg at Shangqiu, and 68.12 mg/kg at Yuanyang. Grain Fe concentration varied from 16.57 to 184.22 mg/kg (mean: 80.56 mg/kg) among the three locations. In Kaifeng, the highest Fe concentration was recorded in genotype KH438 (250.62 mg/kg),

followed by KH445 (245.85 mg/kg), and KH242 (237.74 mg/kg), whereas the lowest Fe concentration was recorded in genotype KH378 (1.33 mg/kg) followed by KH324 (4.67 mg/kg). In Shangqiu, the highest Fe concentration was recorded in genotype SH257 (158.27 mg/kg), followed by SH373 (134.24 mg/kg), and SH272 (133.85 mg/kg), but minimum Fe content was recorded in genotype SH418 (20.23 mg/kg), followed by SH349 (20.24 mg/kg). However, in Yuanyang, the highest Fe concentration was recorded in genotype AH418 (143.78 mg/kg), followed by AH429 (133.76 mg/kg), and AH413 (131.38 mg/kg), while the lowest Fe content was recorded in genotype AH249 (28.15 mg/kg), followed by AH240 (29.59 mg/kg). The mean Fe concentrations of all of the accessions from three locations could be described in decreasing order as follows: Kaifeng (99.57 mg/kg) > Shangqiu (74.00 mg/kg) > Yuanyang (68.12 mg/kg). On the basis of the above statistical analysis, wheat accessions with high Fe concentrations were recommended for breeding cultivars and selected as donors for Fe mineral biofortification in the future. Population distributions of GWAS accessions for grain Fe concentrations were continuous and exhibited a wide range of values for each location (Figure 1), which showed that the inheritance of grain Fe was consistent with the quantitative trait.

Genome-wide association analysis of wheat grain Fe concentration

Across the three locations, a total of 911 SNPs were significantly associated with grain Fe concentration (Figure 2, Supplemental Table S2), which were distributed across all 21 chromosomes. The phenotypic variation explained by each SNP ranged from 5.79% to 25.31%, suggesting that SNPs with moderate and minor effects on grain Fe concentration were detected. Association analysis for grain Fe concentration in three locations was further analyzed. In Yuanyang were identified 48 significantly SNPs on chromosomes 2A, 2B, 2D, 3A, 3B, 3D, 6A, and 6B. However in Kaifeng, 209 significantly SNPs were detected and they distributed on chromosomes 1D, 2A, 2B, 2D, 3A, 3B, 3D, 4A, 4B, 5A, 5B, 5D, 6A, 6B, 6D, 7A, 7B, and 7D. Compared with Yuanyang and Kaifeng, in Shangqiu were identified the more number of significant SNPs. Although total of 446 SNPs were identified in Shangqiu. each SNP had a lower explanation percentage for phenotypic variation. These SNPs were located on chromosomes 2A, 3A, 3D, 4A, 4D, 5A, 5B, 5D, 6A, 6B, 6D, 7A, 7B, and 7D. In addition, SNP number varied greatly across the different chromosomes, of which the highest number of SNPs was found on chromosome 5B (381) and the lowest on chromosome 4D and 6D (1), which suggests that the 5B chromosome is the main genetic region for grain Fe concentration. According to the flanking intervals of SNP, the identified SNPs could be categorized into 46 non-redundant QTLs (Table 2). The number SNPs in each non-redundant QTL covered was different and varied from 1 to 64 SNPs. The number of QTLs varied across different genomes, and the highest number of QTLs

was found in the A genome (18), followed by the B genome (16), while only 12 QTLs were identified in the D genome.

Significantly, two major SNPs were found on chromosome 4A and 3D. One SNP is AX-108912427; its physical position was 699,571,654 bp on chromosome 4A, and it explained 25.31% of the observed variation in Fe grain concentration. Another SNP, AX-94729264, could explain 24.84% of the observed phenotypic variations and was mapped to the physical position of 40,526,440 bp on chromosome 3D. In GWAS analysis, reliable SNPs that were simultaneously detected in more than two environments were considered more relevant inbreeding the new varieties with high grain Fe concentration. In the present study, nine reliable SNPs, including AX-94729264, AX-108912427, AX-94936962, AX-109956643, AX-111493816, AX-111088162, AX-109899864, AX-94702817, and AX-95210102, were detected in Kaifeng and Shangqiu environments. Of particular note, the two significant SNPs (AX-108912427 and AX-94729264) that not only have a more significant effect on grain Fe concentration but have the reliability under the different environments, so further study was conducted for the two SNPs.

Prediction of candidate genes

To understand the molecular mechanisms of Fe accumulation in wheat, candidate gene analysis was conducted for Fe. Candidate genes were predicted for the SNPs that were identified in GWAS analysis. An expression heat map was constructed for these candidate genes using the public database of Wheat Expression Browser (http://www.wheat-expression.com), and genes that were only specifically expressed in grain tissues were predicted as candidate genes for grain Fe concentration. Ten candidate genes were identified for grain Fe concentration (Table 3). On chromosome 3D, two candidate genes were found. One was TraesCS3D01G078500 that encoded for a NAC domain-containing protein, which showed potential relevance to metal remobilization and accumulation in wheat. Another gene is TraesCS3D01G080900, which encodes a defensin-like protein that has biological activities in ion channel blockage. Five candidate genes were found on chromosome 4A, including TraesCS4A01G430000, TraesCS4A01G431200, TraesCS4A01G431800, TraesCS4A01G431900, and TraesCS4A01G432000. Interestingly, TraesCS4A01G430000 and TraesCS4A01G432000 encode plant unknown function DUF581 family proteins that possibly participate in mineral translocation to seeds. However, TraesCS4A01G431800 and TraesCS4A01G431900 encoded an associated family protein (DUF581) that is related to senescence and participates in transporting Fe ions. Only one gene, TraesCS4A01G431200, was associated with acid phosphatase. Acid phosphatase plays a role in providing the desired energy for active transport and possibly affects Fe uptake and transport. TraesCS6A01G403500 on chromosome 6A encodes the remorin protein that may be associated

with the cytoskeleton or membrane skeleton and may play a role in regulating Fe translocation, while TraesCS6B01G449700 on chromosome 6B encodes a ring finger protein that is relevant to the accumulation of more Fe in grains.

Haplotypes associated with grain Fe concentration

Based on the genome-wide association analysis, we detected a significant cluster of 12 SNPs on chromosome 6B. In the corresponding region, 10 SNPs exhibited strong LD and could form two LD blocks (block1 and block2) (Table 4, Figure 3, Supplemental Table S3). The first large LD block spanned approximately 2,908 kb, including six SNPs with two haplotypes (Hap1A and Hap1B), and the second LD block spanned approximately 1,916 kb, including four SNPs with two possible haplotypes (Hap2A and Hap2B). To compare the effect of different haplotypes, we analyzed the effect of each haplotype on grain Fe concentration variation across the three environments. The variation of grain Fe concentration in Hap1 all reached the significant levels in the YY and SQ environments, but in Hap2, the significant level for grain Fe concentration was only observed in the SQ environment. In the KF environment, although Hap1 or Hap2 had non-significant effects on the grain Fe concentrations, the mean of grain Fe concentration with HapB was higher than that of HapA, which coincided with the distribution tendency in YY and SQ. The combined effect of the Hap1 and Hap2 blocks was also analyzed. In the combination of Hap1+Hap2, the variations in grain Fe concentration reached significant levels in YY and SQ. In addition, the combination Hap1A+Hap2A had higher frequency (70.73%) compared to haplotype Hap1B+Hap2B (29.27%).

Comparison with the previous studies

Several previous studies have been conducted for wheat grain Fe, and some QTLs were identified for this trait. We retrieved a total of 94 wheat Fe QTLs that were distributed among all 21 chromosomes except 1D (Supplemental Table S4). To compare the Fe QTLs of this study with that of previous studies, the physical positions of Fe QTLs were determined using the closest linked markers. Ultimately, the physical position of 57 QTLs was successfully determined (*Tong et al., 2020*). When we compared Fe QTLs of this study with those of the previous investigations, major differences in QTL amount or QTL physical position were observed. In the present study, a total of 46 Fe QTLs were identified, whereas total of 94 QTLs were found all of the previous studies. We found that the position of a few QTLs of this study nearly coincided with that of the previous studies. For example, *Gorafi et al. (2016)* and our group found one wheat Fe QTL near the 496 Mb position on chromosome 5D, and *Pu et al. (2014)* and our team also found one wheat Fe QTL near the 616 Mb position on chromosome 7D.

In addition, in the similar location for some QTL of the present study, several QTLs were identified in the previous studies, which suggest that the partly result of this study is accordance with the previous study. Except for the above QTLs, other QTLs of the present study were not identified in the previous studies. For instance, on chromosome 1D, we detected one QTL, but this QTL was found in previous studies.

DISCUSSION

Breeding elite wheat varieties with high microelement content has proven to be an effective biofortification method to address micronutrient malnutrition. The kernel as the most important organ of wheat, and provides useful information for biofortification breeding, and therefore has become the main study objective in Fe content research. Increasing micronutrient content in wheat through breeding requires the existence of substantial genetic variation for this trait. And selecting the accessions with high Fe grain concentration from the diverse germplasms is important in wheat breeding programs for enhancing Fe content. Several previous studies have been conducted for detecting the variation of wheat grain Fe content *(Garnett & Graham, 2005; Xu et al., 2012)*. The present study demonstrated that the association population contained the plentiful genetic variations in grain Fe concentration. The high Fe concentration in some wheat accessions, for example, KH438, KH445, and KH242, as the elite lines, could be used for developing new wheat varieties with high grain Fe concentration.

The understanding of the genetic basis of micronutrients accumulation in the wheat grain and mapping of the quantitative trait loci (QTL) will provide the useful information for improving grain micronutrient concentrations through marker assisted selection (MAS). The present associated mapping population exhibits a quantitative mode of inheritance for grain Fe concentration, which is in agreement with previous reports (*Tiwari et al., 2009; Srinivasa et al., 2014*). A large number of the previous QTL mapping studies have been conducted for wheat grain Fe (*Arora et al., 2019; Alomari et al., 2018; Cu et al., 2020; Gorafi et al., 2016; Rathan et al., 2021; Velu et al., 2017; Wang et al. 2021a; Wang et al. 2021b*). Earlier study conducted by *Pu et al. (2014)* reported five QTLs for grain Fe on chromosomes 2B, 4A, 5A, 5B and 7D in a recombinant inbred wheat lines, these QTLs were coincided with the loci AX-111087936, AX-109551612, AX-110549899, AX-109486399 and AX-95112608 identified in this study. Noteworthy, in our study two major SNPs with the highest phenotypic variation explained value were identified for wheat grain Fe concentration. One SNP (AX-108912427) was mapped to the near genomic region of one QTL on chromosome 4A reported by *Pu et al. (2014)*. Another major SNP (AX-94729264) locate on 3D chromosome that near to *QGFe.co-3D* reported by *Liu et al.(2019)*. Using the two wheat RIL mapping populations, *Velu et al. (2017)* identified four

QTLs for wheat grain Fe on chromosomes 2A, 2B, 5B and 7B. In the near genomic regions of the four QTLs, AX-94513201, AX-109274500, AX-108921926 and AX-110024541 were found in our research. Interestingly, we observed that the loci for wheat grain Fe were seldom found on 1D chromosome before. Until recently, *Rathan et al. (2021)* reported one QTL for wheat grain Fe on chromosome 1D. In our research, we found five significant SNPs for wheat grain Fe on chromosome 1D, which suggested that these SNPs were possible the first time report. In the current study, all significant SNPs scattered across the entire wheat genome. In the comparatively narrow region on 6B chromosome, we detected a significant cluster of 12 SNPs, which spanned approximately 4824 kb. This chromosome segment can be incorporated into breeding efforts in cultivating a high-Fe accumulation germplasm. Compared with the previous studies, we detected more loci for wheat grain Fe concentration, which possible due to the high-density linkage map constructed with a large numbers of gene-based SNP markers based on a 660K SNP array in this study.

Identifying genes within a QTL region can help elucidate trait architecture if gene function can be related to the associated trait. Although several studies have reported QTLs for grain Fe concentration in wheat, only a few candidate genes have previously been identified. The genetic mechanisms of wheat grain Fe concentration are presently unclear. In this study, by candidate gene analysis, we found 10 candidate genes for wheat grain Fe concentration. The functions of these candidate genes are re- -lated to uptake, transport, translocation, remobilization, and accumulation in wheat plants. Of the 10 candidate genes, TraesCS3D01G078500 encoded for a NAC domain containing protein. *Ogo et al. (2008)* isolated an NAC transcription factor, IDEF2, from rice and barley and revealed IDEF2 as a key transcription factor regulating the Fe deficiency response. *Uauy et al. (2006)* also reported one ancestral wild wheat allele, Gpc-B1, that is associated with increased Fe content and encodes an NAC transcription factor that accelerates senescence and increases nutrient remobilization from leaves to developping grains. Interestingly, in the identified ten candidate genes, four genes were related to senescence. nutrients, including most minerals, are mobilized from senescing leaves to other tissues during leaf senescence (*Himelblau & Amasino, 2001*). In wild-type wheat, Fe mobilization from the vegetative parts to the grain during leaf senescence and grain maturation appears to be quite efficient, so most of the total Fe is translocated to the grains (*Garnett & Graham, 2005*). The timing and efficiency of senescence-mediated nutrient mobilization appears to be a primary determinant for grain Fe content. Previous studies have indicated that the NAM-B1 gene, which encodes a member of the NAC transcription factor gene family, could play a central role in senescence and nutrient mobilization (*Uauy et al., 2006*). In this study, TraesCS4A01G430000, TraesCS4A01G432000, TraesCS4A01G431800, and TraesCS4A01G431900 encoded the associated family protein (DUF581). The four candidate genes are all related to senescence and

participate in transporting Fe ions. On 3D chromosome, another gene, TraesCS3D01G080900, encodes a defensin-like protein that has biological activities involving ion channel blockage. More than 500 defensin proteins have been discovered to date. Most plant defensins have been isolated from seeds, and these also been identified in vegetative tissues (*Gachomo et al., 2012*). TraesCS6A01G403500 encodes a remorin protein. Remorin has been found in detergent-insoluble membranes and may be associated with the cytoskeleton or membrane skeleton, which suggests that remorin may play a role in regulating Fe uptake.

Increasing grain Fe content in wheat has received more attention in recent years and become an important quality breeding objective in wheat practice around the world. However, to date, the genetic mechanism controlling grain Fe content is still unclear in wheat. In the present study, a total of 911 significant SNPs associated with grain Fe concentration were identified, and 10 candidate genes were predicted. Collectively, these findings have three potential uses. First, these findings will provide some useful wheat germplasms with high favourable alleles of Fe content for breeding the elite Fe enrich varieties. Second, the markers identified in this study could be utilized in molecular marker-assisted breeding (MAS) for biofortification of wheat to increase kernel-Fe content. The last and perhaps most promising is the utilization of important regions with stacking significant SNPs and candidate genes will facilitate the breeding of Fe enriched wheat varieties. Because most of the significant SNPs and candidate genes in this study were not reported before in wheat, more further studies are needed to validate our findings in the future.

CONCLUSIONS

In briefly, two major SNPs and nine reliable SNPs for wheat grain Fe were identified, and ten candidate genes were predicted. All significant SNPs identified in this study scattered across the entire wheat genome and one loci cluster found on 6B chromosome. The functions of the candidate genes are primarily associated with transport, translocation, remobilization, and accumulation of Fe in wheat plants. Therefore, this study provides useful loci and gene information for improving Fe concentration in wheat grain. In the future, further studies need be conducted for the candidate genes identified herein, which would be helpful to elucidate the molecular mechanisms of Fe content in the wheat grain.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This research was funded by the Joint Fund of National Natural Science Foundation of China

(U1804102), the Scientific and Technological Research Project of Henan Province (202102110027, 212102110065), and the Excellent Youth Fund of Henan Academy of Agricultural Sciences (2020YQ02, 2022YQ13), and China Postdoctoral Science Foundation (2020M682300). There was no additional external funding received for this study.

REFERANCES

Akhunov E, Nicolet C, Dvorak J. 2009. Single nucleotide polymorphism genotyping in polyploid wheat with the Illumina GoldenGate assay. *Theor Appl Genet* **119**: 507-17. DOI: 10.1007/s00122-009-1059-5.

Akpinar BA, Lucas S, Budak H. 2017. A large-scale chromosome-specific SNP discovery guideline. *Funct. Integr. Genomics* **7**: 97-105. DOI: 10.1007/s101420 160536-6.

Alderman H, HoddinottJ, Kinsey B. 2006. Long term consequences of early childhood malnutrition. *Oxford Economic Papers* **58**: 450-474. DOI: 10.1093/oep/gpl008.

Alloway BJ. 2009. Soil factors associated with zinc deficiency in crops and humans. *Environ Geochem Health* **31**: 537-548. DOI: 10.1007/s10653-009-9255-4.

Alomari DZ, Eggert K, Wirén N, Polley A, Plieske J, Ganal MW, Liu F, Pillen K, Röder MS. 2018. Whole-genome association mapping and genomic prediction for iron concentration in wheat grains. *Int. J. Mol. Sci* **20**:76. DOI: 10.3390/ijms20010076.

Arora S, Cheema J, Poland J, Uauy C, Chhuneja P. 2019. Genome-wide association mapping of grain micronutrients concentration in *Aegilops tauschii*. *Front. Plant Sci* **10**: 54. DOI: 10.3389/fpls.2019.00054.

Bouis HE. 2003. Micronutrient fortification of plants through plant breeding: can it improve nutrition in man at low cost? *Proc Nutr Soc* **62**: 403-411. DOI: 10.1079-/PNS2003262.

Cakmak I, Pfeiffer WH, McClafferty B. 2010. Biofortification of durum wheat with zinc and iron. *Cereal Chem* **87**:10-20. DOI: 10.1094/cchem-87-1-0010.

Crespo-Herrera LA, Govindan V, Stangoulis J, Hao YF, Singh RP. 2017. QTL mapping of grain Zn and Fe concentrations in two hexaploid wheat RIL populations with ample transgressive segregation. *Front. Plant Sci*. **8**: 1800. DOI: 10.3389/fpls.2017.01800.

Cui F, Zhang N, Fan XL, Zhang W, Zhao CH, Yang LJ, Pan RQ, Chen M, Han J, Zhao XQ, Ji J, Tong YP, Zhang HX, Jia JZ, Zhao GY, Li JM. 2017. Utilization of a wheat 660K SNP array-derived high-density genetic map for high-resolution mapping of a major QTL for kernel number. *Scientific Reports* **7**: 3788. DOI:10.1038/s41598-017-04028-6.

Cu ST, Guild G, Nicolson A, Velu G, Singh R, Stangoulis J. 2020. Genetic dissection of zinc,

433 iron, copper, manganese and phosphorus in wheat (*Triticum aestivum* L.) grain and rachis at
434 two developmental stages. *Plant Science* **291:** 110338. DOI:10.1016/j.plantsci.2019.110338
435 **Falconer DS, MacKay TFC. 1996.** *Introduction to quantitative genetics*. 4th Edition, Longman
436 Group Ltd., London..
437 **Gachomo EW, Jimenez-Lopez JC, Kayodé APP, Baba-Moussa L, Kotchoni SO. 2012.**
438 Structural characterization of plant defensin protein superfamily. *Mol Biol Rep* **39**: 4461-
439 4469. DOI: 10.1007/s11033-011-1235-y.
440 **Gao FM, Liu JD, Yang L, Wu XX, Xiao YG, Xia XC, He ZH. 2016.** Genome-wide linkage
441 mapping of QTL for physiological traits in a Chinese wheat population using the 90K SNP
442 array. *Euphytica* **209**: 789-804. DOI:10.1007/s10681-016-
443 -1682-6.
444 **Garnett TP, Graham RD. 2005.** Distribution and remobilization of iron and copper in wheat.
445 *Ann Bot* **95**: 817-26. DOI: 10.1093/aob/mci085.
446 **Gorafi YSA, Ishii T, Kim JS, Elbashir AAE, Tsujimoto H. 2016.** Genetic variation and
447 association mapping of grain iron and zinc contents in synthetic hexaploid wheat
448 germplasm. *Plant Genetic Resources* **1**: 1-9. DOI:10.1017/ S1479262116000265.
449 **Groos C, Bervas E, Chanliaud E, Charmet G. 2007.** Genetic analysis of bread-making quality
450 scores in bread wheat using a recombinant inbred line population. *Theor Appl Genet* **115**:
451 313-323. DOI:10.1007/s00122-007-0563-8.
452 **Grzeszczak K, Kwiatkowski S, Kosik-Bogacka D. 2020**. The Role of Fe, Zn, and Cu in
453 Pregnancy. *Biomolecules* **10**:1176. DOI: 10.3390/biom10081176.
454 **Himelblau E, Amasino RM. 2001.** Nutrients mobilized from leaves of arabidopsis thaliana
455 during leaf senescence. *J Plant Physiol* **158**: 1317-1323. DOI: 10.1078/0176-1617-00608.
456 **Khush GS, Lee S, Cho J, Jeon JS. 2012.** Biofortification of crops for reducing malnutrition.
457 *Plant Biotechnol Rep* **6**:195-202. DOI:10.1007/s11816-012-0216-5.
458 **Kumar J, Saripalli G, Gahlaut V, Goel N, Meher PK, Mishra KK, Mishra PC, Sehgal D,**
459 **Vikram P, Sansaloni C, Singh S, Sharma PK, Gupta PK. 2018.** Genetics of Fe, Zn, β-
460 carotene, GPC and yield traits in bread wheat (*Triticum aestivum* L.) using multi-locus and
461 multi-traits GWAS. *Euphytica* **214**:219. DOI:10.1007/s10681-018-2284-2.
462 **Liu J, Wu BH, Singh RP, Velu G. 2019.** QTL mapping for micronutrients concentration and
463 yield component traits in a hexaploid wheat mapping population. *Journal of Cereal Science*
464 **88**: 57-64.DOI: 10.1016/j.jcs.2019.05.008.
465 **Lipka AE, Tian F,Wang QS, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang**
466 **ZW. 2012.** GAPIT: genome association and prediction integrated tool. *Bioinformatics*
467 **28**:2397-2399. DOI:10.1093/bioinformatics/bts444.
468 **Ogo Y, Kobayashi T, Itai RN, Nakanishi H, Kakei Y, Takahashi M, Tok S, Mori S,**

Nishizawa NK. 2008. A Novel NAC transcription factor, IDEF2, that recognizes the iron deficiency-responsive element 2 regulates the genes involved in iron homeostasis in plants. *The Journal of Biological Chemistry* **283**: 13407-13417. DOI:10.1074/jbc.M708732200.

Peleg Z, Cakmak I, Ozturk L, Yazici A, Jun Y, Budak H, Korol AB, Fahima T, Saranga Y. 2009. Quantitative trait loci conferring grain mineral nutrient concentrations in durum wheat × wild emmer wheat RIL population. *Theor Appl Genet* **119**: 353-369. DOI: 10.1007/s00122-009-1044-z.

Pu ZE, Yu M, He QY, Chen GY, Wang JR, Liu YX, Jiang QT, Li W, Dai SF, Wei YM, Zheng YL. 2014. Quantitative trait loci associated with micronutrient concentrations in two recombinant inbred wheat lines. *Journal of Integrative Agriculture* **13**: 2322-2329. DOI:10.1016/S2095-3119(13)60640-1.

Rathan ND, Sehgal D, Thiyagarajan K, Singh R, Singh A-M, Govindan V. 2021. Identification of genetic loci and candidate genes related to grain zinc and iron concentration using a zinc-enriched wheat 'Zinc-Shakti'. *Front. Genet*. **12**:652653. DOI: 10.3389/fgene.2021.652653.

Roshanzamir H, Kordenaeej A, Bostani A. 2013. Mapping QTLs related to Zn and Fe concentrations in bread wheat (*Triticum aestivum*) grain using microsatellite markers. *Iranian journal of genetics and plant breeding* **2**: 10-17.

Shi RL, Tong YP, Jing RL, Zhang FS, Zou CQ. 2013. Characterization of quantitative trait loci for grain minerals in hexaploid wheat (*Triticum aestivum* L.). *Journal of Integrative Agriculture* **12**: 1512-1521. DOI:10.1016/S2095-311-
-9 (13)60559-6

Srinivasa J, Arun B, Mishra VK, Singh GP, Velu G, Babu R, Vasistha NK, Joshi AK. 2014. Zinc and iron concentration QTL mapped in a *Triticum spelta* × *T. aestivum* cross. *Theor Appl Genet* **127**: 1643-1651.DOI: 10.1007/s00122-014-23-
-27-6

Tiwari C, Wallwork H, Arun B, Mishra VK, Velu G, Stangoulis J, Kumar U, Joshi AK. 2016. Molecular mapping of quantitative trait loci for zinc, iron and protein content in the grains of hexaploid wheat. *Euphytica* **207**:563-570. DOI 10.1007/s10681-015-1544-7.

Tiwari VK, Rawat N, Chhuneja P, Neelam K, Aaaarwai R, Randhawa G, Dhaliwal HS, Keller B, Singh K. 2009. Mapping of quantitative trait loci for grain iron and zinc concentration in diploid A genome wheat. *Journal of Heredity* **100**: 771-776. DOI:10.1093/jhered/esp030.

Tong JY, Sun MJ, Wang Y, Zhang Y, Rasheed A, Li M, Xia XC, He ZH, Hao YF. 2020. Dissection of molecular processes and genetic architecture underlying iron and zinc homeostasis for biofortification: from model plants to common wheat. *Int. J. Mol. Sci* **21**:

9280. DOI:10.3390/ijms21239280.

Uauy C, Distelfeld A, Fahima T, Blech A, Dubcovsky J. 2006. A NAC gene regulating senescence improves grain protein, Zn, and Fe content in wheat. *Science* **314**: 1298-301. DOI: 10.1126/science.1133649.

Velu G, Tutus Y, Gomez-Becerra HF, Hao YF, Demir L, Kara R, Crespo-Herrera LA, Orhan S, Yazici A, Singh RP, Cakmak I. 2017. QTL mapping for grain zinc and iron concentrations and zinc efficiency in a tetraploid and hexaploid wheat mapping populations. *Plant Soil* **411**: 81-99. DOI: 10.1007/s-11104-016-
-3025-8.

Wang JS, Liu WH, Wang H, Li LH, Wu J, Yang XM, Li XQ, Gao AN. 2011. QTL mapping of yield-related traits in the wheat germplasm 3228. *Euphytica* **177**: 277-292. DOI:10.1007/s10681-010-0267-z.

Wang W, Guo H, Wu CN, Yu H, Li XK, Chen GF, Tian JC, Deng ZY. 2021a. Identification of novel genomic regions associated with nine mineral elements in Chinese winter wheat grain. *BMC Plant Biol* **21**:311. DOI:10.1186/s12870-021-
-03105-3.

Wang Y, Xu XT, Hao YF, Zhang YL, Liu YP, Pu ZJ, Tian YB, Xu DG, Xia XC, He ZH, Zhang Y. 2021b. QTL Mapping for grain zinc and iron concentrations in bread wheat. *Front. Nutr.* **8:**680391. DOI: 10.3389/fnut.2021.680391.

Welch RM, Graham RD. 1999. A new paradigm for world agriculture: meeting human needs. productive, sustainable, nutritious. *Food Crops Res* **60**:1-10. DOI:10.1016/S0378-4290(98)00129-4

Xu YF, An DG, Liu DC, Zhang AM, Xu HX, 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. DOI: 10.1016/j.fcr.2012.09.0

Zarcinas BA, Cartwright B, Spencer LR. 1987. Nitric acid digestion and multi-elemental analysis of plant material by inductively coupled plasma spectrometry. *Commun Soil Sci Plant Anal* **18**:131–146. DOI: 10.1080/00103628709367806

Table 1(on next page)

Table1 Descriptive statistics for grain Fe concentrations in GWAS population and soil Fe content in three environments

[a] The maximum value of grain Fe concentration. [b]The minimum value of grain Fe concentration. [c] The mean value of grain Fe concentration. Significant differences in the means for different environments (P□0.05) are indicated by different letters. [d] Standard deviation. [e] Kurtosis refer to a measure of the 'tailedness' of the probability distribution of a real-valued random variable. [f] Skewness refer to a measure of the asymmetry of the probability distribution of a real-valued random variable about its mean. [g] The mean value of soil Fe content for different environments.

1

Environment	Max[a](mg/kg)	Min[b](mg/kg)	Mean[c](mg/kg)	SD[d]	Skew[e]	Kurtosis[f]	Mean content for soil (mg/kg)[g]
Kaifeng (KF)	250.62	1.33	99.57[h]	64.50	0.5945	-0.6106	26.1
Shangqiu(SQ)	158.27	20.23	74.00[i]	21.69	0.3568	0.8754	26.7
Yuanyang (YY)	143.78	28.15	68.12[j]	22.38	0.9795	0.8321	26.4
Mean	184.22	16.57	80.56	36.19	0.6436	0.3656	26.4

2

Table 2(on next page)

Table 2 List of significant loci and their detailed information for Fe concentration identified by GWAS

Nu refer to the number of the loci detected in this study; Peak SNP refer to the most significant SNP in the mapping interval; Position refer to the physical position of most significant SNPs in the mapping interval; P Value means P value of the target trait calculated by MLM model; R^2 refer to the percentage of phenotypic variance explained by the locus. KF, SQ, YY refer to Kaifeng, Shangqiu, Yuanyang environment, respectively.

1

Nu	Chromosome	Interval range	No. of SNPs	Environment	Peak SNP	Position	P Value	R^2(%)
1	1A	12192544-504855880	35	YY	AX-110912741	504855880	5.02×10^{-4}	12.19
2	1A	3260219-560890453	17	KF	AX-109487734	560890453	2.29×10^{-4}	11.66
3	1A	338830974-574937064	15	SQ	AX-95178074	377027087	2.05×10^{-4}	7.32
4	1B	328824347-328880060	5	KF	AX-111580083	328879038	6.97×10^{-4}	10.62
5	1B	385161579	1	YY	AX-94632727	385161579	8.66×10^{-4}	11.69
6	1B	336787649-603147101	136	SQ	AX-108903980	409302131	1.42×10-5	10.06
7	1D	5421805-20056510	5	KF	AX-110529533	16132987	9.97×10^{-6}	14.71
8	2A	17032001-182302531	10	KF	AX-95247517	182302531	1.77×10^{-4}	11.91
9	2A	52130032-108517228	4	YY	AX-108896742	105628630	7.26×10^{-4}	11.85
10	2A	83586426-775041568	6	SQ	AX-111455865	775029238	5.07×10^{-4}	6.45
11	2B	27522279-777308839	4	KF	AX-110970921	95738751	3.01×10^{-6}	15.91
12	2B	77961365-475608029	32	YY	AX-108855338	77961365	2.37×10^{-4}	12.89
13	2D	15442231-86079836	5	KF	AX-111611520	15442231	1.79×10-4	11.89
14	2D	50470803	1	YY	AX-95104146	50470803	9.58×10-4	11.60
15	3A	16070421	1	YY	AX-	16070421	7.98×10-4	11.77

					108847051			
16	3A	15348646-144473416	9	SQ	AX-94863805	144473416	6.11×10-4	6.26
17	3A	202587965-731220289	13	KF	AX-94987631	731220289	5.01×10-5	12.87
18	3B	10301481-737764644	3	YY	AX-108861017	679198055	6.95×10^{-4}	11.89
19	3B	263249368-404801032	7	KF	AX-110922471	376625452	9.37×10^{-6}	14.77
20	3B	825341801	1	SQ	AX-110526198	825341801	8.37×10^{-4}	5.96
21	3D	40526440	1	KF	AX-94729264	40526440	5.84×10^{-10}	24.84
22	3D	40526440	1	SQ	AX-94729264	40526440	1.57×10-5	9.95
23	3D	52650282-454744784	5	YY	AX-108814800	57034968	4.46×10-4	12.30

2 **Table 2 (continued)**

No	Chromosome	Interval range	No. of SNPs	Environment	Peak SNP	Position	P Value	R^2(%)
24	4A	451704-699571654	8	KF	AX-108912427	699571654	4.35×10^{-10}	25.31
25	4A	572558552-718836253	4	SQ	AX-108912427	699571654	8.41×10^{-6}	10.60
26	4B	11987531-670399775	26	KF	AX-111479123	14490156	3.06×10^{-5}	13.60
27	4D	500023674	1	SQ	AX-109847855	500023674	7.17×10^{-4}	6.11
28	5A	492890942-553064045	4	SQ	AX-109623019	492890942	6.99×10^{-5}	8.42
29	5A	547505636-650240330	39	KF	AX-109379942	549160416	9.86×10^{-6}	14.72

30	5B	78708064-699558049	2	KF	AX-112289745	78708064	9.33×10^{-7}	17.09
31	5B	387308892-477630916	379	SQ	AX-111708401	473871199	4.09×10^{-5}	8.97
32	5D	218945113-496411812	12	SQ	AX-89438182	238177337	1.15×10^{-4}	7.92
33	5D	431203273	1	KF	AX-95075283	431203273	8.57×10^{-4}	10.43
34	6A	611858549	1	SQ	AX-94936962	611858549	1.39×10^{-4}	7.73
35	6A	601398500	1	YY	AX-109557802	601398500	4.56×10^{-4}	12.28
36	6A	490486081-612124269	8	KF	AX-94975608	612123809	6.31×10^{-5}	12.90
37	6B	556751602	1	YY	AX-110421368	556751602	9.21×10^{-4}	11.64
38	6B	462555585-712475640	48	KF	AX-111084964	462555585	1.37×10^{-5}	14.39
39	6B	479332942-716010197	17	SQ	AX-95210102	708943119	3.04×10^{-4}	9.52
40	6D	471071095	1	SQ	AX-94688853	471071095	2.66×10^{-4}	7.08
41	7A	261687749	1	SQ	AX-111012263	261687749	3.67×10^{-4}	6.76
42	7A	11099182-688980095	21	KF	AX-109282301	36334168	8.23×10^{-5}	12.65
43	7B	47617941-627982714	8	KF	AX-111148246	626954936	3.01×10^{-5}	13.62
44	7B	630565865-711075593	7	SQ	AX-94839775	711067695	$3.63\times10-4$	6.77
45	7D	11666920-58784583	3	KF	AX-110930232	58784583	$3.79\times10-4$	11.19

46	7D	616767211-616767235	2	SQ	AX-94455581	616767235	1.72×10-4	7.52

Table 3(on next page)

Table 3 The candidate genes and their information for grain Fe concentration identified in this study

[a] The number of candidate genes for wheat grain Fe concentration. [b] Physical position of the SNP as reported in the IWGSC Chinese Spring reference genome RefSeq v2.0.

1

Nu[a]	Chromosome	Identified loci in current study	Position(bp)[b]	Candidate genes (closest/nearby)	Annotation
1	3D	AX-94729264	40526440	TraesCS3D01G078500	NAC domain-containing protein
2				TraesCS3D01G080900	defensin-like protein
3	4A	AX-108912427	699571654	TraesCS4A01G430000	DUF581 family protein
4				TraesCS4A01G431200	Acid phosphatase 1
5				TraesCS4A01G431800	senescence-associated family protein, putative (DUF581)
6				TraesCS4A01G431900	senescence-associated family protein, putative (DUF581)
7				TraesCS4A01G432000	DUF581 family protein
8	6A	AX-94936962	611858549	TraesCS6A01G403500	Remorin
9	6B	AX-94702817	708943077	TraesCS6B01G447400	Remorin
10				TraesCS6B01G449700	Ring finger protein, putative

2

3

Table 4(on next page)

Table 4 Variance analysis for haplotypes with different alleles and haplotype combinations in two blocks on chromosome 6B

* Means the significant differences between haplotype A and haplotype B (P□0.05).

1

Block	Haplotypes	Fe_KF	Fe_SQ	Fe_YY	Hap frequency(%)
Block1	Hap1A	98.35±6.52	67.32±2.23	63.32±2.44	34.00
	Hap1B	102.06±5.83	73.45±1.88*	70.59±2.19*	66.00
Block2	Hap2A	92.90±6.02	68.19±2.32	64.28±2.65	70.00
	Hap2B	96.41±10.18	77.27±3.78*	74.20±5.20	30.00
Block1+ Block2	Hap1A+ Hap2A	92.90±6.02	68.19±2.32	64.28±2.65	70.73
	Hap1B+ Hap2B	99.48±10.15	76.91±3.94*	75.49±5.25*	29.27

2

Figure 1

Distribution of grain Fe concentrations in the wheat association analysis population

(A) Box plot for grain Fe concentration in three environments (Shangqiu ,Yuanyang and Kaifeng). (B, C, D) Distribution of grain Fe concentration for the wheat association analysis population in Kaifeng (KF), Shangqiu (SQ), Yuanyang (YY) environment, respectively.

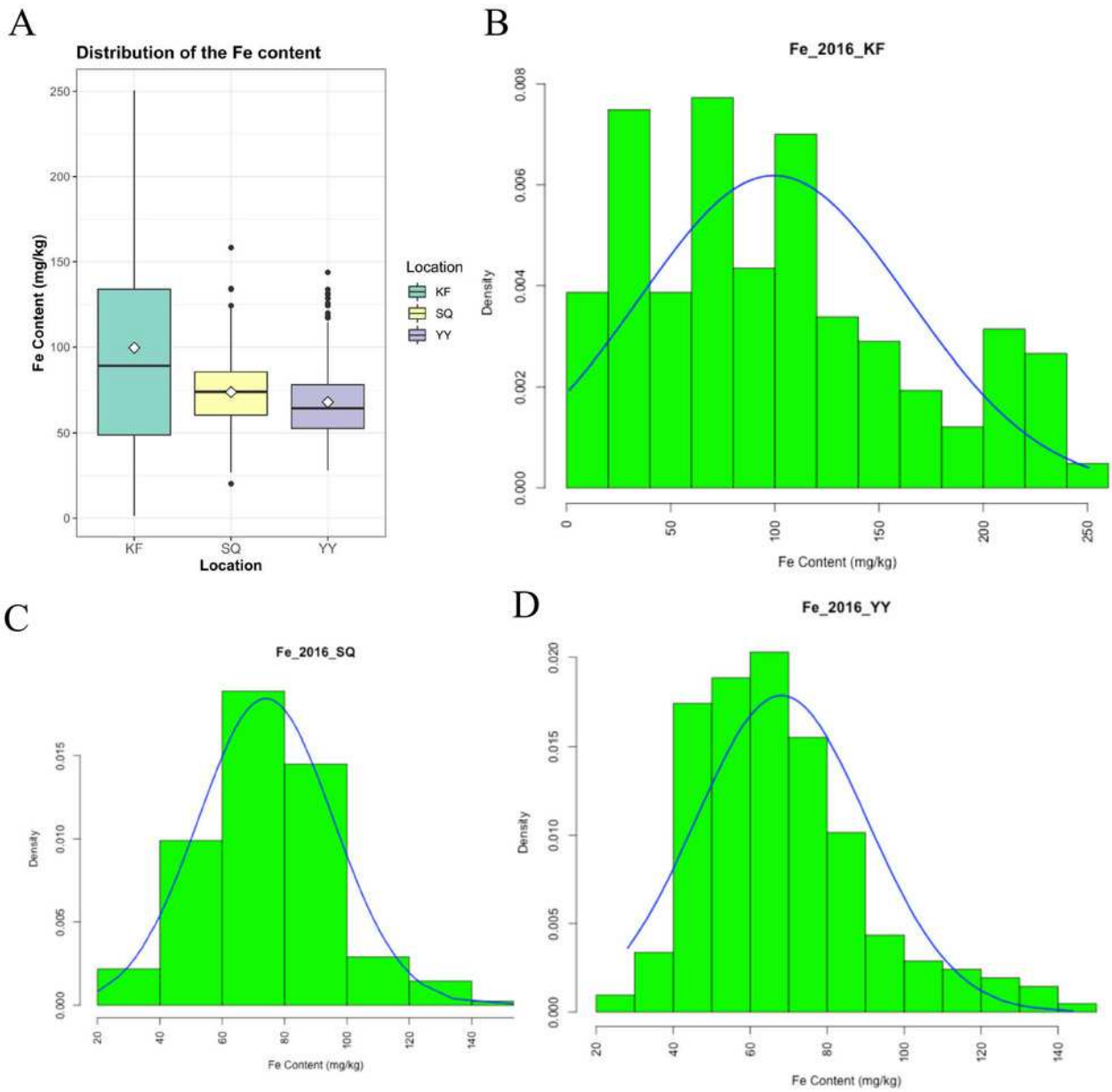

Figure 2

Manhattan and QQ plots of GWAS for wheat grains Fe concentration in the wheat association mapping population based on the mixed linear model (MLM)

The horizontal red color line indicated the genome-wide significant threshold of -$\log_{10}$(*p*-value) of 3.0. The SNPs above the red dotted line are significantly associated with grain Fe variation. Quantile-quantile scale representing expected versus observed -log10 (*p*-value) in three environments (Kaifeng, Shangqiu and Yuanyang)

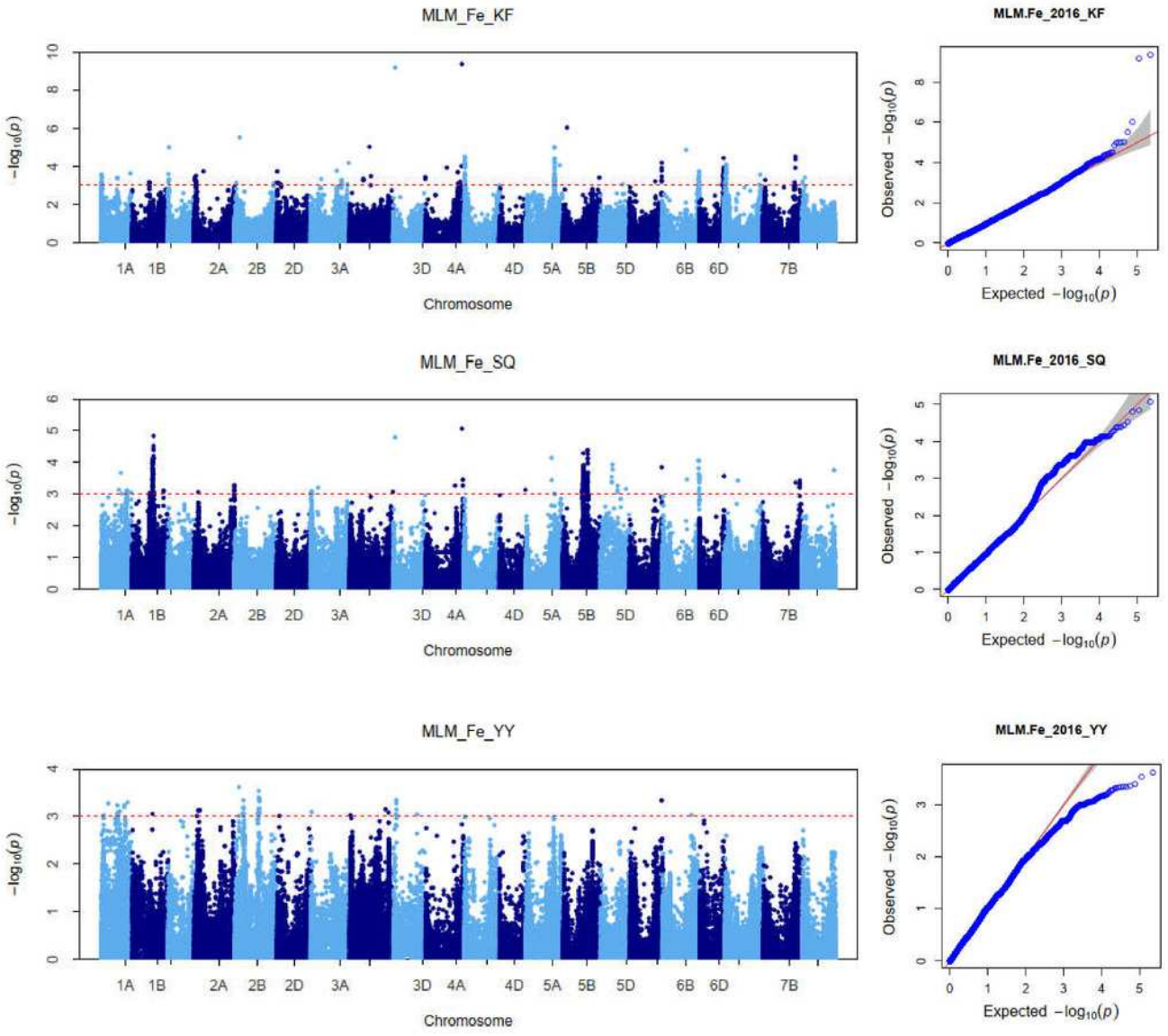
MLM_Fe_KF
MLM.Fe_2016_KF
MLM_Fe_SQ
MLM.Fe_2016_SQ
MLM_Fe_YY
MLM.Fe_2016_YY
−log10(p)
Observed −log10(p)
Expected −log10(p)
1A 1B 2A 2B 2D 3A 3D 4A 4D 5A 5B 5D 6B 6D 7B
Chromosome

Figure 3

Haplotype analysis for the significant SNPs associated with wheat grains Fe concentration on chromosome 6B

Haplotype heatmap surrounding significant SNPs on chromosome 6B