

RESEARCH

Open Access



# Identification of quantitative trait loci (QTLs) and candidate genes of seed Iron and zinc content in soybean [*Glycine max* (L.) Merr.]

Huan Wang<sup>1,2,3†</sup>, Jia Jia<sup>1,2,3†</sup>, Zhandong Cai<sup>1,2,3†</sup>, Mingming Duan<sup>1,2,3</sup>, Ze Jiang<sup>1,2,3</sup>, Qiuju Xia<sup>4</sup>, Qibin Ma<sup>1,2,3</sup>, Tengxiang Lian<sup>1,2,3</sup> and Hai Nian<sup>1,2,3\*</sup>

## Abstract

**Background:** Deciphering the hereditary mechanism of seed iron (Fe) and zinc (Zn) content in soybean is important and sustainable to address the “hidden hunger” that presently affects approximately 2 billion people worldwide. Therefore, in order to detect genomic regions related to soybean seed Fe and Zn content, a recombinant inbred line (RIL) population with 248 lines was assessed in four environments to detect Quantitative Trait Loci (QTLs) related to soybean seed Fe and Zn content.

**Result:** Wide variation was found in seed Fe and Zn content in four environments, and genotype, environment, and genotype × environment interactions had significant influences on both the seed Fe and Zn content. A positive correlation was observed between seed Fe content and seed Zn content, and broad-sense heritability ( $H^2$ ) of seed Fe and Zn content were 0.73 and 0.75, respectively. In this study, five QTLs for seed Fe content were detected with 4.57 - 32.71% of phenotypic variation explained (PVE) and logarithm of odds (LOD) scores ranging from 3.60 to 33.79. Five QTLs controlling the seed Zn content were detected, and they individually explained 3.35 to 26.48% of the phenotypic variation, with LOD scores ranging from 3.64 to 20.4. Meanwhile, 409,541 high-quality single-nucleotide variants (SNVs) and 85,102 InDels (except intergenic regions) between two bi-parental lines were identified by whole genome resequencing. A total of 12 candidate genes were reported in one major QTL for seed Fe content and two major QTLs for seed Zn content, with the help of RNA-Seq analysis, gene ontology (GO) enrichment, gene annotation, and bi-parental whole genome sequencing (WGS) data.

**Conclusions:** Limited studies were performed about microelement of soybean, so these results may play an important role in the biofortification of Fe and Zn and accelerate the development of marker-assisted selection (MAS) for breeding soybeans fortified with iron and zinc.

**Keywords:** Soybean, Iron Content, Zinc content, Biofortification, QTL mapping, WGS

## Background

Micronutrients are essential nutrient elements required for human growth and development, such as Fe and Zn, because of their vital roles in a great deal of biochemical functions and key metabolic reactions. Fe and Zn are considered to have important biological relevance and clinical significance in global public health [1]. Micronutrient malnutrition is also generally known as “hidden hunger” and affects more than 2 billion people worldwide

\*Correspondence: hnian@scau.edu.cn

†Huan Wang, Jia Jia and Zhandong Cai contributed equally to this work.

<sup>3</sup>The Key Laboratory of Plant Molecular Breeding of Guangdong Province, College of Agriculture, South China Agricultural University, 510642 Guangzhou, Guangdong, People's Republic of China  
Full list of author information is available at the end of the article



[2]. For normal human growth and development, the recommended dietary allowances (RDAs) of Fe and Zn are 10–15 mg/d and 12–15 mg/d, respectively [3]. In the world, the majority of people rely on plant foods, which often contain low key micronutrients and do not meet the needs of RDAs, especially for Fe and Zn content [4]. Fe deficiencies can significantly lead to bad influences on human health, such as iron-deficiency anemia, stunting, and heart failure [5, 6]. Zn deficiencies result in the emergence of poor cognitive development, infaust neuronal development, and disordered immunity [7, 8].

Soybean [*Glycine max* (L.) Merr.] is the most prominent food and oil crop between current supply and demand. The micronutrients from soybean seeds are significant portions of the mineral nutrients for humans. Soybean and its products are vital sources of Fe and Zn intake for humans because they have higher content than rice, wheat, and corn [9]. Fortifying the content of Fe and Zn in soybean seeds is significant for alleviating malnutrition in humans whose daily diets lack Fe and Zn. Furthermore, the enrichment of Fe and Zn content in edible parts of crops by biofortification, such as genetic improvement, is considered a sustainable and cost-effective approach to alleviate “hidden hunger” and associated health issues [10, 11].

Understanding the mechanism for the accumulation of Fe and Zn in seeds is vital to increase the seed Fe and Zn content in crops. There are a variety of valuable studies on the seed Fe and Zn content of crops using association analysis and linkage analysis.

Recent years, genome-wide association studies (GWAS) have played a crucial role in identifying significant association locus about seed Fe or Zn content among some crops, such as rice (*Oryza sativa* L.), pearl millet (*Pennisetum glaucum* R. Br.), Mung bean (*Vigna radiata* L.), chickpea (*Cicer arietinum* L.) and so on [12–15]. In contrast, the potential to identify genetic information about seed Fe and Zn content is not underestimated by virtue of linkage analysis. Performing linkage analysis aimed to dissect genetic variability for seed micronutrients by means of RIL population, was demonstrated to be useful and efficient in rice, wheat (*Triticum spelta* L.), common bean (*Phaseolus vulgaris* L.) and pea (*Pisum sativum* L.) [16–21].

However, there are few studies on QTLs related to seed mineral content in soybean, and only several QTLs controlling seed Fe and Zn content have been reported. King et al. [22] found a QTL related to Fe content with LOD score of 4.3 on chromosome 20 (Chr20) using 916 simple sequence repeat (SSR) markers and 92  $F_{2:4}$  lines. Ramamurthy et al. [23] found two QTLs controlling seed Fe content (largest LOD score of 3.79) and three QTLs controlling seed Zn content (largest LOD score of 4.35)

using three RIL populations, and all were independently detected once. and Ning et al. [24] detected two QTLs for seed Fe content and five QTLs for seed Zn content, none of them was mapped twice in single environment and the largest LOD score was 5.55. In summary, due to lack of stable and high efficient QTLs, there is still much work to be done to accelerate the development of MAS for breeding soybean cultivars fortified with Fe and Zn.

Deciphering genetic information on these micronutrient traits is essential for breeding Fe- and Zn-rich cultivars by MAS. Linkage analysis is an effective statistical method to anchor the genomic regions related to a specific trait in a segregating population in different environments [25]. Troublesomely, the environment and genotype  $\times$  environment interaction effects have an influence on the phenotype of the aimed trait in different environments (locations  $\times$  years), and the QTL  $\times$  environment interaction (QEI) could also impact the effect and stability of QTLs [26]. Therefore, investigations of these problems contribute to the detection and utilization of QTLs. In addition, with the development of next-generation sequencing (NGS), it is increasingly affordable to identify single nucleotide polymorphism (SNP) markers, which are used to construct high-density genetic maps for QTL mapping. Currently, high-density genetic maps are widely used for genetic analysis in soybean. Cai et al. [27, 28] identified 15 stable QTLs controlling isoflavone content and found a gene (*GmHADI*) related to low phosphorus tolerance with the help of a high-density genetic map containing 3469 bin markers in soybean. Jiang et al. [29] constructed a high-density map for fine mapping of powdery mildew resistance genes in soybean. Therefore, in this study, we used an RIL population derived from a cross of Guizao 1  $\times$  B13 to identify genetic loci and candidate genes related to seed Fe and Zn content in soybean in four environments by means of QTL mapping and whole genome resequencing approaches. The results can be used to breed Fe- and zinc-rich cultivars with the help of MAS and help reduce the “hidden hunger” caused by Fe and Zn malnutrition worldwide.

## Results

### Evaluation of Phenotypic Variation

The parents and GB RIL population exhibited high diversities in seed Fe and Zn content across each environment (18GZ, 18ZC, 19GZ, and 19ZC) (Table 1; Fig. 1S). The ANOVA results demonstrated that widely significant differences existed between GZ1 and B13 for seed Fe content or seed Zn content in the four environments ( $p < 0.05$ ). In the GB RIL population, the variable range of seed Fe content was 72.52  $\mu\text{g g}^{-1}$  to 185.58  $\mu\text{g g}^{-1}$  with a smallest mean of 108.12, largest mean of 136.37  $\mu\text{g g}^{-1}$ ,

**Table 1** Descriptive statistics of seed Fe and Zn content in parents and GB RIL population

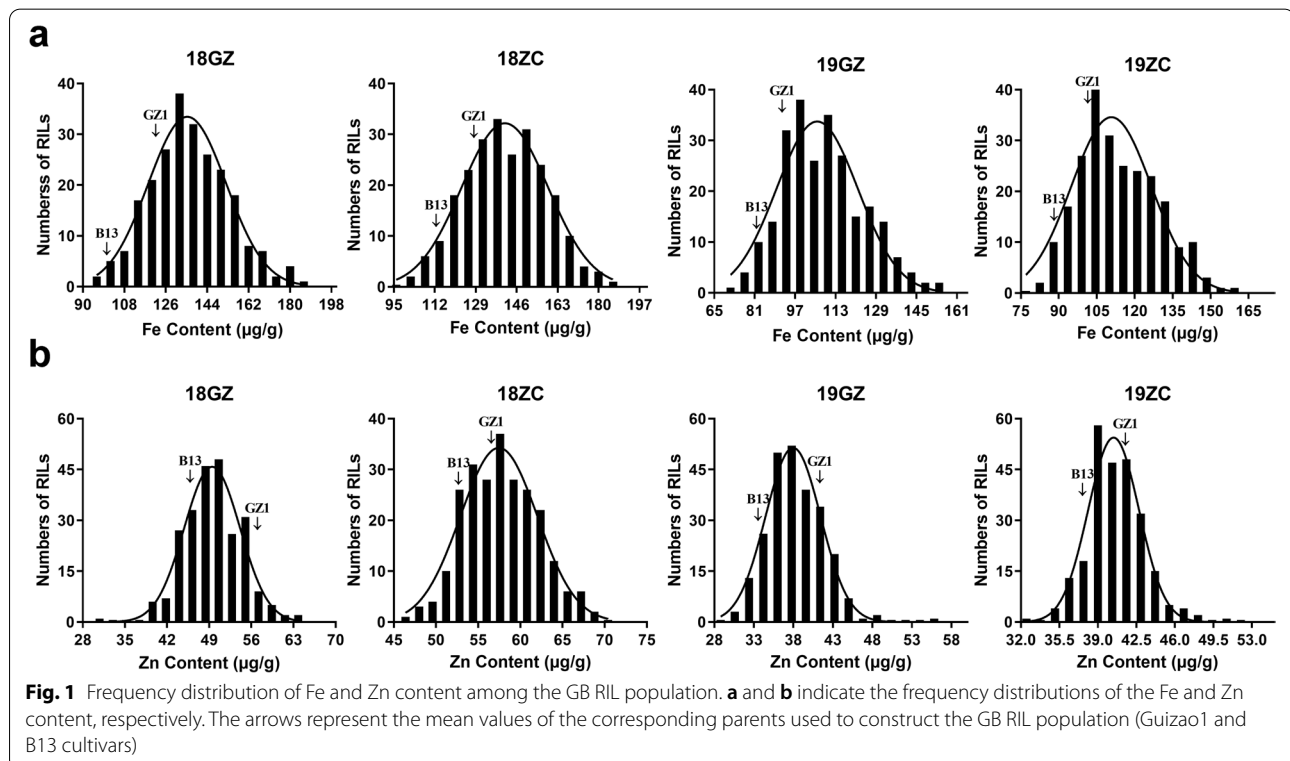
Trait <sup>a</sup>	Envi. <sup>b</sup>	Parents		RIL population								
		GZ1	B13	Mean	SD <sup>c</sup>	Min.	Max.	CV (%) <sup>d</sup>	Skew. <sup>e</sup>	Kurt. <sup>f</sup>	K-S Test	H <sup>2</sup>
Fe Content	18GZ	124.47	100.18**	136.37	17.08	98.17	185.58	12.52	0.23	-0.08	0.200*	0.73
	18ZC	130.51	114.09**	141.39	16.54	100.76	184.15	11.70	0.06	-0.48	0.200*	
	19GZ	91.87	83.01**	108.12	16.12	72.52	152.11	14.91	0.42	-0.19	0.023	
	19ZC	104.70	89.95**	113.36	15.39	82.41	159.37	13.58	0.41	-0.43	0.001	
Zn Content	18GZ	58.87	47.01**	49.71	4.81	30.86	63.52	9.68	0.01	0.66	0.200*	0.75
	18ZC	57.65	53.46**	57.76	4.31	47.00	69.50	7.46	0.18	-0.27	0.200*	
	19GZ	42.41	33.75**	38.32	3.54	30.01	55.52	9.24	0.70	1.74	0.200*	
	19ZC	40.86	38.82**	40.64	2.59	32.92	51.30	6.37	0.45	1.16	0.200*	

<sup>a</sup> The Fe and Zn content is presented as micrograms per gram in the soybean powder. <sup>b</sup> Environment (2018 in Guangzhou, 2018 in Zengcheng, 2019 in Guangzhou, and 2019 in Zengcheng are designated as 18GZ, 18ZC, 19GZ, and 19ZC, respectively). <sup>c</sup> Standard deviation. <sup>d</sup> Coefficient of variation. <sup>e</sup> Skewness. <sup>f</sup> Kurtosis

largest standard deviation (SD) of 17.08  $\mu\text{g g}^{-1}$ , and largest coefficient of variation (CV) of 14.91%. The Zn content varied from 30.01 to 69.50  $\mu\text{g g}^{-1}$ , with a smallest mean of 38.32, largest mean of 57.76  $\mu\text{g g}^{-1}$ , largest SD of 4.81  $\mu\text{g g}^{-1}$ , and largest of CV 9.68%.

Respecting the normality test, the absolute values of kurtosis and skewness for the two traits across four different environments were all less than two. According to the Kolmogorov-Smirnov test (K-S test) for the two traits in the four environments, the *p* values were all greater

than 0.05. Additionally, histograms of the frequency distribution of phenotypic data are depicted in Fig. 1, with wide spans and continuous distributions. All the results above illustrated that the segregation of the seed Fe and Zn content traits fitted normal or skew-normal distribution models, possessing typical quantitative genetic characteristics controlled by multiple genes. Some RILs existed in all environments with lower or higher values than the two parents for seed Fe and Zn content, indicating that transgressive segregation was widely present



in the GB RIL population. These results showed that the population was appropriate for QTL mapping because of the large continuous variation.

The broad-sense heritability ( $H^2$ ) of seed the Fe and Zn content was 73.3% and 75.43%, respectively, in the combined environment (Table 1). High heritability indicated that the genetic variance was superior to other variances, and this population was suitable for high-efficiency selection of high Fe and Zn content cultivars in soybean. The ANOVA results indicated that seed Fe and Zn content were significantly affected by the environment (E), genotype (G), and the interaction between environment and genotype ( $G \times E$  interaction) (Table 2). ANOVA conducted across the environments showed an extremely significant effect ( $P < 0.01$ ) of G, E, and  $G \times E$  interactions on both seed Fe and Zn content. The Pearson correlation analysis revealed that seed Fe and Zn content had a significant positive correlation in each environment except 18ZC (Fig. 2). In 18ZC, the nondistinctive and positive correlation between the seed Fe content and seed Zn content could be caused by environmental factors. There was a significant positive correlation of the seed Fe content of the RILs between all the environments, and the relationship of seed Zn content was the same as for the seed Fe content of the RILs.

### QTLs Mapping

To guarantee the reliability of detecting QTLs associated with seed Fe and Zn content, CIM, ICIM-ADD, and QEI methods were employed in this study. The significant LOD thresholds of the two traits in single-environment analysis and QEI mapping were determined by 1000 permutations at the level of  $p < 0.05$ , and the results are listed in Table S1. Through comprehensive analysis of the calculation results of the mapping, the significant LOD thresholds for the two traits were determined to be the same, and 3.6 was determined to be the minimum LOD score for the CIM and ICIM-ADD methods in all environments. In contrast, 5.3 was

determined as the minimum LOD score for the QEI method.

### QTLs Mapping of Seed Fe and Zn Content by CIM method

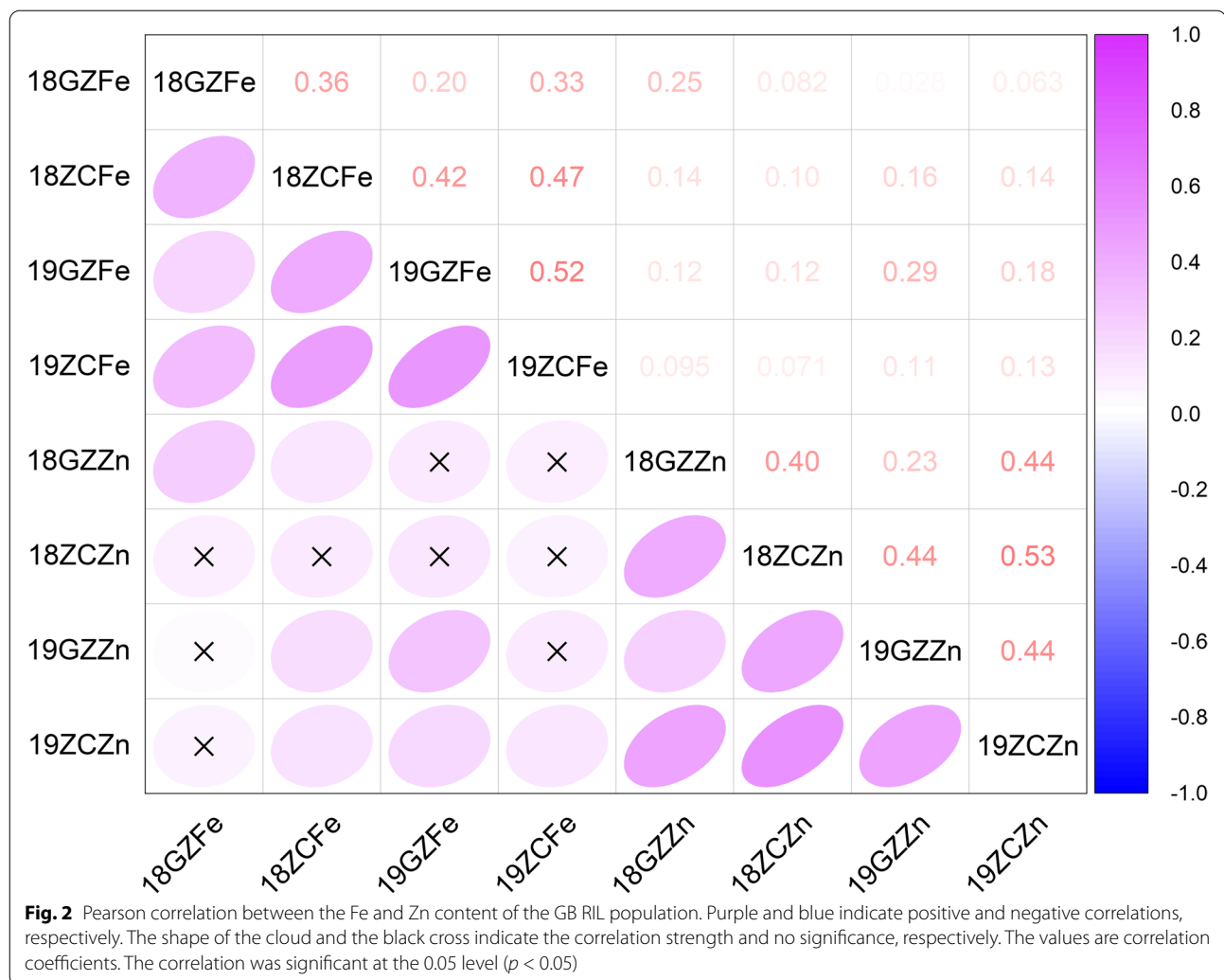
To identify genomic regions associated with the seed Fe and Zn content, a genome-wide scan was performed using the CIM method in Windows QTL Cartographer V2.5. A total of eight QTLs for seed Fe and Zn content were identified in the GB RIL population based on a linkage genetic map with the phenotypic performances of the four environments (Table 3; Fig. 3).

For the seed Fe content, four QTLs, *qFC4*, *qFC5*, *qFC7*, and *qFC13*, were detected on Chr04, Chr05, Chr07, and Chr13, respectively. Their LOD scores ranged from 3.8 to 23.6, with phenotypic variation explained (PVE) ranging from 4.90 to 29.86%. *qFC5* with positive additive effects (ADDs) acquired the favorable allele from the female parent Guizao1. *qFC4*, *qFC7*, and *qFC13* showed negative ADDs and acquired the favorable allele from the male parent B13. Of these QTLs, *qFC4*, *qFC5*, and *qFC13* were only identified in 19GZ, 18ZC, and 19ZC, respectively, while *qFC7* could be detected in all environments and the PVE reached its highest value of 29.86% in 19ZC.

Four QTLs (*qZC3*, *qZC11*, *qZC18*, and *qZC20*) controlling the seed Zn content were detected on Chr03, Chr11, Chr18, and Chr20. They individually explained 5.04 to 26.48% of the phenotypic variation, with LOD scores ranging from 4.4 to 20.4. *qZC11* and *qZC18* were detected in 18GZ and 18ZC, respectively. *qZC3* could be detected in 18GZ and 18ZC, while *qZC20* could be detected in all environments. *qZC3* had the largest PVE of 12.79% in 18GZ, and *qZC20* had the largest PVE of 26.48% in 18ZC. *qZC3* had a negative ADD with a favorable allele from the male parent B13, and the favorable alleles of the remaining three QTLs were contributed by Guizao1.

**Table 2** Analysis of variance of seed Fe and Zn content across environments

Trait	Source	DF	SS	MS	F-value	P-value
Fe Content	Block/Environment	4	104864.36	26216.09	131.96	<0.01
	Environment	3	393427.25	131142.42	660.13	<0.01
	Genotype	247	276086.59	1117.76	5.63	<0.01
	$G \times E$ interaction	720	241040.22	334.78	1.69	<0.01
Zn Content	Block/Environment	4	11528.76	2882.19	321.44	<0.01
	Environment	3	115052.84	38350.95	4277.13	<0.01
	Genotype	247	16772.77	67.91	7.57	<0.01
	$G \times E$ interaction	730	14015.20	19.20	2.14	<0.01



#### QTLs Mapping of Seed Fe and Zn Content by ICIM-ADD method

The ICIM method in the IciMapping program V4.2 was also concurrently employed in this study. Four QTLs for seed Fe content and three QTLs for seed Zn content were identified (Table 4; Fig. 3). Four QTLs (*qFC4*, *qFC7*, *qFC12*, and *qFC13*) for seed Fe content were located on Chr04, Chr07, Chr12, and Chr13, respectively. The LOD scores ranged from 3.6 to 26.83, and the PVE ranged from 4.57 to 32.41%. The ADD of those QTLs was negative, showing that the favorable alleles were contributed by B13. *qFC4*, *qFC12*, and *qFC13* were identified in 19GZ, 19GZ, and 19ZC, respectively. *qFC7* could be detected in all environments, with the largest PVE at 32.41% in 19ZC.

Three QTLs (*qZC3*, *qZC18*, and *qZC20*) controlling the seed Zn content were identified on Chr03, Chr18, and Chr20. They individually explained 4.23 to 22.28% of the phenotypic variation with LOD scores ranging from 3.64

to 15.14. Surprisingly, those QTLs could be identified in multiple environments. *qZC18* was detected in 18ZC and 19GZ; *qZC3* could be identified in 18GZ, 18ZC, and 19ZC; and *qZC20* could be identified in 18GZ, 19GZ, and 19ZC. *qZC3*, *qZC18*, and *qZC20* had the largest PVEs of 11.69%, 6.84%, and 22.28%, respectively. *qZC3* had a negative ADD with a favorable allele from the male parent B13, and the favorable alleles of *qZC18* and *qZC20* were contributed by Guizao1.

#### QEI mapping

Under the MET functionality of the IciMapping program V4.2, three QTLs related to seed Fe and Zn content were identified by the multi-environment QTL analysis method (Table 5; Fig. 3). Three QTLs (*qFC4*, *qFC13*, and *qFC7*) for the seed Fe content present on Chr04, Chr07, and Chr13 had LOD scores ranging from 6.68 to 33.79, and the total PVE ranged from 4.80 to 32.71%. *qFC4* and *qFC13* displayed high phenotypic

**Table 3** Characteristics of QTLs related to seed Fe and Zn content detected using CIM mapping

Trait	Envi. <sup>a</sup>	QTL	Chr <sup>b</sup>	Pos. <sup>c</sup> (cM)	Marker interval	CI (bp) <sup>d</sup>	LOD <sup>e</sup>	ADD <sup>f</sup>	PVE (%) <sup>g</sup>
Seed Fe Content	19GZ	<i>qFC4</i>	Chr04	28.5	bin35~bin36	5311004~5606879	6.50	-4.82	8.44
	18ZC	<i>qFC5</i>	Chr05	96.7	bin108~bin109	35298620~35449207	3.80	3.78	4.90
	18GZ	<i>qFC7</i>	Chr07	79.4	bin87~bin91	8446317~9568439	6.40	-5.41	10.00
	18ZC	<i>qFC7</i>	Chr07	90.0	bin87~bin91	8446317~9568439	16.30	-8.08	23.76
	19GZ	<i>qFC7</i>	Chr07	87.2	bin87~bin91	8446317~9568439	12.40	-6.77	17.48
	19ZC	<i>qFC7</i>	Chr07	84.4	bin87~bin91	8446317~9568439	23.60	-8.47	29.86
	19ZC	<i>qFC13</i>	Chr13	85.7	bin106~bin108	25905367~26941818	4.90	-3.81	5.09
Seed Zn Content	18GZ	<i>qZC3</i>	Chr03	128.5	bin157~bin167	42773578~43911616	8.80	-1.74	12.79
	18ZC	<i>qZC3</i>	Chr03	124.4	bin157~bin167	42773578~43911616	7.90	-1.30	8.94
	18GZ	<i>qZC11</i>	Chr11	89.6	bin78~bin79	11209422~11357194	4.50	1.45	6.28
	18ZC	<i>qZC18</i>	Chr18	124.4	bin205~bin207	51866863~52524554	4.60	0.99	5.04
	18GZ	<i>qZC20</i>	Chr20	55.0	bin56~bin75	13374481~33919927	6.70	1.50	9.57
	18ZC	<i>qZC20</i>	Chr20	56.3	bin56~bin75	13374481~33919927	20.40	2.24	26.48
	19GZ	<i>qZC20</i>	Chr20	53.4	bin56~bin75	13374481~33919927	4.40	0.96	6.98
	19ZC	<i>qZC20</i>	Chr20	55.0	bin56~bin75	13374481~33919927	15.80	1.21	21.43

<sup>a</sup> Environment; <sup>b</sup> Chromosome; <sup>c</sup> Position of the peak of the QTL in chromosomes; <sup>d</sup> Confidence interval; <sup>e</sup> Logarithm of odds; <sup>f</sup> Additive effect of QTL; <sup>g</sup> Percent of phenotypic variance explained



variation contributed by additive × environment effects (PVE (A × E)), which was demonstrated by the high LOD score from the QTL × environment interaction (LOD (A × E)) and the low LOD score from the additive effect (LOD (A)). In contrast, the phenotypic variation contributed by additive effects (PVE (A)) of *qFC7* was more powerful than PVE (A × E), showing its great stability across environments.

Three QTLs (*qZC3*, *qZC6*, and *qZC20*) controlling the seed Zn content were identified on Chr03, Chr06, and Chr20. They individually explained 3.35 to 17.98% of the total phenotypic variation, and the total LOD scores ranged from 5.47 to 20.96. *qZC3* and *qZC6* were equipped with stronger PVE (A) than PVE (A × E), as indicated by the high LOD (A) compared to the low LOD (A × E), which shows their excellent stability

**Table 4** Characteristics of QTLs related to seed Fe and Zn content detected using ICIM mapping

Trait	Envl. <sup>a</sup>	QTL	Chr. <sup>b</sup>	Pos. <sup>c</sup> (cM)	Marker interval	CI (bp) <sup>d</sup>	LOD <sup>e</sup>	ADD <sup>f</sup>	PVE (%) <sup>g</sup>
Fe Content	19GZ	<i>qFC4</i>	Chr04	38.2	bin46~bin47	6194153~6689660	6.13	-4.69	8.00
	18GZ	<i>qFC7</i>	Chr07	84.0	bin87~bin88	8446317~8563902	6.92	-6.08	12.61
	18ZC	<i>qFC7</i>	Chr07	90.0	bin92~bin93	9568440~9745745	15.74	-8.28	25.33
	19GZ	<i>qFC7</i>	Chr07	87.0	bin90~bin91	8633014~9568439	12.69	-7.05	18.19
	19ZC	<i>qFC7</i>	Chr07	84.1	bin87~bin88	8446317~8563902	26.83	-9.01	32.41
	19GZ	<i>qFC12</i>	Chr12	28.0	bin27~bin28	4498589~4568321	3.60	-3.53	4.57
	19ZC	<i>qFC13</i>	Chr13	85.2	bin107~bin108	26336429~26941818	6.07	-3.94	6.19
Zn Content	18GZ	<i>qZC3</i>	Chr03	128.2	bin158~bin159	42959977~43331949	7.75	-1.69	11.69
	18ZC	<i>qZC3</i>	Chr03	128.2	bin158~bin159	42959977~43331949	7.29	-1.31	8.48
	19ZC	<i>qZC3</i>	Chr03	128.2	bin158~bin159	42959977~43331949	3.73	-0.56	4.98
	18ZC	<i>qZC18</i>	Chr18	127.0	bin204~bin205	51698565~51939931	3.78	0.93	4.23
	19GZ	<i>qZC18</i>	Chr18	135.0	bin207~bin208	52009904~52593803	3.64	0.87	6.84
	18GZ	<i>qZC20</i>	Chr20	55.0	bin62~bin63	25069476~27890103	5.98	1.47	8.72
	19GZ	<i>qZC20</i>	Chr20	53.1	bin56~bin57	13374481~22252230	3.82	0.90	7.35
	19ZC	<i>qZC20</i>	Chr20	55.0	bin62~bin63	25069476~27890103	15.14	1.20	22.28

<sup>a</sup> Environment; <sup>b</sup> Chromosome; <sup>c</sup> Position of the peak of the QTL in chromosome; <sup>d</sup> Confidence interval; <sup>e</sup> Logarithm of odds; <sup>f</sup> Additive effect of QTL; <sup>g</sup> Percent of phenotypic variance explained

across environments. *qZC20* exhibited a PVE (A) almost equal to PVE (A × E).

In combination with the results of the CIM and ICIM-ADD mapping, *qFC7* with largest PVE of 32.41% and *qZC20* with largest PVE of 26.48% could be repeatedly detected at least three times in the CIM and ICIM-ADD methods, and they could be detected in the MET analysis with high genetic effects. Although *qZC20* exhibited a slightly smaller PVE (A) than PVE (A × E), it could be detected three and four times in the CIM and ICIM-ADD maps, respectively. *qZC3* with largest PVE of 12.79% was detected three times and two times in the ICIM-ADD and CIM methods, respectively. Additionally, *qZC3* with stronger PVE (A) was mapped in the MET analysis, *qZC3* was a major QTL. Therefore, we will mainly focus on these three major QTLs in the follow-up analysis. From results of three mapping methods, among *qFC7*, *qZC3*, and *qZC20*, every QTL from single environment was overlapping on either one corresponding locus, and peaks of any two QTL from single environment distanced less than 5 cM for *qZC3* and *qZC20*. One or two peaks of QTL slightly distanced more than 5 cM, but they were mostly overlapping. Interestingly, all confidence interval (CI) of QTLs from ICIM-ADD and MET methods were included in corresponding QTL CI from CIM method. Hence, the results of CIM method were mainly used as basic of identifying candidate genes. Although there were larger QTL CIs, it was worthy for high reliability of further analysis.

#### Whole genome resequencing of two parents

To identify accurate candidate genes, the parental lines were resequenced at the whole-genome scale. GC content of Guizao1 was 37.5%, which of B13 was 37.3%. 92.5 and 115.9 million high-quality reads were obtained from Guizao1 and B13, respectively. 98.89% of reads were mapped to reference genome Wm82.a2.v1 (<http://www.soybase.org/>) for Guizao1, and it was 98.34% for B13. The average depths of Guizao1 and B13 were respectively 9.0× and 11.27×. Based on analysis, there were 409,541 high-quality single-nucleotide variants (SNVs) (except intergenic regions) between two bi-parental lines, and nonsynonymous SNVs were 29,548. Similarly, 85,102 InDels were identified and 2490 of which were in coding sequence.

#### Candidate genes prediction of the major QTLs

The release of the soybean reference genome Wm82.a2.v1, whole-genome resequencing, and gene annotations made it efficient to recognize candidate genes in the genetic regions related to the two traits. In this study, three stable and major QTLs; *qFC7*, which was related to seed Fe content; and *qZC3* and *qZC20*, which were related to seed Zn content, were deemed the most promising genomic regions. *qFC7*, *qZC3*, and *qZC20* harbored 108, 151, and 403 genes, respectively, and these all-model genes were downloaded from Phytozome and SoyBase (Table S2). The OmicShare Tools was used to perform GO enrichment analysis, which showed that the main

**Table 5** Characteristics of QTLs related to seed Fe and Zn content detected using QEI mapping

Trait	QTL	Chr. <sup>a</sup>	Pos. (cM) <sup>b</sup>	Marker interval	LOD <sup>c</sup>	LOD (A) <sup>d</sup>	LOD (A × E) <sup>e</sup>	PVE <sup>f</sup>	PVE (A) <sup>g</sup>	PVE (A × E) <sup>h</sup>	ADD <sup>i</sup>	QTL × environment interaction			
												A × E1 <sup>j</sup>	A × E2 <sup>k</sup>	A × E3 <sup>l</sup>	A × E4 <sup>m</sup>
Fe Content	qFC4	Chr04	38.2	bin46~bin47	7.47	0.53	6.94	7.86	0.53	7.33	-0.67	2.54	1.63	-4.00	-0.17
	qFC7	Chr07	84.0	bin87~bin88	33.79	16.46	17.33	32.71	16.94	15.77	-3.80	-1.99	3.20	3.80	-5.02
	qFC13	Chr13	85.2	bin107~bin108	6.68	2.38	4.30	4.80	2.32	2.48	-1.41	1.52	0.71	0.14	-2.37
Zn Content	qZC3	Chr03	128.2	bin158~bin159	20.88	19.49	1.39	17.98	14.82	3.16	-1.01	-0.62	-0.27	0.45	0.45
	qZC6	Chr06	51.0	bin43~bin44	5.47	3.55	1.92	3.35	2.58	0.77	0.42	-0.26	-0.04	0.37	-0.07
	qZC20	Chr20	55.0	bin62~bin63	20.96	8.26	12.71	12.54	6.14	6.40	0.65	0.77	-0.66	-0.66	0.55

<sup>a</sup> Chromosome; <sup>b</sup> Position of the peak of the QTL in chromosome; <sup>c</sup> Overall LOD score of the QTL; <sup>d</sup> LOD score from additive effect; <sup>e</sup> LOD score from QTL × environment interaction; <sup>f</sup> Percent of phenotypic variance explained; <sup>g</sup> Percent of phenotypic variance explained by additive effect; <sup>h</sup> Percent of phenotypic variance explained by QTL × environment interaction; <sup>i</sup> Average additive effect of QTL; <sup>j</sup> QEI effect in E1 (18GZ); <sup>k</sup> QEI effect in E2 (18ZC); <sup>l</sup> QEI effect in E3 (19GZ); <sup>m</sup> QEI effect in E4 (19ZC)



enriched GO terms included cellular process, single-organism process, metabolic process, biological regulation, developmental process, multicellular organismal process, regulation of biological process, cell, signaling, cell part, organelle, membrane, catalytic activity, binding, transporter activity and so on (Table S3).

To reduce the number of candidate genes and identify candidate genes, based on the resequencing analysis of Guizao1 and B13, a total of 41,318 different variants (38,151 SNVs and 3,167 InDels) could be found in three QTL regions concerning Guizao1 and B13 (Tables S4 and S5). Among them, 475 SNVs and 15 InDels were in coding sequence. Further analysis showed that 44, 41, and 125 putative functional mutant genes were obtained in the regions of *qFC7*, *qZC3*, and *qZC20*, respectively.

To further refine these putative candidate genes, the freely available RNA-Seq data of 210 mutant genes from SoyBase were analyzed (Table S6, Fig. 2S) to exclude non-expressed or very low expressed genes. According to the analysis above, in combination with the GO enrichment analysis, different variants, RNA-Seq data, gene annotations, and available literature, three of 41 genes in the *qFC7* genetic region (*Glyma.07G093700*, *Glyma.07G097700*, and *Glyma.07G100700*) were predicted as candidate genes associated with seed Fe content in soybean, while three of 44 genes in the *qZC3* block (*Glyma.03G229400*, *Glyma.03G231200*, and *Glyma.03G240000*), and six of 125 genes in the *qZC20* genetic region (*Glyma.20G063100*, *Glyma.20G068300*, *Glyma.20G067300*, *Glyma.20G076700*, *Glyma.20G088600*, and *Glyma.20G089300*) were predicted as candidate genes related to seed Zn content in soybean (Table S7).

## Discussion

Micronutrient deficiencies, particularly the lack of Fe and Zn, which influence billions of people's health worldwide, are termed "hidden hunger". This is a serious health problem for humans, especially in developing and low-income countries [2]. Effective measures, such as biofortification, need to be taken to obtain rich mineral foods to solve micronutrient malnutrition or hidden hunger due to its promising and cost-efficient characteristics [10, 11]. A variety of teams have interest in pursuing increased seed mineral content of crops such as maize, wheat, rice, and beans [30]. QTL mapping is helpful to dissect the genetic architecture of seed micronutrient content by finding markers or genes associated with seed mineral content in crops. In the current study, we detected one major QTL and three potential candidate genes for seed Fe content and two major QTLs for seed Zn content with nine potential candidate genes. Understanding the genetic

mechanism of increasing the seed Fe and Zn content is a prerequisite to breed mineral-rich cultivars using MAS in soybean.

## Variations of seed Fe and Zn content in soybean

Genetic variation has a great influence on QTL mapping. A number of studies on mineral variation have been reported in many crops, such as wheat, rice, common bean, and soybean. Our wide variation of seed Fe and Zn content corroborated those findings [16–19, 23, 24]. Our study also showed that genotype, environment, and genotype  $\times$  environment interactions had significant influences on both the seed Fe and Zn content (Table 3), which was similar to previous studies [15, 31, 32]. Among them, the influences of environmental and genotype  $\times$  environment interactions increased the difficulty of identifying real QTLs of the desired traits because the phenotype could be inconsistent with the true genotype. In this study, the  $H^2$  values of seed Fe and Zn content were high, that was moderate to high compared with a previous study in soybean [24]. In other crops, there are similar results for  $H^2$  of seed Fe and Zn content [31, 33]. High heritability means that the phenotype is to a great extent shaped by the genotypic effect.

Correlation analysis among quantitative traits is necessary to estimate the feasibility that multiple traits are jointly selected in breeding. Plant breeders can efficiently improve both traits or multiple traits simultaneously with the help of positive genetic correlation between two or more desirable traits [34]. Understanding the genetic correlation between seed Fe and Zn content can assist plant breeders in designing an appropriate breeding approach to increase both seed Fe and Zn content in soybean at the same time. In this study, correlation between seed Fe content and seed Zn content was similar to previous studies [22–24]. We can infer that the accumulation of Fe and Zn in soybean seeds might share the same molecular mechanisms, even in other crops. Similar relationships occurred in rice, wheat, common bean, and spinach [17–19, 35, 36], which indicates that it is realistic to simultaneously improve different mineral content in crop breeding.

## QTL mapping and comparison with previous studies

A previous study explained that the most suitable mapping procedures were different for the presence of different genetic model data in practical experiments [37]. In recent years, as a variety of QTL mapping methods and software have been invented and improved, an increasing number of algorithms for mapping QTLs have been employed. The mapped QTL accuracy is easily influenced by the mapping procedures, and an unsuitable

method could result in the appearance of erroneous or false-positive results [38]. A QTL that can be simultaneously detected by at least two different mapping methods might have high reliability [37].

CIM, ICIM-ADD, and QEI procedures were employed to ensure the reliability of the detected QTLs. Moreover, both traits had significant differences between the two parents that were distant geographical species in the same environment, and the RILs were repeatedly planted in the same location and in different years. This allowed us to identify the environmental stability of the QTLs. Similarly, crop breeders usually identify genetic materials with the desired traits in various environments again to eliminate environmental effects and verify stable genotypes. Several researchers have carried out QTL mapping in multiple environments with locations  $\times$  years by using at least two mapping methods and obtained good results [38, 39].

Non-simultaneous identification of QTLs among procedures or environments demonstrates that there may be the differential expression of genes that are environmentally specific, which results in QTL  $\times$  environment interactions [26]. For seed Fe content, it should be noted that four minor QTLs (*qFC4*, *qFC5*, *qFC12*, and *qFC13*) were only detected in one environment or by one method in single-environment analysis. For the former, those QTLs could be unstable and environment-specific; for the latter, those different QTLs were identified and could be caused by the different detection power of mapping procedures [37]. Although *qFC4* and *qFC13* could be detected in QEI mapping analysis, they possessed a higher PVE ( $A \times E$ ) than PVE ( $A$ ), and they were susceptible to the environment and unstable. For the seed Zn content, *qZC6* had a higher PVE ( $A$ ) than PVE ( $A \times E$ ), and in single-environment analysis, small peaks with LOD scores lower than the threshold of 3.6 could be found adjacent to this locus in several environments. It is possible that *qZC6* had a minor effect and was not identified in the single-environment analysis but was identified in the QEI; such a case was similar to a previous study [40].

A few QTLs related to soybean seed Fe and Zn content have been detected by several researchers in recent years [22–24]. A stable QTL was detected on Chr3 for seed Fe content and a stable QTL was detected on Chr11 for seed Zn content; other QTLs were only detected in the special environment and not repeated in another study. In the current study, three major QTLs (*qFC7* for seed Fe content and *qZC3* and *qZC20* for seed Zn content) were new loci, and other QTLs were not similar to previous results. This may be due to the following reasons: (i) previous studies detected few relevant QTLs on account of the lack of high-quality

linkage maps; (ii) there are immense genetic differences between mapping populations; and (iii) QTL mapping is influenced by environmental effects, population size, and experimental error.

#### Putative candidate genes of interest in three major QTLs

In previous studies, some basic helix-loop-helix (bHLH) transcription factors, myeloblastosis (MYB) transcription factors, zinc transporter proteins (ZIPs), nicotianamine synthases (NASs), YS-like (YSL) transporters, oligopeptide transporters (OPTs), RING zinc finger protein, and hemerythrin motif-containing RING- and zinc-finger (HRZ) proteins were demonstrated to be involved in the uptake, transport, and storage of Fe and Zn in plants [41]. Nonsynonymous sequence variations could change protein sequence, which possibly effect gene or protein function. The identify of SNV plays an importance role in genome-assisted research and QTL mapping [42–44].

Nevertheless, *Glyma.03G231200*, *Glyma.20G063100*, *Glyma.20G068300*, and *Glyma.20G088600* do not have nonsynonymous SNV or InDel mutations that can cause changes in proteins. Interestingly, their homologous genes, *OsNAS2*, *OsZIP9*, *YSL1*, and *bHLH104*, play important roles in Fe and Zn homeostasis in other crops. Zn and Fe concentrations of unpolished rice were significantly increased in *OsNAS2* overexpression populations compared with controls [45]. Yang et al. [46] demonstrated that *OsZIP9* was responsible for transporting Zn from external media into root cells, and *GmZIP1* could facilitate the uptake of Zn in yeast cells [47]. Zn, Cu, and Fe were aberrantly distributed in *Arabidopsis* when *YSL1* and *YSL3* were nonfunctional [48]. *bHLH104* and *bHLH34* were demonstrated to positively regulate Fe homeostasis in *Arabidopsis thaliana* [49]. Most likely, there are some complex structural variations, epigenetic modifications, and enhancers resulting in their differential expression between parents, so *Glyma.03G231200*, *Glyma.20G063100*, *Glyma.20G068300*, and *Glyma.20G088600* could be related to the seed Zn content in soybean. *Glyma.03G240000* is a bHLH transcription factor and *Glyma.03G229400* encodes glucose-6-phosphate 1-dehydrogenase 4, which might participate in Zn homeostasis in soybean. *Glyma.07G100700* and *Glyma.07G097700* is a MYB transcription factor and RING zinc finger protein, respectively. *Glyma.07G093700* encode HRZ proteins, and *OsHRZ1*, the homologous gene of *Glyma.07G093700*, is a putative iron-binding sensor that regulates rice responses to Fe deficiency [50]. Therefore, these genes might be responsible for increasing the seed Fe content in soybean. *Glyma.20G076700* and *Glyma.20G067300* encode RING zinc finger

protein and OPT proteins, respectively. The annotation of *Glyma.20G089300* is responsible for metal ion binding. So, these three candidate genes might play important roles in increasing seed Zn content in soybean. To sum up, all 12 candidate genes may be important genetic resources of improving Fe- or Zn- rich soybean cultivars, however they should be further validated.

## Conclusions

In conclusion, we have used an RIL population derived from a cross of Guizao 1 × B13 to identify QTLs and candidate genes related to seed Fe and Zn content in soybean in four environments by means of three QTL mapping approaches and whole-genome resequencing data of two parents. These QTLs and candidate genes related to seed Fe and Zn content can provide vital genetic information to breed Fe- and Zn-fortified cultivars with the help of MAS in soybean. Promisingly, Fe and Zn biofortification projects are being conducted, and several candidate genes are in the process of functional validation to ascertain their effects for increasing seed Fe and Zn content in soybean.

## Methods

### Plant Materials and Field Experiments

The mapping population (GB RIL population) was developed by crossing Guizao 1 and BRSMG68 ('B13' hereafter) and consisted of 248 lines. Guizao1 is a cultivar bred by the Cash Crops Research Institute, Guangxi Academy of Agricultural Sciences, which is rich in mineral elements. B13 was introduced from Brazil with a low content of mineral elements. Both parents and GB RILs were provided by the Guangdong Subcenter of the National Center for Soybean Improvement and South China Agricultural University, Guangzhou, China.

The F<sub>8:11–12</sub> generations of GB RILs together with the two parents were planted at two the following different locations in July 2018 and in July 2019: Guangzhou Experimental Station (N23° 15', E113° 34') in Guangdong and Zengcheng Experimental Station (N23° 24', E113° 64') in Guangdong. The four environments (2018 at Guangzhou Experimental Station, 2019 at Guangzhou Experimental Station, 2018 at Zengcheng Experimental Station, and 2019 at Zengcheng Experimental Station) were assigned as 18GZ, 19GZ, 18ZC, and 19ZC, respectively. All GB RILs and their two parents were planted in a single-line plot of 2 m in length, 0.5 m in width, and 0.1 m between plants in a randomized complete block design with two replicates. In each environment, field management followed normal soybean production practices for the area. In the end, this study absolutely complies with relevant institutional, national, and international guidelines.

### Determination and Analysis of Phenotypic Data

The seed Fe and Zn content in soybean were determined by atomic absorption spectrometry (AAS, Model AA-6800, Shimadzu, Japan) using the wet digestion method. In brief, approximately 100 whole seeds that were randomly selected from the full plot representing each line were ground into a fine powder by a grinder, and then each sample was overdried at 40 °C for approximately 72 h to a constant mass. Then, 0.5 g of dried soybean powder was coldly digested with 10 ml of a diacidic mixture of HNO<sub>3</sub>:HClO<sub>4</sub> (4:1 v/v) for 12 h, followed by heat treatment. The next digestion procedure was set as follows: raising from room temperature to 120 °C and holding at 120 °C for one hour; raising from 120 to 180 °C and holding at 180 °C for 30 min; and then finally raising from 180 to 240 °C and holding at 240 °C to thorough digestion. The resulting solutions were diluted to 50 ml with ultrapure water and then filtered after cooling to room temperature. The Fe and Zn content were determined using AAS.

All phenotypic data were analyzed. Descriptive statistics and Pearson correlations were performed by the Origin21 software (OriginLab Corporation, Northampton, USA.). Frequency distribution graphs of seed Fe and Zn content were depicted by GraphPad Prism 9.0 (GraphPad Software, Inc., USA.). The broad-sense heritability ( $H^2$ ) estimation was calculated using the following equation:

$$H^2 = \sigma_G^2 / (\sigma_G^2 + \sigma_{GE/E}^2 + \sigma_e^2 / rE) \times 100\%$$

where  $\sigma_G^2$  is the genotypic variance,  $\sigma_{GE}^2$  is the genotype × environment interaction variance,  $\sigma_e^2$  is the error variance,  $E$  is the number of environments, and  $r$  is the number of replicates within an environment.

### QTL Detection

The genetic map employed in the current study was constructed in our previous study, including 3748 bins and 3031.9 cM in length, with an average distance of 0.81 cM between adjacent markers on 20 chromosomes [29]. QTL analysis was carried out by composite interval mapping (CIM) in Windows QTL Cartographer V2.5 and inclusive composite interval mapping (ICIM-ADD) in IciMapping V4.2 [51, 52]. QTL × environment interaction (QEI) mapping for multi-environmental trials was performed using ICIM-ADD [40]. The LOD thresholds of significant QTLs were determined by performing 1000 permutations using a Type I error set at  $p < 0.05$ . QTL mapping results were comprehensively compared to SoyBase (<http://www.soybase.org/>). The physical locations of QTLs were drawn on related chromosomes using the LinkageMapView

package in R [53]. Standard nomenclature was involved in naming QTLs [54]. QTL detected at least three times by either one of CIM and ICIM-ADD methods was termed as a stable QTL. QTL that was detected with PVE of more 10% in CIM or ICIM-ADD methods and could be simultaneously detected in MET method, which was categorized a major QTL [55, 56]. In this study, QTLs with large shared position or peaks of adjacent QTL distanced to less than 5 cM were regarded as the same QTL [38].

### Whole genome resequencing of two parents

WGS approach was implemented in bi-parental genome using Illumina HiSeq2000 at the Beijing Genome Institute (BGI) Tech, Shenzhen, China [57], details of technique and analysis are as described in Jiang et al. [29]. In brief, using the method of CTAB extracted the genomic DNA of two parent [58]. Illumina libraries were built according to Illumina protocol, and 90-bp paired-end reads were produced. Using Burrows-Wheeler Alignment tool (BWA), sequenced reads were mapped the Wm82. a2. v1 (soybean reference genome) from SoyBase (<http://www.soybase.org/>) [59, 60]. Aligned reads were dealt with samtools, and Picard software was used to remove PCR duplicates of aligned reads. Finally, high-confidence SNVs and InDels were obtained by using (GATK) [61].

### Candidate Genes Identification

In the current study, the gene models that fell into mapping intervals of detected QTLs were retrieved from Wm82.a2.v1 (soybean reference genome) in SoyBase (<http://www.soybase.org/>). We carried out Gene Ontology (GO) enrichment analysis for all the genes within the three major QTL intervals using OmicShare Tools (<http://www.omicshare.com/tools>). The Guizao1 and B13 resequencing data were used to identify candidate genes possessing nonsynonymous SNV or InDel mutations in coding regions. The available RNA-Seq dataset from SoyBase (<http://www.soybase.org/soyseq/>) was downloaded and used to analyze the expression of candidate genes in different developmental stages of soybean tissues [62]. In summary, these methods were used in combination to identify credible candidate genes.

### Abbreviations

Fe: Iron; Zn: Zinc; QTL: Quantitative trait loci; RIL: Recombinant inbred line; MAS: Marker-assisted selection; NGS: Next-generation sequencing; SD: Standard deviation; CV: Coefficient of variance;  $H^2$ : Broad-sense heritability; Chr: Chromosome; CIM: Composite interval mapping method; ICIM-ADD: Inclusive composite interval mapping of the additive method; QE: QTL × environment interaction; LOD: Logarithm of odds; RNA-Seq: RNA sequencing; GO: Gene Ontology; SNP: Single nucleotide polymorphism; SNV: Single-nucleotide variants; InDel: Insertion and deletion.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-022-08313-1>.

**Additional file 1: Table S1.** The significant LOD thresholds were determined by 1000 permutations ( $p < 0.05$ ). **Table S2.** Overall genes related to seed Fe and Zn content in three major QTL regions. **Table S3.** GeneOntology (GO) enrichment analysis of overall genes in three major QTL regions. **Table S4.** Significant SNV mutations and mutant genes between Guizao1 and B13 in three major QTLs. **Table S5.** Significant InDel mutations and mutant genes between Guizao1 and B13 in three major QTLs. **Table S6.** Expression analysis of mutant genes in soybean in different developmental stages and tissues. **Table S7.** Variations in putative candidate genes within *qFC7*, *qZC3*, and *qZC20*. **Table S8.** Raw phenotyping data of GB RIL population and two parents collected in four environments.

**Additional file 2: Fig. S1.** The significant differences between Guizao1 and B13 for seed Fe and Zn content Histogram showing the significant differences between Guizao1 and B13 for seed Fe and Zn content in the four environments, according to the ANOVA analysis denoted as follows: \*\* $p < 0.01$ . **Fig. S2.** The expression patterns analysis of mutant genes from three major QTLs Heatmap showing the expression pattern of mutant genes from three major QTLs among the different tissues during soybean development stages based on public RNA-seq data from SoyBase. DAF: days after flowering. a, b, and c represent mutant genes from *qZC3*, *qFC7*, and *qZC20*, respectively.

### Acknowledgements

Not applicable.

### Authors' contributions

H. N., H.W., J. J., Z.C., M. D., Z. J., and T.L. collected the plant materials used in this study and carried out the seed Fe and Zn content analyses. Q.X. and H.W. analyzed the genome-wide resequencing and QTL mapping data. H.W. and H. N. prepared the first draft of the manuscript. H.N., Q.M., and T.L. planned, supervised and financed this work and edited the manuscript. All authors have read and approved the final version of the manuscript.

### Funding

This work was supported by the R & D Project in Key Areas of Guangdong Province (2020B020220008), the China Agricultural Research System (CARS-04-PS09), the Guangdong Agricultural Research System (2019KJ136), and the National Key R&D Program of China (2017YFD0101505).

### Availability of data and materials

The data that support the findings of this study are available from the Genome Sequence Archive database at the National Genomics Data Center, Beijing Institute of Genomics (BIG), Chinese Academy of Sciences, with accession number CRA004753 (<https://bigd.big.ac.cn/gsa/browse/CRA004753>) and CRA004754 (<https://bigd.big.ac.cn/gsa/browse/CRA004754>). The phenotype dataset used during the current study is provided in the Additional file 1: Table S8.

### Declarations

#### Ethics approval and consent to participate

This study complies with relevant institutional, national, and international guidelines.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>The State Key Laboratory for Conservation and Utilization of Subtropical Agro-bioresources, South China Agricultural University, 510642 Guangzhou,

Guangdong, People's Republic of China. <sup>2</sup>Guangdong Laboratory for Lingnan Modern Agriculture, 510642 Guangzhou, Guangdong, People's Republic of China. <sup>3</sup>The Key Laboratory of Plant Molecular Breeding of Guangdong Province, College of Agriculture, South China Agricultural University, 510642 Guangzhou, Guangdong, People's Republic of China. <sup>4</sup>Rice Molecular Breeding Institute, GRANLUX ASSOCIATED GRAINS, 518024 Shenzhen, Guangdong, People's Republic of China.

Received: 3 August 2021 Accepted: 13 January 2022

Published online: 19 February 2022

## References

- Garcia-Oliveira AL, Chander S, Ortiz R, Menkir A, Gedil M. Genetic Basis and Breeding Perspectives of Grain Iron and Zinc Enrichment in Cereals. *Front Plant Sci.* 2018;9:937.
- FAO I, WFP. The State of Food Insecurity in the World 2015. Meeting the 2015 international hunger targets: taking stock of uneven progress. Rome, Italy: FAO; 2015.
- Welch RM, Graham RD. Breeding for micronutrients in staple food crops from a human nutrition perspective. *J Exp Bot.* 2004;55(396):353–364.
- Waters BM, Grusak MA. Quantitative trait locus mapping for seed mineral concentrations in two *Arabidopsis thaliana* recombinant inbred populations. *New Phytol.* 2008;179(4):1033–1047.
- Jankowska EA, von Haehling S, Anker SD, Macdougall IC, Ponikowski P. Iron deficiency and heart failure. diagnostic dilemmas and therapeutic perspectives. *Eur Heart J.* 2013;34(11):816–829.
- Ejaz MS, Latif N. Stunting and micronutrient deficiencies in malnourished children. *J Pak Med Assoc.* 2010;60(7):543–547.
- Salgueiro MJ, Zubillaga MB, Lysionek AE, Caro RA, Weill R, Boccio JR. The role of zinc in the growth and development of children. *Nutrition.* 2002;18(6):510–519.
- Plum LM, Rink L, Haase H. The essential toxin: impact of zinc on human health. *Int J Env Res Pub He.* 2010;7(4):1342–1365.
- Ma G, Jin Y, Piao J, Kok F, Guusje B, Jacobsen E. Phytate, Calcium, Iron, and Zinc Contents and Their Molar Ratios in Foods Commonly Consumed in China. *J Agr Food Chem.* 2005;53(26):10285–10290.
- Bouis HE. Plant breeding. A new tool for fighting micronutrient malnutrition. *J Nutr.* 2002;132(3):491s–494s.
- Bouis HE, Saltzman A. Improving nutrition through biofortification: A review of evidence from HarvestPlus, 2003 through 2016. *Glob Food Secur-Agr.* 2017;12:49–58.
- Huang Y, Sun C, Min J, Chen Y, Tong C, Bao J. Association Mapping of Quantitative Trait Loci for Mineral Element Contents in Whole Grain Rice (*Oryza sativa* L.). *J Agric Food Chem.* 2015;63(50):10885–10892.
- Pujar M, Gangaprasad S, Govindaraj M, Gangurde SS, Kanatti A, Kudapa H. Genome-wide association study uncovers genomic regions associated with grain iron, zinc and protein content in pearl millet. *Sci Rep.* 2020;10(1):19473.
- Wu X, Islam A, Limpot N, Mackasmil L, Mierzwa J, Cortes AJ, Blair MW. Genome-Wide SNP Identification and Association Mapping for Seed Mineral Concentration in Mung Bean (*Vigna radiata* L.). *Front Genet.* 2020;11:656.
- Upadhyaya HD, Bajaj D, Das S, Kumar V, Gowda CL, Sharma S, Tyagi AK, Parida SK. Genetic dissection of seed-iron and zinc concentrations in chickpea. *Sci Rep.* 2016;6:24050.
- Yu YH, Shao YF, Liu J, Fan YY, Sun CX, Cao ZY, Zhuang JY. Milled Rice (*Oryza sativa* L.). *J Agric Food Chem.* 2015;63(35):7813–7818.
- Crespo-Herrera LA, Govindan V, Stangoulis J, Hao Y, Singh RP. QTL Mapping of Grain Zn and Fe Concentrations in Two Hexaploid Wheat RIL Populations with Ample Transgressive Segregation. *Front Plant Sci.* 2017;8:1800.
- Blair MW, Medina JI, Astudillo C, Rengifo J, Beebe SE, Machado G, Graham R. QTL for seed iron and zinc concentration and content in a Mesoamerican common bean (*Phaseolus vulgaris* L.) population. *Theor Appl Genet.* 2010;121(6):1059–1070.
- Blair MW, Astudillo C, Rengifo J, Beebe SE, Graham R. QTL analyses for seed iron and zinc concentrations in an intra-genepool population of Andean common beans (*Phaseolus vulgaris* L.). *Theor Appl Genet.* 2011;122(3):511–521.
- Ma Y, Coyne CJ, Grusak MA, Mazourek M, Cheng P, Main D, McGee RJ. Genome-wide SNP identification, linkage map construction and QTL mapping for seed mineral concentrations and contents in pea (*Pisum sativum* L.). *BMC Plant Biol.* 2017;17(1):43.
- Suman K, Neeraja CN, Madhubabu P, Rathod S, Bej S, Jadhav KP, Kumar JA, Chaitanya U, Pawar SC, Rani SH, et al. Identification of Promising RILs for High Grain Zinc Through Genotype x Environment Analysis and Stable Grain Zinc QTL Using SSRs and SNPs in Rice (*Oryza sativa* L.). *Front Plant Sci.* 2021;12:587482.
- King KE, Peiffer GA, Reddy M, Lauter N, Lin SF, Cianzio S, et al. Mapping of Iron and Zinc Quantitative Trait Loci in Soybean for Association to Iron Deficiency Chlorosis Resistance. *J Plant Nutr.* 2013;36(14):2132–2153.
- Ramamurthy RK, Jedlicka J, Graef GL, Waters BM. Identification of new QTLs for seed mineral, cysteine, and methionine concentrations in soybean [*Glycine max* (L.) Merr.]. *Mol Breeding.* 2014;34(2):431–445.
- Ning L, Sun P, Wang Q, Ma D, Hu Z, Zhang D, Zhang G, Cheng H, Yu D. Genetic architecture of biofortification traits in soybean (*Glycine max* L. Merr.) revealed through association analysis and linkage mapping. *Euphytica.* 2014;204(2):353–369.
- Collard BCY, Jahufer MZZ, Brouwer JB, Pang ECK. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica.* 2005;142(1-2):169–196.
- Boer MP, Wright D, Feng L, Podlich DW, Luo L, Cooper M, van Eeuwijk FA. A mixed-model quantitative trait loci (QTL) analysis for multiple-environment trial data using environmental covariables for QTL-by-environment interactions, with an example in maize. *Genetics.* 2007;177(3):1801–1813.
- Cai Z, Cheng Y, Ma Z, Liu X, Ma Q, Xia Q, Zhang G, Mu Y, Nian H. Fine-mapping of QTLs for individual and total isoflavone content in soybean (*Glycine max* L.) using a high-density genetic map. *Theor Appl Genet.* 2018;131(3):555–568.
- Cai Z, Cheng Y, Xian P, Ma Q, Wen K, Xia Q, Zhang G, Nian H. Acid phosphatase gene *GmHAD1* linked to low phosphorus tolerance in soybean, through fine mapping. *Theor Appl Genet.* 2018;131(8):1715–1728.
- Jiang B, Li M, Cheng Y, Cai Z, Ma Q, Jiang Z, Ma R, Xia Q, Zhang G, Nian H. Genetic mapping of powdery mildew resistance genes in soybean by high-throughput genome-wide sequencing. *Theor Appl Genet.* 2019;132(6):1833–1845.
- Garg M, Sharma N, Sharma S, Kapoor P, Kumar A, Chunduri V, Arora P. Biofortified Crops Generated by Breeding, Agronomy, and Transgenic Approaches Are Improving Lives of Millions of People around the World. *Front Nutr.* 2018;5:12.
- Gu R, Chen F, Liu B, Wang X, Liu J, Li P, Pan Q, Pace J, Soomro AA, Lubberstedt T, et al. Comprehensive phenotypic analysis and quantitative trait locus identification for grain mineral concentration, content, and yield in maize (*Zea mays* L.). *Theor Appl Genet.* 2015, 128(9):1777–1789.
- Phuke RM, Anuradha K, Radhika K, Jabeen F, Anuradha G, Ramesh T, Hariprasanna K, Mehtre SP, Deshpande SP, Anil G, et al. Genetic Variability, Genotype x Environment Interaction, Correlation, and GGE Biplot Analysis for Grain Iron and Zinc Concentration and Other Agronomic Traits in RIL Population of Sorghum (*Sorghum bicolor* L. Moench). *Front Plant Sci.* 2017;8(712).
- Kumar S, Hash CT, Thirunavukkarasu N, Singh G, Rajaram V, Rathore A, Senapathy S, Mahendrakar MD, Yadav RS, Srivastava RK. Mapping Quantitative Trait Loci Controlling High Iron and Zinc Content in Self and Open Pollinated Grains of Pearl Millet [*Pennisetum glaucum* (L.) R. Br.]. *Front Plant Sci.* 2016;7:1636.
- Rai KN, Govindaraj M, Rao AS. Genetic enhancement of grain iron and zinc content in pearl millet. *Qual Assur Saf Crop.* 2012;4(3):119–125.
- Calayugan MIC, Formantes AK, Amparado A, Descalsota-Empelo GI, Nha CT, Inabangan-Asilo MA, Swe ZM, Hernandez JE, Borromeo TH, Lalusin AG, et al. Genetic Analysis of Agronomic Traits and Grain Iron and Zinc Concentrations in a Doubled Haploid Population of Rice (*Oryza sativa* L.). *Sci Rep.* 2020;10(1):2283.
- Qin J, Shi A, Mou B, Grusak MA, Weng Y, Ravelombola W, Bhattarai G, Dong L, Yang W. Genetic diversity and association mapping of mineral element concentrations in spinach leaves. *BMC Genomics.* 2017;18(1):941.
- Su C, Qiu X, Ji Z. Study of strategies for selecting quantitative trait locus mapping procedures by computer simulation. *Mol Breeding.* 2013;31(4):947–956.

38. Yang Y, Xuan L, Yu C, Wang Z, Xu J, Fan W, Guo J, Yin Y. High-density genetic map construction and quantitative trait loci identification for growth traits in (*Taxodium distichum* var. *distichum* x *T. mucronatum*) x *T. mucronatum*. *BMC Plant Biol.* 2018;18(1):263.
39. Yu M, Liu ZX, Jiang SS, Xu N, Chen QS, Qi ZM, Lv WH. QTL mapping and candidate gene mining for soybean seed weight per plant. *Biotechnol Biochem Eq.* 2018;32(4):908–914.
40. Li S, Wang J, Zhang L. Inclusive Composite Interval Mapping of QTL by Environment Interactions in Biparental Populations. *PLoS One.* 2015;10(7):e0132414.
41. Tong J, Sun M, Wang Y, Zhang Y, Rasheed A, Li M, Xia X, He Z, Hao Y. Dissection of Molecular Processes and Genetic Architecture Underlying Iron and Zinc Homeostasis for Biofortification: From Model Plants to Common Wheat. *Int J Mol Sci.* 2020;21(23):9280.
42. International Barley Genome Sequencing C, Mayer KF, Waugh R, Brown JW, Schulman A, Langridge P, Platzer M, Fincher GB, Muehlbauer GJ, Sato K, et al. A physical, genetic and functional sequence assembly of the barley genome. *Nature.* 2012;491(7426):711–6.
43. Yeo S, Hodgkinson CA, Zhou Z, Jung J, Leung M, Yuan Q, Goldman D. The abundance of cis-acting loci leading to differential allele expression in F1 mice and their relationship to loci harboring genes affecting complex traits. *BMC Genomics.* 2016;17(1):620.
44. Zhang H, Wang X, Pan Q, Li P, Liu Y, Lu X, Zhong W, Li M, Han L, Li J, et al. QTL-Seq Accelerates QTL Fine Mapping through QTL Partitioning and Whole-Genome Sequencing of Bulk Segregant Samples. *Mol Plant.* 2019;12(3):426–437.
45. Johnson AA, Kyriacou B, Callahan DL, Carruthers L, Stangoulis J, Lombi E, Tester M. Constitutive overexpression of the *OsNAS* gene family reveals single-gene strategies for effective iron- and zinc-biofortification of rice endosperm. *PLoS One.* 2011;6(9):e24476.
46. Yang M, Li YT, Liu ZH, Tian JJ, Liang LM, Qiu Y, Wang GY, Du QQ, Cheng D, Cai HM, et al. A high activity zinc transporter *OsZIP9* mediates zinc uptake in rice. *Plant J.* 2020;103(5):1695–1709.
47. Moreau S, Thomson RM, Kaiser BN, Trevaskis B, Guerinot ML, Udvardi MK, Puppo A, Day DA. *GmZIP1* encodes a symbiosis-specific zinc transporter in soybean. *J Biol Chem.* 2002;277(7):4738–4746.
48. Waters BM, Grusak MA. Whole-plant mineral partitioning throughout the life cycle in *Arabidopsis thaliana* ecotypes Columbia, Landsberg erecta, Cape Verde Islands, and the mutant line ysl1ysl3. *New Phytol.* 2008;177(2):389–405.
49. Li XL, Zhang HM, Ai Q, Liang G, Yu DQ. Two bHLH Transcription Factors, bHLH34 and bHLH104, Regulate Iron Homeostasis in *Arabidopsis thaliana*. *Plant Physiol.* 2016;170(4):2478–2493.
50. Zhang HM, Li Y, Yao XN, Liang G, Yu DQ. POSITIVE REGULATOR OF IRON HOMEOSTASIS1, OsPRI1, Facilitates Iron Homeostasis. *Plant Physiol.* 2017;175(1):543–554.
51. Wang S, CJB, and Z.-B. Zeng. Windows QTL Cartographer 2.5. In: Department of Statistics, North Carolina State University, Raleigh, NC; 2012.
52. Meng L, Li H, Zhang L, Wang J. QTL IciMapping: Integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *The Crop Journal.* 2015;3(3):269–283.
53. Ouellette LA, Reid RW, Blanchard SG, Brouwer CR. LinkageMapView-rendering high-resolution linkage and QTL maps. *Bioinformatics.* 2018;34(2):306–307.
54. McCouch SR, Chen X, Panaud O, Temnykh S, Xu Y, Cho YG, Huang N, Ishii T, Blair M. Microsatellite marker development, mapping and applications in rice genetics and breeding. *Plant Mol Biol.* 1997;35(1-2):89–99.
55. Hu J, Wang X, Zhang G, Jiang P, Chen W, Hao Y, Ma X, Xu S, Jia J, Kong L, et al. QTL mapping for yield-related traits in wheat based on four RIL populations. *Theor Appl Genet.* 2020;133(3):917–933.
56. Shen Y, Yang Y, Xu E, Ge X, Xiang Y, Li Z. Novel and major QTL for branch angle detected by using DH population from an exotic introgression in rapeseed (*Brassica napus* L.). *Theor Appl Genet.* 2018;131(1):67–78.
57. Shendure J, Ji H. Next-generation DNA sequencing. *Nat Biotechnol.* 2008;26(10):1135–1145.
58. Duan M, Sun Z, Shu L, Tan Y, Yu D, Sun X, Liu R, Li Y, Gong S, Yuan D. Genetic analysis of an elite super-hybrid rice parent using high-density SNP markers. *Rice (N Y).* 2013;6(1):21.
59. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25(16):2078–2079.
60. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009;25(14):1754–1760.
61. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9):1297–1303.
62. Severin AJ, Woody JL, Bolon YT, Joseph B, Diers BW, Farmer AD, Muehlbauer GJ, Nelson RT, Grant D, Specht JE, et al. RNA-Seq Atlas of Glycine max: a guide to the soybean transcriptome. *BMC Plant Biol.* 2010;10:160.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

