

RESEARCH PAPER

Toward identification of a putative candidate gene for nutrient mineral accumulation in wheat grains for human nutrition purposes

Dalia Z. Alomari^{1,*}, Ahmad M. Alqudah², Klaus Pillen², Nicolaus von Wirén¹ and Marion S. Röder¹

¹ Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Corrensstrasse 3, D-06466 Stadt Seeland OT Gatersleben, Germany

² Institute of Agricultural and Nutritional Sciences, Martin-Luther-University Halle-Wittenberg, Betty-Heimann-Str. 3, 06120 Halle/Saale, Germany

* Correspondence: alomari@ipk-gatersleben.de or alamridalia@gmail.com

Received 13 February 2021; Editorial decision 15 June 2021; Accepted 16 June 2021

Editor: Greg Rebetzke, CSIRO Agriculture and Food, Australia

Abstract

A multilocus genome-wide association study of a panel of 369 diverse wheat (*Triticum aestivum*) genotypes was carried out in order to examine the genetic basis of variations in nutrient mineral concentrations in the grains. The panel was grown under field conditions for three consecutive years and the concentrations of Ca, K, Mg, Mn, P, and S were determined. Wide ranges of natural variation were detected among the genotypes. Strong positive correlations were found among the minerals except for K, which showed negative correlation trends with the other minerals. Genetic association analysis detected 86 significant marker–trait associations (MTAs) underlying the natural variations in mineral concentrations in grains. The major MTA was detected on the long arm of chromosome 5A and showed a pleiotropic effect on Ca, K, Mg, Mn, and S. Further significant MTAs were distributed among the whole genome except for chromosomes 3D and 6D. We identified putative candidate genes that are potentially involved in metal uptake, transport, and assimilation, including *TraesCS5A02G542600* on chromosome 5A, which was annotated as a Major Facilitator Superfamily transporter and acted on all the minerals except K. *TraesCS5A02G542600* was highly expressed in seed coat, and to a lesser extent in the peduncle, awns, and lemma. Our results provide important insights into the genetic basis of enhancement of nutrient mineral concentrations that can help to inform future breeding studies in order to improve human nutrition.

Keywords: Grain minerals, GWAS, Major Facilitator Superfamily transporter, ICP-OES, *Triticum aestivum*, wheat.

Introduction

Cereal-based foods represent the largest proportion of the human daily diet consumed worldwide and wheat (*Triticum aestivum*) is the primary protein source in developing countries, with 2.5 billion consumers in 89 countries (<https://wheat.org/>). However, wheat grain is inherently relatively low in nutrients, where ‘nutrients’ refers to a broad range of minerals and vitamins that play important roles in the biological functioning of the human body (Tapiero *et al.*, 2003; Pujar *et al.*, 2020).

Many wheat research programs have focused on increasing yields whilst ignoring grain quality, leading to depletion in the nutrient composition and decreased nutritional value. This can lead to high incidences of nutrient deficiencies and malnutrition occurring in countries that rely on a cereal-based diet and where bread wheat forms the majority of the daily intake of calories. For example, the World Health Organization has reported that in 2012, 162 and 99 million children were stunted and underweight, respectively, mainly because of insufficient intake of essential nutrients (WHO, 2013). The importance of developing wheat genotypes with improved nutrient contents in the grains is therefore clear.

Wheat is the second most important staple crop worldwide: it is cultivated on almost 200×10^6 ha, feeds more than a third of the world's population, and provides a fifth of the total calories consumed worldwide (Godfray *et al.*, 2010). Ray *et al.* (2013) anticipated that by 2050 the food demand for wheat in particular would increase by 35–40%. In response, recent research has begun to focus on exploiting natural variations among genotypes to identify those that have not only good yields but also high quality in terms of nutritional value (Alomari *et al.*, 2017; Yu and Tian, 2018).

Recent developments in molecular markers and genome sequencing technologies combined with the release of the high-quality wheat reference genome sequence allow plant researchers to characterize the genetic basis of complex phenotypic traits by using hundreds of thousands of genetic markers in association mapping and in the detection of quantitative trait loci (QTLs) (Appels *et al.*, 2018). Applying high-density marker arrays for single-nucleotide polymorphisms (SNPs) combined with a suitable approach such as genome-wide association studies (GWAS) can identify robust QTLs and help detect genes underlying complex phenotypic traits (Rasheed and Xia, 2019; Alqudah *et al.*, 2020). Nutrient accumulation in wheat grains is one such example, as it is a function of complex inherited traits that are controlled by several different factors, including nutrient uptake by the roots from the soil, translocation, assimilation, and remobilization to the grains (Sperotto *et al.*, 2014).

Another approach that has recently become popular in plant breeding research is based on using molecular markers covering the whole genome, the so-called genomic prediction (GP) approach. This uses genome-wide marker information to predict the breeding value of complex traits in order to accelerate breeding programs (Meuwissen *et al.*, 2001). The accuracy of GP can vary according to the prediction method that is used, as different assumptions and treatments of marker effects and models are utilised (Desta and Ortiz, 2014).

In the current study, we aimed to examine the natural variations in nutrient mineral accumulation in grains of elite European hexaploid winter wheat genotypes. Using high-density SNP arrays, we attempted to detect stable genomic regions associated with natural variations in calcium (Ca), magnesium (Mg), manganese (Mn), phosphorus (P), potassium (K), and sulphur (S) based on multi-locus GWAS, and to assess their value in terms of genomic prediction and breeding potential. The ultimate goal was to identify the most relevant

candidate genes potentially involved in controlling mineral accumulation in wheat grains with a view to assisting future breeding programs aimed at crop improvement.

Materials and methods

Plant material

The wheat (*Triticum aestivum*) germplasm that we used consisted of 369 European elite registered varieties, made up of 355 winter varieties and 14 spring varieties (Supplementary Table S1). The genotypes were mostly from Germany and France. Field experiments were conducted at IPK, Gatersleben, Germany (51°49'N, 11°17'0"E), during the growing seasons of 2014/2015, 2015/2016, and 2016/2017, with the full set of genotypes being grown each year in plots of 2×2 m with six rows spaced 0.20 m apart. All varieties were subjected to standard agronomic wheat management practices.

Grain sampling and measurement of nutrients

For each genotype, ears were harvested by hand from the whole plot and a randomly selected sample of kernels were placed in a MARViN digital seed analyser/counter (GTA Sensorik GmbH, Neubrandenburg, Germany) and 50 were collected for mineral analysis. These kernels were also used for determination of thousand-kernel weight (TKW). The samples for mineral analysis were ground using a Mixer Mill MM 300 (Retsch GmbH, Germany) and dried at 40 °C in an incubator overnight. The concentrations of Ca, K, Mg, Mn, P, and S were measured by inductively coupled plasma optical emission spectrometry (ICP-OES) using an iCAP 6000 (ThermoFisher Scientific) combined with an ASXPRESS PLUS Rapid Sample Introduction system and an ASX-560 autosampler (both Teledyne CETAC Technologies, Omaha, NE, USA). Each ground sample was subjected to wet digestion in 2 ml 69% nitric acid (HNO₃) using a high-performance microwave reactor (UltraClave IV, MLS, Germany). The digested samples were made up to 15 ml final volume with de-ionized distilled water. Element standards were prepared from Bernd Kraft GmbH multi-element standard solution (Duisburg, Germany). Standard solutions were used for determination of the minerals, and yttrium was used as an internal standard for matrix correction (ICP Standard Certipur, Merck, Germany).

Genotyping and quality control for markers

The wheat panel was genotyped using two marker arrays: a 90K iSELECT Infinium array including 7761 markers and a 35K Affymetrix SNP array including 7762 markers (Wang *et al.*, 2014) (Axiom® Wheat Breeder's Genotyping Array, <https://www.cerealsdb.uk.net/cerealgenomics/>). Both were genotyped by SGS-TraitGenetics GmbH, Gatersleben, Germany (www.traitgenetics.com) as detailed previously (Alomari *et al.*, 2019). The ITMI-DH population was used as a reference map to anchor the SNP markers of the two arrays (Sorrells *et al.*, 2011; Poland *et al.*, 2012). To obtain high-quality makers, the SNP markers of the 90K and 35K chips were subjected to a quality control and filtration process by removing those with ≤3% missing values, a minor allele frequency (MAF) of ≤3%, and those with unknown chromosomal positions. We then used the physical position of wheat genome sequence RefSeq v.1.0 for the SNPs.

GWAS and genomic prediction

The GWAS analysis was carried out using the Genomic Association and Prediction Integrated Tool (GAPIT) in the R software (Lipka *et al.*, 2012). First, the GWAS analysis was computed using a mixed linear model (MLM) that took into account the variance-covariance kinship matrix and principle component analysis (PCA), and conducted by

incorporating the phenotypic and genotypic dataset. The kinship matrix was calculated using the VanRaden method (VanRaden, 2008) to determine relative kinship among the sampled individuals. Both PCA and kinship matrix were used for population correction and stratification.

We also applied a GAPIT model known as ‘fixed and random model circulating probability unification’ (FarmCPU) to our data analysis. This was applied by considering the incorporation of multiple markers simultaneously as a covariate in a fixed-effects model and optimization on the associated covariate markers separately in a random effects model, which enabled us to avoid any false-negatives and to control the false-positive associations by preventing model over-fitting (Liu *et al.*, 2016). Thus, FarmCPU is a powerful tool with less false-positives than MLM. The selection of an appropriate model and thresholds are important steps in identifying markers that are truly associated with specific traits and that might be located within or very close to genes that control the trait variation while also controlling both false-positive and false-negative associations. To determine which of the tested models best fitted the data, we plotted quantile-quantile (Q-Q) plots based on the observed and expected $-\log_{10}(P)$ values. Based on the GWAS output of three models (GLM, MLM, and FarmCPU), the number of significant associations, and the resulting Q-Q plots, we then selected the FarmCPU model. A threshold P -value 0.001 equal to $-\log_{10}(P) \geq 3$ was used to determine the significance of marker-trait associations, and then Bonferroni correction at $P < 0.05$ was used to adjust the $-\log_{10}(P)$ threshold to a value of 5.49 for the studied traits.

Marker effects (positive or negative) and phenotypic variance explained by the associated markers (R^2) were removed from the GWAS results.

Ridge regression–best linear unbiased prediction (RR-BLUP) and Bayes-C π methods were used to evaluate the genomic predictions (Meuwissen *et al.*, 2001; Habier *et al.*, 2011), with both methods being implemented in the R software. Five-fold cross-validation was applied to the complete set, with the set being randomly divided into five subsets, four of which were used as estimation sets and the remaining ones were used as the test set. The prediction ability was calculated from the correlation between the observed and predicted values. The whole process was repeated 100 times to obtain a mean value.

Gene identification, annotation, and expression analysis

Significant markers and the markers located within the linkage disequilibrium region ($r^2 \geq 0.2$) were considered for BLAST. The sequences of the identified markers were obtained from the wheat 90K (Wang *et al.*, 2014) and 35K databases (Allen *et al.*, 2017). Marker sequences were BLASTed against the recently released IWGSC RefSeq v.1.1 genome using EnsemblPlants (http://plants.ensembl.org/Triticum_aestivum) to obtain their gene annotations. The expression profiles of all the putative candidate genes associated with the identified SNPs obtained from the published RNA-seq expression database for wheat in the WheatGmap web tool (<https://www.wheatgmap.org>; Zhang *et al.*, 2020).

Statistical analysis

Significant differences in minerals among genotypes and years were determined using ANOVA and Pearson correlation coefficients were determined to evaluate strengths of relationships among the measured parameters, both using GenStat v.19. This software was also used to determine best linear unbiased estimates (BLUEs) over the three years of the experiments by restricted maximum likelihood (REML) analysis with a mixed linear model (MLM) where the genotype considered as a fixed effect and the environment as a random effect. BLUEs were calculated for each genotype of each trait across the years.

Broad-sense heritability was calculated for each trait using the formula $H^2 = V_G / [V_G + (G_e/nE)]$, where V_G is the variance of the genotype, G_e is the variance of the residual, and nE is the number of years.

Results

Genetic variation and correlations

The wheat genotypes showed wide variations in mineral concentrations in the grains across the three years (Fig. 1a) and BLUEs that followed approximately normal distributions (Supplementary Fig. S1). The variation ranged from 208–797 $\mu\text{g g}^{-1}$ for Ca, 3495–6727 $\mu\text{g g}^{-1}$ for K, 963–1988 $\mu\text{g g}^{-1}$ for Mg, 23.3–62.2 $\mu\text{g g}^{-1}$ for Mn, 2943 to 5807 $\mu\text{g g}^{-1}$ for P, and 974–2368 $\mu\text{g g}^{-1}$ for S (Supplementary Table S2). Concentrations of Ca and S were highest in 2015, K, P, and Mn concentrations were highest in 2016, and Mg concentration did vary appreciably over the three years. This variation among years could be explained by environmental influences (Supplementary Table S3). Correlation analysis was performed for the nutrient traits and for thousand-kernel weight (TKW) based on BLUE values (Fig. 1B). Strong positive correlations were found between Mg and P, and between Mg and Mn ($r=0.69$ and $r=0.63$, respectively, $P < 0.05$), while moderate positive correlations were detected between Ca and Mg ($r=0.25$), Mn ($r=0.43$), P ($r=0.25$), and S ($r=0.270$ (all $P < 0.01$). Negative correlations were detected between K and Ca, Mg, Mn, and S, but this was not significant for both Ca. All the minerals showed only very weak correlations with TKW. In this study, the five genotypes with highest measured mineral concentrations were Isengrain, Inoui, Nirvana, Exotic, and Lona.

ANOVA indicated that significant differences concentrations existed among the genotypes ($P < 0.001$) and the years ($P < 0.001$) for all the minerals, except Mn where there was no effect of year (Supplementary Table S3). The broad-sense heritability (H^2) values were high and ranged from 0.72 for Mn to 0.87 for Ca. These results indicated that there was considerable natural variation in mineral concentrations that were predominantly genetically controlled with a relatively low influence of the environment. This suggested that it might be possible to detect shared genomic regions for the mineral traits.

Association analysis and genomic predictions

GWAS analysis was performed to determine the genetic basis for the accumulation of the minerals in the grains using FarmCPU for 15 523 SNP markers. Most of the markers were mapped on the B-genome, followed by the A- and D-genomes. In brief, the marker-trait associations (MTAs) were distributed across the whole genome, except for 3D and 6D where 86 significant associations were detected [$-\log_{10}(P) \geq 3$], and the highest numbers of significant markers were found on 2B (12) followed by 5B (nine) and 3B (eight). The MTAs were identified using estimated BLUE values across the three years (Supplementary Table S4A). A total of 17 MTAs were above the Bonferroni threshold of $-\log_{10}(P\text{-value})=5.42$ (Fig. 2A, Supplementary Table S4B). A total of 50 MTAs had positive effects and 36 had negative effects on the mineral concentrations (Supplementary

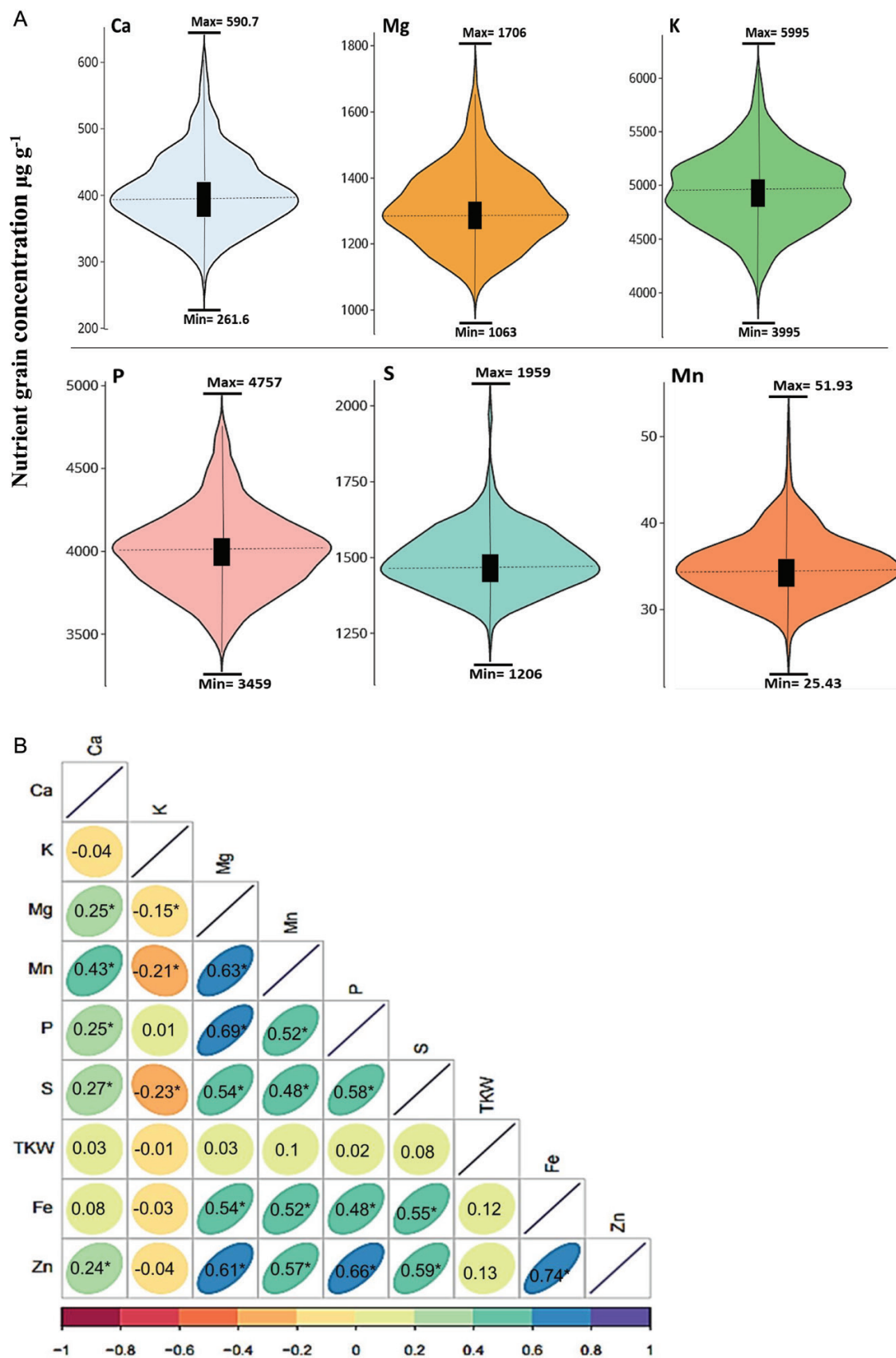


Fig. 1. Mineral nutrient concentrations in the grains of 369 European wheat genotypes grown in over three seasons. (A) Violin plots mineral concentrations. (B) Pearson correlations among mineral concentrations and thousand-kernel weight (TKW) based on best linear unbiased estimates (BLUEs). Significant correlations were determined using ANOVA ($P \leq 0.05$).

Table S4A). The phenotypic variance explained by each SNP ranged from 0.04–10.54% (Supplementary Table S4A, B). RAC875_c8642_231 showed the highest $-\log_{10}(P\text{-value})$ of Ca, Mg, Mn, and K (Figs 2A, 3A) and had a positive effect on Ca, Mn, Mg, and S (26, 1.8, 49, and 28 $\mu\text{g g}^{-1}$, respectively) and a negative effect on K ($-117 \mu\text{g g}^{-1}$) (Supplementary Table S4B). This marker could explain $\sim 10\%$ of the variation in Ca and $\sim 3.4\%$ in K. This indicated that there was allelic variation at this marker among the genotypes that lead to variation in the mineral concentrations (Supplementary Table S5). The Q-Q plots showing the observed associations between SNPs and grain nutrient concentrations compared to expected associations are presented in Fig. 2B.

Statistical methods for developing prediction models for the breeding value of studied traits include ridge regression–best linear unbiased prediction (RR-BLUP) and the Bayes-C π method. We found that the mean of prediction ability values

that resulted from using these two different methods were in close agreement for the individual minerals. The highest values were obtained for Mg followed by Mn (Fig. 4).

Detection of candidate genes and expression analysis

The significant SNPs that we identified were used to predict gene models located on the wheat genome using the reference Chinese Spring RefSeqv1.1. Candidate gene transcripts and their corresponding annotation information were obtained from EnsemblPlants (http://plants.ensembl.org/Triticum_aestivum). The results showed that the genes were annotated as transmembrane transporter activity, protein kinase, ATPase-coupled cation transmembrane transporter, metal ion binding, and magnesium ion binding. All the potential candidate genes and their corresponding annotations within all detected loci are listed in Table 1. It was notable

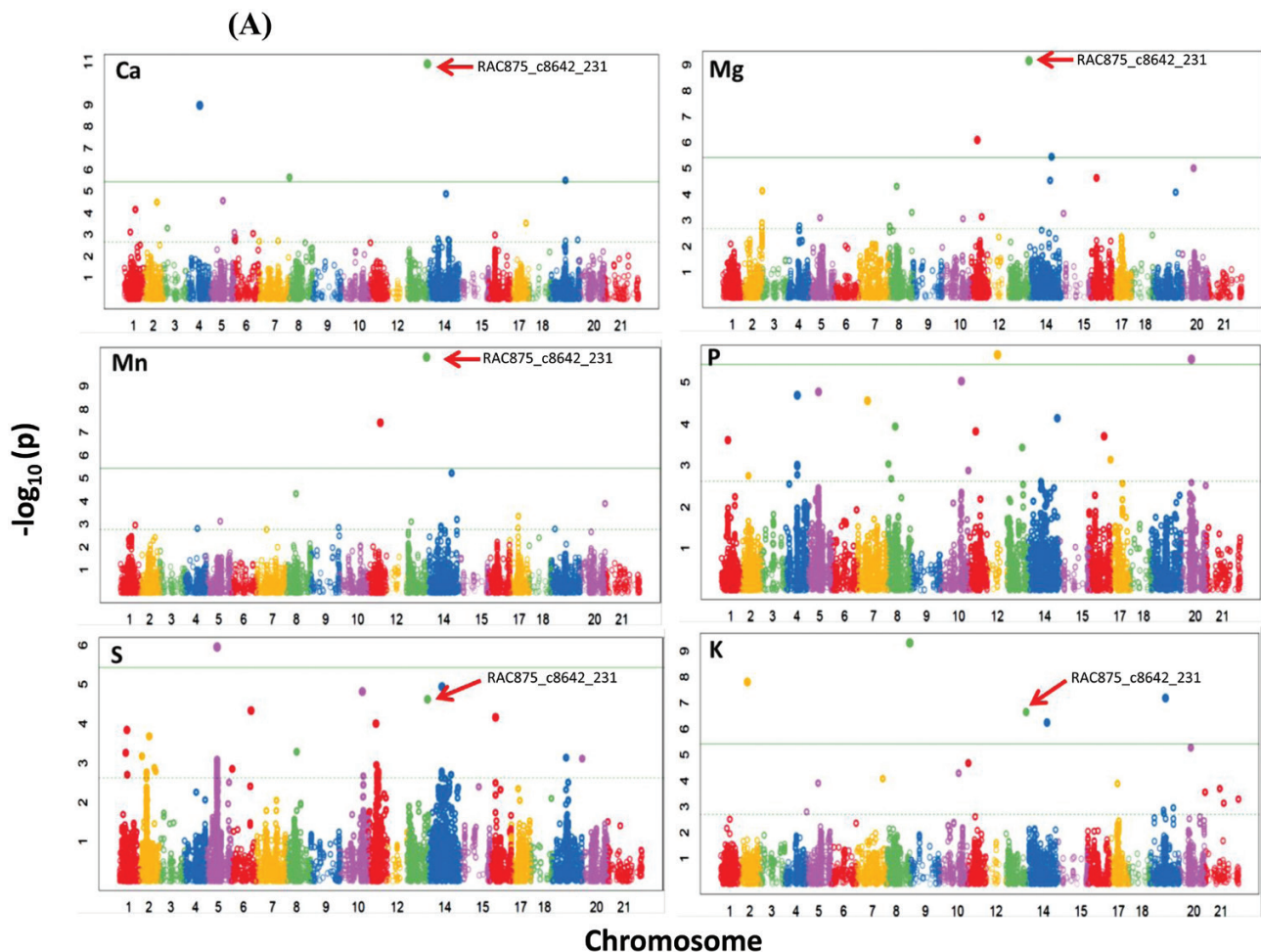


Fig. 2. Genome-wide association study (GWAS) of mineral nutrient concentrations in the grains of 369 European wheat genotypes grown over three seasons. Data were analysed using two marker arrays: a 90K iSELECT Infinium array and a 35K Affymetrix SNP array based on best linear unbiased estimates (BLUES). (A) Summary of the GWAS data for each mineral. The horizontal lines indicate the Bonferroni threshold (5.42). RAC875_c8642_231 is indicated. (B) Quantile-quantile plots showing the observed associations between SNPs and grain nutrient concentrations compared to expected associations.

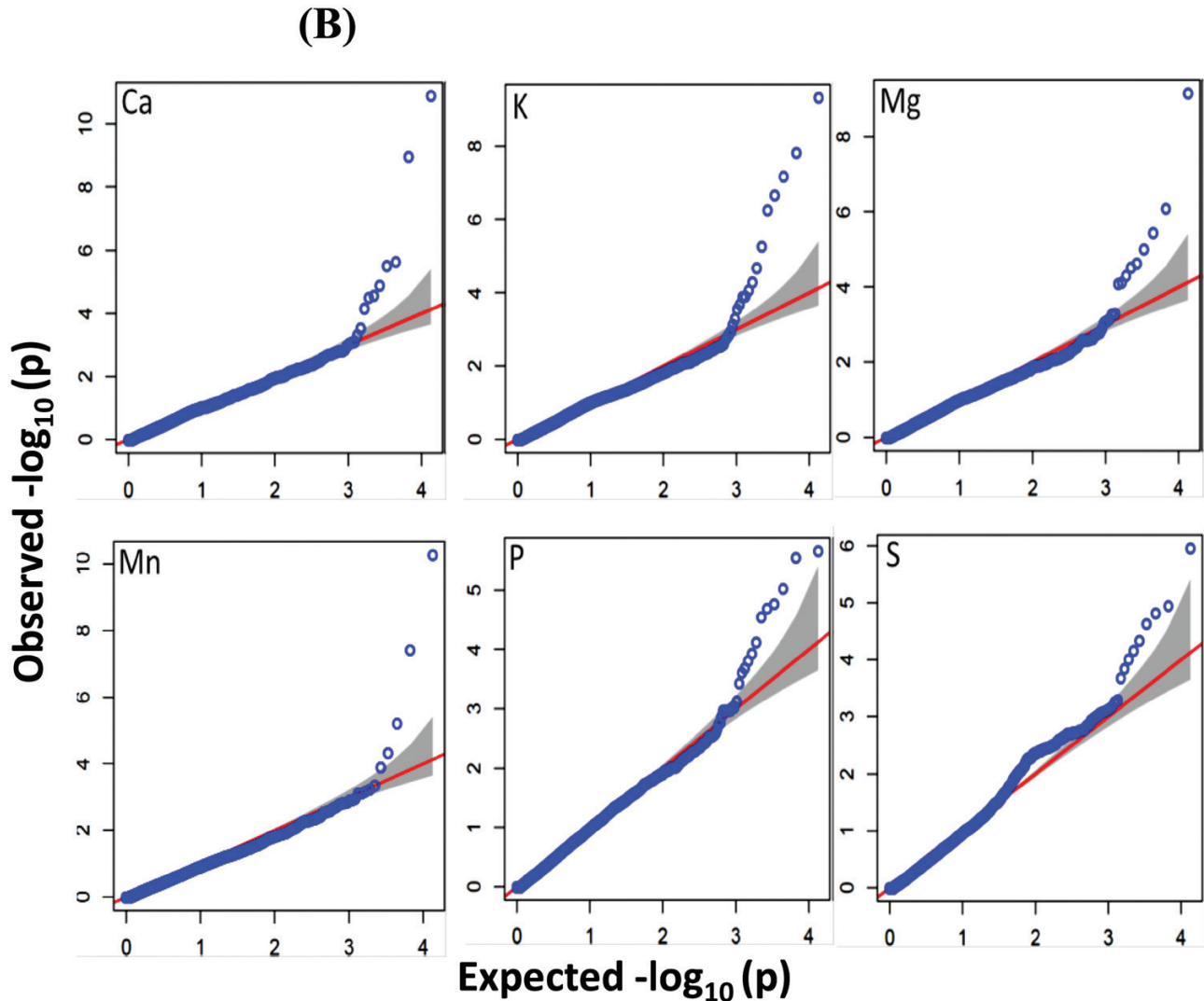


Fig. 2. Continued.

that *TraesCS5A02G542600*, which was annotated as transmembrane transporter activity and that underlay the natural variation of Ca, Mn, Mg, S, and K, was located on chromosome 5A (698 507 247–698 511 217 bp; Fig. 3A). Interestingly, RAC875_c8642_231 that had the highest association value among all the SNPs was also located on chromosome 5A (698 510 016 bp), inside *TraesCS5A02G542600* (Fig. 3A). Based on the gene structure, the RAC875_c8642_231 marker was physically located inside exon 3 of *TraesCS5A02G542600* (Fig. 3C). Polymorphism analysis of RAC875_c8642_231 alleles showed that the genotypes of the population either carried the C allele (337 genotypes) or the T allele (32 genotypes) (Supplementary Table S5). The genotypes with the T allele had significantly higher ($P < 0.05$) mineral concentration values in the grains except for K (Fig. 3D). A positive effect of the T

allele on TKW was also found. The genotypes that carried the T allele originated from France (Fig. 3F) and showed the highest minerals concentration (Fig. 3E).

We found a wide range of expression for the candidate genes in different grain tissues and at different developmental stages using the WheatGmap web tool (Fig. 5). Generally, *TraesCS2B02G202600* at 2B and *TraesCS5A02G542600* at 5A showed the highest expression in most of the organs and tissues, indicating that they play vital roles during growth, filling, and development. *TraesCS3B02G006700* was highly expressed in the lemma and seed coat, and also in the flag-leaf blade, leaf ligule, and grain. One of the significantly associated genes, *TraesCS6B02G002500*, showed very low expression in the tissues compared to the other genes and was removed from the expression analysis.

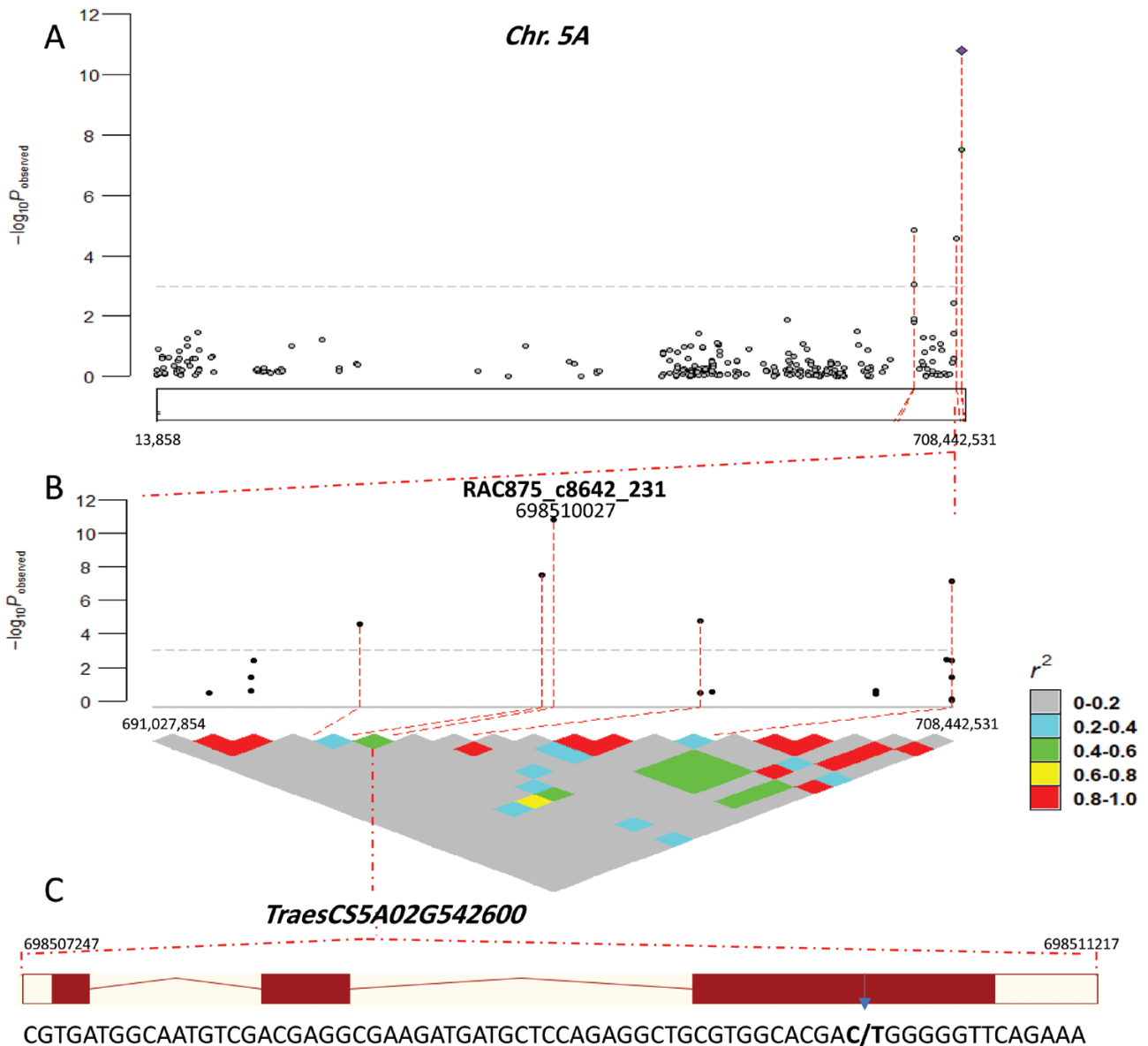


Fig. 3. Marker–trait associations (MTAs) in chromosome 5A for nutrient mineral concentrations in grains of 369 European wheat genotypes grown over three seasons. (A) Manhattan plot for associations and (B) detail of the locus of the highly associated SNP RAC875_c8642_231. The heatmap below shows linkage disequilibrium for SNPs within 15 Mbp. (C) Structure of the candidate gene *TraesCS5A02G542600* showing the C/T sequence variation in third exon. (D) Box plots showing the allele effects on nutrient concentrations and thousand-kernel weight (TKW) in the genotypes, and (E) table showing the corresponding data for some of the traits. (F) Frequency distribution of countries of origin of the genotypes carrying the T allele.

Discussion

Phenotypic variations in grain mineral across genotypes

The use of genomics breeding tools to increase the contents of essential nutrients in wheat grain is a promising approach for crop improvement, and the detection of natural phenotypic variation in existing genotypes is the primary step in this process (Yu and Tian, 2018). Overall, we observed wide natural variation in grain concentrations of Ca, Mg, Mn, P, K, and S in our panel of 369 European elite registered varieties (Fig.

1A, Supplementary Fig. S1), which was consistent with previous results (Bhatta *et al.*, 2018). We found high broad-sense heritability values for the mineral concentrations ($H^2 > 0.7$; Supplementary Table S3), indicating that the major part of the variation was genetically controlled. Similar findings Ca, Mg and S for have been reported in recombinant inbred lines of tetraploid wheat (Peleg *et al.*, 2009), while lower H^2 values have been found for Mn, P, and K in a wild barley NAM population (Herzig *et al.*, 2019). The positive correlations that we found among the minerals (Fig. 1B) suggested the presence of common genetic factors affecting the accumulation of these

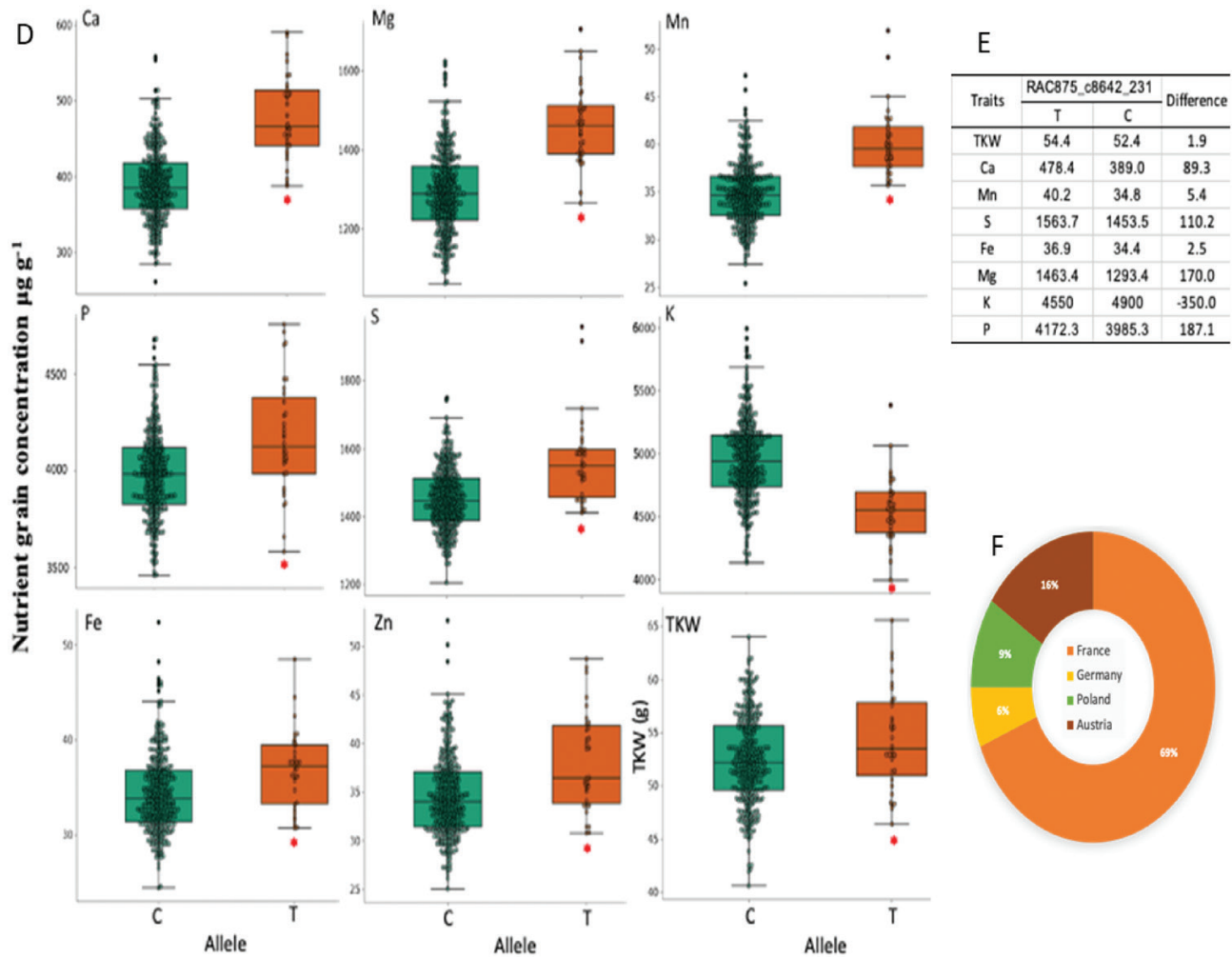


Fig. 3. Continued.

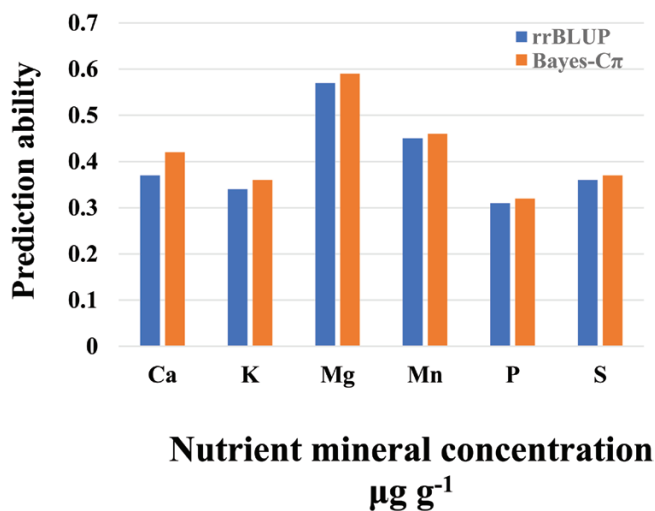


Fig. 4. Genomic prediction ability for nutrient mineral concentrations in grains of 369 European wheat genotypes grown over three seasons according to best linear unbiased estimates (BLUEs) values as determined using two different statistical models, ridge regression–best linear unbiased prediction (RR-BLUP) and the Bayes-C π method.

nutrients in the grains, and similar trends of correlations have been observed previously in hexaploid wheat (Shi *et al.*, 2013). The lack of correlations between thousand-kernel weight and the mineral contents implied that there was no clear effect of grain weight on the content.

GWAS analysis and genomic prediction

Most of the previous genetic studies in wheat have used QTL mapping to investigate the genetic basis of nutrient accumulation (Shi *et al.*, 2013; Pu *et al.*, 2014). In these studies, bi-parental crosses have been used to identify QTLs and genes associated with the mineral concentrations. However, the resulting mapping is relatively low resolution and hence GWAS is increasingly being used as an alternative that can identify alleles within a relatively broader set of germplasm with high resolution (Alomari *et al.*, 2017, 2018, 2019; Alqudah *et al.*, 2020). In addition, GWAS analysis is useful for identifying molecular markers that are tightly linked to the genomic regions that underlie natural variation in nutrients, which can then

Table 1. Candidate genes for the significant marker–trait associations for nutrient mineral concentrations in a panel of 369 European wheat genotypes

Chromosome	Mineral	Gene ID	Description	Putative functionality
5A	Ca, Mn, Mg, S, K	<i>TraesCS5A02G542600</i>	Transmembrane transporter activity	Major facilitator superfamily
3B	Ca	<i>TraesCS3B02G006700</i>	Transferase activity	Diacylglycerol O-acyltransferase
7A	Ca	<i>TraesCS7A02G169100</i>	Transmembrane transporter activity	WAT1-related protein
5B	Ca	<i>TraesCS5B02G403400</i>	Aspartate-semialdehyde dehydrogenase activity	Semialdehyde dehydrogenase
6B	Ca	<i>TraesCS6B02G428500</i>	NAD ⁺ kinase activity	NAD kinase/diacylglycerol kinase-like domain superfamily
4B	K	<i>TraesCS4B02G380200</i>	Cysteine-type endopeptidase activity	Type II CAAX prenyl endopeptidase Rce1-like
5B	K	<i>TraesCS5B02G301100</i>	Protein binding	WD40 repeat
7D	K	<i>TraesCS7B02G478200</i>	DNA binding	Homeobox superfamily
	K	<i>TraesCS3B02G590500</i>	Protein kinase activity	Protein kinase superfamily
4A	K	<i>TraesCS4A02G352200</i>	Ubiquitin protein ligase activity	E3 ubiquitin-protein ligase RNF170
2D	K	<i>TraesCS2D02G190600</i>	Protein kinase activity	Protein kinase domain-containing protein
7D	K	<i>TraesCS7D02G540700</i>	DNA binding	Homeobox superfamily
3B	Mg	<i>TraesCS3B02G125400</i>	Membrane	Uncharacterized protein
3B	Mg	<i>TraesCS3A02G514300</i>	Protein binding	F-box superfamily
4A	Mg	<i>TraesCS4A02G369500</i>	Catalytic activity	Alkaline-phosphatase
4B	Mg	<i>TraesCS4B02G293600</i>	ATPase-coupled cation transmembrane transporter activity	P-type ATPase
5B	Mg	<i>TraesCS5B02G427800</i>	Protein binding	IQ motif, EF-hand binding site
7A	Mg	<i>TraesCS7A02G498400</i>	Metal ion binding	Ubiquitin specific protease
4B	Mn	<i>TraesCS4B02G024300</i>	Protein binding	Tetratricopeptide-like helical domain superfamily
5B	Mn	<i>TraesCS5B02G012300</i>	Magnesium ion binding	Phosphopyruvate hydratase
5B	Mn	<i>TraesCS5B02G042900</i>	Protein binding	Tetratricopeptide-like helical domain superfamily
6B	Mn	<i>TraesCS6B02G181100</i>	Integral component of membrane	Uncharacterized protein
2A	P	<i>TraesCS2A02G130200</i>	Hydrolase activity	Haloacid Dehalogenase
2A	P	<i>TraesCS2A02G123400</i>	Oxidoreductase activity	FAD-binding PCMH-type domain-containing protein
2B	P	<i>TraesCS2B02G202600</i>	Integral component of membrane	GPI transamidase subunit PIG-U
3B	P	<i>TRAES_3BF053100070CFD_c1</i>	Cytoskeleton	Targeting protein for Xk1p2 domain containing protein, expressed
3B	P	<i>TraesCS3B02G013300</i>		Uncharacterized protein
3B	P	<i>TraesCS3B02G125400</i>	Integral component of membrane	Uncharacterized protein
4A	P	<i>TraesCS4A02G369500</i>	Catalytic activity	Alkaline-phosphatase
5A	P	<i>TraesCS5A02G486100</i>	Calmodulin binding	CALMODULIN-BINDING PROTEIN60
6A	P	<i>TraesCS6A02G352300</i>	Protein binding	Ankyrin repeat
6B	P	<i>TraesCS6B02G002500</i>	Protein kinase activity	Serine-threonine/tyrosine-protein kinase
1A	S	<i>TraesCS1A02G292100</i>	Protein kinase activity	Serine-threonine/tyrosine-protein kinase
1A	S	<i>TraesCS1A02G322300</i>	Serine-type endopeptidase activity	Peptidase S9
1B	S	<i>TraesCS1B02G338500</i>	Cysteine-type peptidase activity	Papain-like cysteine peptidase superfamily
1B	S	<i>TraesCS1B02G033200</i>	Cysteine-type peptidase activity	Papain-like cysteine peptidase superfamily
2B	S	<i>TraesCS2B02G166300</i>	Regulation of DNA methylation	SAC3 family protein B
2B	S	<i>TraesCS2B02G169300</i>	Lipid binding	Synaptotagmin-like mitochondrial-lipid-binding domain
2B	S	<i>TraesCS2B02G169500</i>	Nucleic acid binding	Cold-shock protein, DNA-binding
2B	S	<i>TraesCS2B02G169400</i>	Actin binding	Stomatal closure-related actin-binding protein
2B	S	<i>TraesCS2B02G169400</i>	Actin binding	Stomatal closure-related actin-binding protein
2B	S	<i>TraesCS2B02G169500</i>	Nucleic acid binding	Cold-shock protein, DNA-binding
2D	S	<i>TraesCS2D02G074700</i>	Hydrolase activity	Amidohydrolase
4B	S	<i>TraesCS4B02G278100</i>	DNA-binding transcription factor activity	Heat shock transcription factor HsfA2-8
5B	S	<i>TraesCS5B02G392500</i>	DNA binding	B3 DNA binding domain
6A	S	<i>TraesCS6A02G037800</i>	Metal ion binding	Ribonuclease Nob1
7A	S	<i>TraesCS7A02G419800</i>	ADP binding	P-loop containing nucleoside triphosphate hydrolase

The significant SNPs identified in the current study were used to predict gene models located on the wheat genome using the reference Chinese Spring RefSeqv1.1. Candidate gene transcripts and their corresponding annotation information were obtained from the website of EnsemblPlants (http://plants.ensembl.org/Triticum_aestivum).

be used in genomics-assisted breeding for enhancing the efficiency of biofortification (Collard *et al.*, 2005).

Most of the association mapping studies that have been undertaken for different complex traits have used single-locus

GWAS models (GLM, MLM) and these require additional corrections, for example multiple testing corrections to control false positives (Yu *et al.*, 2006; Price *et al.*, 2010; Pujar *et al.*, 2020). However, the corrections are often either not stringent

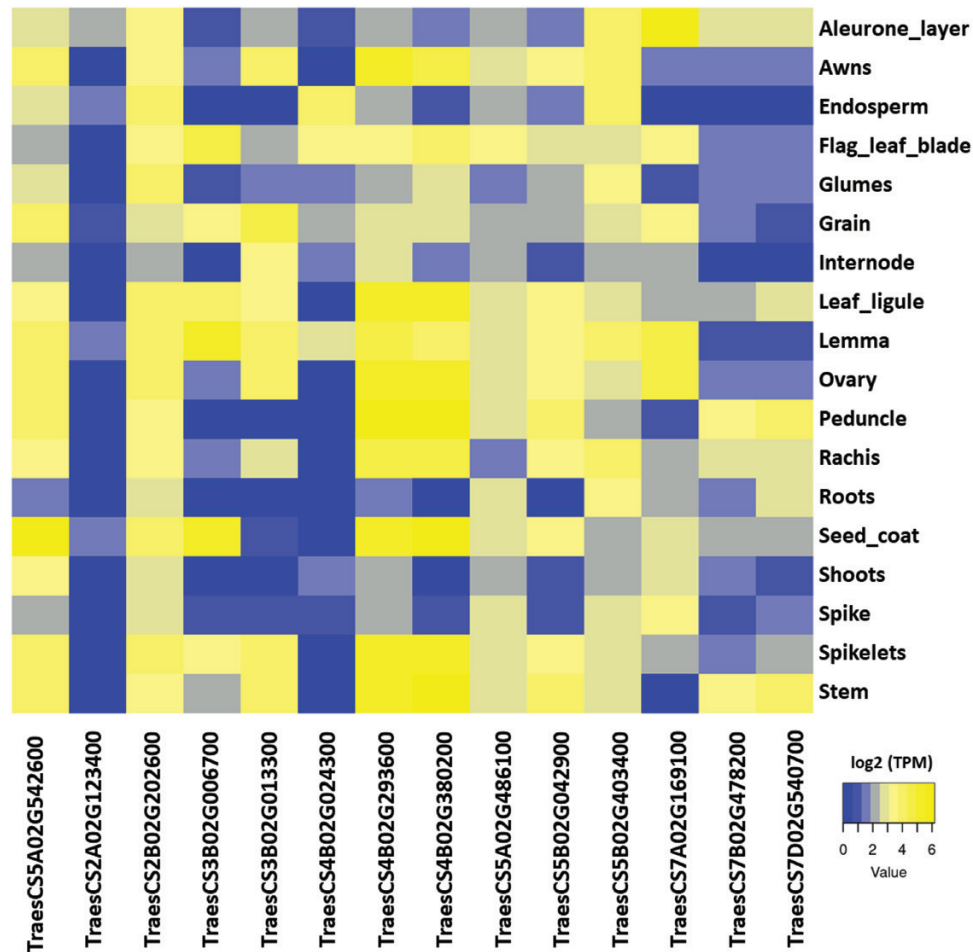


Fig. 5. Expression patterns of selected candidate genes for mineral nutrient concentrations within different tissues of wheat. Expression data were obtained from the Wheat Gmap database (<https://www.wheatgmap.org/expression/search/gene/>) and are presented as a heatmap of transcripts per kilobase million (TPM) values.

enough, or too stringent and overcompensate for population structure and kinship, which can lead to over-correction that results in some potentially important marker–trait associations (MTAs) being missed (Alqudah *et al.*, 2020). To help overcome this issue, a GAPIT model known as ‘fixed and random model circulating probability unification’ (FarmCPU) has been developed and has been shown to be much more powerful, efficient, and superior to previous approaches in reducing false positive/negative associations, particularly in complex traits, by incorporating multiple markers simultaneously as covariates in a stepwise MLM to partially remove confounding effects between testing markers and kinship (Liu *et al.*, 2016). For this reason, we applied the FarmCPU model in our GWAS analysis.

We observed many genomic regions harboring markers that were associated with nutrient concentrations. In total, we detected 19 MTAs for S, 15 for K, 15 for P, 14 for Mg, 13 for Ca, and 10 for Mn (Supplementary Table S4A). The most significant association was detected at the end of the long arm of chromosome 5A and was linked with the RAC875_c8642_231 marker at the position ~114.5 cM (698 5100 16 bp; Fig. 3). This

was linked with all the minerals that we focused on except for P. P is positively correlated with phytate content in plants and it is known as an ‘anti-nutrient’ compound that negatively influences the absorption of other minerals in the human body (Stangoulis *et al.*, 2007). The RAC875_c8642_231 marker therefore appears to be a good candidate in breeding wheat to improve the concentrations of many minerals within the grain simultaneously. We also observed that the K concentration in the grain was negatively correlated with the concentrations of Ca, Mn, Mg, and S (Fig. 1B). MTAs for grain K concentration were also negatively associated with those for Ca, Mn, Mg, and S at the 5A locus (Supplementary Table S4A).

Co-localization of QTLs on chromosome 5A of wheat have previously been reported for N, Fe, Cu, Mg, and K (Shi *et al.*, 2013), and Fe, Zn, Cu, and Mg but not P (Cu *et al.*, 2020), and the latter study detected a co-located locus for Zn, Fe, Cu, P, and Mn on chromosome 5B at ~95 cM. It therefore seems that the QTL on chromosome 5A plays a vital role in mineral accumulation in wheat, and this needs further functional validation. Our results showed that most of the genotypes with the

highest nutrient contents originated from France (Fig. 1E), and these were characterized by having awns. The top five of these genotypes were Isengrain, Inoui, Nirvana, Exotic, and Lona (Supplementary Table S1), and including varieties such as this in breeding programs might help in enhancing the accumulation of minerals in wheat grain. We extended our analysis to include the genomic prediction (GP) method, which has a practical role in improving the breeding efficiency of quantitative and complex traits. The predictability results for our wheat panel showed low-to-moderate values (Fig. 4), which was in agreement with a previous study of macro- and micro-nutrients in wheat (Manickavelu *et al.*, 2017). GP might be considered as a promising approach for enhancing nutrient minerals in wheat in situations where relatively large germplasm panels are used with high numbers of markers, thus giving more accurate estimates of breeding values.

Candidate genes and their expression patterns

The recently released IWGSC RefSeq v1.1 genome by EnsemblPlants enabled us to further investigate the candidate genes potentially responsible for the variations in the grain nutrient concentrations (listed in Table 1). In the following sections we only consider high-confidence candidate genes that are linked to transcription factor regulators, transporters, and grain development.

Ca.

The most significant MTA detected in the GWAS output was found to be associated with Ca, K, Mg, Mn, and S, and was located on chromosome 5A (114.5 cM; Fig. 3). It was linked to the gene *TraesCS5A02G542600*, which encodes a transmembrane transporter belonging to the Major Facilitator Superfamily (MSF). This gene family is known to be one of the two largest superfamilies of membrane transporters, and they act as uniporters or symporters for different substances (Niño-González *et al.*, 2019). Interestingly, three recent studies in wheat have explored the role of *TraesCS5A02G542600* in the inhibition of awn formation; however, whilst it is located at the genetic locus of awn suppression, a closely linked gene has been considered as the most likely candidate (DeWitt *et al.*, 2020; Huang *et al.*, 2020; Würschum *et al.*, 2020). Further studies are needed to shed light on the role of *TraesCS5A02G542600* in nutrient mineral uptake and accumulation in wheat grains. We found that the expression of *TraesCS5A02G542600* was very high in grain tissues such as the seed coat, lemma, aleurone, and endosperm, and in tissues closely associated with the grain such as the peduncle, spikes, and spikelets (Fig. 5), which would be consistent with its involvement in mineral accumulation in the grains.

Another candidate gene for Ca accumulation was *TraesCS7A02G169100*, (125 712 948–125 715 936 bp), which has an annotation as a transmembrane transporter and encodes a WALLS ARE THIN 1 (WAT1)-related protein. This gene

has been found to be involved in secondary cell wall formation (Kaur *et al.*, 2017), and it has been demonstrated that Ca^{2+} plays a major role in determining the structural rigidity of the cell wall, with high concentrations making the wall more rigid and less plastic whilst low concentrations make the wall more pliable and easily ruptured (Hepler, 2005). *TraesCS7A02G169100* was highly expressed in the aleurone layer, the lemma, and within the grain (Fig. 5).

TraesCS3B02G006700 (3 601 450–3 607 208 bp) was annotated as transferase activity and it encodes a diacylglycerol O-acyltransferase that catalyses the final step of the triacylglycerol (TAG) synthesis. TAG shows a significant increase in accumulation in Arabidopsis seedlings during nitrogen deprivation (Yang *et al.*, 2011), and it is therefore possible that it might have a role in the accumulation of other nutrient minerals such as Ca in wheat grains. Expression of *TraesCS3B02G006700* was found to be high for the lemma, seed coat, flag-leaf blade, leaf ligule, and grain (Fig. 5), which potentially supports this hypothesis.

TraesCS5B02G403400 (580 100 440–580 104 246bp) encodes semialdehyde dehydrogenase, which is one of three enzymes constituting the gamma-aminobutyric acid shunt, a metabolic pathway that has been associated with abiotic stress responses in durum wheat (Carillo, 2018). The first enzyme in the pathway is glutamate decarboxylase (GAD), which is a calcium/calmodulin-binding protein (Baum *et al.*, 1996; Busch and Fromm, 1999).

K.

Two Homeobox superfamily genes, *TraesCS7B02G478200* (733 527 123–733 530 917 bp) and *TraesCS7D02G540700* (630 527 242–630 530 005 bp) were linked with K, and this superfamily is one of the transcription factor families that are involved in plant development, growth, and in the response to diverse stresses (Wei *et al.*, 2019). These genes showed almost the same expression pattern with the highest value in the peduncle followed by the stem, rachis, aleurone layer, seed coat, leaf ligule, and spikelets (Fig. 5).

TraesCS3B02G590500 (815 476 561–815 480 015 bp) and *TraesCS2D02G190600* (134 638 904–134 645 251 bp) were both annotated as protein kinase, which is a large superfamily that plays vital roles in plant development and stress tolerance; however, the functions of only a limited number of protein kinases have been studied in wheat (Wei and Li, 2019). Several studies have shown that the E3 ubiquitin-protein ligase gene *TraesCS4A02G352200* (627 813 928–627 816 836 bp) is associated with the uptake and accumulation of various minerals such as Mn, Zn, and P in wheat and pearl millet (Cu *et al.*, 2020; Pujar *et al.*, 2020), and thus it might have a role in K uptake as well.

Mg.

We detected *TraesCS4B02G293600* (579 391 172–579 398 644 bp) encoding a P-type ATPase on chromosome 4B, and

interestingly this is considered as a transmembrane protein that plays a crucial role in the transport of a wide variety of cations across membranes and is vital for ion homeostasis and detoxification of heavy metals (Axelsen and Palmgren, 2001). Very high expression levels were detected for in the peduncle, seed coat, spikelets, leaf ligule, and stem (Fig. 5).

TraesCS7A02G498400 (688 960 269–688 968 295 bp) encoding ubiquitin specific protease was detected on chromosome 7A and annotated as metal ion binding.

Mn.

Genes encoding tetratricopeptide-like helical domain superfamily proteins, namely *TraesCS4B02G024300* (173 621 92–173 688 28 bp) and *TraesCS5B02G042900* (475 842 53–475 880 87 bp), were found on chromosomes 4B and 5B, respectively. These are involved in root development and auxin polar transport as well as gibberellin signal transduction in Arabidopsis (Jacobsen *et al.*, 1996; Zhang *et al.*, 2015), which might indicate their involvement in Mn accumulation in wheat plants. These two genes had their highest expression in the endosperm, flag-leaf blade, lemma, grain, and glumes.

Another gene was identified on chromosome 5B as encoding phosphopyruvate hydratase and annotated as magnesium ion binding, namely *TraesCS5B02G012300* (12 324 287–12 328 841 bp). (Silva *et al.*, 2018) have reported an association between phosphopyruvate hydratase and enhanced nitrogen metabolism in maize seedlings.

P.

TraesCS2A02G130200 (77 938 466–77 941 846 bp) encoding a haloacid dehalogenase (HAD) was found on chromosome 2A. HAD enhances phosphate accumulation, and *LePS2* in tomato was the first low-Pi inducible gene in this superfamily to be characterized (Baldwin *et al.*, 2001, 2008). Another HAD gene, *PvHAD1*, shows specific induction at low P and encodes a functional serine/threonine phosphatase (Liu *et al.*, 2012). Two other HAD genes responsive to low P, *AtPPsPase1* and *AtPECP1*, have been reported to encode functional pyrophosphatase and phosphoethanolamine phosphatase in Arabidopsis (Pandey *et al.*, 2017). In rice, only two HAD genes have been shown to be up-regulated under P deficiency, namely *OsACP1* (Hur *et al.*, 2007) and *OsHAD1* (Pandey *et al.*, 2017). These studies therefore suggest HAD superfamily members have important functions in P accumulation in various plants; however, there are no studies available on wheat.

TraesCS2A02G123400 (729 311 06–729 331 77 bp) encoding a FAD-binding PCMH-type domain-containing protein was detected on chromosomes 2A, and *TraesCS2B02G202600* (182 355 759–182 361 233 bp) encoding GPI transamidase subunit PIG-U was detected on chromosome 2B, which may be involved in P accumulation in wheat.

Interestingly, *TraesCS3B02G01330* (5 952 761–5 955 519 bp) encoding a targeting protein for Xklp2 (TPX2) was detected on chromosome 3B. This gene is important for phosphorylation/

dephosphorylation of regulatory proteins by phosphatase. The *TPX2* gene is not well studied, and it is only recently that two genes, *TPXL2* and *TPXL3*, have been functionally characterized in Arabidopsis (Dvořák Tomašíková *et al.*, 2020). Another gene involved in phosphorylation, *TraesCS6B02G002500* (1 963 516–1 970 774 bp), was detected on chromosome 6B and encodes as serine-threonine/tyrosine-protein kinase (Dvořák Tomašíková *et al.*, 2020). Another important gene, *TraesCS4A02G369500* (641 504 088–641 508 259 bp), was detected on chromosome 4A and belongs to the alkaline-phosphatase family, which is and associated with phosphate efflux from arbuscules (Aono *et al.*, 2004). In addition, *TraesCS5A02G486100* (657 191 801–657 195 889 bp) encoding CALMODULIN-BINDING PROTEIN60 has a significant role in protein phosphorylation/dephosphorylation (Bergey *et al.*, 2014).

S.

Two SNP markers were detected on chromosome 1B as papain-like cysteine peptidase (PLCP) superfamily genes, namely *TraesCS1B02G338500* (566 601 144–566 605 225 bp) and *TraesCS1B02G033200* (162 164 83–162 176 35 bp). PLCPs have important functions in plant growth, seed germination, anther development, and senescence. Thus, they might play a role in S accumulation in wheat grains (Liu *et al.*, 2018).

Conclusions

There have been very few GWAS analyses identifying significant loci associated with nutrient mineral accumulation in wheat grains. Hence, the marker-trait associations and candidate genes identified in this study will be useful for the future genetic improvement of wheat nutritional quality through marker-assisted selection. This study also provides useful information on the range of phenotypic variation encountered within European wheat germplasm. However, further research is needed for a better understanding of the relationships between the individual mineral nutrients in wheat and to provide more detailed information that can lead to efficient breeding to overcome malnutrition problems.

Supplementary data

The following supplementary data are available at [JXB online](#).

Fig. S1. Frequency distributions of nutrient mineral concentrations in wheat grains based on BLUEs.

Table S1. Details of the 369 European wheat varieties used in the study.

Table S2. Phenotypic variation for grain nutrient concentrations across the three study years and BLUE values.

Table S3. ANOVA results and heritability values for the nutrient concentrations.

Table S4. Significant marker-trait associations of the SNP markers based on BLUE data, and those above the Bonferroni correction threshold.

Table S5. Allelic variation analysis for the RAC875_c8642_231 marker across all the genotypes.

Table S6. Expression analysis for the putative candidate genes underlying the accumulation of nutrient minerals in the grains.

Acknowledgements

This project was funded by internal financial support of Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany. We thank Ellen Weiß and Yudelys Antonia Tandron Moya of Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) for their excellent technical assistance.

Author contributions

DZA analysed the data, interpreted the results, and wrote the manuscript; AMA contributed to the gene structure analysis; NvW participated in the nutrient mineral measurements; KP and MSR designed the experiment; MSR conceived the research; DZA, AMA, and MSR helped in the manuscript editing; all authors read and approved the final manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Data availability

All data supporting the findings of this study are available within the paper and within its supplementary materials published online.

References

- Allen AM, Winfield MO, Burridge AJ, *et al.* 2017. Characterization of a Wheat Breeders' Array suitable for high-throughput SNP genotyping of global accessions of hexaploid bread wheat (*Triticum aestivum*). *Plant Biotechnology Journal* **15**, 390–401.
- Alomari DZ, Eggert K, von Wirén N, Alqudah AM, Polley A, Plieske J, Ganal MW, Pillen K, Röder MS. 2018. Identifying candidate genes for enhancing grain Zn concentration in wheat. *Frontiers in Plant Science* **9**, 1313.
- Alomari DZ, Eggert K, von Wirén N, Pillen K, Röder MS. 2017. Genome-wide association study of calcium accumulation in grains of European wheat cultivars. *Frontiers in Plant Science* **8**, 1797.
- Alomari DZ, Eggert K, von Wirén N, Polley A, Plieske J, Ganal MW, Liu F, Pillen K, Röder MS. 2019. Whole-genome association mapping and genomic prediction for iron concentration in wheat grains. *International Journal of Molecular Sciences* **20**, 76.
- Alqudah AM, Sallam A, Baenziger PS, Börner A. 2020. GWAS: fast-forwarding gene identification and characterization in temperate cereals: lessons from barley – a review. *Journal of Advanced Research* **22**, 119–135.
- Aono T, Maldonado-Mendoza IE, Dewbre GR, Harrison MJ, Saito M. 2004. Expression of alkaline phosphatase genes in arbuscular mycorrhizas. *New Phytologist* **162**, 525–534.
- Appels R, Eversole K, Feuillet C, *et al.* 2018. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* **361**, eaar7191.
- Axelsen KB, Palmgren MG. 2001. Inventory of the superfamily of P-type ion pumps in Arabidopsis. *Plant Physiology* **126**, 696–706.
- Baldwin JC, Karthikeyan AS, Cao A, Raghothama KG. 2008. Biochemical and molecular analysis of *LePS2;1*: a phosphate starvation induced protein phosphatase gene from tomato. *Planta* **228**, 273–280.
- Baldwin JC, Karthikeyan AS, Raghothama KG. 2001. LEPS2, a phosphorus starvation-induced novel acid phosphatase from tomato. *Plant Physiology* **125**, 728–737.
- Baum G, Lev-Yadun S, Fridmann Y, Arazi T, Katsnelson H, Zik M, Fromm H. 1996. Calmodulin binding to glutamate decarboxylase is required for regulation of glutamate and GABA metabolism and normal development in plants. *The EMBO Journal* **15**, 2988–2996.
- Bergey DR, Kandel R, Tyree BK, Dutt M, Dhekney SA. 2014. The role of calmodulin and related proteins in plant cell function: an ever-thickening plot. *Springer Science Reviews* **2**, 145–159.
- Bhatta M, Baenziger PS, Waters BM, Poudel R, Belamkar V, Poland J, Morgounov A. 2018. Genome-wide association study reveals novel genomic regions associated with 10 grain minerals in synthetic hexaploid wheat. *International Journal of Molecular Sciences* **19**, 3237.
- Busch KB, Fromm H. 1999. Plant succinic semialdehyde dehydrogenase. Cloning, purification, localization in mitochondria, and regulation by adenine nucleotides. *Plant Physiology* **121**, 589–597.
- Carillo P. 2018. GABA shunt in durum wheat. *Frontiers in Plant Science* **9**, 100.
- Collard BCY, Jahufer MZZ, Brouwer JB, Pang ECK. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: the basic concepts. *Euphytica*, **142**, 169–196.
- Cu ST, Guild G, Nicolson A, Velu G, Singh R, Stangoulis J. 2020. Genetic dissection of zinc, iron, copper, manganese and phosphorus in wheat (*Triticum aestivum* L.) grain and rachis at two developmental stages. *Plant Science* **291**, 110338.
- Desta ZA, Ortiz R. 2014. Genomic selection: genome-wide prediction in plant improvement. *Trends in Plant Science* **19**, 592–601.
- DeWitt N, Guedira M, Lauer E, *et al.* 2020. Sequence-based mapping identifies a candidate transcription repressor underlying awn suppression at the *B1* locus in wheat. *New Phytologist* **225**, 326–339.
- Dvořák Tomaštková E, Rutten T, Dvořák P, Tugai A, Ptošková K, Petrovská B, van Damme D, Houben A, Doležel J, Demidov D. 2020. Functional divergence of microtubule-associated TPX2 family members in *Arabidopsis thaliana*. *International Journal of Molecular Sciences* **21**, 2183.
- Godfray HC, Beddington JR, Crute IR, Haddad L, Lawrence D, Muir JF, Pretty J, Robinson S, Thomas SM, Toulmin C. 2010. Food security: the challenge of feeding 9 billion people. *Science* **327**, 812–818.
- Habier D, Fernando RL, Kizilkaya K, Garrick DJ. 2011. Extension of the Bayesian alphabet for genomic selection. *BMC Bioinformatics* **12**, 186.
- Hepler PK. 2005. Calcium: a central regulator of plant growth and development. *The Plant Cell* **17**, 2142–2155.
- Herzig P, Backhaus A, Seiffert U, von Wirén N, Pillen K, Maurer A. 2019. Genetic dissection of grain elements predicted by hyperspectral imaging associated with yield-related traits in a wild barley NAM population. *Plant Science* **285**, 151–164.
- Huang D, Zheng Q, Melchardt T, *et al.* 2020. Dominant inhibition of awn development by a putative zinc-finger transcriptional repressor expressed at the *B1* locus in wheat. *New Phytologist* **225**, 340–355.
- Hur YJ, Lee HG, Jeon EJ, Lee YY, Nam MH, Yi G, Eun MY, Nam J, Lee JH, Kim DH. 2007. A phosphate starvation-induced acid phosphatase from *Oryza sativa*: phosphate regulation and transgenic expression. *Biotechnology Letters* **29**, 829–835.
- Jacobsen SE, Binkowski KA, Olszewski NE. 1996. SPINDLY, a tetratricopeptide repeat protein involved in gibberellin signal transduction in Arabidopsis. *Proceedings of the National Academy of Sciences, USA* **93**, 9292–9296.
- Kaur S, Zhang X, Mohan A, Dong H, Vikram P, Singh S, Zhang Z, Gill KS, Dhugga KS, Singh J. 2017. Genome-wide association study

reveals novel genes associated with culm cellulose content in bread wheat (*Triticum aestivum*, L.). *Frontiers in Plant Science* **8**, 1913.

Lipka AE, Tian F, Wang Q, Peiffer J, Li M, Bradbury PJ, Gore MA, Buckler ES, Zhang Z. 2012. GAPIT: genome association and prediction integrated tool. *Bioinformatics* **28**, 2397–2399.

Liu H, Hu M, Wang Q, Cheng L, Zhang Z. 2018. Role of papain-like cysteine proteases in plant development. *Frontiers in Plant Science* **9**, 1717.

Liu J, Wang X, Huang H, Wang J, Li Z, Wu L, Zhang G, Ma Z. 2012. Efficiency of phosphorus utilization in *phyA*-expressing cotton lines. *Indian Journal of Biochemistry & Biophysics* **49**, 250–256.

Liu X, Huang M, Fan B, Buckler ES, Zhang Z. 2016. Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS Genetics* **12**, e1005767.

Manickavelu A, Hattori T, Yamaoka S, Yoshimura K, Kondou Y, Onogi A, Matsui M, Iwata H, Ban T. 2017. Genetic nature of elemental contents in wheat grains and its genomic prediction: toward the effective use of wheat landraces from Afghanistan. *PLoS ONE* **12**, e0169416.

Meuwissen TH, Hayes BJ, Goddard ME. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* **157**, 1819–1829.

Niño-González M, Novo-Uzal E, Richardson DN, Barros PM, Duque P. 2019. More transporters, more substrates: the Arabidopsis major facilitator superfamily revisited. *Molecular Plant* **12**, 1182–1202.

Pandey BK, Mehra P, Verma L, Bhadouria J, Giri J. 2017. OsHAD1, a haloacid dehalogenase-like APase, enhances phosphate accumulation. *Plant Physiology* **174**, 2316–2332.

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. *Theoretical and Applied Genetics* **119**, 353–369.

Poland JA, Brown PJ, Sorrells ME, Jannink JL. 2012. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE* **7**, e32253.

Price AL, Zaitlen NA, Reich D, Patterson N. 2010. New approaches to population stratification in genome-wide association studies. *Nature Reviews Genetics* **11**, 459–463.

Pu Z, Yu M, He Q, et al. 2014. Quantitative trait loci associated with micronutrient concentrations in two recombinant inbred wheat lines. *Journal of Integrative Agriculture* **13**, 2322–2329.

Pujar M, Gangaprasad S, Govindaraj M, Gangurde SS, Kanatti A, Kudapa H. 2020. Genome-wide association study uncovers genomic regions associated with grain iron, zinc and protein content in pearl millet. *Scientific Reports* **10**, 19473.

Rasheed A, Xia X. 2019. From markers to genome-based breeding in wheat. *Theoretical and Applied Genetics* **132**, 767–784.

Ray DK, Mueller ND, West PC, Foley JA. 2013. Yield trends are insufficient to double global crop production by 2050. *PLoS ONE* **8**, e66428.

Shi R, Tong Y, Jing R, Zhang F, Zou C. 2013. Characterization of quantitative trait loci for grain minerals in hexaploid wheat (*Triticum aestivum* L.). *Journal of Integrative Agriculture* **12**, 1512–1521.

Silva IT, Abbaraju HKR, Fallis LP, Liu H, Lee M, Dhugga KS. 2018. Biochemical and genetic analyses of N metabolism in maize testcross seedlings: 2. Roots. *Theoretical and Applied Genetics* **131**, 1191–1205.

Sorrells ME, Gustafson JP, Somers D, et al. 2011. Reconstruction of the synthetic W7984 × Opata M85 wheat reference population. *Genome* **54**, 875–882.

Sperotto RA, Ricachenevsky FK, Williams LE, Vasconcelos MW, Menguer PK. 2014. From soil to seed: micronutrient movement into and within the plant. *Frontiers in Plant Science* **5**, 438.

Stangoulis JCR, Huynh B-L, Welch RM, Choi E-Y, Graham RD. 2007. Quantitative trait loci for phytate in rice grain and their relationship with grain micronutrient content. *Euphytica* **154**, 289–294.

Tapiero H, Townsend DM, Tew KD. 2003. Trace elements in human physiology and pathology. Copper. *Biomedicine & Pharmacotherapy* **57**, 386–398.

VanRaden PM. 2008. Efficient methods to compute genomic predictions. *Journal of Dairy Science* **91**, 4414–4423.

Wang S, Wong D, Forrest K, et al. 2014. Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnology Journal* **12**, 787–796.

Wei K, Li Y. 2019. Functional genomics of the protein kinase superfamily from wheat. *Molecular Breeding* **39**, 141.

Wei M, Liu A, Zhang Y, Zhou Y, Li D, Dossa K, Zhou R, Zhang X, You J. 2019. Genome-wide characterization and expression analysis of the HD-Zip gene family in response to drought and salinity stresses in sesame. *BMC Genomics* **20**, 748.

WHO. 2013. The world health report. Geneva, Switzerland: World Health Organization. <https://www.who.int/publications/i/item/9789240690837>.

Würschum T, Jähne F, Phillips AL, Langer SM, Longin CFH, Tucker MR, Leiser WL. 2020. Misexpression of a transcriptional repressor candidate provides a molecular mechanism for the suppression of awns by *Tipped 1* in wheat. *Journal of Experimental Botany* **71**, 3428–3436.

Yang Y, Yu X, Song L, An C. 2011. ABI4 Activates *DGAT1* expression in Arabidopsis seedlings during nitrogen deficiency. *Plant Physiology* **156**, 873–883.

Yu J, Pressoir G, Briggs WH, et al. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics* **38**, 203–208.

Yu S, Tian L. 2018. Breeding major cereal grains through the lens of nutrition sensitivity. *Molecular Plant* **11**, 23–30.

Zhang L, Dong C, Chen Z, et al. 2020. WheatGmap: a comprehensive platform for wheat gene mapping and genomic studies. *Molecular Plant* **14**, 187–190.

Zhang M, Wang C, Lin Q, et al. 2015. A tetratricopeptide repeat domain-containing protein SSR1 located in mitochondria is involved in root development and auxin polar transport in Arabidopsis. *The Plant Journal* **83**, 582–599.