Molecular Genetics and Genomics
https://doi.org/10.1007/s00438-023-02074-6

ORIGINAL ARTICLE

Single- and multi-trait genomic prediction and genome-wide association analysis of grain yield and micronutrient-related traits in ICARDA wheat under drought environment

Wuletaw Tadesse[1] · Zakaria El Gataa[1] · Fatima Ezzahra Rachdad[1] · Adil El Baouchi[2] · Zakaria Kehel[1] · Admas Alemu[3]

Received: 15 July 2023 / Accepted: 28 September 2023
© The Author(s) 2023

Abstract

Globally, over 2 billion people suffer from malnutrition due to inadequate intake of micronutrients. Genomic-assisted breeding is identified as a valuable method to facilitate developing new improved plant varieties targeting grain yield and micronutrient-related traits. In this study, a genome-wide association study (GWAS) and single- and multi-trait-based genomic prediction (GP) analysis was conducted using a set of 252 elite wheat genotypes from the International Center for Agricultural Research in Dry Areas (ICARDA). The objective was to identify linked SNP markers, putative candidate genes and to evaluate the genomic estimated breeding values (GEBVs) of grain yield and micronutrient-related traits.. For this purpose, a field trial was conducted at a drought-prone station, Merchouch, Morocco for 2 consecutive years (2018 and 2019) followed by GWAS and genomic prediction analysis with 10,173 quality SNP markers. The studied genotypes exhibited a significant genotypic variation in grain yield and micronutrient-related traits. The GWAS analysis identified highly significantly associated markers and linked putative genes on chromosomes 1B and 2B for zinc (Zn) and iron (Fe) contents, respectively. The genomic predictive ability of selenium (Se) and Fe traits with the multi-trait-based GP GBLUP model was 0.161 and 0.259 improving by 6.62 and 4.44%, respectively**,** compared to the corresponding single-trait-based models. The identified significantly linked SNP markers, associated putative genes, and developed GP models could potentially facilitate breeding programs targeting to improve the overall genetic gain of wheat breeding for grain yield and biofortification of micronutrients via marker-assisted (MAS) and genomic selection (GS) methods.

Keywords Bread wheat · Micronutrients · GWAS · Multi-trait genomic prediction

Introduction

Wheat is the world's leading food crop and the cradle of human civilization with an average annual production of 770 million tons in about 220 million hectares (FAOSTAT 2021). It has been playing a fundamental role in improving global food security by providing an average of 19% of calories and 21% of the protein needed for daily human diets (Shewry 2009). In addition to carbohydrates and proteins, wheat grains also contain several essential micronutrients, such as iron (Fe), zinc (Zn), and selenium (Se), which play a significant role in human growth and development, cognitive and immune function, and gene regulation. However, reports have shown that more than 2 billion people worldwide, mainly women and children in developing countries, suffer from Fe and Zn malnutrition (Lim et al. 2013). This leads to anemia, stunted physical growth, neuron motor retardation, fatigue, and reduced productivity (Umamaheswari et al. 2011). At the global level, 43% of children, 38% of pregnant and 29% of non-pregnant women of reproductive ages are anemic (World Health Organization 2015). Although diversification of diets, provision of mineral supplements, and food fortification are potential solutions, none of these

Communicated by Bing Yang.

✉ Admas Alemu
admas.alemu.abebe@slu.se

[1] The International Center for Agricultural Research in the Dry Areas (ICARDA), Rabat, Morocco

[2] AgroBioSciences, Mohammed VI Polytechnic University (UM6P), Ben Guerir, Morocco

[3] Department of Plant Breeding, Swedish University of Agricultural Sciences, Alnarp, Sweden

Published online: 18 October 2023

approaches is sustainable for solving micronutrient malnutrition (Gletsu-Miller and Wright 2013). Identification, development, and deployment of wheat varieties with high micronutrients content is the most economical, feasible, and sustainable approach to prevent malnutrition, especially in the developing world (Shewry 2009; Alomari et al. 2018; Rathan et al. 2022). Genomic-assisted breeding techniques including QTL mapping via GWAS and recently genomic prediction are widely applied methods to assist the development of new crop varieties with an improved micronutrient content.

Climate change reduces rainfall, exacerbates water stress, and has a significant negative impact on wheat production. Drought induced by climate change affects wheat grain production and quality at any growth stage and reduces wheat grain production up to 30% (Lesk et al. 2016). In addition to increasing yield and improve stress tolerance/ resistance, enhancing wheat quality under drought conditions has been identified as the key objective of wheat breeding programs (Suliman et al. 2021). Research findings have shown that wheat wild relatives are identified as an important source for micronutrients and tolerance/resistance for many biotic and abiotic stresses. Many efforts have been carried out to transfer these genes into adapted cultivars through the utilization of synthetic hexaploid wheat through the hybridization of tetraploid wheat ($2n = 4x = 28$, AABB) with the diploid wild relative, *Aegilops tauschii* ($2n = 2x = 14$, DD) followed by chromosome doubling of the corresponding diploids (Li et al. 2018). Synthetic hexaploids are great sources for improving wheat qualities including macronutrients and micronutrients and other economically relevant traits (Tadesse et al. 2015). According to recent reports, more than 30% of the elite spring bread wheat genotypes in the CIMMYT and ICARDA breeding programs contain synthetic wheat of different percentages in their pedigrees (Tadesse et al. 2019). Understanding the genetic variation and dissecting the genetic basis of traits in these elite germplasms is very important to further improve wheat quality traits, especially Fe and Zn composition in breeding programs.

Previous studies using biparental populations have reported QTLs for Fe concentrations on wheat chromosomes 6A, 3B, 7B, and 4D (Wang et al. 2021). Marker-trait associations (MTAs) linked to Fe and Zn concentrations in wheat grains have also been reported through GWAS analysis (Tadesse et al. 2019). In addition to GWAS, genomic prediction (GP), also called genomic selection (GS) when applied in practical breeding to select individual plants, have been applied to determine the genome-estimated breeding values (GEBVs) of quantitatively inherited traits from the genomic data of selection candidates without undertaking field phenotyping or expensive quality analysis (Alemu et al. 2021; Alomari et al. 2018). The GEBVs of candidate or breeding individuals are determined using genomic prediction models developed from the genotypic and phenotypic data of the training population with several statistical models (Goiffon et al. 2017).

The current study was carried out using 252 elite spring bread wheat genotypes developed from ICARDA with the following objectives: (I) to evaluate the available genotypic variation in micronutrient content and grain yield; (II) to identify significantly linked MTAs with targeted traits via GWAS and localize putative candidate genes; and (IV) to estimate the GEBVs of target traits through single-trait- and multi-trait-based genomic prediction analysis. For this purpose, the panel was evaluated for 2 consecutive years (2018 and 20,219) under field condition and genotyped with 15 K SNP array.

Materials and methods

Plant material and field experimental conditions

The current study used a panel comprising 252 elite wheat genotypes developed from ICARDA (Supplementary Table S1). The field experiment was carried out at the ICARDA experimental station in Merchouch, Morocco (33°36′24.3"N 6°42′50.0"W, 430 masl) for two consecutive cropping seasons (2017–2018 and 2028–2019). The field trial was conducted testing all genotypes in a 3 m^2 plot size following an alpha lattice design with two replications. Planting was done on the first week of December during the two trial seasons similarly at a seeding rate of 100 kg per hectare. Merchouch station is characterized with a cambisol soil and was received an annual precipitation of about 320 mm over the two cropping seasons. Merchouch station has a moderate humidity with annual temperature ranging between 10 and 40 °C.

Phenotyping of grain yield and micronutrients

The currently used panel was phenotyped for four traits including grain yield (GY), iron (Fe), zinc (Zn), and selenium (Se). The micronutrient content of 252 elite wheat samples was evaluated following the standard protocol explained by Pequerul et al. (1993). Briefly, a total of 0.5 g of whole grain flour sample was placed in a digestion tube (QBlock series, Horiba) and kept overnight after adding 8 ml of nitric acid (HNO3). Samples were exposed for heating at 90 °C for 60 min followed by adding 3–4 ml of 30% hydrogen peroxide (H_2O_2). Then samples were filtered followed by diluting with hydrochloric acid (HCl) after the digestion process was completed and the solution became colorless. The determination of micronutrient content was performed using a simultaneous multi-element inductively coupled plasma emission spectrometer (iCAP-7000Duo, Thermo

Fisher Scientific) at the Cereal and Legume Quality Laboratory, ICARDA, Morocco. Grain yield was measured in kg/plot after threshing the whole plot.

The ANOVA, standard deviation as well as the coefficient of variation was determined using the *RcmdrMisc* package in R (Fox et al. 2022). The broad-sense heritability of traits was estimated using the formula:

$$H^2 = \mathrm{Vg}/\mathrm{Vp}$$

where H^2 represents the broad-sense heritability, Vg is the genotypic variance, and Vp is the phenotypic variance. The PCA and biplot analysis of phenotypic data for the four traits were conducted using the *FactoMineR* in R environment.

DNA extraction and genotyping

Genomic DNA was extracted from fresh leaf samples of 2-week-old seedlings following the procedures described by Ogbonnaya et al. (2001). Genotyping was carried out using the 15 K SNP array at SGS Institute Fresenius GmbH, Trait Genetics in Gatersleben, Germany. A total of 10,173 SNP markers were applied for the current genomic prediction, linkage disequilibrium, population structure, and marker-trait association analysis after filtering out below-standard-quality SNP markers with less than 5% minor allelic frequency (MAF) and above 20% missing values per individual.

Linkage disequilibrium and population structure analyses

The pairwise SNP markers linkage disequilibrium (LD) was calculated as r^2 using Trait Analysis by Association, Evolution, and Linkage (TASSEL) software (Bradbury et al. 2007) with a full size making a total of 786,219,856 marker comparisons. The LD decay trend through generations was estimated using the procedure developed by Remington et al. (2001) in R language following the nonlinear regression model developed by Hill and weir (Hill and Weir 1988) with the formula:

$$E\left(r^2\right) = \left[\frac{10+C}{(2+C)(11+C)}\right]\left[1+\frac{(3+C)\left(12+12C+C^2\right)}{n(2+C)(11+C)}\right]$$

where $E(r^2)$ is the expected value of r^2 under drift-recombination equilibrium; c is the recombination fraction between sites and n is the sample size.

Genome-wide association analysis and gene annotation

Genome-wide association analysis of the current panel from the absolute mean values of grain yield and micronutrient traits was performed using the mixed linear model (Q + K) in TASSEL including the kinship matrix (K) and principal components (Q) to control possible false-positive results raised from kinship and population structure, respectively. Manhattan plots were generated to display significant markers across chromosomes using *CMplot* package in the R environment using $-\log10(p) > 3.0$ as exploratory threshold (LiLin-Yin 2018). The physical position of SNP markers was retrieved from the International Wheat Genome Sequence Consortium (IWGSC) v1.1 (IWGS 2018). The Ensemble Plants database (https://plants.ensembl.org/Triticum_aestivum/Info/Index) was explored to identify putative functional genes-associated SNP markers identified as MTAs with studied traits following the variant effect predictor method, and the UniProt (https://www.uniprot.org) database was used to determine associated protein functions.

Single- and multi-trait genomic prediction and cross-validation analysis

Both the single- and multi-trait genomic prediction analysis was conducted using the genomic best unbiased prediction (GBLUP) model. The single-trait GBLUP-based genomic prediction analysis was done following the mixed linear model formula:

$$Y = 1\beta + \mathrm{Z}u + \varepsilon$$

where Y is the single-trait phenotypic observations of wheat genotypes for yield and micronutrient traits, β is the overall mean, Z is the design matrix for the random effects of SNP markers, u is the vector of additive effects derived from SNP markers following the multivariate normal $\mu \sim MN\,(\mathbf{0}, \mathbf{G}\sigma^2_g)$, where **G** is the marker-based genomic relationship matrix calculated using the method developed by VanRaden (2008) as $G = \frac{ZZ'}{p}$, where Z is the centered and standardized SNP marker matrix and p is the number of markers and σ^2_g is the genetic variance, ε is the vector of the error term derived similarly following the multivariate normal $(\mathbf{0}, \mathbf{I}\sigma^2_e)$, where **I** is the identity matrix and σ^2_e is the residual variance.

The micronutrient-related traits genomic prediction analysis was further analyzed with multi-trait based genomic prediction models following the mixed model formula:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} Z_1 & 0 & 0 \\ 0 & Z_2 & 0 \\ 0 & 0 & Z_3 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}$$

where Y_{ij} $i = 1, 2, 3$, and $j = 1, \ldots, n$ is a matrix of the response of the $_i$th trait related to micronutrients in the $_j$th wheat genotype, β_1- β_3 is a matrix from intercepts of each of the three micronutrient-related traits, Z_1- Z_3 the matrices of random SNP markers effect, u_1—u_3 is a matrix of

vector derived from markers effect, and $\varepsilon_1 - \varepsilon_3$ is a matrix of the error terms from each traits.

A fivefold cross-validation analysis was employed to evaluate the single- and multi-trait genomic prediction models where 80% of the panel was used to develop the prediction model. The developed model was used to estimate the genomic breeding values (GEBVs) of the remaining 20% individuals. The prediction accuracy was derived from the correlation between the GEBVs with recorded phenotypic values of yield and micronutrient-related traits. The cross-validation analysis was done for 100 iterations.

All the single- and multi-trait genomic prediction analysis was done in BGLR software package (Pérez and Los 2014) computed with Markov Chains Monte Carlo (MCMC) sampler with chain length of 12,000 iterations and 10 thinning intervals with the first 2000 was used as burn-in.

Results

Phenotypic analysis

The highly performed elite genotypes for grain yield can be found in Supplementary Table S1 with their pedigree information. The two genotypes G-163 and G-233 showed the highest grain yield scoring with 6.13 and 4.85 t/ha, respectively. A normal frequency distribution was observed from the phenotypic values for all four studied traits (Fig. 1). Elite lines showed a significant genotypic variation with a wide range from 1.82 to 6.13 t/ha with an average 3.16 t/ha while the broad-sense heritability was 0.42 (Table 1). All the micronutrient-related traits showed a significant variation among genotypes. Elite genotypes exhibited a wide range of variation for Zn, Fe, and Se with a score ranging from 15.33 to 45.62, 15.68 to 56.06, and 0.08 to 0.39 and with average values of 26.04, 45.62, and 0.22 mg/kg, respectively. The

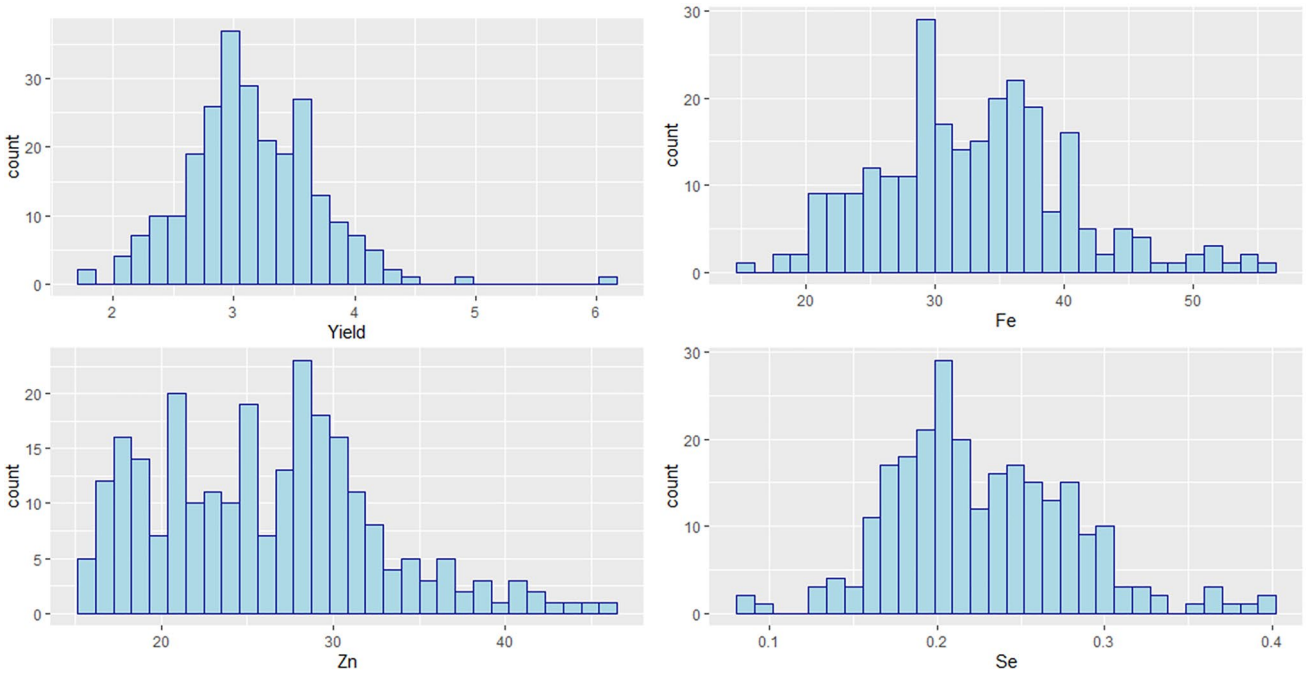


Fig. 1 Frequency distribution for grain yield and micronutrient-related traits recorded from the 252 elite wheat genotypes tested for 2 years in Merchouch, Morocco

Table 1 Analysis of variance for grain yield and micronutrient-related traits from 252 elite wheat genotypes recorded from 2 years field trials in Merchouch

Trait	Min	Max	Mean	Sd	CV	H^2
Fe	15.68	56.06	32.95	7.44	22.7	0.64
Zn	15.33	45.62	26.13	6.5	24.8	0.47
Se	0.08	0.39	0.22	0.05	23.7	0.68
GY	1.82	6.13	3.16	0.54	17.2	0.39

H^2 broad-sense heritability, *Sd* standard error, *CV* coefficient of variation

Se trait exhibited the highest broad-sense heritability with 0.68 followed by Fe and Zn contents with 0.64 and 0.47, respectively, while GY had showed only 0.39.

Pearson correlation and principal component analyses

Principal component analysis (PCA) with phenotypic scorings clustered the 252 genotypes into 4 distinct groups (Fig. 2). The first and second principal components accounted for 34.9% and 28.1% of the total phenotypic variation observed in grain yield and micronutrient-related traits, respectively. Eigenvalue for cluster one was trended to the right while clusters two, three, and four were located to the left, top, and bottom side from the biplot origin, respectively. The vectors of the biplot analysis for Fe and Zn were moved on the positive direction above the origin. In contrast, the vector value for grain yield was above the origin but in the left direction, while the vector

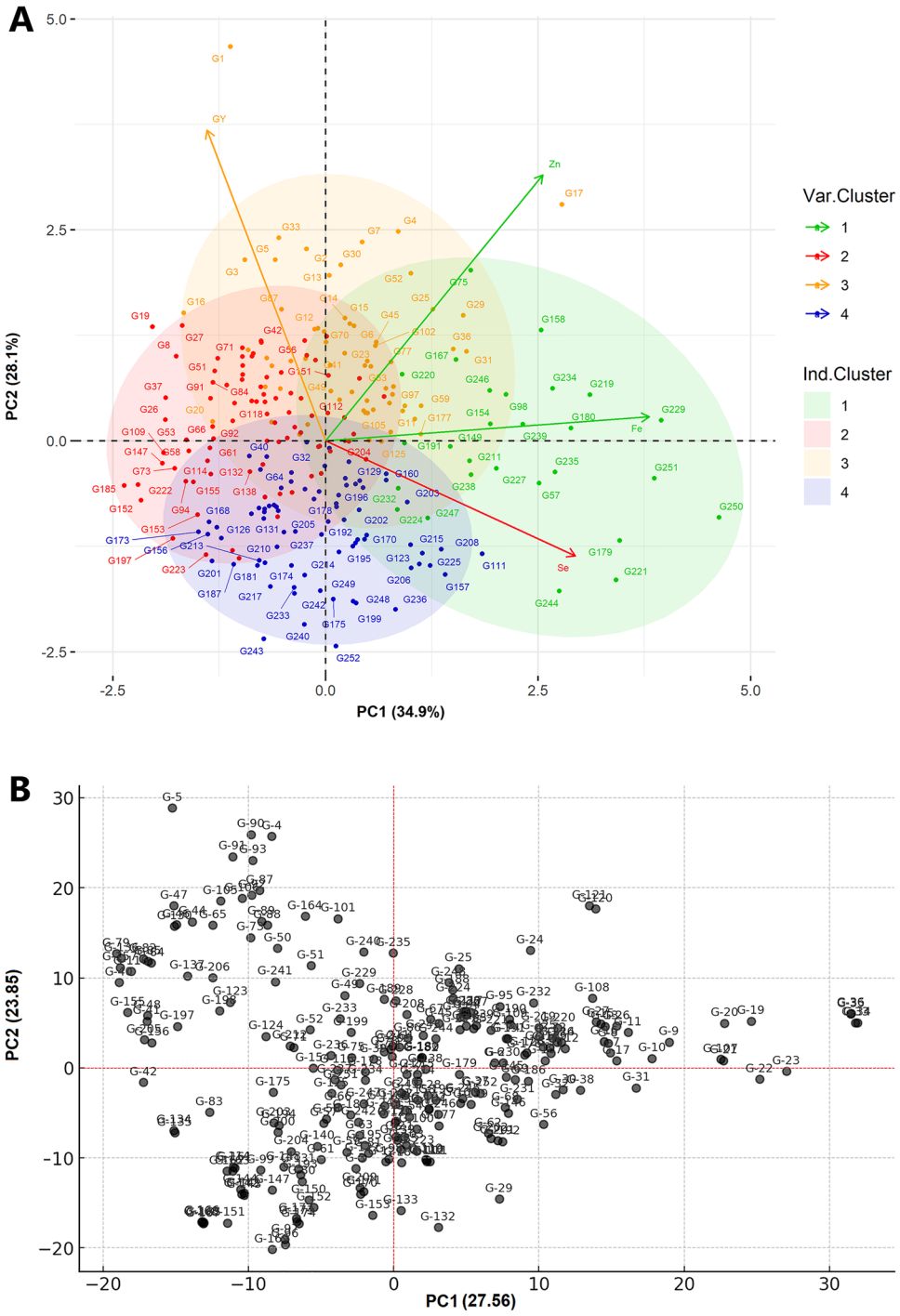
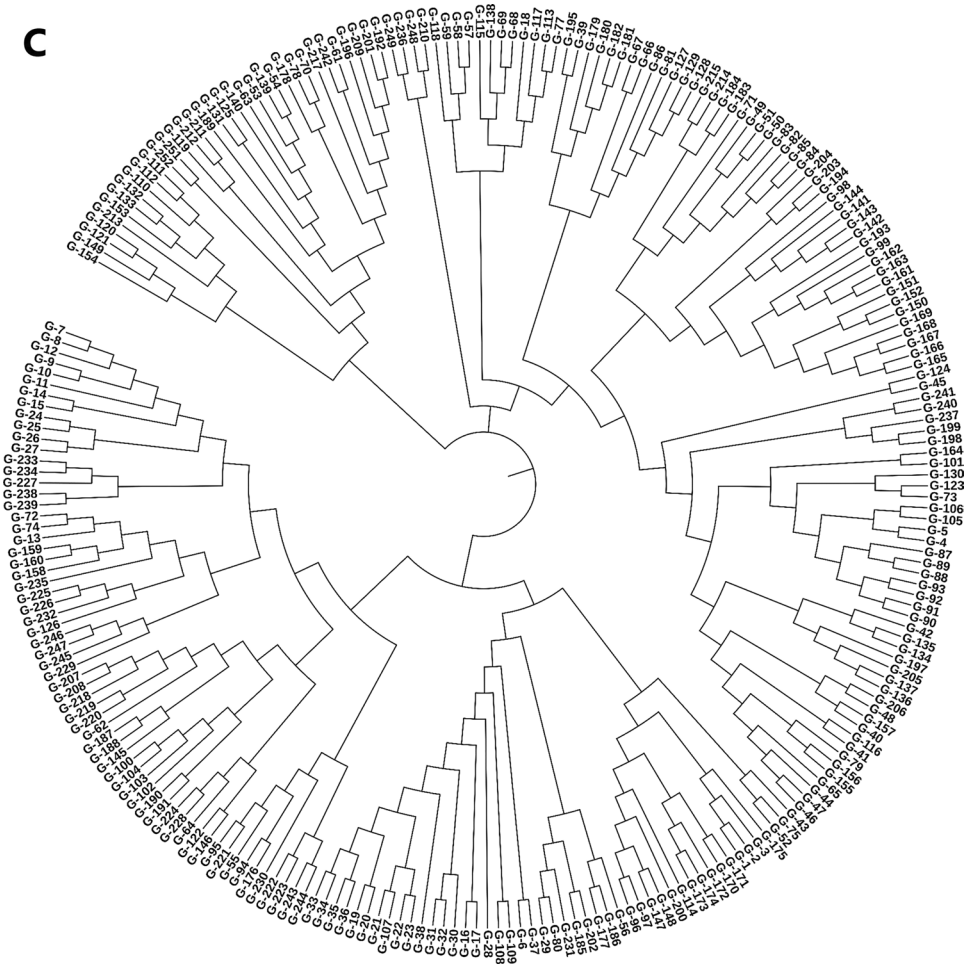


Fig. 2 Cluster and principal component biplot analysis of tested bread wheat genotypes in Merchouch (Morocco). Biplot analysis of genotypes according to their phenotypic performance (**A**), principal component (**B**) and neighbor_joining tree (NJ) analysis (**C**) of the 252 genotypes with their SNP marker profile

Fig. 2 (continued)



for Se was below the origin and in the right direction. Fe and Zn contents were grouped in cluster 1, Se in cluster 2, and grain yield grouped in cluster 3. The SNP markers files with the neighbor-joining (NL) tree and principal component algorithms grouped elite lines with their respective pedigree similarity (Fig. 2b and c). For instance, the genotypes G-1, G-2, and G-3 clustered together on the NJ tree are linked to each other as sister lines with the same pedigree (CROC-1/AE.SQUARROSA(224)//OPATA/4/SERI.1B*2/3/KAUZ*2/BOW//KAUZ) while the G-142 (SERI.1B*2/3/KAUZ*2/BOW//KAUZ/4/ATTILA/3*BCN) and G-143 (SERI.1B*2/3/KAUZ*2/BOW//KAUZ/4/TRAP#1/BOW//PFAU/3/MILAN) are sharing the same node in their pedigree information (Supplementary Table S1). Furthermore, Fe showed a significant and positive correlation with Zn and Se with Pearson's correlation coefficients r^2) 0.23 and 0.26, respectively (Fig. 3).

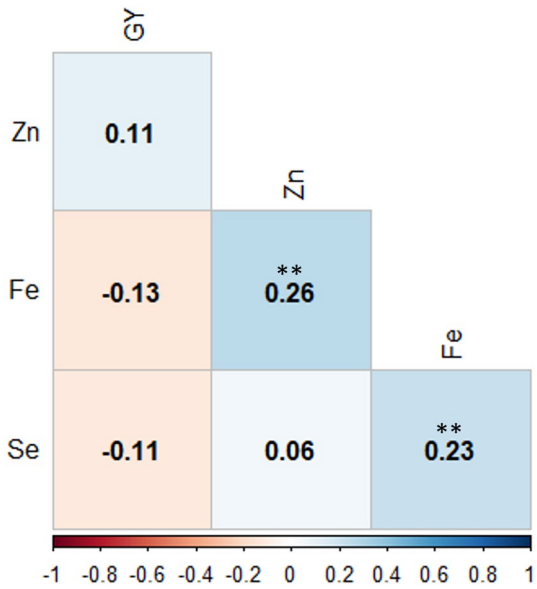


Fig. 3 Pearson's correlation between grain yield and the three micronutrient-related traits. **Significant at $p < 0.01$

Linkage disequilibrium, marker-trait association, and gene annotation

The linkage disequilibrium analysis exhibited the D subgenome with the highest linkage disequilibrium having an

average r^2 value of 0.21 followed by the sub-genomes B and A with 0.20 and 0.19, respectively. The genome-wide linkage disequilibrium revealed the LD decay started at r^2 value of 0.46. The half-LD decay ($r^2 = 0.23$) intersected with the LD decay trend at 8.77 Mbp (Fig. 4).

A total of 19 marker-trait associations (MTAs) was identified significantly ($P < 0.001$) linked to the 4 tested traits (Table 2). Six MTAs were identified as significantly linked to grain yield under drought condition. The marker *IAAV2271* located on chromosome 4B was particularly highly significantly ($P < 0.000001$) linked with grain yield with $-\log_{10}{}^{P}$ value of 6.06 followed by the markers

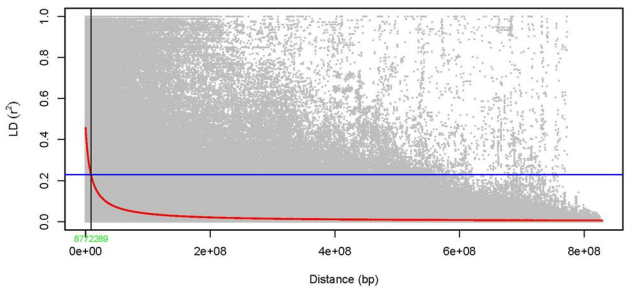


Fig. 4 Scatter plot for the pairwise SNPs genome-wide linkage disequilibrium (LD) decay measured for 252 elite wheat genotypes in ICARDA. Genetic distance in Mbp is plotted against pairwise LD r^2 values. The red solid curve represents the genome-wide LD decay generated from the smoothing spline regression model. The horizontal blue line represents the half-LD decay e ($r^2 = 0.23$), and the vertical black solid line represents the genetic distance (8.77) at which the half-LD decay intersects with the LD decay curve

BobWhite_c4502_252 and *RAC875_c46194_201* which are both located on chromosome 3B with $-\log_{10}{}^{P}$ value of 4.8 and 3.6, respectively. Other six MTAs were identified for Fe located on chromosomes 3A, 4A, 7A, 2B, and 1D (Table 2, Fig. 5). *BS00049927_51* located on chromosome 1B was the SNP marker identified as highly significantly linked with Fe with $-\log_{10}{}^{P}$ value of 3.58. For zinc content, four significantly linked MTAs were identified on chromosomes 3A, 7A, 1B, and 5B. Marker *RAC875_c66649_186* located on chromosome 1B was identified as significantly linked with Fe with $-\log_{10}{}^{P}$ value of 5.0. The marker *Excalibur_c10046_579* on chromosome 5D was detected to be linked with selenium content.

A total of 16 putative candidate genes were discovered from the 19 identified SNP markers significantly associated to Fe, Zn, Se, and GY traits (Table 3). The marker *RAC875_c66649_186* highly significantly associated with Zn was linked with *TraesCS1B02G395000* gene which encodes for the zinc finger protein *BRUTUS* and is involved in zinc ion binding. The SNP marker *wsnp_Ku_c4299_7814936* identified on chromosome 7A had a significant association with Fe and is linked with the gene *TraesCS7A02G024800*. This gene encodes for the WAT1-related protein involved in transmembrane transporter activity. The SNP *Kukri_c33670_506* having a MTA with Se content from GWAS analysis is linked with the gene *TraesCS1D02G041400* which encodes for the Protein *gar2* involved in RNA binding. For grain yield, the marker *BobWhite_c4502_252* was linked to the gene *TraesCS3B02G086400*, which encodes for the *E3*

Table 2 List of identified marker-trait associations significantly linked with Fe, Zn, Se, and grain yield identified through genome-wide association analysis

Trait	Marker	Chromosome	Position (bp)	R^2	–log10(*p*)
Fe	*BS00049927_51*	1B	108,837,029	0.06	3.58
Fe	*wsnp_Ku_c4299_7814936*	7A	9,753,520	0.06	3.38
Fe	*BS00077914_51*	2B	42,281,519	0.06	3.22
Fe	*Kukri_rep_c92967_369*	2B	15,882,612	0.05	3.05
Fe	*wsnp_Ex_c13031_20625900*	4A	604,209,019	0.04	3.01
Fe	*wsnp_Ex_c45877_51547406*	3A	743,190,499	0.05	3.00
Zn	*RAC875_c66649_186*	1B	626,707,491	0.09	5
Zn	*RAC875_c36922_829*	3A	9,852,634	0.04	3.19
Zn	*BS00079237_51*	7A	3,906,831	0.04	3.11
Zn	*Kukri_c41408_233*	5B	684,612,171	0.04	3.02
Se	*Excalibur_c10046_579*	5D	160,064,658	0.06	3.51
Se	*RAC875_rep_c70595_321*	5D	154,539,742	0.05	3.13
Se	*Kukri_c33670_506*	1D	20,056,459	0.04	3.09
Yield	*IAAV2271*	4B	17,469,995	0.11	6.06
Yield	*BobWhite_c4502_252*	3B	54,752,111	0.09	4.8
Yield	*RAC875_c46194_201*	3B	59,064,804	0.06	3.60
Yield	*wsnp_Ra_c26091_35652620*	5B	483,000,113	0.05	3.57
Yield	*RAC875_c12959_869*	4B	17,469,492	0.06	3.19
Yield	*Ra_c6065_1145*	6A	9,327,369	0.05	3.09

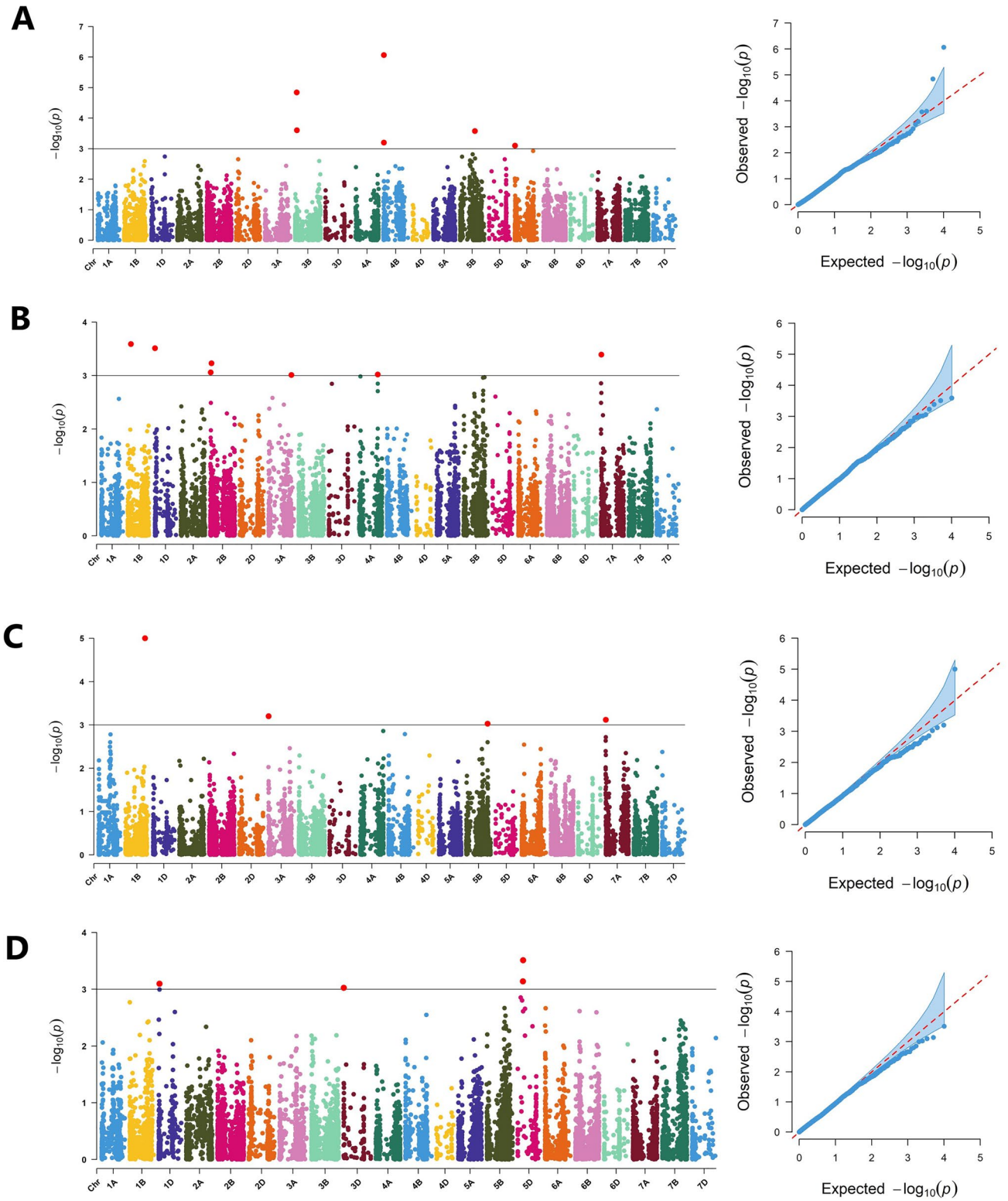


Fig. 5 Manhattan (left) and Q–Q (right) plots for grain yield (**A**), Fe (**B**), Zn (**C**), and Se (**D**) traits identified through GWAS analysis. The horizontal solid line indicates the exploratory significant threshold applied at P of 0.001 or $-\log P < 3$ to identify MTAs

Table 3 List of putative genes and their function identified from marker-trait associations for Fe, Zn, Se, and grain yield with the genome-wide association analysis

Trait	Gene	Protein name	Molecular function	Biological process
Fe	*TraesCS2B02G033000*	Tyrosine N-monooxygenase-like	Iron ion binding, monooxygenase activity	–
Fe	*TraesCS3A02G530000*	AP-3 complex subunit beta	–	Intracellular protein transport
Fe	*TraesCS4A02G313900*	Protein kinase domain-containing protein	Protein kinase activity	Protein phosphorylation
Fe	*TraesCS7A02G024800*	WAT1-related protein	Transmembrane transporter activity	–
Zn	*TraesCS1B02G395000*	Zinc finger protein BRUTUS	Zinc ion binding	Protein ubiquitination
Zn	*TraesCS3A02G013800*	Fatty acyl-CoA reductase	Alcohol-forming fatty acyl-CoA reductase activity	Long-chain fatty-acyl-CoA metabolic process
Zn	*TraesCS5B02G523800*	AAA domain-containing protein	Metalloendopeptidase activity	Proteolysis
Zn	*TraesCS7A02G008900*	Monocopper oxidase-like protein SKU5	Copper ion binding	–
Se	*TraesCS1D02G041300*	Mitoc_mL59 domain-containing protei	–	–
Se	*TraesCS1D02G041400*	Protein gar2	RNA binding	–
Se	*TraesCS5D02G118000*	HECT-type E3 ubiquitin transferase	Ubiquitin protein ligase activity	Protein ubiquitination
GY	*TraesCS3B02G086400*	E3 ubiquitin-protein ligase	Ubiquitin protein ligase activity	Regulation of gene expression
GY	*TraesCS3B02G089200*	F-box domain-containing protein	–	–
GY	*TraesCS4B02G024500*	Dirigent protein	Carbohydrate binding	Phenylpropanoid biosynthetic process
GY	*TraesCS5B02G299400*	Transcription factor 25	–	–
GY	*TraesCS6A02G019400*	Protein transport protein sec16	–	Golgi organization

ubiquitin-protein ligase involved in the regulation of gene expression and ubiquitin-protein ligase activity.

Genomic prediction and cross-validation analysis

The single-trait-based genomic prediction analysis of grain yield with the GBLUP model estimated an average predictive ability of 0.32 ranging from 0.186 to 0.502 (Table 4, Fig. 6). The genomic prediction for the other three micronutrient-related traits was explored with both single- and multi-trait-based genomic prediction methods in GBLUP model. The predictive ability of Se and Fe traits with multi-trait-based genomic prediction GBLUP model was 0.161 and 0.259 improving by 6.62 and 4.44%, respectively, compared to the single-trait-based models. However, the predictive ability of Zn remained constant in both methods (Table 4).

Discussion

Grain yield and quality-related traits of crops are quantitatively inherited and are governed by many genes (polygenes). Environment, management, and the interaction between genotypes with these factors further affect the performance of wheat genotypes for these traits. The current study assessed the genotypic variation of 252 elite bread wheat genotypes for yield and micronutrients under rainfed conditions followed by GWAS and genomic prediction analysis. A significant variation was observed among evaluated genotypes for grain yield and micronutrient-related traits agreeing with several previous studies (Arora et al. 2019; Rathan

Table 4 Single- and multi-trait-based genomic prediction analysis for grain yield and micronutrient-related traits

Trait	Single-trait			Multi-trait		
	Minimum	Maximum	Average	Minimum	Maximum	Average
GY	0.186	0.502	0.32	–	–	–
Fe	0.216	0.301	0.248	0.243	0.35	0.259
Zn	0.28	0.379	0.341	0.278	0.378	0.34
Se	0.11	0.183	0.151	0.11	0.248	0.161

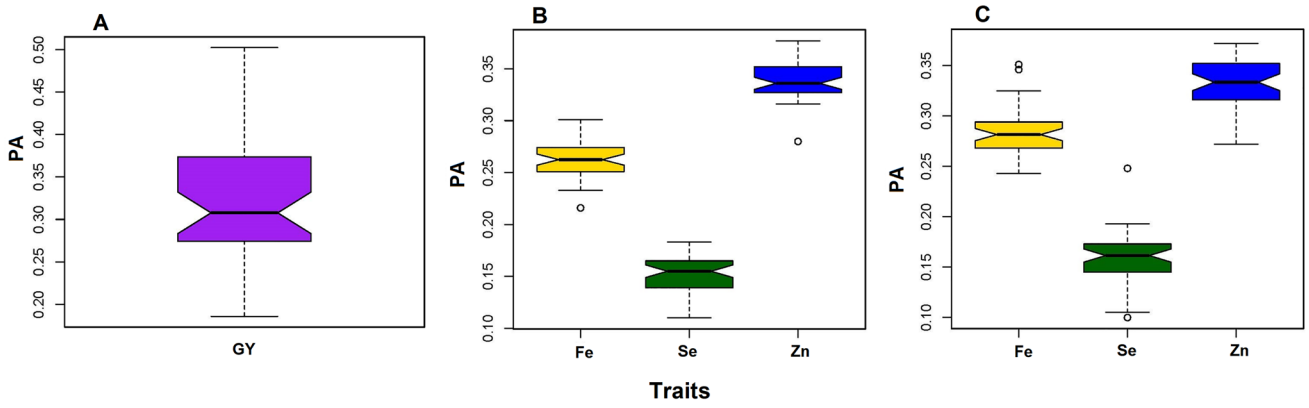


Fig. 6 Predictive ability results from the single- and multi-trait-based genomic prediction analysis using the GBLUP model. The predictive ability for grain yield (**A**) was measured on a single-trait-based GBLUP model while the micronutrient-related traits were tested with both single-trait (**B**) and multi-trait (**C**) GBLUP models

et al. 2022; Gupta et al. 2022). According to Tadesse et al. (2015), the performance of yield and yield-related traits decreases under drought conditions while quality traits such as grain protein content and gluten content show significant increments. Agreed with previous studies, grain yield demonstrated a negative correlation with the currently studied quality traits (Kaya and Akcura 2014). In the current study, Fe exhibited a slightly significant positive correlation with Zn and Se inconsistent with previous reports (Bhatta et al. 2018). However, it is noteworthy to know that environmental factors play an important role in the performance of these traits and is important to consider both genotype and environment when making breeding decisions. The significant genotype x environment interaction further emphasizes the need to evaluate genotypes in multiple environments to identify those with broad adaptation.

GWAS has been used to identify unique MTAs and QTLs associated with valuable traits in different crops. In this study, we found a total of 19 significant markers associated with micronutrient-related traits and grain yield across different wheat chromosomes. Similarly, a higher number of MTAs was previously reported for grain yield, yield-related traits, and quality traits on the A and B genomes while the D genome relatively contributed less (Wang et al. 2021). We identified a highly significant marker for Fe content on chromosome 1B with $-\log_{10}{}^{P}=3.58$ located at 108.8 Mbp. Several studies have detected many MTAs and QTLs for Fe content on the same chromosomes (Tong et al. 2022; Devate et al. 2022), but slightly on different positions indicating that these MTA could possibly be a novel marker discovered in our germplasm.

In this study, we identified four MTAs for Zn content located on chromosomes 1B, 3A, 5B, and 7A. The most significant marker was *RAC875_c66649_186* located on chromosome 1B. A previous study found the same marker on chromosome 1B linked to the grain filling rate (Yang et al. 2020), indicating that this marker is effective across different sets of materials tested in various environments and could be used for germplasm screening via MAS in breeding programs. In addition, we identified a marker highly associated with Se content on chromosome 5D with $-\log_{10}{}^{P}=3.51$. Several MTAs and QTLs on chromosome 5D have been identified previously linked to Se under dry conditions (Ma et al. 2022). In this study, we found a total of six significant MTAs on chromosomes 3B, 4B, 5B, and 6A associated with grain yield. Previous studies have also reported MTAs linked to grain yield on the same chromosomes (Maphosa et al. 2014; Lesk et al. 2016; Li et al. 2018). Using a population of 127 RILs derived from the cross Ning7840 X Clark, Li et al. (2016) discovered the QTL QGpc.hwwgr-5BL on chromosome 5B.

Candidate genes linked to the MTAs with studied grain yield and micronutrient-related traits were mapped using the bread wheat genome reference database at Ensemble plant and UniProt. Previous studies have identified 16 genes involved in relevant biosynthetic pathways, including pathways related to zinc ion binding, Fe ion binding, ATP binding, monooxygenase activity, protein kinase activity, and ubiquitin-protein ligase activity (Raza et al. 2014; Michaletti et al. 2018). The candidate gene *TraesCS2B02G033000* located close to the marker *BS00077914_51*, which is associated with the Fe content, encodes for the *Tyrosine N-monooxygenase-like* enzyme which is involved in Fe ion binding and monooxygenase activity (Robertson and Biaggioni 2012). The marker *RAC875_c66649_186* located close to the gene *TraesCS1B02G395000* encodes for zinc finger protein BRUTUS and this protein is involved in zinc ion binding and protein ubiquitination (Rodríguez-Celma et al. 2019). *TraesCS4B02G024500* closely found with the SNP marker

RAC875_c12959_869 is one of the five identified putative genes linked to grain yield. This gene encodes for Dirigent protein which is involved in carbohydrate binding and phenylpropanoid biosynthetic process (Li et al. 2019).

Genomic prediction can be a powerful tool for plant breeding and has the potential to significantly accelerate the development of new improved crop varieties (Alemu et al. 2023). The use of genomic prediction, genomic selection when the method applied practically in breeding, has been identified as the best strategy to increase the breeding efficiency of genetically complicated low-heritable traits in breeding (Alomari et al. 2018). Our genomic prediction results showed moderate prediction accuracies of mineral traits, which is consistent with previous studies that found low to moderate genomic predictability values of micronutrient-related traits (Kristensen et al. 2018; Mérida-García et al. 2019). The multi-trait-based genomic prediction methods could improve the accuracy of genomic selection, especially highly correlated traits, such as several micronutrient-related traits in crops (Atanda et al. 2022). Comparing the previous genomic prediction studies of micronutrient-related traits in wheat to other crops, rice shows a low to moderate prediction accuracy values (0.21 to 0.52) (Muvunyi et al. 2022) while a moderate prediction accuracy in maze (0.34 to 0.47) (Mageto et al. 2020). A previous genomic prediction study in pea quality traits shows low to moderate prediction accuracy (Atanda et al. 2022).

To conclude, the current study investigated the potential of ICARDA's elite wheat genotypes for grain yield and three micronutrient-related traits followed by identification of MTAs and estimating the GEBVs of individual genotypes via GWAS and genomic prediction studies, respectively. Nineteen MTAs and sixteen associated putative genes were identified with the potential to facilitate the MAS process in the search for developing new improved varieties. The multi-trait-based genomic prediction analysis leads to an improved prediction accuracy for highly correlated micronutrient-related traits.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00438-023-02074-6.

Author contribution WT, ZE, and AE are involved in material preparation, field data collection, phenotypic and GWAS analysis. AA performed genomic prediction analysis. WT, ZE, and AA prepared the draft manuscript. All authors reviewed, edited, and approved the final version of the manuscript.

Funding Open access funding provided by Swedish University of Agricultural Sciences. This work was supported by the International Center for Agricultural Research in the Dry Areas (ICARDA).

Data availability The datasets generated for this study are included in the article and as supplementary files.

Declarations

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.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Alemu A, Suliman S, Hagras A et al (2021) Multi-model genome-wide association and genomic prediction analysis of 16 agronomic, physiological and quality related traits in ICARDA spring wheat. Euphytica 217:205. https://doi.org/10.1007/s10681-021-02933-6

Alemu A, Batista L, Singh PK, Ceplitis A, Chawade A (2023) Haplotype-tagged SNPs improve genomic prediction accuracy for Fusarium head blight resistance and yield-related traits in wheat. Theor Appl Genet. https://doi.org/10.1007/s00122-023-04352-8

Alomari DZ, Eggert K, von Wirén N et al (2018) Identifying candidate genes for enhancing grain Zn concentration in wheat. Front Plant Sci 9:1313. https://doi.org/10.3389/fpls.2018.01313

Arora S, Cheema J, Poland J et al (2019) Genome-wide association mapping of grain micronutrients concentration in Aegilops tauschii. Front Plant Sci 10:54. https://doi.org/10.3389/fpls.2019.00054

Atanda SA, Steffes J, Lan Y et al (2022) Multi-trait genomic prediction improves selection accuracy for enhancing seed mineral concentrations in pea. Plant Genome. https://doi.org/10.1002/tpg2.20260

Bhatta M, Baenziger P, Waters B et al (2018) Genome-wide association study reveals novel genomic regions associated with 10 grain minerals in synthetic hexaploid wheat. IJMS 19:3237. https://doi.org/10.3390/ijms19103237

Bradbury PJ, Zhang Z, Kroon DE et al (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23:2633–2635. https://doi.org/10.1093/bioinformatics/btm308

Devate NB, Krishna H, Sunilkumar VP et al (2022) Identification of genomic regions of wheat associated with grain Fe and Zn content under drought and heat stress using genome-wide association study. Front Genet 13:1034947. https://doi.org/10.3389/fgene.2022.1034947

FAOSTAT (2021) Food and Agriculture Organization. FAO

Fox J, Muenchen R, Putler D (2022) Package 'RcmdrMisc.' CRAN 40

Gletsu-Miller N, Wright BN (2013) Mineral malnutrition following bariatric surgery. Adv Nutr 4:506–517. https://doi.org/10.3945/an.113.004341

Goiffon M, Kusmec A, Wang L et al (2017) Improving response in genomic selection with a population-based selection strategy: optimal population value selection. Genetics 206:1675–1682. https://doi.org/10.1534/genetics.116.197103

Gupta OP, Singh AK, Singh A et al (2022) Wheat biofortification: utilizing natural genetic diversity, genome-wide association

mapping, genomic selection, and genome editing technologies. Front Nutr 9:826131. https://doi.org/10.3389/fnut.2022.826131
Hill WG, Weir BS (1988) Variances and covariances of squared linkage disequilibria in finite populations. Theor Popul Biol 33:54–78. https://doi.org/10.1016/0040-5809(88)90004-4
IWGSC (2018) Shifting the limits in wheat research and breeding using a fully annotated reference genome. Science 361(eaar7191):2018. https://doi.org/10.1126/science.aar7191
Kaya Y, Akcura M (2014) Effects of genotype and environment on grain yield and quality traits in bread wheat (*T. aestivum* L.). Food Sci Technol (campinas) 34:386–393. https://doi.org/10.1590/fst.2014.0041
Kristensen PS, Jahoor A, Andersen JR et al (2018) Genome-wide association studies and comparison of models and cross-validation strategies for genomic prediction of quality traits in advanced winter wheat breeding lines. Front Plant Sci 9:69. https://doi.org/10.3389/fpls.2018.00069
Lesk C, Rowhani P, Ramankutty N (2016) Influence of extreme weather disasters on global crop production. Nature 529:84–87. https://doi.org/10.1038/nature16467
Li C, Bai G, Chao S et al (2016) Single nucleotide polymorphisms linked to quantitative trait loci for grain quality traits in wheat. Crop J 4:1–11. https://doi.org/10.1016/j.cj.2015.10.002
Li A, Liu D, Yang W et al (2018) Synthetic hexaploid wheat: yesterday, today, and tomorrow. Engineering 4:552–558. https://doi.org/10.1016/j.eng.2018.07.001
Li F, Wen W, Liu J et al (2019) Genetic architecture of grain yield in bread wheat based on genome-wide association studies. BMC Plant Biol 19:168. https://doi.org/10.1186/s12870-019-1781-3
LiLin-Yin (2018) CMplot: Circle Manhattan Plot. CRAN
Lim K, Riddell L, Nowson C et al (2013) Iron and zinc nutrition in the economically-developed world: a review. Nutrients 5:3184–3211. https://doi.org/10.3390/nu5083184
Ma J, Qi S, Yuan M et al (2022) A genome-wide association study revealed the genetic variation and candidate genes for grain copper content in bread wheat (*Triticum aestivum* L.). Food Funct 13:5177–5188. https://doi.org/10.1039/D1FO04173H
Mageto EK, Crossa J, Pérez-Rodríguez P et al (2020) Genomic Prediction with genotype by environment interaction analysis for kernel zinc concentration in tropical maize germplasm. G3 Genes|genom|genet. 10:2629–2639. https://doi.org/10.1534/g3.120.401172
Maphosa L, Langridge P, Taylor H et al (2014) Genetic control of grain yield and grain physical characteristics in a bread wheat population grown under a range of environmental conditions. Theor Appl Genet 127:1607–1624. https://doi.org/10.1007/s00122-014-2322-y
Mérida-García R, Liu G, He S et al (2019) Genetic dissection of agronomic and quality traits based on association mapping and genomic selection approaches in durum wheat grown in Southern Spain. PLoS ONE 14:e0211718. https://doi.org/10.1371/journal.pone.0211718
Michaletti A, Naghavi MR, Toorchi M et al (2018) Metabolomics and proteomics reveal drought-stress responses of leaf tissues from spring-wheat. Sci Rep 8:5710. https://doi.org/10.1038/s41598-018-24012-y
Muvunyi BP, Zou W, Zhan J et al (2022) Multi-trait genomic prediction models enhance the predictive ability of grain trace elements in rice. Front Genet 13:883853. https://doi.org/10.3389/fgene.2022.883853
Ogbonnaya FC, Seah S, Delibes A et al (2001) Molecular-genetic characterisation of a new nematode resistance gene in wheat. Theor Appl Genet 102:623–629. https://doi.org/10.1007/s001220051689
Pequerul A, Pérez C, Madero P et al (1993) A rapid wet digestion method for plant analysis. In: Fragoso MAC, Van Beusichem ML, Houwers A (eds) Optimization of plant nutrition. Springer, Dordrecht, pp 3–6
Pérez P, De Los CG (2014) Genome-wide regression and prediction with the BGLR statistical package. Genetics 198:483–495. https://doi.org/10.1534/genetics.114.164442
Rathan ND, Krishna H, Ellur RK et al (2022) Genome-wide association study identifies loci and candidate genes for grain micronutrients and quality traits in wheat (*Triticum aestivum* L.). Sci Rep 12:7037. https://doi.org/10.1038/s41598-022-10618-w
Raza MAS, Saleem MF, Shah GM et al (2014) Exogenous application of glycinebetaine and potassium for improving water relations and grain yield of wheat under drought. J Soil Sci Plant Nutr. https://doi.org/10.4067/S0718-95162014005000028
Remington DL, Thornsberry JM, Matsuoka Y et al (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proc Natl Acad Sci USA 98:11479–11484. https://doi.org/10.1073/pnas.201394398
Robertson D, Biaggioni I (eds) (2012) Primer on the autonomic nervous system, 3rd edn. Elsevier, Amsterdam, Boston
Rodríguez-Celma J, Connorton JM, Kruse I et al (2019) Arabidopsis BRUTUS-LIKE E3 ligases negatively regulate iron uptake by targeting transcription factor FIT for recycling. Proc Natl Acad Sci USA 116:17584–17591. https://doi.org/10.1073/pnas.1907971116
Shewry PR (2009) Wheat. J Exp Bot 60:1537–1553. https://doi.org/10.1093/jxb/erp058
Suliman S, Alemu A, Abdelmula AA et al (2021) Genome-wide association analysis uncovers stable QTLs for yield and quality traits of spring bread wheat (*Triticum aestivum*) across contrasting environments. Plant Gene 25:100269. https://doi.org/10.1016/j.plgene.2020.100269
Tadesse W, Ogbonnaya FC, Jighly A et al (2015) Genome-wide association mapping of yield and grain quality traits in winter wheat genotypes. PLoS ONE 10:e0141339. https://doi.org/10.1371/journal.pone.0141339
Tadesse W, Bishaw Z, Assefa S (2019) Wheat production and breeding in Sub-Saharan Africa: challenges and opportunities in the face of climate change. IJCCSM 11:696–715. https://doi.org/10.1108/IJCCSM-02-2018-0015
Tong J, Zhao C, Sun M et al (2022) High resolution genome wide association studies reveal rich genetic architectures of grain zinc and iron in common wheat (*Triticum aestivum* L.). Front Plant Sci 13:840614. https://doi.org/10.3389/fpls.2022.840614
Umamaheswari K, Bhaskaran M, Krishnamurthy G et al (2011) Effect of iron and zinc deficiency on short term memory in children. Indian Pediatr 48:289–293. https://doi.org/10.1007/s13312-011-0060-7
VanRaden PM (2008) Efficient methods to compute genomic predictions. J Dairy Sci 91:4414–4423. https://doi.org/10.3168/jds.2007-0980
Wang W, Guo H, Wu C et al (2021) Identification of novel genomic regions associated with nine mineral elements in Chinese winter wheat grain. BMC Plant Biol 21:311. https://doi.org/10.1186/s12870-021-03105-3
World Health Organization (2015) The global prevalence of anaemia in 2011. World Health Organization, Geneva
Yang Y, Chai Y, Zhang X et al (2020) Multi-locus GWAS of quality traits in bread wheat: mining more candidate genes and possible regulatory network. Front Plant Sci 11:1091. https://doi.org/10.3389/fpls.2020.01091

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