Check for updates

OPEN ACCESS

EDITED BY Soren K. Rasmussen, University of Copenhagen, Denmark

REVIEWED BY Ana M. Casas, Spanish National Research Council (CSIC), Spain Dionysia A. Fasoula, Agricultural Research Institute, Cyprus

*CORRESPONDENCE Joanne Russell Joanne.russell@hutton.ac.uk

RECEIVED 03 December 2024 ACCEPTED 28 March 2025 PUBLISHED 24 April 2025

CITATION

Dinsa GF, Russell J, Steffenson B, Halpin C and Waugh R (2025) Variation at the major facilitator superfamily *ZIFL1* gene influences zinc concentration of barley grain. *Front. Plant Sci.* 16:1539029. doi: 10.3389/fpls.2025.1539029

COPYRIGHT

© 2025 Dinsa, Russell, Steffenson, Halpin and Waugh. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Variation at the major facilitator superfamily *ZIFL1* gene influences zinc concentration of barley grain

Girma Fana Dinsa¹, Joanne Russell^{2*}, Brian Steffenson³, Claire Halpin¹ and Robbie Waugh^{1,2}

¹University of Dundee, Division of Plant Sciences at the James Hutton Institute, Dundee, United Kingdom, ²The James Hutton Institute, Cell and Molecular Science Department, Dundee, United Kingdom, ³Department of Plant Pathology, University of Minnesota, St. Paul, MN, United States

Food and nutritional security are global challenges exacerbated by an increasing human population and impacted by climate change. Barley is among the top cereal crops grown worldwide and is a strategic crop for food and nutrition security in several geographical domains. However, barley grains are generally limited in iron and zinc, two major micronutrient deficiencies affecting billions of people around the world, but particularly women and children in developing countries. One promising strategy to enhance crop micronutrient status is via biofortification, the identification and use of nutrient-rich natural variants in crop genetic improvement. Germplasm assessed as being rich in essential nutrients are used as parental materials in traditional breeding strategies. While simple in theory, directly assessing grain nutrient concentration as a phenotype in a crop breeding program is not trivial, particularly in lesser developed geographies. As an alternative, genetic diagnostics can simplify the identification of desirable progenies and accelerate the breeding process. Here we explored natural variation for grain zinc concentration within 296 Ethiopian and Eritrean barley landraces using a genome-wide association study. We found strong associations with two SNPs, both of which were located within the barley ortholog of a tonoplast-associated major facilitator superfamily (MFS) transporter gene, Zinc induced facilitator-like 1 (ZIFL1) of Arabidopsis thaliana (AtZIFL1). Sequencebased haplotype analysis of the barley gene (HvZIFL1) extended this association to a 153-162 bp deletion in a non-coding region. The favourable haplotype, associated with higher grain Zn concentration, was found in ~20% of Ethiopian and Eritrean barley germplasm. Markers are designed to the diagnostic SNPs for use as molecular diagnostics in breeding for genotypes with enhanced grain Zn.

KEYWORDS

micronutrient, genome-wide association study, quantitative trait loci, zinc, barley

01

Introduction

An estimated 3.7 billion people worldwide suffer from iron and zinc deficiencies (Karakoy et al., 2012; De Groote et al., 2021). With our general reliance on only a few crops to supply these essential micronutrients, it is considered increasingly important that we try to develop crop-specific solutions to improving nutritional quality (Stangoulis and Knez, 2022). Genetic biofortification, the natural enrichment of foods through breeding, has been suggested as a sustainable and viable approach (Stangoulis and Knez, 2022). A complementary strategy, supplementing micronutrients in fertilisers or flour (agronomic biofortification), has also been proposed (Cu et al., 2021) but is considered costly, largely ineffective and less easy to implement in subsistence farming situations. Screening germplasm to identify accessions that have increased micronutrient concentration in the grain or fruit, then understanding and incorporating the responsible beneficial alleles through marker assisted selection, is perhaps the most promising approach towards eradicating the 'hidden hunger' of micronutrient deficiency (Karakoy et al., 2012; Hefferon, 2019; Stangoulis and Knez, 2022; Devate et al., 2022). This global issue is particularly important in developing countries, like Ethiopia, where cereal crops such as barley are consumed as staples, despite being unable to provide sufficient critical micronutrients considered a necessary component of a healthy diet. There, 38% of children under five years of age were recently classed as stunted, a condition directly linked to low dietary zinc intake. Malnutrition and stunting are particularly prevalent in the extreme highlands where barley and faba bean are the main sources of food and nutrition. These are the principal crops grown by 4.5 million smallholder farmers with over 20 million dependent individuals. Micronutrient deficient soils are common in these regions (Karakoy et al., 2012). Indeed, it is estimated that 50% of agricultural soils in Ethiopia are zinc deficient (De Groote et al., 2021). As dietary intake is often the main source of Zn nutrition in these regions, Zn deficiency frequently reflects the widespread consumption of staple cereals that contain low zinc concentrations (Menguer et al., 2017). In barley, the endosperm and embryo generally constitute ~85-90% of the whole-grain biomass but only ~50% of the accumulated Zn, although these figures are considered genotype dependent (Menguer et al., 2017; Lombi et al., 2011).

The Consultative Group on International Agricultural Research (CGIAR) is developing biofortified crops through its *HarvestPlus* program together with an interdisciplinary alliance of research institutions, targeting incremental increases in micronutrients without compromising yield productivity or agronomic performance (Virk et al., 2021). In the context of Ethiopia, the grain Zn concentration breeding target is an increase of +10 mg/kg above a baseline of 28 mg/kg, which is based on the average grain Zn concentration of elite cultivars. The average daily *per capita* zinc requirement has been estimated as 8-11 mg for adult females and males respectively (Institute of Medicine, 2001).

As a 'Vavilovian centre of genetic diversity', Ethiopia has an extensive collection of barley genetic resources. The unique diversity of Ethiopian barley germplasm has long been recognised, from early studies using isozymes (Bekele, 1983) to more recent reports using gene-based SNP assays (Mamo et al., 2014; Dido et al., 2022; Yirgu et al., 2023; Caproni et al., 2023). Similarly, the range of phenotypic variation is well documented, often cited as unprecedented and unique to Ethiopia (Fantahun et al., 2023). Landraces have evolved with high levels of genetic variability, and these represent the largest component (44%) of germplasm in gene banks worldwide (van Hintum and Menting, 2003; Bockelman and Valkoun, 2010). In situ, these have been under continuous natural selection and human conservation for both desirable agronomic traits and phenotypic variants whose full expression is shaped by environmental influences (Mamo et al., 2014). Highly variable local landraces have been identified with large within-population diversity varying along altitudinal gradients (Hadado et al., 2009). This genetic diversity is generally determined by the type and number of genes involved, the type of trait assayed, and the influence of the environmental, amongst other factors. Several genes of small effect may be involved in expression of a quantitative trait (Yule, 2018).

Despite the importance of grain micronutrients, only one study of zinc and iron concentration in Ethiopian germplasm has been reported. Mamo et al. (2014) identified considerable variation for grain micronutrient concentration within a collection of 298 Ethiopian and Eritrean landraces that were grown in both field and glasshouse environments. Using a 9K SNP genotyping array developed for barley (Comadran et al., 2012) for a genome wide association scan (GWAS), they identified two quantitative trait loci (QTLs) on chromosome 6H associated with grain zinc concentration (Mamo et al., 2014). Here, we enhance the power of this study by increasing marker density through use of a 50K SNP array (Bayer et al., 2017). We used 296 accessions and ionomic data from field and glasshouse studies of Mamo et al. (2014) in a multilocus GWAS model which highlighted natural variation in a Zinc Induced Facilitator-Like 1 gene as influencing the zinc concentration of barley grain. Using a large global collection of diverse barley germplasm, we then developed a simple Kompetitive Allele Specific PCR (KASP) marker for use in marker assisted breeding.

Materials and methods

Phenotypic data and statistical analyses

This study was based on the Ethiopian and Eritrean Barley Collection (EEBC), a germplasm assembled by Mamo et al. (2014) from accessions held at USDA-ARS National Small Grains Collection, the International Centre for Agricultural Research in the Dry Areas (ICARDA) and N.I. Vavilov All-Russian Institute of Plant Genetic Resources (VIR). The EEBC was previously used to identify novel genes for disease resistance and other agronomic traits including grain Zn (GZC) and Fe (GFeC) concentrations. All genotypes were originally phenotyped in 2011 glasshouse and 2012 field trials in Saint Paul, Minnesota (Mamo et al., 2014; Mamo and Steffenson, 2015). The datasets were downloaded from the online

archive [T3/Barley (triticeaetoolbox.org)]. Before applying any genetic analysis, missing phenotypic values (5%) were imputed using the MICE (Multivariate Imputation via Chained Equations) package in R software [CRAN - Package mice (r-project.org)] with a maximum of 50 iterations and 5 replications. MICE assumes that the missing data are Missing at Random (MAR), which means that the probability that missing data are related to observed values and can be predicted using linear regression for continuous missing values and logistic regression for categorical missing values. MICE imputes data on a variable-by-variable basis by specifying an imputation model per variable. Data were then checked for a normal distribution and descriptive statistics were used to show the spread. Assumptions of ANOVA were subsequently checked while fitting the data and estimated Best Linear Unbiased Predictor (BLUP) and Best Linear Unbiased Estimate (BLUE) models were assessed when the assumptions were violated. Biological replicates of three seeds per plot were treated as environments to estimate BLUEs and BLUPs. The BLUEs are solutions (or estimates) associated with the fixed effects and BLUPs are the solutions (identified as predictions) associated with the random effects of a model. The BLUEs, BLUPs and correlations (genetic and phenotypic) between genotypes were analysed using META-R software (Alvarado et al., 2015; 2020). Descriptive statistics such as mean, minimum, maximum, and standard deviation were obtained using the 'stat.desc' function of 'pastecs' package (Grosjean et al., 2014) in the latest version of R software (R Core Team, 2020). Analysis of variance (ANOVA) was conducted using the 'aov' function, available in the stats package of R software. Normality and homogeneity of variance were checked by Shapiro-Wilk (Shapiro and Wilk, 1965) and Levene's test (Levene, 1960), respectively.

Genetic and population structure analyses

Genetic analyses were performed on the DNA extracted from 7-10 days old leaf samples using QIAamp 96 DNA kit (Qiagen GmbH) on an automated nucleic acid purification robot. DNA quality was assessed using a Nanodrop 2000 (Thermo Fisher Scientific, Waltham, MA, United States) with a requirement for 260/280 and 260/230 ratios > 1.8. DNA was then quantified using Picogreen (Thermo Fisher Scientific, Waltham, MA, United States). 300ng lyophilised DNA per sample was sent to Geneseek (Neogen Europe, Ltd., Auchincruive, United Kingdom) for genotyping using a high throughput Barley 50K SNP array (Bayer et al., 2017) by Illumina HTS processing and HiScan array imaging (Illumina, San Diego, CA, United States). R and Theta scores were extracted from resulting idat files using Genome Studio Genotyping Module v2.0.2 (Illumina, San Diego, CA, United States) and allele scores were created using paRsnps (an R package for clustering, visualising, and comparing Illumina SNP genotyping data). The data were visualised with, and exported from, Flapjack, a genotype visualisation software for further analysis (Milne et al., 2010). The genotypic data were assessed for missing, heterozygous, and monomorphic markers and markers with more than 10% missing data were removed. Polymorphic markers were maintained for cluster analysis, population structure and GWAS. For cluster analysis, the R package 'SelectionTools' [Index of/~software/(uni-giessen.de)] was used to read in marker data to construct Principal Components Analysis (PCA) for population structure. Euclidean distance was used to differentiate between genotypes on the PCA axes based on their genetic distance. The optimum number of clusters (K) was determined using the Elbow method which is inferred at the highest value beyond which the total within sum of squares (WSS) is no longer significantly reduced, and the gap statistic (GS) which is increasing at decreasing rate. Population structure analysis was also performed using STRUCTURE software (Pritchard et al., 2000) with a hypothetical number of sub-groups (K) ranging from 1-10 and 5000 burn-in iterations followed by Markov Chain Monte Carlo, MCMC) iterations of 5000. Five independent runs were performed for each value of K. Molecular variances (within and among population variances) were analysed in GenAlex software (Peakall and Smouse, 2012) using 10% of the markers (5,000) randomly selected from all chromosomes. The fixation index (Fst) among the EEBC panel, also known as Wright's population differentiation statistics (Wright, 1951) was analysed in R software.

GWAS approaches to identify associations with GZC

GWAS was initially run to compare with published results on grain colour and the QTL interval was set based on the extent of linkage disequilibrium (LD). Genome-wide and within chromosome LD (r²) decay were analysed in Trait Analysis by aSSociation, Evolution, and Linkage (TASSEL) software (Bradbury et al., 2007) with a sliding window of 500 markers around the significant marker to obtain the extent of inter-chromosomal LD decay. All possible candidate genes were identified based on an LD score of $r^2 \ge 0.2$ between the most significant marker associated with the trait and neighbouring SNPs. The most likely candidate genes from the list were highlighted according to their predicted functions. GWAS was performed using the grain Zn concentration data from Minnesota as described in 2.1 above and 36,431 polymorphic SNPs, and the result compared to the GWAS in the study by Mamo et al. (2014) that used 7,842 polymorphic markers. Results from several GWAS (single locus, SL_GWAS versus multilocus, ML-GWAS) models were compared to validate the observed results. A Q-Q plot produced by the best fit model was used to determine whether the GWAS model was robust and detected only the most significantly associated SNP markers with the highest observed versus expected [-log10P] values. Too many markers below the Q-Q proportional line means the model is overfitting the data with false negative results, while those above the line indicate false positive results. The first three principal components (PCs) were added as covariates in the model to reasonably capture and correct for the confounding effect of population structure. GWAS was performed with the GAPIT3 package in R software (Wang and Zhang, 2021) using the BLUE data for the phenotypic trait. The results were compared between the multi-locus BLINK and the

single-locus MLM models. QTL intervals were determined based on significant LD of squared correlation value $r^2 \ge 0.20$ and possible candidates were prioritised based on genes annotated with relevant functions in other species.

Characterisation and validation of associated SNPs

Genomic sequences of the HvZIFL1 candidate gene identified in the genome-wide association (GWA) analyses were downloaded from the genome assembly of Morex v1 (Mascher et al., 2017)(HORVU4Hr1G081570.1). Both version 2 (Monat et al., 2019) (HORVU.MOREX.r2.4HG0341120.1) and version 3 (Mascher et al., 2021) (HORVU.MOREX.r3.4H409580) were released during this study and used for subsequent genetic characterisation of the gene. The latest reference transcriptome database [Morex Gene Atlas - Hv_Mx_chr4HG27982 (hutton.ac.uk)] was used to identify potential transcripts and compared to EnsemblPlants which maintains 5 transcripts of HvZIFL1 (https://plants.ensembl.org/Hordeum_vulgare/Gene/ Summary?g=HORVU.MOREX.r3.4HG0409580;r=4H:586550476-586579926;t=HORVU.MOREX.r3.4HG0409580.1;db=core). A Kompetitive Allele Specific PCR (KASP) assay was designed (Supplementary Table 6) to the most significantly associated marker for technology independent validation across the EEBC germplasm as well as for other barley germplasm. Haplotype analysis was performed for 194 georeferenced EEBC landrace collections using 19 SNPs from the 50K SNP platform and a locus haplotype was also plotted using 80 polymorphic SNPs across the gene obtained from exome capture data of the European Wheat and Barley Legacy for Breeding Improvement (WHEALBI) project (Russell et al., 2016; Bustos-Korts et al., 2019).

Amino acid sequences of ZIFL1 paralogues were downloaded from *EnsemblPlants* and BLAST searched in the National Centre of Biotechnology institute (NCBI) database. The top 100 sequences were downloaded and used for multiple alignment in Jalview 2.11.2.1 software (Waterhouse et al., 2009) and sequences were uploaded in the *Multiple Expectation maximisation for Motif Elicitation (MEME)* online database [https://memesuite.org/ meme/tools/meme] to explore protein motifs and hence possible similar functions (Bailey and Elkan, 1994). Known conserved motifs and new signatures were discovered using detailed analysis in the latest version of Jalview (Waterhouse et al., 2009) and by multiple alignment using Muscle (Edgar, 2004).

Results

Marker dense GWAS directly reveals a highly plausible candidate gene associated with grain zinc concentration

Population structure is one of the confounding factors limiting efficiency of detecting quantitative trait loci in genome-wide association studies. It is hence crucial to characterise and determine the diversity and genetic structure existing within the germplasm panel. To this end, we used the same extensive EEBC panel assembled by University of Minnesota (Mamo et al., 2014) to determine the level of population structure. We genotyped the accessions using the barley 50K SNP assay (Bayer et al., 2017) and after filtering to remove failed, missing (>10% removed), monomorphic and low frequency (< 5%) markers, 36431 polymorphic gene-based and physically mapped SNPs remained. We observed four major clusters of the EEBC germplasm with 72% present in subpopulation 1 (SP1) with no partitioning by row type or by geography (Figures 1A, C). Molecular variance explained 69% and 31% (Figure 1B) of genetic diversity within and among subpopulations of the EEBC, and the first two principal components (PC1 and PC2) explained 82% and 7% of the genetic variation between the sub-populations, respectively. To examine diversity of the EEBC in a global context 50K SNP data was combined with the 1000 accession core barley collection in the German Federal Gene bank IPK Gatersleben which had been genotyped using the same platform (Darrier et al., 2019). Principal Component Analysis (PCA) uncovered the level of diversity and population structure with the first (PC1) and second (PC2) axes explaining 41% and 25% of the variation, respectively identifying four separate genotypic clusters (Supplementary Figure 1A). The Ethiopian (EEBC) germplasm was shown to form tight clusters with 238 lines (80%) of the 296 accessions forming a single main cluster (highlighted in green in Supplementary Figure 1A). Although 46 Ethiopian lines were among the core panel, only half grouped within the EEBC cluster. Molecular variance depicted 57% of the variation within and 43% between sub-populations (Supplementary Figure 1B).

The phenotypic data (Mamo et al., 2014) was analysed and descriptive statistics used to summarise grain Zn concentration of the EEBC panel in both Minnesota glasshouse and field trials (Supplementary Table 1). Grain Zn concentration ranged between 6-72 mg/kg in the glasshouse and 20-87 mg/kg in the field. Analysis of variance (ANOVA) depicted significant difference between genotypes (G), locations/environment (E) and genotype-bylocation/environment interaction [GEI] (Supplementary Table 2). Genotype, location/environment (E) and genotype-by-environment interaction (GEI) contributed 32%, 36%, and 21% to the total variance (% total sum of squares, %TSS) (Supplementary Table 2). In contrast, within-population analysis revealed no statistical difference among the sub-populations (Figure 1D) in the 2011 glasshouse trial with a mean GZC ranging from 28 mg/ kg for SP4 to 32 mg/kg for SP2. The 2012 field trial, however, depicted significantly higher GZC of 51 mg/kg for the SP1 compared to the lowest grain Zn concentration of 41 mg/kg for SP4. The two sub-populations SP2 and SP3 had statistically similar grain Zn concentration of 44 mg/kg and 45 mg/kg, respectively. The SP4 is comprised of 20 two-rowed elite cultivars or breeding lines including the UK standard check variety known as Concerto.

We then performed GWAS, accounting for the observed population structure, and compared several single locus (SL-GWAS) and multi-locus (ML-GWAS) models with the objective



of identifying consistent associations regardless of the statistical model. GWAS for grain zinc concentration (BLUE data, Mamo et al., 2014) from glasshouse and field trial was analysed using three different statistical models (Figures 2A, B; Supplementary Table 4; Supplementary Figures 2A, B). For the glasshouse data GWAS consistently revealed a significant association at the bottom of chromosome 4HL. From the field data, only BLINK detected a significant association at this locus (Supplementary Figure 2B).

Given grain Zn concentration is likely to be under complex genetic control, on weight of evidence we decided to investigate the terminal 4H region in more detail. High SNP marker density allowed the association boundaries to be defined accurately based on the drop in LD (Supplementary Figure 3) between the most highly associated and physically adjacent non-significant markers. The most highly associated SNP (JHI-Hv50-2016-265280) was located within a gene annotated as a member of the Major



Facilitator Superfamily (MFS) of transporters that includes Zinc Induced Facilitator (ZIF) and ZIF-Like (ZIFL) proteins. As no other annotated genes within the region appeared strong candidates, we prioritised HvZIFL1 as a strong causal gene candidate for more detailed analyses.

We extracted the complete *HvZIFL1* (HORVU4Hr1G081570.1) gene sequence from Morex reference genome assembly version 1 (Mascher et al., 2017) to investigate molecular polymorphism, and again when version 2 became available (Monat et al., 2019). The protein encoded by HvZIFL1 consists of 493 amino acid residues and its secondary structure was predicted to contain a typical MFS protein structure with 12 membrane spanning domains known as transmembrane helices (TMH) that facilitate substrate transport across the biological membrane (Supplementary Figure 4). We predicted that a mutation changing Glutamine to Histidine (Q330H) and an adjacent Valine to Phenylalanine (V331F) respectively) located at the gate of TMH8 (Figure 3B; Supplementary Figure 4) may be functionally disruptive (positive mutation because it increase grain GZC). Further examination of the 50K SNP data revealed that 19 polymorphic SNPs were located within HvZIFL1 and haplotype analysis revealed eight haplotypes (Table 1) across the 297 (296 EEBC plus Concerto) accessions with frequencies ranging from 0.007 (2 accessions) to 0.581 (172 accessions). In the combined EEBC and IPK dataset (Darrier et al., 2019), we observed 10 unique haplotypes with frequencies ranging from 0.003 to 0.536. The most significantly associated GWAS SNP was a T to G transversion that was unique to Haplotype 1 (Tables 1, 2) which had significantly higher grain zinc concentration (47.6 mg/kg) when compared to all other haplotypes (42.8 mg/kg for haplotype 5 to 34.1 mg/kg for haplotype 4).

Ethiopian and Eritrean barley germplasm possess unique variation in *HvZIFL1*

We designed 12 pairs of primers (Supplementary Table 5) to amplify the complete HvZIFL1 based on Morex version 1 and PCRsequenced the gene in twelve genotypes that represented high and low grain Zn concentration. A total of 95 SNP polymorphisms were identified, almost all of which are silent/synonymous changes occurring within non-coding regions (black inverted triangle marks, Figure 3A). The two exceptions were the diagnostic SNP marker JHI-Hv50k-2016-265280 identified by GWAS, and an adjacent SNP discovered by Sanger sequencing using primer pair 9 (Supplementary Table 5; indicated with branching lines in Figure 3). Sequencing and alignment also revealed a variable INDEL of either 162 bp or 153 bp in length within the noncoding region between exons six and seven (Figure 3B). This INDEL was confirmed using agarose gel electrophoresis and was present in 10 of the 12 Ethiopian barley sequences, with EEBC_005 having a 162 bp INDEL and EEBC_114 the same 153bp INDEL as the Morex reference annotation (Figure 3B). Assaying the INDEL across the EEBC identified 51 lines with the 153bp insertion identical to the reference cultivar Morex (Supplementary Table 8). The highly significant markers from GWAS were found in Exon 11 changing a CAA (Glutamine, Q) to CAC (Histidine, H) with the adjacent mutation changing GTT (Valine, V) to TTT (Phenylalanine, F). Somewhat surprisingly, the most significantly associated GWAS SNP was not in complete LD with the 153 bp deletion. Further investigation of the HvZIFL1 gene sequence revealed duplication of ~7 kilobase pairs with 15.5 kbp interspaced sequence (Figure 4).

Because the 50K SNP assay has been widely used by the barley research community as a platform for genetic analysis and marker



alignment of HvZIFL1 among 12 EEBC lines (top section). Gene model with blue blocks indicate exons interspaced with lines indicating introns and polymorphisms positions (middle section). The two important polymorphism sites are indicated with 153-162 bp INDEL sequence and the SNP changes at exon-11 with codons indicated in the table (bottom section). The last SNP of the first codon (CAA) is the wild type allele coding for the amino acid Glutamine (Q) while the mutation (CAC) causes amino acid change to Histidine (H). similarly, the first SNP of the second codon (GTT) is the wild type allele coding for Valine (V) while the mutation changes to Phenylalaine (F).

assisted selection (MAS), we were able to examine the distribution of the highly associated marker JHI-Hv50k-2016-265280 across a wide range of germplasm including 575 cultivars, 197 landraces and 258 wild barley accessions (Table 2). The frequency of the minor allele associated with increased grain zinc is 0.202 in the EEBC compared to 0.047 in cultivars and 0.004 in wild barley accessions. These frequencies observed in both the landrace and wild germplasm are below 0.05 which is usually chosen as the threshold for robust marker:trait detection in association studies and would therefore not have been significant in any of these germplasm sets. We also assembled haplotype sequences from the genetically diverse WHEALBI collection (N=432) (Bustos-Korts

TABLE 1 Sequence of haplotypes around the diagnostic marker of the EEBC accessions.

Haplotype	EEBC Frequency	EEBC + IPK frequency	Sequence
Haplotype 1	0.199	0.181	GAGGTTAAAGGTTGAGGCC
Haplotype 2	0.034	0.003	AGGTGAGGATGCCGAACCT
Haplotype 3	0.017	0.003	AAGGTAGAATGTCGGAGCT
Haplotype 4	0.128	0.163	AGATGAGGGTACCGGACTT
Haplotype 5	0.010	0.038	AGATGAGGATACCGGACTT
Haplotype 6	0.581	0.536	GAGGTTAAATGTTGAGGCC
Haplotype 7	0.007	0.038	GAGGTTAAATGTTTAGGCC
Haplotype 8	0.024	0.012	GAGGTTGAATGTTGAGGCC
Haplotype 9	-	0.015	AGGTGAGGATGCCGGACCT
Haplotype 10	-	0.012	AGGTGAGGGTACCGAACTT

The EEBC panel contained 297 accessions (296 EEBC lines + the standard line Concerto) and the IPK panel contained 46 collections from Ethiopia (Darrier et al., 2019).

TABLE 2	Frequency	of	associated	SNP	across	barley
germplası	m collectio	ns.				

Germplasm	Number of lines	JHI-Hv50k- 2016-265280	
		G*	т
Cultivar	575	0.047	0.953
Landraces	197	0.046	0.954
Wild barley	258	0.004	0.996
EEBC	292	0.202	0.798

*Denotes diagnostic SNP for high grain zinc concentration on the forward strand (C on the reverse strand).

et al., 2019) using 80 polymorphic SNPs discovered in exome capture data. We identified 38 unique haplotypes where the favourable JHI-Hv50k-2016-265280 'C' allele was found only in Haplotype 18 (Hap_18), comprised of 40 lines of which 30% were from Ethiopia and Eritrea (Supplementary Tables 7, 9, 10). Analysis of grain Zn concentration showed a significant effect associated with both the 153bp INDEL and the diagnostic marker, individually and in combination (Figures 5A-E). The boxplots and t-tests showed that genotypes with the 153 bp deletion in ZIFL1 had a mean grain Zn concentration of 31 mg/kg compared to 28 mg/kg for those without the deletion (t=2.71**, Figure 5A) using the Minnesota (glasshouse) data. The difference between these groups in the field was 8 mg/kg (t= 5.23**, Figure 5B). The JHI-Hv50k-2016-265280 SNP effect for the glasshouse and the field trials showed an 8 mg/kg (t= 5.94***) and 12 mg/kg (t= 7.55***) difference between alleles (Figure 5C). Variability in grain Zn concentration was also apparent in the data across the two locations/years (Figures 5D, E). In the 2011 glasshouse trial, deletion had an average effect of 1 mg/kg (tstatistics = 1.12 ns) while the SNP marker had a mean effect of 8 mg/ kg (t=5.67***, Figure 5D). In the 2012 field trial, deletion had an average effect of 5 mg/kg (t= 3.50***) while the SNP effect was 11 mg/kg (t= 6.53^{***}). Interestingly, the combined effect of the 153bp INDEL and the favourable allele increased grain Zn concentration level by 9 and 16 mg/kg in the glasshouse and field conditions respectively in Minnesota.

HvZIFL1 on 4H is the only one of the 34 annotated ZIFL genes associated with grain zinc concentration

A BLAST search of the amino acid sequence of the candidate HvZIFL1 gene identified 33 paralogues distributed across the barley genome (Supplementary Table 3). For most of these genes, SNPs were present on the 50K chip and almost all were polymorphic with minimum allele frequency (MAF) values ranging from 0.021 to 0.476. However, none of the SNPs were associated with grain zinc concentration except the HvZIFL1 on 4H with an MAF of 0.202. We then compared the HvZIFL1 protein sequence with its cereal orthologs and identified 30 short conserved sequence motifs that may be related to both conserved and unique functions of the protein. The motifs could be categorised into three major and three minor groups (Supplementary Figure 5). The first major group (G-I) contained HvZIFL1 and 10 other ZIFL genes and exhibited 11 conserved motifs. The second major group (G-II) consisted of 9 ZIFL proteins with 10 unique motifs while the third group (G-III) contained 9 ZIFL genes with 12 unique sequence motifs including motif 6 and 15 which were conserved in all groups. The three minor groups consisted of 1-2 ZIFL genes with four conserved motifs.

HvZIFL1 gene expression supports a role in micronutrient uptake

We used the cv. Morex EoRNA gene expression atlas to explore the temporal and spatial transcript abundance and post transcriptional variation of *HvZIFL1* (https://ics.hutton.ac.uk/morexgeneatlas/



index.html). HvZIFL1 (transcript Hv_Mx_chr4HG27982 in MorexGeneAtlas) (Guo et al., 2025) was detected as multiple alternative transcripts (Supplementary Figure 6) with all, including the most abundant, expressed exclusively in root tissues with no indication of expression in any above ground tissues, including the developing grain (Supplementary Figure 7). These observations are consistent with HvZIFL1 contributing to variation in total Zn uptake from soil. Whether subsequent transport to the developing grain is passive (e.g. via the transpiration stream) or active (involving other Zn transporters) could not be determined



FIGURE 5

Box plots and t-tests of grain zinc concentration associated with the diagnostic SNP marker and the INDEL (153-162 bp). (A, B) The INDEL effect on GZC of glasshouse and field trials at Minnesota in 2011 and 2012, respectively; (C) Marker effect on GZC of glasshouse and field trials at Minnesota; (D, E) Combined INDEL + marker effects on GZC of glasshouse and field trials at Minnesota;

but does not apparently involve HvZIFL1. Furthermore, there was little variation in the level of HvZIFL1 expression in roots when Zn (or other heavy metals) were added to the growth medium suggesting its expression is not responsive to Zn availability (Kintlová et al., 2017) but rather is genetically pre-determined in different genotypes.

Development of a Kompetitive Allele Specific PCR assays for breeder's toolbox

The diagnostic SNP marker (JHI-Hv50k-2016-265280) identified by GWAS was selected for validation as a tool for marker-assisted selection (MAS) using a KASP assay. Alternative alleles in all EEBC accessions were then classified using the KASP assay. 237 lines had the JHI-Hv50k-2016-265280 'A' allele associated with lower grain zinc concentration and 59 lines the JHI-Hv50k-2016-265280 'C' allele associated with higher grain zinc concentration (Figures 5A, B). Comparing the data from the KASP assay and the 153bp deletion PCR-based size marker, the JHI-Hv50k-2016-265280 'C' allele was only associated with the deletion whereas the 'A' allele was associated with and without the deletion in a frequency of 0.212 ('A' allele, no deletion) and 0.787 ('A' allele, with deletion). The impact of using both markers increased the differences of grain zinc concentration to 11 mg/kg and 16 mg/kg for glasshouse and field data, respectively (Figures 5C, D).

Discussion

Since its domestication over 10,000 years ago, barley landraces have been cultivated widely as a dependable crop in areas with diverse and extreme environmental conditions (Grando and Macpherson, 2005). These landraces have evolved across a wide geographical range and exhibit high genetic variability. Today, due largely to their geographical and morphological diversity, landraces comprise the largest proportion of germplasm (44%) in worldwide gene banks (van Hintum and Menting, 2003; Bockelman and Valkoun, 2010). In Ethiopia, Hadado et al. (2010) confirmed the high variability of local populations of barley landraces with large within-population diversity varying along an altitudinal gradient. This rich phenotypic diversity (Mamo et al., 2014) is the result of variation in both major genes and quantitative characters comprised of multiple genes of small effect (Yule, 2018). In a recent gene bank genomics study based on extensive Genotyping by Sequencing (GBS) data, Milner et al. (2019) observed that 5000 Ethiopian lines out of > 20,000 geographically diverse accessions held by the German Federal Gene bank at IPK-Gatersleben formed a distinct and unique genetic cluster. Using the barley 50K SNP array (Bayer et al., 2017), Darrier et al. (2019) genotyped 1,000 representatives of the core of this collection including 46 accessions from Ethiopia. The SNP data reflected the observations using GBS, with a similar outcome reported using exome capture data (Russell et al., 2016; Chen et al., 2022). Here, using the 50K SNP platform we genotypically and phenotypically characterised 296 Ethiopian and

Eritrean genotypes (the EEBC panel) to identify associations between molecular markers and phenotypic traits. The EEBC is somewhat unique as it is assembled from germplasm that has experienced an extended period of geographic isolation from the rest of the global barley population. Our genotypic data again show that the EEBC accessions are genetically closer to each other than to other global barley accessions (Milner et al., 2019). Genotype By Sequencing (GBS) data revealed three major clusters (K=3) in the world barley collection (Milner et al., 2019) described as two-row spring type, six-row mixed type, and Asian six-row types (Darrier et al., 2019). Our pooled analysis of the two datasets combined was enabled by use of the 50K SNP array and identified four clusters, with EEBC forming two tightly clustered groups (Van Leur and Gebre, 2003) emphasizing the uniqueness of the Ethiopian barley germplasm, and thus its potential as a source of potentially valuable alleles for trait improvement.

One such trait could be biofortification, and biofortification of food crops with enhanced mineral nutrient concentration has become a research priority as a means to combat micronutrient deficiencies in human diets (Cu et al., 2021). One strategy is to identify genetic markers associated with high grain nutrient concentration and then use marker assisted breeding to develop nutrient-rich varieties (Karakoy et al., 2012; Hefferon, 2019). Identifying linked genetic markers and genes using GWAS has been successful for grain Zn concentration in rice (Cu et al., 2021) and wheat (Velu et al., 2018), Na in barley (Houston et al., 2020), and other nutrients in wheat (Alomari et al., 2021). Here we focused our analyses on Zinc given it remains the most widespread micronutrient deficiency in agricultural lands around the world, causing significant yield reduction and loss of nutritional quality (Welch and Graham, 2004). Indeed, cereal-based diets with low Zn concentration are considered a major reason for widespread human Zn deficiency, especially in developing countries (Biesalski, 2013).

Cereal species and genotypes vary widely in their grain Zn concentration, and this is affected by both genetic and environmental factors. For example, soil pH affects nutrient availability for plant growth. A soil pH of between 5.5 and 6.5 is optimal for Zn availability for plant uptake and in acidic soils, increasing soil pH from 5.2 to 6.8 by liming resulted in an approximate 10-fold decrease in plant Zn concentration (Parker and Walker, 1986). In addition, Zinc availability is often reported to be low in soils with high organic matter content due to increased adsorption of Zn by organic ligands and components. Low soil moisture and low temperature also contribute to Zn deficiency (Moraghan and Mascagni, 1991). However, the physiological and molecular mechanisms of Zn deficiency tolerance are only just beginning to be understood (Hacisalihoglu and Kochian, 2003). The importance of transporters is reflected in reports that suggest about 7% of E. coli proteome, 2.5% of that of Arabidopsis thaliana, and 7% of the predicted genes in Vitis vinifera, are transporter proteins (Patil et al., 2019). Zn-efficient genotypes may produce high amounts of chelators that bind Zn in the rhizosphere and increase its physiological availability, while Zn transporters in the root help sustain Zn uptake during periods of Zn deficiency. Consequently, grain Zn concentration levels are influenced by genotype, environment and genotype-by-environment interactions (GEI).

Based on the data from two locations in Minnesota, these factors respectively contributed 32%, 36% and 21% of the total variation of GZC in the EEBC panel. Different levels of baseline grain Zn concentration have been established for commercial varieties of staple crops growing in different regions ranging from 16 mg/kg for rice to 25 mg/kg in maize and wheat with their corresponding breeding targets for improvement (Virk et al., 2021). The barley baseline grain Zn concentration in the EEBC has been set here to an average grain Zn concentration of 28 mg/kg and the breeding target of +10 mg/kg increase to develop elite barley varieties that meet the average daily human Zn requirement of 10 mg. Descriptive statistics showed within population difference in grain Zn concentration for the Minnesota 2012 field data but not for the 2011 glasshouse trial. For example, the average grain Zn concentration was significantly higher for SP1 (51 mg/kg) compared to the other three EEBC sub-populations.

Here we used GWAS, incorporating different statistical models to robustly map the association of phenotypic grain Zn concentration data in the EEBC to a specific location on barley chromosome 4H. We then used reference genomic information from the cultivar Morex (Mascher et al., 2017; 2021; Monat et al., 2019) to identify natural gene-level variations in HvZIFL1 arguing that this gene is likely involved in Zn homeostasis in the EEBC germplasm collection. Sequence characterisation revealed two important types of polymorphism: intronic deletions and point mutations. The majority of the latter were silent mutations. The INDEL and the highly associated SNP marker identified by GWAS together differentiated the EEBC germplasm into 20% high and 80% low Zn haplotypes. We also observed that the significance of the genetic association with these markers was environmentally sensitive. However, despite this, the combined effect of both SNP and INDEL markers resulted in an increase of 9 and 16 mg/kg in the glasshouse and field conditions at Minnesota, respectively.

Zn is a key structural component of many plant proteins including transcription factors and metalloenzymes (Figueiredo et al., 2012). Lira-Morales et al. (2019) reviewed several families of proteins implicated in Zn homeostasis in barley which include metal tolerance proteins (MTPs) that are involved in vacuolar Zn storage (Desbrosses-Fonrouge et al., 2005); heavy-metal ATPases (HMAs), which are involved in heavy metal tolerance and Zn storage in the vacuole (Morel et al., 2009); Natural Resistance Associated Macrophage Protein 4 (NRAMP4) which is involved in the root to shoot Zn translocation (Verret et al., 2004); the yellow stripe-like (YSL) family that are associated with Zn transport to reproductive tissues and leaves (Waters et al., 2006); and Plant Cadmium Resistance 2 (PCR2) which is associated with longdistance Zn transport (Song et al., 2010). Of particular relevance to the work we describe here, Haydon and Cobbett (2007) described a Zinc Induced Facilitator (ZIF1) gene in Arabidopsis thaliana along with two ZIF1-like (ZIFL1 and ZIFL2) genes. A detailed analysis of the ZIFL family of plant genes was given by Ricachenevsky et al. (2011).

Transport of Zn from the root to the shoot and its translocation within the plant has been reported to involve proteins belonging to different families (Sharma et al., 2019). ZIFL genes are members of the largest group of secondary active transporters (carrier proteins) that transport solutes against concentration gradients using active energy sources (Patil et al., 2019). They are members of the Major Facilitator Superfamily (MFS) of proteins that transport a broad range of molecules through three distinct kinetic mechanisms: uniport, antiport or symport. ZIFL family proteins described in Arabidopsis thaliana were expressed in the tonoplast (vacuolar membrane) and implicated in providing tolerance to excess Zn (Haydon and Cobbett, 2007). Our BLAST searches identified 33 paralogs of HvZIFL1 distributed across the barley genome and 345 homologous genes (or gene fragments) in other species including 13 in rice (Ricachenevsky et al., 2011) and 35 in wheat (Sharma et al., 2019). All belong to the MFS protein family where crystal structures suggest a distinctive protein constitution consisting of two similar (N-terminus and C-terminus) domains each containing multiple transmembrane-spanning α -helices (TMH domains) (Niño-González et al., 2019), along with an MFS_1 domain (Pfam ID: PF07690) and an antiporter domain (Meena et al., 2022). The latter surround a substrate translocation pore that operates via a rockerswitch movement of the two halves of the protein (Law et al., 2008). MFS proteins differentiate between a wide range of substrates and while we suggest HvZIFL1 is involved in Zn uptake, the functions of all 33 HvZIFL1 paralogs distributed across the barley genome are not yet elucidated. There is evidence in Arabidopsis that AtZIFL2 is involved in Potassium and Cesium homeostasis (Remy et al., 2015).

In summary, by applying a GWAS on the EEBC we identified highly significant genetic associations between barley grain Zn concentration and SNPs embedded in HvZIFL1 on the long arm of chromosome 4H. HvZIFL1 is exclusively expressed in barley root tissues, supporting its role in Zn uptake from the environment, and appears to be unaffected by the concentration of Zn or other tested heavy metals in the growth medium (Kintlová et al., 2017). Using barley reference and pan-genome assemblies (Jayakodi et al., 2020), we characterised DNA sequence polymorphisms in HvZIFL1 and found that an intronic deletion of 153 bp along with two adjacent point mutations in exon 11 were associated with variation in grain zinc concentration. We found that the favourable haplotype was present in approximately 20% of Ethiopian and Eritrean barley germplasm and the two types of polymorphism together were associated with higher grain Zn accumulation. As is common for multigenic traits, HvZIFL1 has a variable effect size in different growing environments. In a world barley collection, variant HvZIFL1 types were mainly found in East African germplasm while it is largely genetically fixed in global landraces and wild barley accessions. The unique and specific distribution of the desirable variant of HvZIFL1 clearly illustrates the value of using different gene pools for GWAS as the observed associations would not have been found in germplasm from outside of Ethiopia and Eritrea. We designed KASP markers to the polymorphic SNPs and validated them for potential use in marker-assisted selection in breeding programs inside and outside of their regional origin. At the genomic level, we identified 33 ZIFL paralogs distributed within

10.3389/fpls.2025.1539029

the barley genome whose functions are not yet elucidated. Based on the function of characterised ZIFL proteins it seems likely that they will be redundantly involved in functions associated with micronutrient transport from the environment to the root, from the root to the shoot and from the shoot to the grain. Their full functional characterisation remains largely to be described.

In conclusion, eliminating both overt and hidden hunger is top of the global food and nutrition security agenda. Zn deficiency affects one-third of the global population and is most widespread in regions with Zn-deficient soils. By exploring naturally occurring variation in micronutrient concentration, we took a genetic approach to develop genetic markers associated with grain Zinc concentration. We, therefore, propose the combined use of the simple deletion and mutation genetic markers we detected here, to be used as a fast-track approach to barley breeding for enhanced Zn concentration. This could significantly improve Zn nutrition of smallholder farmers in certain geographies where barley is a daily staple.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

JR: Data curation, Methodology, Writing – original draft, Writing – review & editing. GD: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. BS: Conceptualization, Data curation, Resources, Writing – original draft, Writing – review & editing. CH: Conceptualization, Funding acquisition, Project administration, Writing – original draft, Writing – review & editing. RW: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This research was funded by Global Challenges Research Fund (GCRF) of the UK government

References

through the Scottish Funding Council (SFC) to the University of Dundee and from the Rural and Environment Science and Analytical Services (RESAS) Division of the Scottish Government.

Acknowledgments

The authors thank GCRF for funding, University of Dundee and the James Hutton Institute for hosting and providing logistics for successful execution of the project. The authors would like to acknowledge Nicola McCallum, Richard Keith, Christopher Warden, and Mairi Waugh for their help in field trial and data collection; Malcolm Macaulay and Clare Macaulay for genotyping and sequencing; Craig Simpson, Micha Bayer, Kelly Houston, Abdella Barakate, and Jean-Noel Thauvin for their help in bioinformatics and data analysis.

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.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2025.1539029/ full#supplementary-material

Repository Network, V23). Available at: https://hdl.handle.net/11529/10201 (Accessed November 10, 2022).

Alvarado, G., Rodríguez, F. M., Pacheco, A., Burgueño, J., Crossa, J., Vargas, M., et al. (2020). META-R: A software to analyze data from multi-environment plant breeding trials. *Crop J.* 8, 745–756. doi: 10.1016/j.cj.2020.03.010

Bailey, T. L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 2, 28–36.

Alomari, D. Z., Alqudah, A. M., Pillen, K., von Wirén, N., and Röder, M. S. (2021). Toward identification of a putative candidate gene for nutrient mineral accumulation in wheat grains for human nutrition purposes. *J. Exp. Bot.* 72, 6305–6318. doi: 10.1093/ jxb/erab297

Alvarado, G., López, M., Vargas, M., Pacheco-Gil, R.Á., Rodríguez, F., Burgueño, J., et al. (2015). "META-R (Multi Environment Trail Analysis with R for Windows) Version 6.04" (Carretera México-Veracruz: CIMMYT Research Data & Software

Bayer, M. M., Rapazote-Flores, P., Ganal, M., Hedley, P. E., Macaulay, M., Plieske, J., et al. (2017). Development and evaluation of a barley 50k iSelect SNP array. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.01792

Bekele, E. (1983). The neutralist-selectionist debate and estimates of allozyme multilocus structure in conservation genetics of the primitive land races of Ethiopian barley. *Hereditas* 99, 73–88. doi: 10.1111/j.1601-5223.1983.tb00734.x

Biesalski, H. K. (2013). *Hidden hunger*. Hidden hunger (Berlin, Heidelberg: Springer), 25-50.

Bockelman, H. E., and Valkoun, J. (2010). "Barley germplasm conservation and resources," in *Barley: production, improvement, and uses.* Ed. S. E. Ullrich (Wiley-Blackwell, North America), 144–159.

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308

Bustos-Korts, D., Dawson, I. K., Russell, J., Tondelli, A., Guerra, D., Ferrandi, C., et al. (2019). Exome sequences and multi-environment field trials elucidate the genetic basis of adaptation in barley. *Plant J.* 99, 1172–1191. doi: 10.1111/tpj.14414

Caproni, L., Lakew, B. F., Kassaw, S. A., Miculan, M., Ahmed, J. S., Grazioli, S., et al. (2023). The genomic and bioclimatic characterization of Ethiopian barley (Hordeum vulgare L.) unveils challenges and opportunities to adapt to a changing climate. *Global Change Biol.* 29, 2335–2350. doi: 10.1111/gcb.16560

Chen, Y.-Y., Schreiber, M., Bayer, M. M., Dawson, I. K., Hedley, P. E., Lei, L., et al. (2022). The evolutionary patterns of barley pericentromeric chromosome regions, as shaped by linkage disequilibrium and domestication. *Plant J.* 111, 1580–1594. doi: 10.1111/tpj.15908

Comadran, J., Kilian, B., Russell, J., Ramsay, L., Stein, N., Ganal, M., et al. (2012). Natural variation in a homolog of *Antirrhinum CENTRORADIALIS* contributed to spring growth habit and environmental adaptation in cultivated barley. *Nat. Genet.* 44, 1388–1392). doi: 10.1038/ng.2447

Cu, S. T., Warnock, N. I., Pasuquin, J., Dingkuhn, M., and Stangoulis, J. (2021). A highresolution genome-wide association study of the grain ionome and agronomic traits in rice Oryza sativa subsp. indica. *Sci. Rep.* 11, 1–17. doi: 10.1038/s41598-021-98573-w

Darrier, B., Russell, J., Milner, S. G., Hedley, P. E., Shaw, P. D., Macaulay, M., et al. (2019). A comparison of mainstream genotyping platforms for the evaluation and use of barley genetic resources. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00544

De Groote, H., Tessema, M., Gameda, S., and Gunaratna, N. S. (2021). Soil zinc, serum zinc, and the potential for agronomic biofortification to reduce human zinc deficiency in Ethiopia. *Sci. Rep.* 11, 8770. doi: 10.1038/s41598-021-88304-6

Desbrosses-Fonrouge, A. G., Voigt, K., Schröder, A., Arrivault, S., Thomine, S., and Krämer, U. (2005). Arabidopsis thaliana MTP1 is a Zn transporter in the vacuolar membrane which mediates Zn detoxification and drives leaf Zn accumulation. *FEBS Lett.* 579, 4165–4174. doi: 10.1016/j.febslet.2005.06.046

Devate, N. B., Krishna, H., Sunilkumar, V. P., Manjunath, K. K., Mishra, C. N., Jain, N., et al. (2022). Identification of genomic regions of wheat associated with grain Fe and Zn concentration under drought and heat stress using genome-wide association study. *Front. Genet.* 13. doi: 10.3389/fgene.2022.1034947

Dido, A. A., Krishna, M. S. R., Assefa, E., Degefu, D. T., Singh, B. J. K., and Tesfaye, K. (2022). Genetic diversity, population structure and relationship of Ethiopian barley (Hordeum vulgare L.) landraces as revealed by SSR markers. *J. Genet.* 101, 9. doi: 10.1007/s12041-021-01346-7

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Fantahun, B., Woldesemayate, T., and Shiferaw, E. (2023). The association between hordein polypeptide banding and agronomic traits in partitioning genetic diversity in six-rowed Ethiopian barley lines (*Hordeum vulgare* L.). *BMC Plant Biol.* 23, 102. doi: 10.1186/s12870-023-04117-x

Figueiredo, D. D., Barros, P. M., Cordeiro, A. M., Serra, T. S., Lourenço, T., Chander, S., et al. (2012). Seven zinc-finger transcription factors are novel regulators of the stress responsive gene OsDREB1B. *J. Exp. Bot.* 63, 3643–3656. doi: 10.1093/jxb/ers035

Grando, S., and Macpherson, H. G. (2005). Food barley: importance, uses and local knowledge (Aleppo, Syria: ICARDA).

Grosjean, P., Ibanez, F., and Etienne, M. (2014). Pastecs: Package for Analysis of Space-Time Ecological Series. *R Package* Version 1.3-18. Available online at: http://CRAN.R-project.org/package=pastecs. (Accessed June 3, 2022)

Guo, W., Schreiber, M., Marosi, V. B., Bagnaresi, P., Jørgensen, M. E., Braune, K. B., et al. (2025). A barley pan-transcriptome reveals layers of genotype-dependent transcriptional complexity. *Nat. Genet.* 57, 441–450. doi: 10.1038/s41588-024-02069-y

Hacisalihoglu, G., and Kochian, L. V. (2003). How do some plants tolerate low levels of soil zinc? Mechanisms of zinc efficiency in crop plants. *New Phyt.* 159, 341–350. doi: 10.1046/j.1469-8137.2003.00826.x

Hadado, T. T., Rau, D., Bitocchi, E., and Papa, R. (2009). Genetic diversity of barley (Hordeum vulgare L.) landraces from the central highlands of Ethiopia: comparison between the Belg and Meher growing seasons using morphological traits. *Genet. Res. Crop Evol.* 56, 1131–1148. doi: 10.1007/s10722-009-9437-z

Hadado, T. T., Rau, D., Bitocchi, E., and Papa, R. (2010). Adaptation and diversity along an altitudinal gradient in Ethiopian barley (Hordeum vulgare L.) landraces

revealed by molecular analysis. BMC Plant Biol. 10, 121. doi: 10.1186/1471-2229-10-121

Haydon, M. J., and Cobbett, C. S. (2007). A novel major facilitator superfamily protein at the tonoplast influences zinc tolerance and accumulation in Arabidopsis. *Plant Phys.* 143, 1705–1719. doi: 10.1104/pp.106.092015

Hefferon, K. (2019). Biotechnological approaches for generating zinc-enriched crops to combat malnutrition. *Nutrients* 11, 253. doi: 10.3390/nu11020253

Houston, K., Qiu, J., Wege, S., Hrmova, M., Oakey, H., Qu, Y., et al. (2020). Barley sodium concentration is regulated by natural variants of the Na+ transporter HvHKT1;5. *Commun. Biol.* 3, 258. doi: 10.1038/s42003-020-0990-5

Institute of Medicine (2001). "Dietary reference intakes for vitamin A, vitamin K, arsenic, boron, chromium, copper, iodine, iron, manganese, molybdenum, nickel, silicon, vanadium, and zinc," in *Food and Nutrition Board* (National Academy Press, Washington, DC).

Jayakodi, M., Padmarasu, S., Haberer, G., Bonthala, V. S., Gundlach, H., Monat, C., et al. (2020). The barley pan-genome reveals the hidden legacy of mutation breeding. *Nature* 588, 284–289. doi: 10.1038/s41586-020-2947-8

Karakoy, T., Erdem, H., Baloch, F. S., Toklu, F., Eker, S., Kilian, B., et al. (2012). Diversity of macro- and micronutrients in the seeds of lentil landraces. *Sci. World J.* 2012, 710412. doi: 10.1100/2012/710412

Kintlová, M., Blavet, N., Cegan, R., and Hobza, R. (2017). Transcriptome of barley under three different heavy metal stress reaction. *Genomics Data* 13, 15–17. doi: 10.1016/j.jhazmat.2022.128430

Law, C. J., Maloney, P. C., and Wang, D. N. (2008). Ins and outs of major facilitator superfamily antiporters. *Annu. Rev. Microbiol.* 62, 289–305. doi: 10.1146/annurev.micro.61.080706.093329

Levene, H. (1960). "Robust tests for equality of variances," in *Contributions to Probability and Statistics*. Ed. I. Olkin (Stanford University Press, California), 278–292.

Lira-Morales, J. D., Varela-Bojórquez, N., Montoya-Rojo, M. B., and Sañudo-Barajas, J. A. (2019). The role of ZIP proteins in zinc assimilation and distribution in plants: current challenges. *Czech. J. Genet. Plant Breed.* 55, 45–54. doi: 10.17221/54/2018-CJGPB

Lombi, E., Smith, E., Hansen, T. H., Paterson, D., de Jonge, M. D., Howard, D. L., et al. (2011). Megapixel imaging of (micro)nutrients in mature barley grains. *J. Exp. Bot.* 62, 273–282. doi: 10.1093/jxb/erq270

Mamo, B. E., Barber, B. L., and Steffenson, B. J. (2014). Genome-wide association mapping of zinc and iron concentration in barley landraces from Ethiopia and Eritrea. *J. Cereal Sci.* 60, 497–506. doi: 10.1016/j.jcs.2014.08.007

Mamo, B. E., and Steffenson, B. J. (2015). Genome-wide association mapping of fusarium head blight resistance and agromorphological traits in barley landraces from Ethiopia and Eritrea. *Crop Sci.* 55, 1494–1512. doi: 10.2135/cropsci2014.06.0428

Mascher, M., Gundlach, H., Himmelbach, A., Beier, S., Twardziok, S. O., Wicker, T., et al. (2017). A chromosome conformation capture ordered sequence of the barley genome. *Nature* 544, 427–433. doi: 10.1038/nature22043

Mascher, M., Wicker, T., Jenkins, J., Plott, C., Lux, T., Koh, C. S., et al. (2021). Longread sequence assembly: a technical evaluation in barley. *Plant Cell.* 33, 1888–1906. doi: 10.1093/plcell/koab077

Meena, V., Sharma, S., Kaur, G., Singh, B., and Pandey, A. K. (2022). Diverse functions of plant zinc-induced facilitator-like transporter for their emerging roles in crop trait enhancement. *Plants* 11, 102. doi: 10.3390/plants11010102

Menguer, P. K., Vincent, T., Miller, A. J., Brown, J. K. M., Vincze, E., Borg, S., et al. (2017). Improving zinc accumulation in cereal endosperm using HvMTP1, a transition metal transporter. *Plant Biotechnol. J.* 16, 63–71. doi: 10.1111/pbi.12749

Milne, I., Shaw, P., Stephen, G., Bayer, M., Cardle, L., Thomas, W. T., et al. (2010). Flapjack—graphical genotype visualization. *Bioinformatics* 26, 3133-3134. doi: 10.1093/bioinformatics/btq580

Milner, S. G., Jost, M., Taketa, S., Mazón, E. R., Himmelbach, A., Oppermann, M., et al. (2019). Genebank genomics highlights the diversity of a global barley collection. *Nat. Genet.* 51, 319–326. doi: 10.1038/s41588-018-0266-x

Monat, C., Padmarasu, S., Lux, T., Wicker, T., Gundlach, H., Himmelbach, A., et al. (2019). TRITEX: chromosome-scale sequence assembly of Triticeae genomes with open-source tools. *Genome Biol.* 20, 284. doi: 10.1186/s13059-019-1899-5

Moraghan, J. T., and Mascagni, H. J. Jr. (1991). "Environmental and soil factors affecting micronutrient deficiencies and toxicities," in *Micron. Agric.* Eds. J. J. Mordvedt, F. R. Cox, L. M. Shuman and R. M. Welch (Soil Science Society of America, Madison), 371–425.

Morel, M., Crouzet, J., Gravot, A., Auroy, P., Leonhardt, N., Vavasseur, A., et al. (2009). AtHMA3, a P_{1B}-ATPase allowing cd/zn/co/pb vacuolar storage in arabidopsis. *Plant Phys.* 149, 894–904. doi: 10.1104/pp.108.130294

Niño-González, M., Novo-Uzal, E., Richardson, D. N., Barros, P. M., Duque,, and More, P. (2019). Transporters, more substrates: the arabidopsis major facilitator superfamily revisited. *Mol. Plant* 12, 1182–1202. 1674-2052. doi: 10.1016/j.molp.2019.07.003

Parker, M. B., and Walker, M. E. (1986). Soil pH and manganese effects on manganese nutrition of peanut 1. Agron. J. 78, 614-620. doi: 10.2134/agronj1986.00021962007800040011x

Patil, S. S., Prashant, R., Kadoo, N. Y., Upadhyay, A., and Gupta, V. S. (2019). Global study of MFS superfamily transporters in Arabidopsis and grapes reveals their functional diversity in plants. *Plant Gene.* 18, 100179. doi: 10.1016/j.plgene.2019.100179

Peakall, R., and Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* 28, 2537–2539. doi: 10.1093/bioinformatics/bts460

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multi-locus genotype data. *Genet* 155, 945–959. doi: 10.1093/genetics/ 155.2.945

R Core Team (2020). R: A language and Environment for Statistical Computing (Vienna: R Foundation for Statistical Computing). Available at: https://www.r-project. org. (Accessed June 3, 2022)

Remy, E., Cabrito, T. R., Batista, R. A., Teixeira, M. C., Sá-Correia, I., and Duque, P. (2015). The major facilitator superfamily transporter ZIFL2 modulates cesium and potassium homeostasis in arabidopsis. *Plant Cell Phys.* 56, 1, 148–1, 162. doi: 10.1093/ pcp/pcu157

Ricachenevsky, F. K., Sperotto, R. A., Menguer, P. K., Sperb, E. R., Lopes, K. L., and Fett, J. P. (2011). ZINCINDUCED FACILITATOR-LIKE family in plants: lineagespecific expansion in monocotyledons and conserved genomic and expression features among rice (*Oryza sativa*) paralogs. *BMC Plant Biol.* 11, 20. doi: 10.1186/1471-2229-11-20

Russell, J., Mascher, M., Dawson, I. K., Kyriakidis, S., Calixto, C., Freund, F., et al. (2016). Exome sequencing of geographically diverse barley landraces and wild relatives gives insights into environmental adaptation. *Nat. Genet.* 48, 1024–1030. doi: 10.1038/ng.3612

Shapiro, S. S., and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika* 52, 591–611. doi: 10.1093/biomet/52.3-4.591

Sharma, S., Kaur, G., Kumar, A., Meena, V., Kaur, J., and Pandey, A. K. (2019). Overlapping transcriptional expression response of wheat zinc-induced facilitator-like transporters emphasize important role during Fe and Zn stress. *BMC Mol. Biol.* 20, 22. doi: 10.1186/s12867-019-0139-6

Song, W., Choi, K. S., Kim, D. Y., Geisler, M., Park, J., Vincenzetti, V., et al. (2010). *Arabidopsis* PCR2 is a Zinc exporter involved in both Zinc extrusion and long-distance Zinc transport. *Plant Cell* 22, 2237–2252. doi: 10.1105/tpc.109.070185

Stangoulis, J. C. R., and Knez, M. (2022). Biofortification of major crop plants with iron and zinc - achievements and future directions. *Plant Soil* 474, 57–76. doi: 10.1007/s11104-022-05330-7

van Hintum, T., and Menting, F. (2003). Diversity in *ex situ* gene bank collections of barley. *Dev. Plant Genet. Breed.* 7, 247–257. doi: 10.1016/S0168-7972(03)80014-2

Van Leur, J. A. G., and Gebre, H. (2003). Diversity between some Ethiopian farmer's varieties of barley and within these varieties among seed sources. *Genet. Res. Crop Evol.* 50, 351–357. doi: 10.1023/A:1023966702389

Velu, G., Singh, R. P., Crespo-Herrera, L., Juliana, P., Dreisigacker, S., Valluru, R., et al. (2018). Genetic dissection of grain zinc concentration in spring wheat for mainstreaming biofortification in CIMMYT wheat breeding. *Sci. Rep.* 8, 1–10. doi: 10.1038/s41598-018-31951-z

Verret, F., Gravot, A., Auroy, P., Leonhardt, N., David, P., Nussaume, L., et al. (2004). Overexpression of AtHMA4 enhances root-to-shoot translocation of zinc and cadmium and plant metal tolerance. *FEBS Lett.* 576, 306–312. doi: 10.1016/ j.febslet.2004.09.023

Virk, P. S., Andersson, M. S., Arcos, J., Govindaraj, M., and Pfeiffer, W. H. (2021). Transition from targeted breeding to mainstreaming of biofortification traits in crop improvement programs. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.703990

Wang, J., and Zhang, Z. (2021). GAPIT Version 3: boosting power and accuracy for genomic association and prediction. *Genom. Proteome. Bioinform.* 19, 629–640. doi: 10.1016/j.gpb.2021.08.005

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., and Barton, G. J. (2009). Jalview Version 2 - A multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191. doi: 10.1093/bioinformatics/btp033

Waters, B. M., Chu, H. H., DiDonato, R. J., Roberts, L. A., Eisley, R. B., Lahner, B., et al. (2006). Mutations in Arabidopsis yellow stripe-like1 and yellow stripe-like3 reveal their roles in metal ion homeostasis and loading of metal ions in seeds. *Plant Phys.* 141, 1446–1458. doi: 10.1104/pp.106.082586

Welch, R. M., and Graham, R. D. (2004). Breeding for micronutrients in staple food crops from a human nutrition perspective. J. Exp. Bot. 55, 353-364. doi: 10.1093/jxb/erh064

Wright, S. (1951). The genetical structure of populations. Ann. Eugen. 15, 323–354. doi: 10.1111/j.1469-1809.1949.tb02451.x

Yirgu, M., Kebede, M., Feyissa, T., Lakew, B., Woldeyohannes, A. B., and Fikere, M. (2023). Single nucleotide polymorphism (SNP) markers for genetic diversity and population structure study in Ethiopian barley (*Hordeum vulgare L.*) germplasm. *BMC Genom. Data.* 24, 7. doi: 10.1186/s12863-023-01109-6

Yule, G. U. (2018). Statistics of production and the census of production Act, (1906). J. R. Stat. Soc. 70, 52–87. doi: 10.1111/j.2397-2335.1907.tb02392.x