Theoretical and Applied Genetics (2025) 138:111
https://doi.org/10.1007/s00122-025-04892-1

ORIGINAL ARTICLE

Genetic structure of Ethiopian finger millet landraces and genome-wide association mapping for agronomic and nutritional traits

Adane Gebreyohannes[1,2] · Hussein Shimelis[1] · Davis M. Gimode[3,8] · Prasad Grandham[4,9] · Vinod Kumar Valluri[4] · Habte Nida[2,5] · Susan M. Moenga[3] · Chris O. Ojiewo[6] · Benjamin Kilian[7] · Damaris A. Odeny[3]

Received: 30 May 2024 / Accepted: 26 March 2025
© The Author(s) 2025

Abstract

Finger millet (*Eleusine coracana* subsp. *coracana*) (2n = 4x = 36) remains one of the most important millets in East Africa (EA), where it was most likely domesticated along the highlands of Ethiopia and Uganda. The goal of the current study was to understand the population structure of the Ethiopian finger millet landraces and identify quantitative trait nucleotides (QTNs) and haplotypes associated with agronomic and nutritional traits. In a field evaluation across three environments, 448 genotypes were assessed for days to flowering (DTF), days to maturity (DTM), thousand seed weight (TSW), grain yield (GY), stay-green score (STG), and drought score (DrtSc). The harvested grain was analyzed for Fe and Zn contents. A subset of 391 genotypes was skim-sequenced, generating 24,112 high-quality SNPs that were employed for population structure, association mapping, and haplotype analysis. Seventy marker-trait associations were detected including 15 major QTNs with more than 30% phenotypic variance explained (PVE) for all traits except STG and GY. Pleiotropic major QTNs were identified for DTM/DTF and Fe/Zn on chromosomes 9B and 2B, respectively. Haplotype analysis of major QTNs identified 54 significant haplotype blocks and 2 additional haplotypes for a *multidrug ABC transporter gene family like* protein on chromosome 4A that was associated with PTH. Favorable haplotypes from pleiotropic DTM/DTF and Fe/Zn QTNs were present in 13 and 12 genotypes respectively, majority from Tigray region. Two genotypes from Tigray and one from Amhara harbored favorable haplotypes for DTM/DTF and Fe/Zn. These findings provide invaluable insights for targeted breeding to enhance finger millet resilience, nutritional profile, and yield.

Introduction

Finger millet {2n = 4x = 36; *Eleusine coracana* (L.) Gaertn} is an important cereal crop cultivated on ~ 2.1 m ha with an annual production of ~ 3.7 m tons globally (Indiastat 2020; FAOSTAT 2021). It represents 20% of the

Communicated by Hai-Chun Jing.

✉ Hussein Shimelis
shimelish@ukzn.ac.za

✉ Damaris A. Odeny
damaris.odeny@icrisat.org

1 African Centre for Crop Improvement, School of Agricultural, Earth and Environmental Sciences, University of KwaZulu-Natal, Scottsville 3209, South Africa

2 Ethiopian Institute of Agricultural Research, Melkassa Agricultural Research Center, P.O. Box 436, Adama, Ethiopia

3 International Crops Research Institute for the Semi-Arid Tropics, P.O Box 39063, Nairobi 00623, Kenya

4 Center of Excellence in Genomics and Systems Biology, ICRISAT – Patancheru, Hyderabad 502324, India

5 Department of Agronomy, Purdue University, West Lafayette, IN 47907, USA

6 International Maize and Wheat Improvement Center (CIMMYT), ICRAF House United Nations Avenue, Gigiri, Village Market, P.O. Box 1041, Nairobi 00621, Kenya

7 Global Crop Diversity Trust, Bonn, Germany

8 CGIAR Standing Panel on Impact Assessment, Alliance of Bioversity International and CIAT, Cali, Colombia

9 School of Plant, Environmental and Soil Sciences, Louisiana State University Agricultural Center, Baton Rouge, LA, USA

Published online: 08 May 2025

global millet area and 26% of the total global millet production, respectively (Gebreyohannes et al. 2024). It is the most important millet in East Africa, where it is believed to have been domesticated along the sub-humid highlands (Hilu et al. 1979; Tsehaye et al. 2006; Tesfaye and Mengistu 2017). Its adaptability to extreme weather conditions, extended storage quality, rich nutritional profile, high seed multiplication rate, and diverse health benefits make finger millet a desirable climate resilient crop for food and nutrition security (Ceasar et al. 2018; Kudapa et al. 2023; Gebreyohannes et al. 2024).

Finger millet exhibits outstanding adaptability due to its C4 photosynthetic pathway. This metabolic advantage, as documented by Yogeesh et al. (2016), Hittalmani et al. (2017), Ceasar et al. (2018), and Kudapa et al. (2023), allows finger millet to thrive in diverse and often challenging climates characterized by high temperatures, intense heat, and limited moisture availability. Beyond its climatic resilience, finger millet demonstrates additional stress tolerance. Studies by Hatti et al. (2018) report its ability to perform well in depleted soils, while Krishnamurthy et al. (2014) highlight its salinity tolerance. This inherent adaptability is further emphasized by Dida and Devos (2006), who describe finger millet's success in various agro-ecological zones, performing well under low rainfall in drylands and on a wide range of tropical soils.

Finger millet boasts an exceptional nutritional profile, which is particularly rich in a range of nutrients such as calcium (50–589 mg/100 g), potassium (430–669 mg/100 g), iron (3–53 mg/100 g), phosphorus (130–250 mg/100 g), sulfur (106–149 mg/100 g), sodium (6–42 mg/100 g), and magnesium (78–201 mg/100 g) (Backiyalakshmi et al. 2023; Teklu et al. 2024). This impressive array of nutrients, along with rich sources for essential amino acids, dietary fiber (18%), and phenolic compounds (0.3–3%), has earned finger millet the title as a"superfood"(Devi et al. 2014; Krishnan 2024). The grain is gluten-free and has extended storage quality due to the polyphenol content (Xiang et al. 2019; Odeny et al. 2020; Gebreyohannes et al. 2021), thus an important cereal for famine reserves. After grain harvest, the finger millet stover is an important source of fodder for livestock, enhancing feed security in areas that rely on crop and livestock farming. Micronutrient deficiencies, often referred to as the hidden hunger, affect over 2 billion people worldwide, primarily in low-income developing countries (Gupta et al. 2017; Sheoran et al. 2022; Teklu et al. 2023a). Micronutrient deficiencies are particularly severe in sub-Saharan African countries, where many populations lack access to a diverse diet. A recent study by Gebremedhin et al. (2021) reported that in Ethiopia, 37.3% of children aged 6–23 months had not received any recommended micronutrient sources. Therefore, finger millet is a healthy alternative for managing nutrition deficiencies, especially when the affected families adopt and cultivate high micronutrient-containing varieties.

While finger millet boasts climate resilience and impressive nutritional value, its production is severely hampered by several biotic and abiotic challenges. Yield losses exceeding 50% have been reported due to drought stress, impacting both grain quality and key yield components (Maqsood and Ali 2007). Furthermore, the genetic control of the important agronomic traits is not fully understood. Variability for the desirable nutritional traits and how they are affected by different soils and environments are equally not well studied. There have been hypotheses suggesting that the micronutrient content may be negatively correlated with yield (Upadhyaya et al. 2010; Backiyalakshmi et al. 2023) but there are no convincing and reproducible experiments reporting the same. The current study employed 448 diverse finger millet genotypes to enhance the understanding of genetic factors controlling agronomic and nutritional traits. The genotypes were evaluated in three different environments for both agronomic and nutritional traits. High throughput genotypic data was generated through skim-sequencing of 391 genotypes and a subsequent comprehensive genome-wide association study (GWAS) of all traits undertaken using different mixed models. The GWAS identified significant high-effect quantitative trait nucleotides (QTNs) for agronomic, drought, and tolerance related traits and Fe and Zn contents.

Materials and methods

Description of the experimental sites

Three different sites were used in the study; Arsinegelle (E1), Maitsebri (E2), and Meiso (E3) in Ethiopia (Supplementary Fig. S1). E1 is characterized by loam soils, slightly acidic pH, higher total organic carbon and nitrogen content, moderate clay content, and relatively high annual rainfall. E2 experiences moderate annual rainfall. E3 is characterized by clay loam soil, neutral pH, lower total organic carbon and nitrogen content, highest clay content, and lowest annual rainfall (Supplementary Table S1). The study areas varied in altitudes, which spanned from 1323 to 1933 m above sea level. Before planting, soil samples were taken across the three sites from various depths ranging from 0 to 60 cm. The soil samples were thoroughly mixed and transported to Melkassa Agricultural Research Center (MARC, Adama, Ethiopia) for physicochemical analysis employing the Estefan et al. (2013) method. Nutritional composition analysis was also carried out at MARC, at the food and nutrition laboratory immediately following harvest.

Plant material

A comprehensive collection of 448 finger millet accessions, including 425 diverse landraces and 23 improved varieties was assembled for the present study (Supplementary Table S2). The germplasm was sourced from various regions and research centers across Ethiopia, representing the country's rich diversity in finger millet (Supplementary Fig. S1). Tigray contributed the highest number of germplasm collection (168 entries), followed closely by Amhara (145 entries) and Oromia (86 entries), the Southern Nation Nationalities People Region (SNNPR; 17 entries) and Benishangul and Gumuz (9 entries) (Supplementary Table S2). The study additionally included 23 released crop varieties developed by the Ethiopian regional and national research institutes, encompassing research centers, namely, Bako, Melkassa, Pawe, Adet, Jinka, Shire-Maitsebri, and Axum (Supplementary Table S2).

Experimental design, planting, and managing the field experiments

A row-column design with two replicates was employed at each location with each replicate consisting of 64 rows and 7 columns. Each plot consisted of three meters of a single row with a row spacing of 0.4 m and a plant-to-plant spacing of 10 cm. To minimize lodging risk, only Di-ammonium phosphate (DAP) fertilizer was applied during sowing at the rate of 100 kg/ha. No additional nitrogen fertilizer (urea), herbicide, and insecticide were applied during the subsequent growth stages. Weeding was performed manually.

Phenotyping for agro-morphological traits

Agro-morphological and drought related data were collected following the procedures already described in IBPGR (1985) with minor modifications. Plant height (PTH; cm) was measured from ground level to the tip of the inflorescence (ear) at the dough stage on five randomly selected plants within each plot. Days to 50% flowering (DTF) was recorded as the number of days from sowing to when half of the main tillers in a plot had emerged ears. Similarly, days to maturity (DTM) were recorded as the number of days from sowing to the day when the main tillers in a plot had fully mature ears. Thousand seed weight (TSW) was determined by drying 1000 randomly chosen seeds to a standard moisture content of 12.5% and recording their weight in grams (g). Grain yield (GY; t ha^{-1}) was measured as the total grain weight of a plot (1.2 m^2) harvested at 12.5% moisture content and then converted to t ha^{-1}. Stay-green score (STG) was rated using a visual score on a scale of 1 to 5 per plot basis representing plant canopy greenness at maturity, with 1 indicating fully stay green (drought tolerant) (Supplementary Fig. S2), and 5 indicating completely senesced (drought sensitive). Finally, drought score (DrtSc) was assessed using a visual score on a scale of 1–5 per plot basis, representing overall drought tolerance, where 1 refers to highly tolerant, and 5 indicating highly sensitive. It is essential to note that drought tolerance is a complex trait encompassing multiple physiological and morphological attributes, including stay green, earliness, and late maturity with associated stay-green characteristics.

Phenotyping for nutrient content

Nutritional composition analysis of the crop was conducted at MARC, at the food and nutrition laboratory immediately after harvest. A representative grain sample of 150–200 g per accession was carefully collected from each replicated plot at each testing site during the 2019/2020 harvesting period. The seeds were cautiously inspected to ensure they were clean and free from any inert matter, dust, or grains from other crops. To prevent moisture loss and contamination, the clean seeds were stored in airtight containers. Prior to near infra-red (NIR) analysis, a standardization step was performed using a reference plate made of polystyrene, a material with a known NIR spectrum. This step calibrated the instrument to match the NIR spectrum of the reference plate, ensuring accurate NIR spectrum measurements of the samples. For non-destructive NIR analysis, a representative amount of the seed sample (150–200 g) was added into the sample hopper of the NIR analyzer (Perten Instruments AB, Sweden). The NIR analysis process was then initiated, and the Perten IM 9500 measured the NIR spectrum of the sample. This measurement process was repeated twice, and the average of the two readings was used for each sample to enhance accuracy. The machine recorded data on iron (Fe) (mg/100 g), and zinc (Zn) (mg/100 g) contents.

Leaf tissue sampling, DNA extraction, library preparation and sequencing

To guarantee seed purity and uniformity, three rounds of single-seed descent were conducted for each genotype before phenotyping. Subsequently, 3–4 seeds were sown in six '72 cavity seedling trays' at MARC, Ethiopia. Thinning was done to ensure only one healthy and vigorous seedling was retained per cavity. Leaf tissue samples were collected 3–4 weeks after planting and placed in individual wells of a 96-well plate for efficient sample handling and storage. The collected samples were deep-frozen at − 80 °C for 24 h, before being subjected to freeze-drying (lyophilization) using the Martin Christ Alpha 1–2 LDplus (Donau lab, Ukraine) for 24 h. The lyophilized leaf samples were carefully packaged and shipped to the International Crops

Research Institute for the Semi-Arid Tropics (ICRISAT) genomics laboratory unit in Nairobi, Kenya for DNA extraction. DNA was extracted using the Isolate II plant DNA extraction kits (Bioline Pty Ltd, Nottingham, UK) according to manufacturer's instructions. The purity and quantity of the extracted DNA were determined using gel electrophoresis and a Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, CA) respectively with final dilution to 50 ng/µl. DNA samples were sent to Psomagen (USA) for library preparation and next-generation sequencing. Libraries were constructed using the RIPTIDE kit (Twist Bioscience, San Francisco, CA, USA) according to the manufacturer's protocol. The libraries were subjected to paired-end sequencing on the Illumina Novaseq6000.

Quality control and SNP calling

Raw sequencing reads were trimmed to remove adapter sequences and poor quality reads using Trimmomatic (version 0.39) as per Bolger et al. (2014). The high-quality reads were aligned to the chromosome scale finger millet reference genome (Devos et al. 2023) and SNPs called using GATK (version 4.2.5.0) (McKenna et al. 2010). Haplosweep, a tool specifically designed for filtering high-quality SNPs from polyploid crops was used to eliminate homeologous SNPs (Clevenger et al. 2018). Additional filtering was done to retain SNPs with a minor allele frequency (MAF) > 5%, < 20% missing data, and heterozygosity of less than 25%. Only SNPs that mapped to finger millet's 18 chromosomes were retained for downstream analysis.

Population structure and linkage disequilibrium

Filtered SNPs were used to undertake principal component analysis (PCA) in the SNP & Variation Suite (SVS version 8.9.0). To establish the differences and likely effects of each of the finger millet sub-genomes, whole-genome SNPs were first used for a separate PCA analysis. Subsequently, additional analyses were conducted using SNPs generated from sub-genomes A and B. Ten principal components (PCs) and the additive model were used to generate Eigen values. The first three principal components of the variation were plotted and visualized in R software using the scatterplot3 d 0.3–41 package (Ligges and Machler 2003). The Discriminant Analysis of Principal Components (DAPC) was done using the adegenet v. 2.1.10 package in R software by retaining the first three principal components (Jombart 2010). Linkage disequilibrium (LD) decay was also done for whole genome, sub-genomes A and B SNPs independently. LD decay was estimated by using LD analysis from Tassel v5 (Bradbury et al. 2007) program with a sliding window of 10 kb. The resulting LD R^2 values were plotted using R script. The r^2 threshold was set to 0.2.

Phenotypic data analysis

Analysis of variance (ANOVA), means, and variances for each quantitative trait was done using R (R core team 2021). Best linear unbiased predictions (BLUPs) and thereafter variance components within environments were estimated in the lme4 package (Bates et al. 2015) in R (R core team 2021) by manipulating the REstricted Maximum Likelihood (REML) method using the model:

$$Y_{ijk} = \mu + r/\alpha i + r/\beta_j + \varepsilon_{ijk}$$

where Y_{ijk} is the kth observation for the ith row and jth column; μ is the overall mean; $r/\alpha i$ and r/βj are the effect of ith and jth rows and columns nested in replicates, respectively.

ε_{ijk}: Random error term.

BLUP variance components estimated within environments were appropriated to calculate broad-sense heritability (H^2) for respective traits per location using the formula:

H^2 scores were classified according to Robinson et al. (1949) as follows: 0–30% = low; 30–60% = moderate; > 60% = high. BLUPs were further used to generate frequency distribution curves and input as trait in GWAS. Pearson's correlation coefficients between all possible trait pairs were analyzed using the Proc corr function in SAS (SAS 2008).

Marker-trait associations

Marker-trait associations (MTAs) were calculated by combining the filtered SNP dataset of the genotypes with the corresponding BLUPs in R software using the Genome Association and Prediction Integrated Tool (GAPIT) version 2 package (Tang et al. 2016). A combination of statistical models was employed, including BLINK, CMLM, and MLM. All the models used the first three principal components (PCs) as the fixed effect to correct for population structure. Manhattan and quantile–quantile (Q–Q) plots were generated using the R package qqman (Turner 2018) to visualize the results. Manhattan plots provide a scatter plot representation of the GWAS results, displaying the distribution of SNPs across the chromosome, while Q–Q plots assess the effectiveness of the GWAS model in accounting for population structure and familial relationships. A stringent Bonferroni-corrected P values (0.05/no. of markers) were used as the threshold for retaining markers with significant associations. Putative candidate genes within LD distance of the significant SNPs were identified based on the KNE796-1 reference genome available in Phytozome (https://phytozomenext.jgi.doe.gov/info/Ecoracana_v1).

Haplotype analysis

The most rapid LD decay distance obtained after LD analysis was used to build haplotype blocks around all major quantitative trait nucleotides (QTNs). All markers that were within the LD decay distance of 1.2 Mb (600 Kb up- and downstream) of the major QTN were considered for the analysis based on the minimum LD distance calculated (Supplementary Fig. S3). Genes that were colocalized with a minimum of two SNPs around the major QTN haplotype blocks were analyzed for gene-specific haplotypes. Phenotypic data was categorized based on the identified haplotypes and used to test for haplo-pheno association in individual and across environments. One way ANOVA with Duncan's test as a post hoc test was used to identify significant associations and measure specific differences between pairs of means in R using package DescTools (Andri et al. 2021). Only haplotypes that were present in at least five genotypes were considered for the statistical analysis.

Results

Phenotypic variation

Highly significant differences were observed across locations (Loc), genotypes (Gen), and their interactions (Gen x Loc) for all traits under study at $P < 0.001$ (Table 1). DTF, DTM, GY, and DrtSc demonstrated high heritability across all environments (E1, E2, E3), with values ranging from 0.7 to 0.9. PTH and Fe content also displayed consistent moderate to high broad sense heritability across all environments (0.5–0.6). Mean DTF demonstrated marked environmental dependence, with E1 (120.8 days) and E2 (90.4 days) reporting contrasting extremes. The highest range and mean (0.2–6.6 tons/ha; 3.2 tons/ha) for GY was reported in E1 and the lowest (0.5–3.3 tons/ha; 1.2 tons/ha) from E3 (Table 1). Mean Fe and Zn reported similar and increasing patterns of variations from E3 to E1 (Table 1). Strong negative correlations were reported between phenological traits (DTF, DTM, STG, DrtSc) and yield components (TSW, GY) across all environments (Supplementary Table S3). Fe and Zn displayed consistent and location-independent positive association with correlation coefficients ranging from 0.76 to 0.81 (Supplementary Table S3). However, there was no correlation between Fe and Zn with all the agronomic traits, except for a once-off weak positive correlation (0.09) between Fe and GY at E3 (Supplementary Table S3). Frequency distribution histograms for all traits from each environment and across all environments are presented in Fig. 1.

None of the top 10 best yielders in each of the environments was an improved variety except Mecha (229371) and Meba (GBK011119 A) at Meiso (Table 2). The top three best yielders in E1 surpassed the potential yield of 6 tons/ha and were from Tigray and Amhara (Table 2). Most of the top ten yielders in each of the locations were from Amhara and Tigray (Table 2). All the top 10 genotypes having the most abundant Fe and Zn were landraces, majority from Tigray (Table 3).

Linkage disequilibrium, genetic diversity, and population structure

Skim-sequencing generated 17,401,572 raw SNPs from 391 genotypes and retained 24,112 high-quality SNPs after filtering and LD pruning. Out of the 24,112 markers, 10,246 markers were from sub-genome A, while 13,866 were from sub-genome B. Assuming a genome size of 1.5 Gb, the overall density of the markers across the genome would be 16 SNPs/Mbp. LD varied depending on the SNP sets used. Whole-genome SNPs reported moderate LD decay at 2.1 Mb, with sub-genome B reporting the slowest decay point (3.2 Mb), and sub-genome A revealing the most rapid decay at 1.2 Mb (Supplementary Fig. 3). A detailed analysis of LD across markers of each chromosome is provided in Supplementary Table S4.

PCA using all markers and markers from each of the sub-genomes confirmed that the markers retained were informative and were able to differentiate the genotypes (Fig. 2). Sub-genome B markers were the most informative, explaining 51.5% genetic variation across the genotypes within the first 3 principal components (PCs), followed by whole-genome SNPs (45.6%) and lastly the sub-genome A SNPs (43.3%) (Fig. 2C). While most genotypes clustered as admixtures, there was relatively distinct clustering of genotypes from Oromia and Tigray regions (Fig. 2). Genotypes from SNNP region always clustered together with Oromia region genotypes, which is consistent with the neighboring locations of these two regions (Fig. 2 and Supplementary Fig. S1).

To gain a deeper understanding of population structure, DAPC analysis was subsequently employed. The three categories of markers clustered the 391 genotypes into 4 sub-populations with different memberships (Fig. 3). Sub-populations generated by whole-genome SNPs were based on regions of origin, with clusters 2 and 3 being the most distinct and having genotypes predominantly from Amhara and Tigray regions, respectively (Fig. 3A), while cluster 1 was predominantly consisting of genotypes from Amhara (69 out of 145) (Supplementary Table S5). The most distinct sub-populations generated by the sub-genome A SNPs (Fig. 3B) was cluster 2 that consisted of majority genotypes from Tigray (100 genotypes), while genotypes from

Table 1 Descriptive statistics for all traits measured across the three environments

Location		DTF (days)	PTH (cm)	DTM (days)	GY (tons/ha)	TSW (g)	STG (1–5)	DrtSc (1–5)	Fe (mg/100 g)	Zn (mg/100 g)
Arsinegelle (E1)	Range	96.5–140.0	55.0–132.5	122.5–183.0	0.2–6.6	1.5–3.9	1–5	1–5	1.0–7.7	4.5–8.1
	Mean	120.8	86.6	155.1	3.2	2.3	1.9	3.3	5.5	7.4
	SE	1.9	7.0	2.9	0.2	0.1	0.4	0.4	0.7	0.2
	H^2	0.8	0.6	0.8	0.9	0.8	0.8	0.8	0.5	0.7
Maitsebri (E2)	Range	76.0–103.5	62.3–146.5	103.5–145.5	0.3–4.5	1.0–3.7	1.0–4.7	1.0–5.0	2.6–8.2	5.2–8.3
	Mean	90.4	101.3	129.7	2.1	2.0	2.2	2.8	6.5	7.6
	SE	1.6	7.7	2.5	0.2	0.2	0.4	0.3	0.4	0.2
	H^2	0.9	0.6	0.9	0.9	0.6	0.6	0.9	0.6	0.6
Meiso (E3)	Range	78.5–107.5	80.0–182.5	101.0–149.5	0.5–3.3	1.2–3.0	1.0–5.0	1.0–5.0	5.4–8.1	6.9–8.3
	Mean	93.5	136.4	128.6	1.3	1.9	2.9	3	6.9	7.7
	SE	93.5	136.4	128.6	1.3	1.9	2.9	3	6.9	7.7
	H^2	0.8	0.6	0.8	0.7	0.7	0.8	0.8	0.6	0.5
Source of variation	Location	***	***	***	***	***	***	***	***	***
	Genotypes	***	***	***	***	***	***	***	***	***
	Gen x Loc	***	***	***	***	***	***	***	***	***

H^2 = broad sense heritability; means were calculated based on best linear unbiased predictions (BLUPS); SE = standard error; Gen x Loc = genotype by location interaction, *** = significant difference at $P < 0.001$.

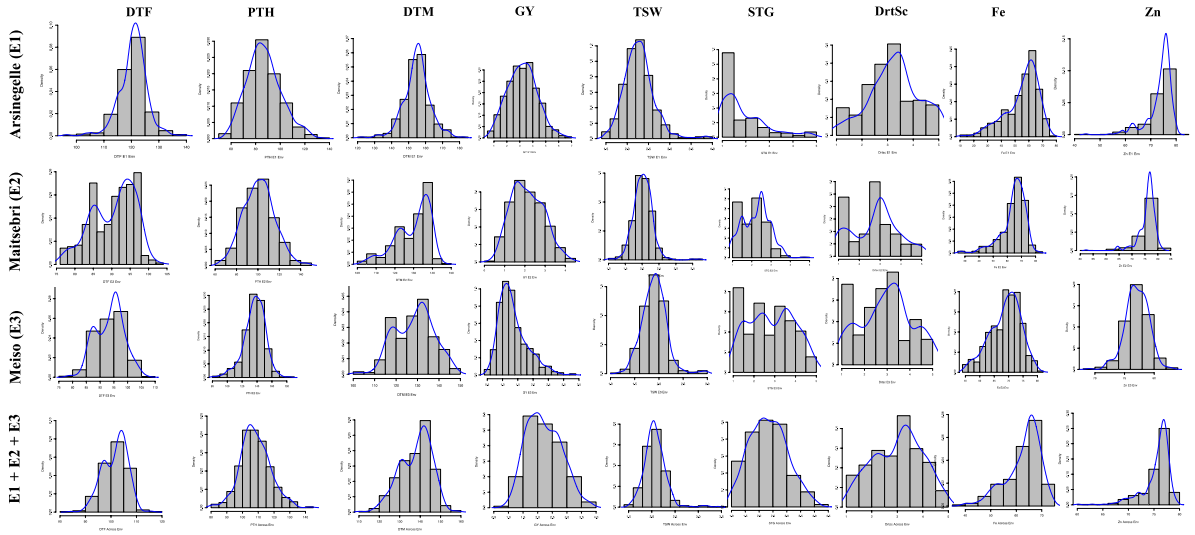

Fig. 1 Frequency distribution graphs for each of the traits across all environments. The graphs were drawn using BLUPs

Table 2 A summary of the top ten best yielders in each of the environments and their origin

Arsinegelle (E1)			Maitsebri (E2)			Meiso (E3)		
Accession	Origin	Yield (tons/ha)	Accession	Origin	Yield (tons/ha)	Accession	Origin	Yield (tons/ha)
237468	Tigray	6.61	AxumARColl-22	Tigray	4.46	219819	Tigray	3.24
234170	Tigray	6.42	AxumARColl-21	Tigray	4.25	AxumARColl-37	Tigray	2.86
243617	Amhara	6.18	242115	Amhara	4.12	238326	Tigray	2.81
238325	Tigray	5.97	234186	Tigray	4.07	237449	Tigray	2.67
242122	Amhara	5.97	AxumARColl-28	Tigray	3.97	207897	Tigray	2.62
238297	Tigray	5.85	235835	Amhara	3.92	Mecha*	Amhara	2.62
234162	Tigray	5.73	243631	Amhara	3.89	234165	Tigray	2.56
238321	Tigray	5.69	AxumARColl-33	Tigray	3.88	234206	Tigray	2.53
237452	Tigray	5.64	235828	Amhara	3.80	Meba*	Oromia	2.48

*Improved varieties

Amhara were split into 2 sub-populations (cluster 1 and 3) (Supplementary Table S5). Sub-genome B SNPs were the most informative in distinguishing genotypes from Tigray, which were grouped into cluster 1 (108 genotypes), but the same set of SNPs were not informative for genotypes from Amhara that were clustered in each of the four sub-populations (Fig. 3C; Supplementary Table S5). An overview of sub-population membership is presented in Supplementary Table S5.

Genetic diversity of released varieties

Although the different marker categories clustered most of the released varieties uniquely, they all revealed that the released varieties have been developed using a very narrow set of genotypic diversity, and almost always overlapped with genotypes from SNNP (Fig. 4). Whole genome SNPs showed a single cluster (16 out of the 22) of the released varieties that overlapped with genotypes from SNNP and was flanked by genotypes predominantly from Oromia (Fig. 4A). Sub-genome A SNPs followed a similar pattern and clustered most of the released varieties in between a sub-cluster of Oromia genotypes that also contained SNNP genotypes (Fig. 4B). Sub-genome B SNPs clustered the released varieties with an admixture set of genotypes from Oromia, SNNP, Amhara, and Benishangul & Gumuz (Fig. 4C). There were hardly any released varieties clustering under the predominant clusters from Tigray (pink; Fig. 4) and Amhara (red; Fig. 4) for both whole-genome and sub-genome A SNP markers. Although a cluster of Amhara genotypes was not as distinct using sub-genome B markers, there were not many overlaps between released varieties and Amhara genotypes (red; Fig. 4). The same scenario was observed for genotypes from Tigray, which

Table 3 A summary of the top ten accessions with the highest Fe and Zn content in each of the environments

Environment	Accession	Origin	Fe (mg/100 g)	Accession	Origin	Zn (mg/100 g)
Arsinegelle (E1)	223145	Tigray	7.66	234159	Tigray	8.08
	242612	Tigray	7.41	234203	Tigray	7.99
	238319	Tigray	7.36	215982	Amhara	7.97
	235700	SNNP	7.35	243639	Amhara	7.97
	215805	Oromia	7.34	235700	SNNP	7.97
	238316	Tigray	7.33	238316	Tigray	7.97
	221697	Tigray	7.30	215805	Oromia	7.96
	238304	Tigray	7.24	221697	Tigray	7.95
	228306	Amhara	7.20	238319	Tigray	7.94
	229734	Amhara	7.16	223145	Tigray	7.94
Maitsebri (E2)	229735	Amhara	8.20	229735	Amhara	8.26
	215875	Amhara	8.10	215805	Oromia	8.16
	207963	Oromia	7.94	208730	Oromia	8.14
	234183	Tigray	7.80	238333	Tigray	8.13
	238323	Tigray	7.75	245085	Oromia	8.09
	238297	Tigray	7.68	238342	Amhara	8.09
	238317	Tigray	7.67	229728	Bennishangul & Gumuz	8.08
	234171	Tigray	7.63	219503	Amhara	8.05
	215805	Oromia	7.63	238313	Tigray	8.05
	223145	Tigray	7.61	238335	Tigray	8.02
Meiso (E3)	216048	Oromia	8.10	241769	SNNP	8.31
	234202	Tigray	7.94	242118	Amhara	8.25
	208730	Oromia	7.87	216032	Oromia	8.20
	238337	Tigray	7.84	100095	Oromia	8.18
	242122	Amhara	7.84	234154	Tigray	8.12
	211029	Tigray	7.84	215867	Amhara	8.11
	234207	Tigray	7.83	238313	Tigray	8.09
	234175	Tigray	7.81	242108	Amhara	8.09
	234171	Tigray	7.80	234193	Tigray	8.07
	243637	Amhara	7.80	216057	Oromia	8.04

might suggest that most of the collections from Tigray and Amhara are yet to be used in breeding programs for crop improvement.

Marker-trait associations (MTA)

Seventy (70) MTAs were detected across the nine traits from at least one of the environments using the stringent Bonferroni threshold of $P < 2.07366E-06$ for whole-genome SNPs, $P < 4.8E-06$ for sub-genome A, and $P < 3.6E-06$ for sub-genome B (Table 4). Sixty-two of the 70 MTAs were detected using whole-genome SNPs while the remaining 8 MTAs were major (≥ 30% phenotypic variation explained) QTNs detected using sub-genome A and sub-genome B SNPs independently (Table 4). Out of the 70 MTAs, 8 markers were significant for more than one trait, resulting in a non-redundant set of 62 QTNs.

Whole genome MTAs

Fifty-five of the 62 non-redundant QTNs were detected using whole-genome SNPs, of which 8 were major QTNs detected in association with DTM (40.6% phenotypic variation explained (PVE)), TSW (35.8%, 38.1%, 56.1% & 60.6% PVE), DrtSc (33.8% PVE), Zn (44.5% PVE), and Fe and Zn (30.6% PVE; 32.9% PVE) (Table 4). The 4 major QTNs detected for TSW were on different chromosomes; 1B (SNP 1B_64659309; 56.1% PVE), 5B (SNP 5B_3150126; 35.8% PVE), and two on chromosome 6B (SNP 6B_24190829, 60.6% PVE; SNP 6B_40602961, 38.1% PVE) (Fig. 5; Table 4). The highest number of MTAs (13) was reported for TSW, followed by 10 MTAs for Zn and nine for PTH (Table 4). The lowest number was reported for STG, for which the QTN was specifically detected in E3 (Table 4). MTAs were detected in at least one environment for all the traits except STG. One major

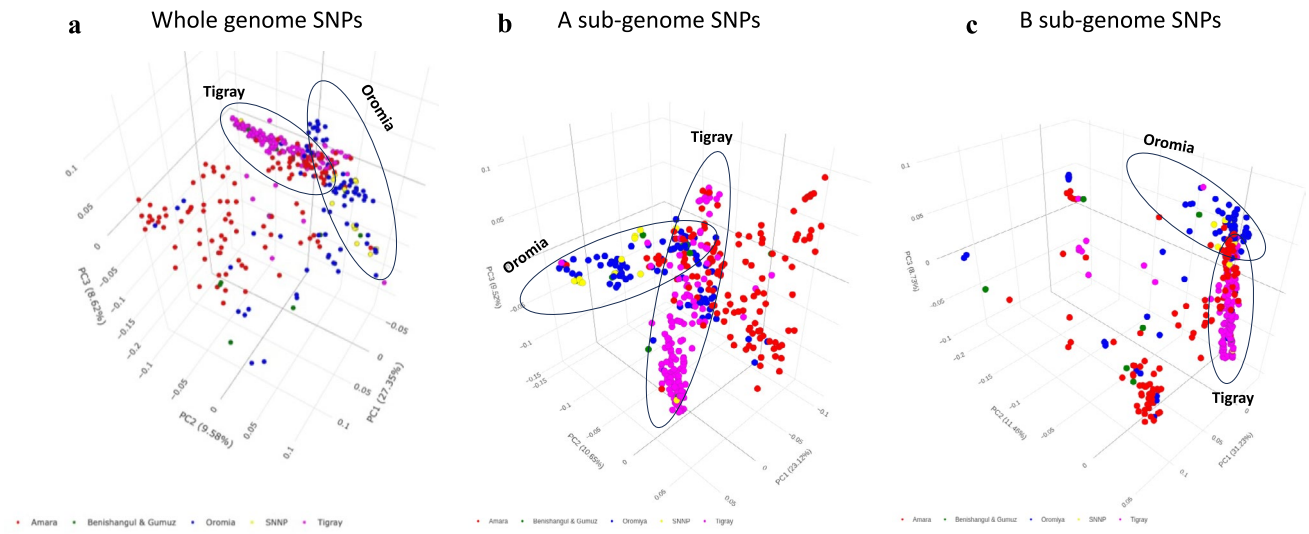

Fig. 2 PCA using the retained markers across 391 genotypes. **a.** PCA done using whole-genome SNPs in which the first 3 PCs explain 45.6% genetic variation. **b.** PCA done using sub-genome A SNPs across the 391 genotypes where the first 3 PCs explain 43.3% genetic variation. **c.** PCA done using sub-genome B SNPs. The first 3 PCs explain 51.5 genetic variation across individuals

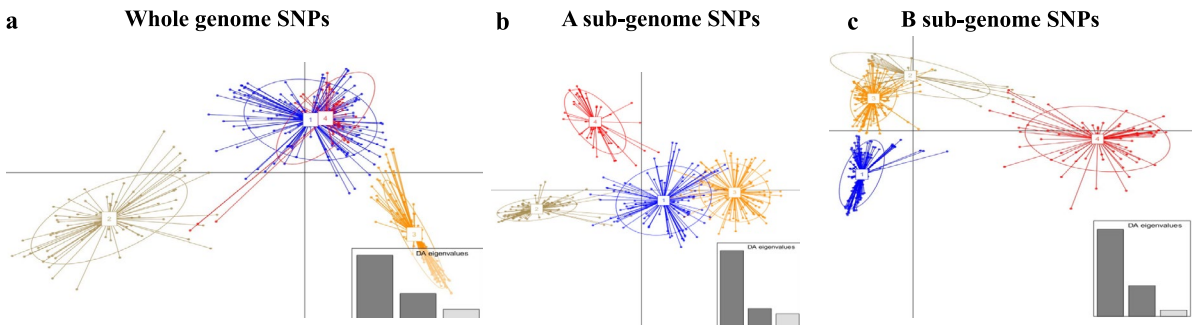

Fig. 3 Discriminant analysis of principal components (DAPC) analysis using different marker categories. **a.** Whole-genome SNPs separated the genotypes into 4 sub-populations but sub-populations 1 and 4 were not distinct. **b** Sub-genome A SNP markers successfully separated the 4 sub-populations into distinct clusters. **c** Sub-genome B SNP markers distinguished clearly the sub-population 4 from the rest of the sub-populations. Membership in each of the clusters is described in Supplementary Table S5

QTN (SNP 2B_35281485) was detected for both Fe and Zn (Table 4).

Sub-genome-specific MTAs

Only major QTNs were reported for the sub-genome-specific MTAs in Table 4, while the remaining sub-genome MTAs are reported in Supplementary Table S6. Sub-genome A SNPs detected five additional major QTNs, 3 localized on chromosome 8 A and 1 on chromosome 4 A (Table 4, Supplementary Table S6). The 5 major QTNs were detected for DTF (SNP 8 A_18375210; 55.6% PVE); PTH (SNP 4 A_3954933; 33.8% PVE); TSW (SNP 8 A_22146023; 33.8% and 62.8% PVE), DrtSc (SNP 8 A_23730370; 38.7% PVE) and Fe quality (SNP 8 A_14715050; 42% PVE) (Table 4, Supplementary Table S7). All the major QTNs detected using sub-genome A markers were not detected as major QTNs using whole-genome SNPs (Table 4). Despite having fewer markers than sub-genome B (13,866), the use of sub-genome A markers (10,246) alone added more information and new major QTNs that were not detected while using whole-genome SNPs (Supplementary Table S6).

Sub-genome B SNPs detected additional 2 major QTNs, same SNP (9B_20268455) for DTF (35.9% PVE) and DTM (34.7% PVE), and another SNP (6B_27369367; 72.3% PVE) for TSW (Table 4). The two major QTNs for DTF were only detected when the sub-genome SNPs were used and were totally missed when whole-genome SNPs were deployed for analysis (Table 4). STG MTAs were better detected using sub-genome B or whole-genome SNPs as opposed to

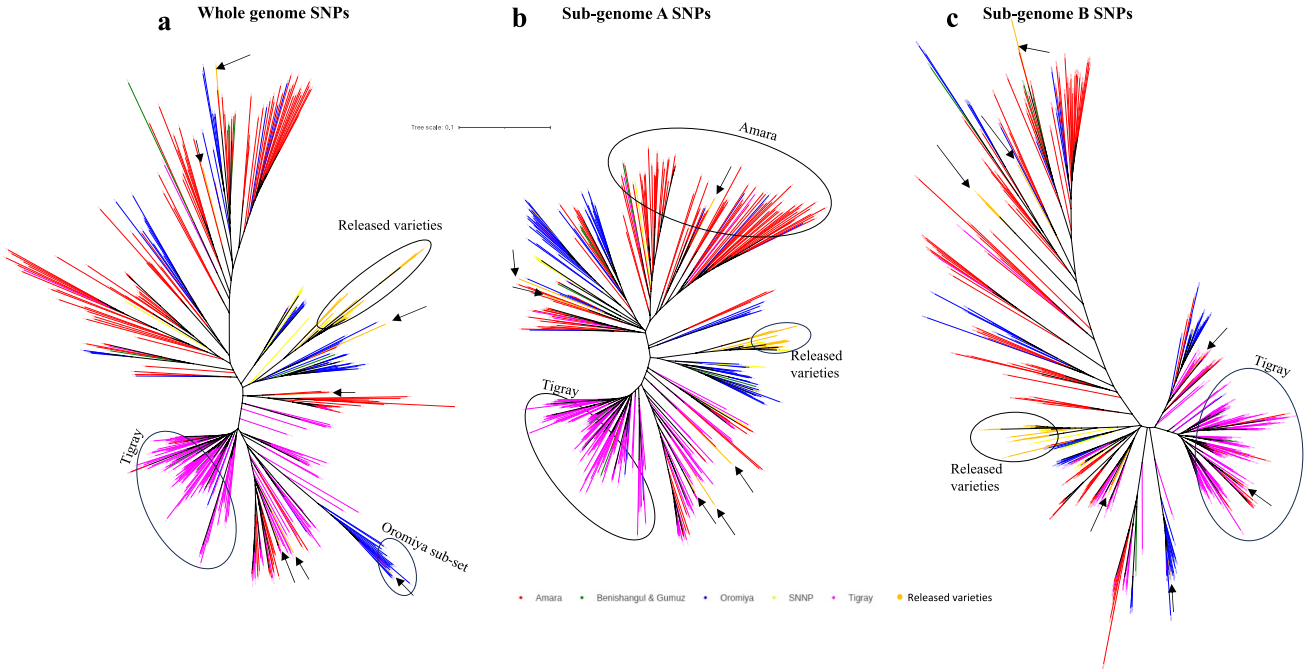

Fig. 4 Dendrograms showing the clustering of released varieties in the background of all germplasm used in the current study. All released varieties clustering individually are shown with arrows. **a.** Clustering of released varieties using whole-genome SNPs. **b.** Clustering of released varieties using sub-genome A SNPs. **c.** Clustering of released varieties using sub-genome B SNPs

sub-genome A SNPs (Supplementary Table S6). The highest PVE by a marker was reported on chromosome 6B (SNP 6B_27369367) for TSW in E1 at 72.3% PVE followed by another marker downstream on the same chromosome (SNP 6B_24190829) for the same trait in E3 at about 60.0% PVE using both CMLM and MLM (Fig. 5).

A QTN hotspot for Zn and Fe was detected on chromosome 2B using CMLM and MLM (Fig. 6; Table 4). The hotspot QTN was detected across E1 and E2 only (Fig. 6; Table 4) and not in E3.

A major QTN on chromosome 9B (SNP 9B_20268455) for DTF (35.9% PVE) and DTM (34.7% PVE) was detected across the three environments using sub-genome B specific markers (Table 4). All major QTNs were plotted on the finger millet marker density map as shown on Fig. 7. Supplementary Table S6 provides more details of all the significant MTAs with corresponding *P* values, minor allele frequencies (MAF), models, QQ, and Manhattan plots.

Haplotype analysis

Haplotype analysis of all major QTNs revealed 54 haplotype blocks having significant association with the respective traits of interest (Table 5; Supplementary Table S7). The maximum number of haplotypes detected per locus was 5, while the longest haplotype comprised of 16 nucleotides (Table 5). Detailed information on the haplotypes, the locations of the SNPs, and their frequencies are provided in Supplementary Table S7.

One way ANOVA of different haplotype pairs detected 195 significant associations with target traits, majority (136) of which were significant at $P < 0.001$, 35 at $P < 0.05$, and 24 at $P < 0.01$ (Supplementary Table S8). Box plots showing comparisons between pairs of haplotypes at each test location and across environments revealed that most of the favorable haplotypes were rare and present in a smaller percentage of individuals (Supplementary Fig. S4). Pleiotropic haplotypes associated with both DTM and DTF identified on chromosome 9B revealed the largest block containing 16 SNPs (Fig. 8). The most favorable haplotype for DTM and DTF at the pleiotropic QTN on chromosome 9B (9B_20268455) was H3 (CACCAACAGGAAACTT), which was present in 13 individuals (Fig. 8C and D).

The other pleiotropic haplotype associated with Fe and Zn was revealed on chromosome 2B (2B_35281485), was having two nucleotides, with the most favorable haplotype being GG that was present in 12 individuals (Fig. 9).

Regional distribution of genotypes having favorable haplotypes around the pleiotropic QTNs

Using data across the 3 environments, majority of the 13 and 12 genotypes harboring favorable haplotypes for DTM/DTF and Fe/Zn respectively were from Tigray region (Table 6),

Table 4 Marker-trait associations detected using whole-genome and sub-genome SNPs across different models

Trait	SNP	*P* values			PVE (%)	Locations
		BLINK	CMLM	MLM		
Using whole-genome SNPs						
DTF	1 A_41691781	1.48E−06	–	–	11.8	E3
	6B_17245025	7.72E−07; 1.70E−07	–	–	2.0; 1.2	E2; E1 + E2 + E3
	7 A_42035020	3.38E−07	–	–	17.8	E1
	7B_29006305	1.31E−06	–	–	1.6	E1 + E2 + E3
	8 A_6485125	8.23E−10	–	–	6.7	E1
PTH	4B_50601501	6.51E−09; 1.30E−06	–	–	1.5; 4.8	E3, E1 + E2 + E3
	5B_27829128	1.29E−07	–	–	2.5	E3
	6 A_29685470	5.27E−07	–	–	16.1	E2
	6B_2712127	2.97E−08	–	–	2.1	E3
	7 A_35748697	9.979E−07	–	–	16.7	E1 + E2 + E3
	8B_59128846	9.16E−07	–	–	1.3	E3
	9 A_40793867	4.13E−07	–	–	1.0	E3
	9B_15553707	9.84E−09	–	–	9.0	E3
	9B_19836556	2.59E−09	–	–	8.4	E3
DTM	4 A_3946408	5.11E−08	–	–	7.3	E1
	5B_65558510	1.33E−06	–	–	40.6	E3
	6B_16874563	1.17E−06	–	–	10.3	E2
	6B_17245025	6.30E−08	–	–	1.4	E1 + E2 + E3
	7 A_42035020	2.72E−07	–	–	16.4	E1
	7B_29006305	8.81E−08	–	–	1.8	E1 + E2 + E3
GY	3 A_22037946	3.72E−09; 1.02E−06	–	–	7.2; 4.7	E1; E1 + E2 + E3
	4 A_40121282	1.25E−06	–	–	1.9	E1 + E2 + E3
	5B_55857122	2.61E−07	–	–	1.3	E1 + E2 + E3
	6 A_14427382	3.18E−11; 3.26E−11			10.6; 5.9	E1; E1 + E2 + E3
TSW	**1B_64659309**	–	3.77E−07	3.77E−07	56.1	E1 + E2 + E3
	1B_7786365	1.95E−06; 4.60E−08	–	–	4.8; 4.3	E1; E1 + E2 + E3
	2 A_34244148	1.23E−07	–	–	1.8	E3
	3B_35970259	1.26E−09	–	–	1.4	E1 + E2 + E3
	3B_45474371	1.40E−06; 4.51E−08	–	–	15.1; 29.6	E1; E1 + E2 + E3
	5B_28252956	2.04E−07	–	–	2.5	E1
	5B_3150126	1.51E−15; 1.55E−07; 4.16E−14	–	–	22.1; 35.8; 10.6	E1; E2; E1 + E2 + E3
	6 A_40922437	1.92E−07; 1.50E−06	–	–	5.4; 1.9	E2; E1 + E2 + E3
	6B_24190829	–	1.66E−06	1.66E−06	60.6	E3
	6B_40602961	1.35E−16	–	–	38.1	E3
	7 A_13129745	1.74E−07	–	–	0.8	E1
	7B_49045901	1.61E−06	–	–	6.7	E3
	8B_12513208	2.389E−07	–	–	1.1	E1 + E2 + E3
STG	6B_10864215	1.58E−07	9.90E−08	7.54E−07	21.3	E3
DrtSc	1 A_41691781	2.71E−10	–	–	6.3	E3
	1 A_47690818	1.47E−06	–	–	20.2	E2
	4 A_2499716	5.34E−07	–	–	19.4	E2
	5B_35573546	1.02E−07	–	–	33.8	E3
	7 A_42035020	1.72E−07	–	–	10.1	E1
	9 A_6603556	5.02E−09	–	–	13.5	E1

Table 4 (continued)

Trait	SNP	*P* values			PVE (%)	Locations
		BLINK	CMLM	MLM		
Fe	2B_35126324	–	1.04E-06	1.04E-06	4.6	E2
	2B_35281485*	2.01E−11; 2.38E−08	1.14E-07	1.14E-07	1.9–32.9	E1; E2
	2B_37326032	–	1.05E-06	1.05E-06	2.6	E2
	3B_49895136	3.61E−07	–	–	2.5	E1
	7B_22082060	4.37E−07	–	–	5.0	E1
	8 A_14715050	5.87E−07	–	–	2.8	E1
	8 A_34625287	1.11E−08	–	–	6.2	E1
	9 A_44001813	5.65E−08	–	–	4.5	E1
Zn	2B_35126324	–	1.19E−06	–	12.0	E2
	2B_35281485*	2.76E−27; 2.50E−08	6.31E−07; 5.65E−07	6.31E−07	8.1–30.6; 2.1–8.9	E1; E2
	2B_35342493	-	7.74E−07	–	0.1	E2
	3 A_38743459	3.30E−07	–	–	2.1	E1
	5 A_62694353	9.73E−08	–	–	1.9	E1
	5 A_65700827	2.50E−07	–	–	2.2	E1
	7B_14800902	5.00E−13	–	1.61E-06	7.4–44.5	E2
	8B_30935289	1.60E−06	–	–	4.6	E2
	9 A_42427372	9.97E−07	–	–	1.5	E1
	9B_11179336	2.78E−07	–	–	7.3	E1
Additional Major QTNs detected using Sub-genome A specific MTAs						
DTF	**8 A_18375210**	3.55E−06	–	–	55.6	E1
PTH	**4 A_3954933**	–	2.01823E-06	2.01823E-06	33.8	E1 + E2 + E3
TSW	**8 A_22146023**	4.92E-09; 5.00E−09	–	–	62.8; 33.8	E1; E1 + E2 + E3
DrtSc	**8 A_23730370**	1.85E−15	–	–	38.7	E3
Fe	**8 A_14715050**	5.56E−07	–	–	42	E1
Additional Major QTNs detected using Sub-genome A specific MTAs						
DTF	**9B_20268455***		1.37E−06	1.37E−06	35.9	E1 + E2 + E3
DTM	**9B_20268455***		1.67E−06	1.67E−06	34.7	E1 + E2 + E3
TSW	**6B_27369367**	–	2.49E−06	2.49E−06	72.3	E1

*Same QTN is reported for Fe and Zn, and for DTF and DTM. SNPs in bold are considered major QTNs reporting > = 30% PVE in at least one environment

consistent with the enrichment of superior performing genotypes in the same region (Table 2 and 3). Three genotypes (243635, AxumARCColl-25 and AxumARCColl-31) uniquely harbored both the DTM/DTF and Fe/Zn favorable haplotypes (Table 6). Only one released variety (Tekeze-1) was among the list of genotypes having superior pleiotropic haplotypes (Table 6).

Functional QTNs and gene-specific haplotypes

Ten genes were found to co-localize with major QTNs, two of which contained more than 1 SNP (Chromosomes 4 A and 9B; Table 7). Four of the genes were associated with PTH, one gene (*glycosyltransferase*) with TSW, two genes with Fe and one gene with DTM/DTF (Table 7). The two genes that harbored at least two SNPs were *ABC transporter family protein* (ELECO.r07.4 AG0303600) and *ankyrin repeat family protein* (ELECO.r07.9BG0703480) and were associated with PTH and TSW (Table 7). The study identified 2 haplotypes (H1-ATA, H2-GCG) for ELECO.r07.4 AG0303600 and 4 haplotypes (H1-CCG GCGAG, H2-TTAACAGA, H3- CCGGTGAG, H4-CTA ACAGA) for ELECO.r07.9BG0703480 (Supplementary Table S9). Mean comparison tests between haplotypes of ELECO.r07.9BG0703480 revealed no significant differences ($P < 0.05$) between genotypes (Supplementary Table S10) and were therefore dropped from further analysis. Significant association between ELECO.r07.4 AG0303600 haplotypes were observed in E1 ($P < 0.05$),

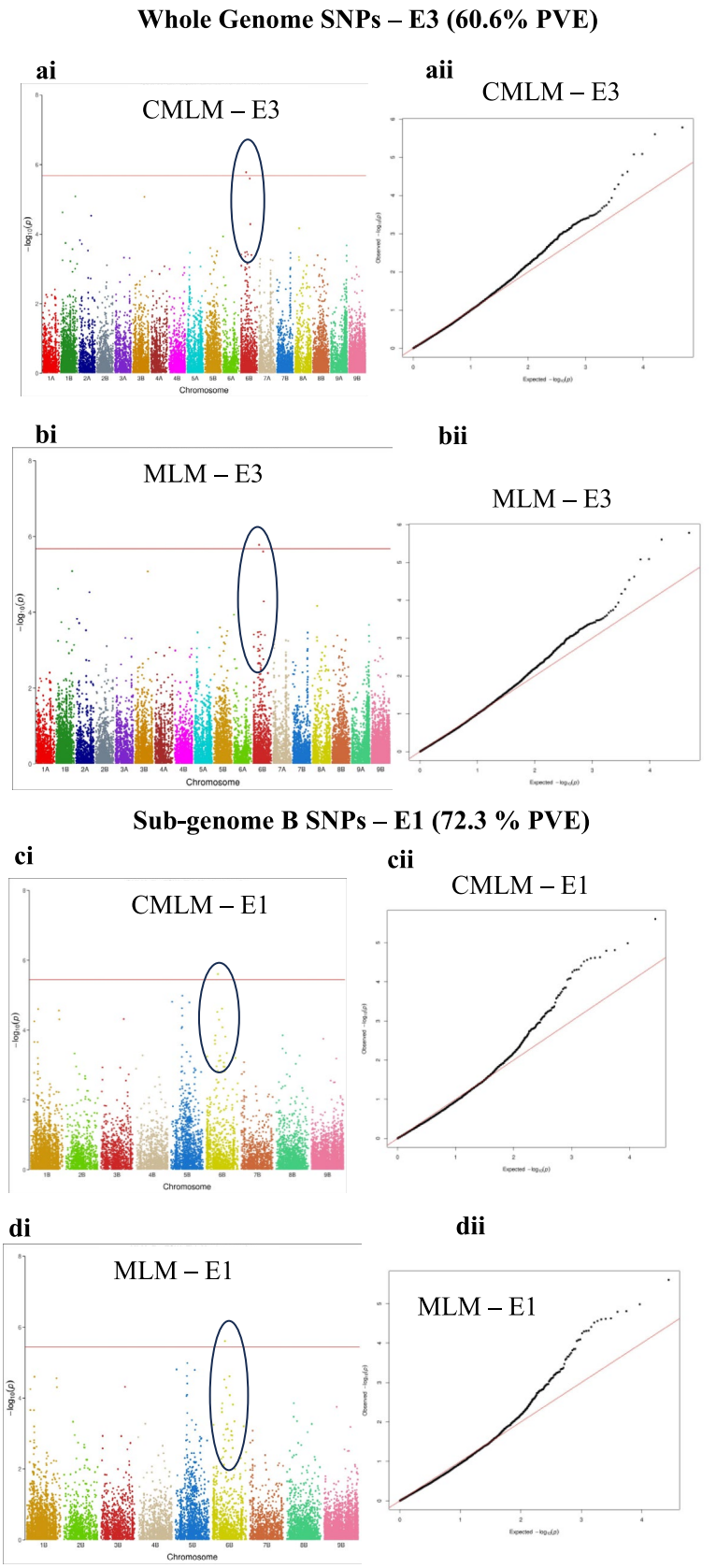

Fig. 5 Manhattan and QQ plots showing the TSW major QTNs explaining ~60.6% and 72.3% phenotypic variation. **ai & bi.** Manhattan plots drawn using CMLM and MLM models with whole-genome SNPs and showing significant association with TSW for E3. **aii & bii.** The corresponding QQ plot to the Manhattan plot in Ai and Bi plotted based on phenotypic data generated from E3. **ci & di**. Manhattan plots showing the detected major QTN for TSW (72% PVE) for E1 using CMLM and MLM models. **cii & dii**. The corresponding QQ plots to the Manhattan plots in Cii and Dii. Sub-genome A chromosomes are represented by odd numbers, while sub-genome B are represented by even numbers

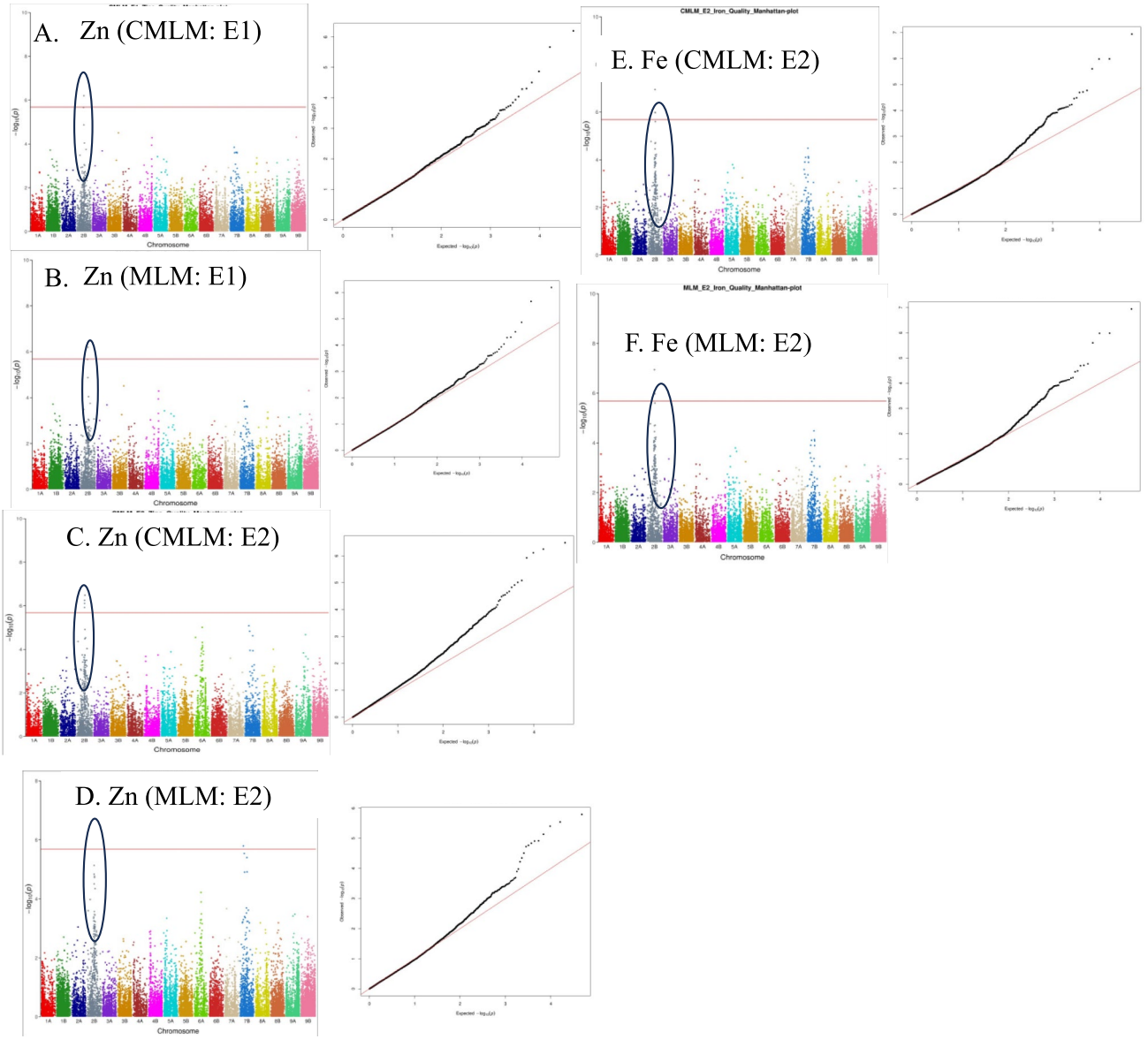

Fig. 6 Manhattan and QQ plots revealing the QTN hotspot for Fe and Zn on chromosome 2B. **A–D** Manhattan plots and corresponding QQ plots drawn using Zn quality data from E1 and E2 as shown using CMLM or MLM GWAS models. **E–F**. Manhattan plots and corresponding QQ plots drawn using Fe quality data from E2 as shown using CMLM and MLM GWAS models

E2 ($P < 0.05$), E3 ($P < 0.01$) and across environments ($P < 0.001$) (Supplementary Table S9; Fig. 10).

Identification of all other candidate genes from major QTNs using whole-genome and sub-genome A SNPs

All other candidate genes identified within major QTNs were retrieved and assembled in Supplementary Table S10. Three hundred and nineteen (319) genes were retrieved from chromosomes 1B (TSW), 2B (Fe and Zn), 5B (DrtSc and DTM), 6B (TSW), and 7B (Zn content) based on whole-genome SNP analysis, and a further one hundred and forty nine (149) candidate genes (15 for DTF; 4 for TSW; 8 for DrtSc, 23 for Fe and 99 for PTH) from using sub-genome A SNPs only and 10 genes (6 for DTF and DTM; 4 for TSW) from using sub-genome B SNPs only (Supplementary Table S10). The four major QTNs for TSW detected using whole-genome SNPs revealed 269 candidate genes, with the first QTN on chromosome 1B having a cluster of *Late embryogenesis abundant (LEA) hydroxyproline-rich glycoprotein* family and *Katanin p80 WD40 repeat-containing subunit B1 homolog* (Supplementary Table S10). The second TSW major QTN on chromosome 5B reported an abundance of

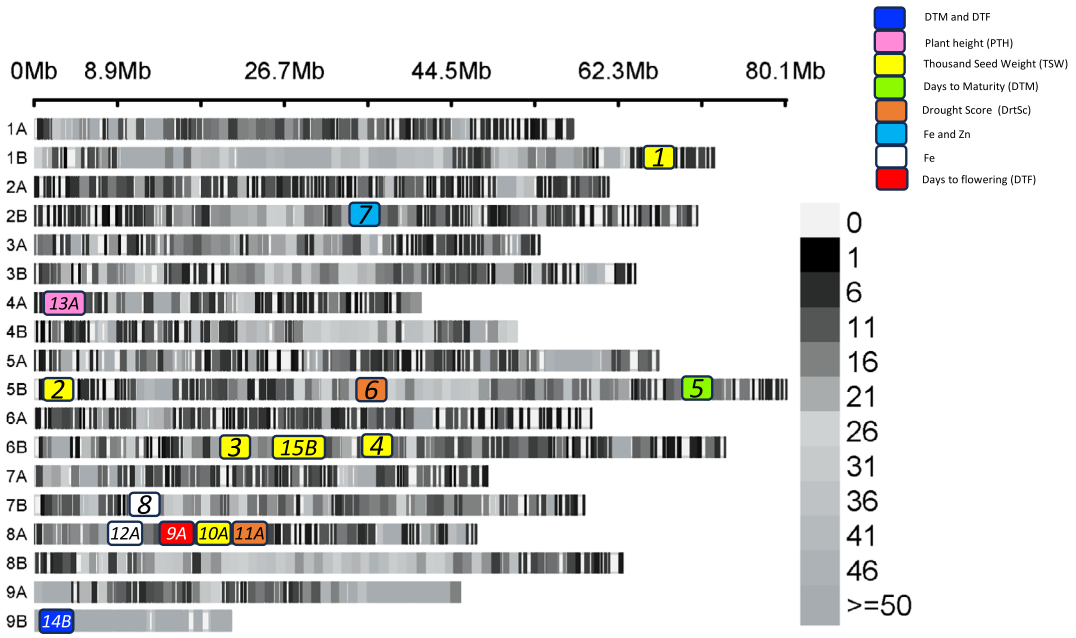

Fig. 7 The overall SNP density and the locations of major QTNs for which candidate genes were detected. The SNP density is indicated by the gray to black scale. The color key for each QTN is plotted within the figure. Major QTN locations are estimates and not drawn to scale. Major QTNs detected using sub-genome A SNPs only are denoted with "*A*" and those detected using sub-genome B SNPs only denoted with "*B*"

Ubiquitin carboxyl-terminal hydrolase 12, *O-methyltransferase-like protein* and *glycosyltransferases,* while the third TSW QTN revealed abundance of *serpins* (Supplementary Table S10). An additional *Ubiquitin family protein* was also identified in the TSW major QTN on chromosome 8 A that was detected using the sub-genome A SNPs only (Supplementary Table S10). The additional candidate genes detected for TSW on chromosome 6B contained four genes, two of which had been detected in other TSW QTNs. GDSL esterase/lipase gene detected on chromosome 6B for TSW was also detected on chromosome 5B major QTN for TSW; GRF zinc finger family protein was also detected upstream within LD of another TSW QTN (6B_40602961).

The DTM QTN showed abundance of *Calmodulin-binding protein,* as well as *Transmembrane protein* families. Two additional *Transmembrane protein* families were identified for DTF on chromosome 8 A using sub-genome A SNPs. The major QTN for Fe detected using sub-genome A markers revealed a cluster of 3 *Cysteine protease genes* that were ~ 0.5 Mb upstream of the significant marker (Supplementary Table S10). The same Fe major QTN on chromosome 8 A also harbored 2 *cytochrome P450* genes and one *Cytochrome P450 reductase* (Supplementary Table S10). One *Receptor protein kinase* was reported for Fe major QTN on chromosome 7B, while 2 *Receptor protein kinases* were reported for the same trait on chromosome 8 A (Supplementary Table S10). All enriched candidate genes have been highlighted in yellow in Supplementary Table S10 for ease of detection.

Discussion

The findings of this research demonstrate tremendous genetic variation for both agronomic and nutrition-related traits for Ethiopian finger millet germplasms evaluated across three different environments. Our study is the first comprehensive GWAS for both agronomic and nutrition-related traits for finger millet in Africa using more than 300 genotypes. Beyond identifying significant MTAs, we also identified promising candidate genes and favorable haplotypes for the target traits that could be deployed right away into breeding programs. Overall, results confirm earlier studies that there is adequate genetic variation among finger millet germplasm to select for several traits such as drought, agro-morphological, and micronutrient content (Lule et al. 2012; Assefa et al. 2013; Backiyalakshmi et al. 2023; Chandra et al. 2024).

Phenotypic variation for the target traits

The present investigation revealed highly significant phenotypic variation between genotypes for all the traits studied across and within each of the locations. The current

Table 5 Haplotypes detected from major QTNs with significant association with the traits

Trait	Major QTN	Haplotypes
DrtSc	1. 5B_35573546	1. GG 2. AG 3. GT
	2. 8 A_23730370	1. GCG 2. ACG 3. GCA 4. GTG
DTF	8 A_18375210	1. CCG 2. TCC 3. CAG 4. CCC
DTF & DTM	9B_20268455	1. CGCCAA CAGGAA ACTT 2. TGTAGC TGTAGT TTCC 3. CACCAA CAGGAA ACTT
DTM	5B_65558510	1. CT 2. CG 3. TT
Fe	8 A_14715050	1. AGC 2. CGT 3. CGC 4. CAC
Fe & Zn	2B_35281485	1. AG 2. GA 3. GG
Zn	7B_14800902	1. CGG 2. TGA 3. TAA
PTH	4 A_3954933	1. TTCC 2. CCGT 3. TTGT
TSW	1. 1B_64659309	1. GAG 2. AAG 3. GAA 4. GGG
	2. 5B_3150126	1. TGCGG 2. TGTAA 3. GGCGG 4. TACGG
	3. 6B_24190829	1. GAGGT 2. GGGAG 3. GGGGG 4. AAGGT 5. GGAGG
	4. 6B_27369367	1. GGC 2. AAC 3. GGG
	5. 6B_40602961	1. ACGCG 2. GCGTG 3. GTGCG 4. ACACG 5. GCGCG
	6. 8 A_22146023	1. AGCC 2. AGCT 3. GATC

findings confirm that the assembled set of genotypes were ideal for trait discovery and paves the way for their utilization in other studies. Ethiopia is considered a primary center of diversity for finger millet (Purseglove 1985; Tsehaye et al. 2006; Tesfaye and Mengistu 2017), therefore likely to harbor important agronomic diversity required for crop improvement. Confirming this high extent of agronomic diversity among Ethiopian germplasm was also important for finger millet, in which the yields are extremely low and the gap between potential yield (6–8 tons/ha) and farmers' yields (< 2.5 tons/ha) is extremely high (Gebreyohannes et al. 2021, 2024). Our study identified unimproved material yielding much more than the released cultivars, especially in Arsinegelle, suggesting that the existing native agronomic potential of Ethiopian finger millet has not been exploited to the farmers' advantages. Furthermore, most of the top ten best yielders and those harboring the highest concentration of micronutrient content were from Amhara and Tigray, the two regions from which the study could confirm minimal clustering with released varieties. Therefore, future efforts will need to introduce these promising landraces into breeding programs and design relevant agronomic and management packages that are appropriate for farmers in enhancing their yields (Beres et al. 2020).

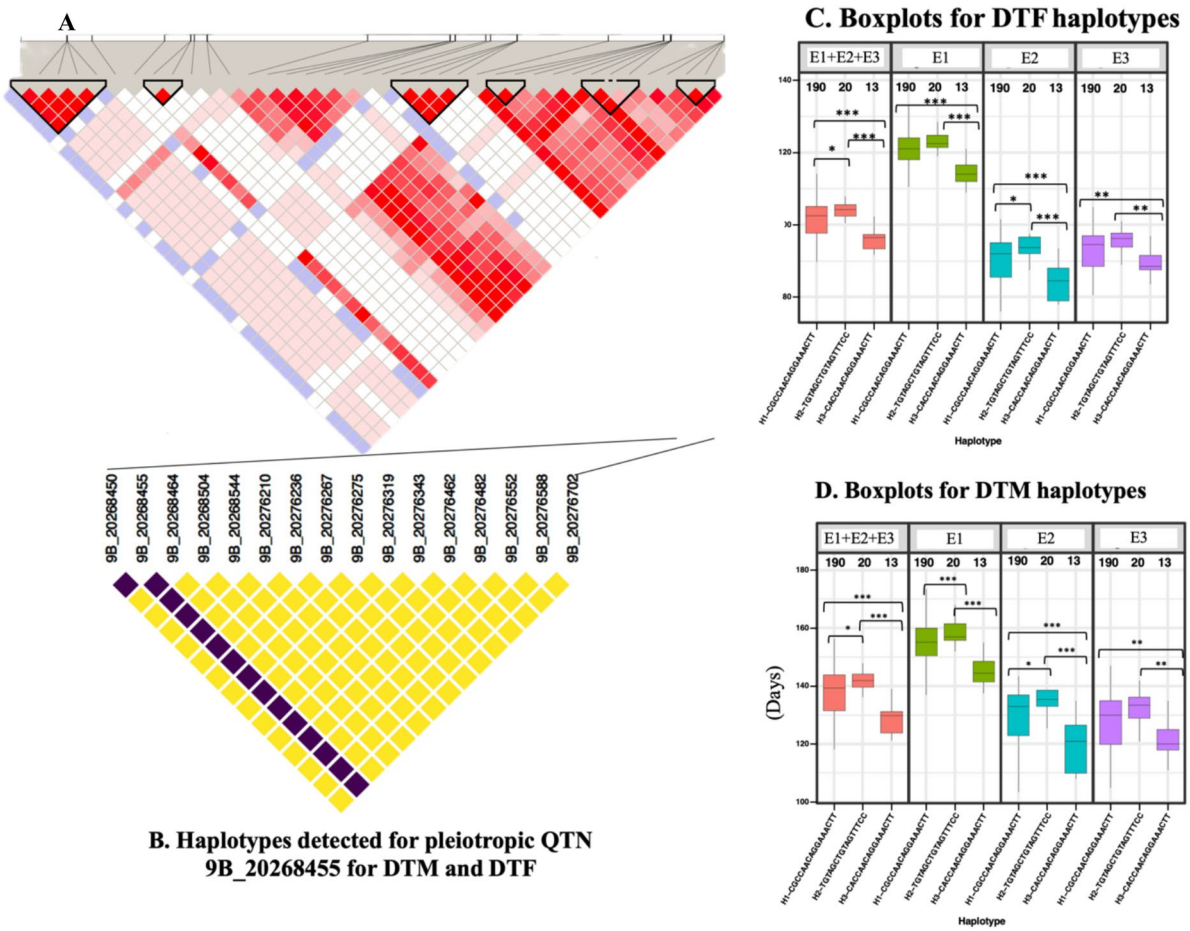

Fig. 8 Graphical representation of haplotype blocks across chromosome 9 and around the pleiotropic QTN 9B_20268455 for DTM and DTF. **A.** Haplotype blocks across the whole chromosome 9B. **B.** Haplotypes detected for pleiotropic QTN 9B_20268455 for DTM and DTF. **C.** Boxplots showing the performance of genotypes harboring various haplotypes with respect to DTF. **D.** Boxplots showing the performance of individual genotypes harboring various haplotypes with respect to DTF. All haplotypes are significantly different at $*P < 0.05$; $**P < 0.01$; $***P < 0.001$ (Supplementary Table S8)

Relatively high broad sense heritability (H^2) for most of the traits was identified in each of the environments, an indication of greater genetic control for the traits studied. Previous studies in finger millet that included at least 100 genotypes also reported similar H^2 values for some of the traits such as DTF (Lule et al. 2012; Eric et al. 2016), DTM (Anteneh et al. 2019), GY (Owere et al. 2015; Eric et al. 2016), PTH (Lule et al. 2012; Eric et al. 2016), and TSW (Lule et al. 2012). To the best of our knowledge, this was the first reported large-scale screening for the stay-green trait (STG) in finger millet, for which very high H^2 values were recorded. The stay-green trait has been used in several crops, including sorghum (Borell et al. 2014), to select for drought tolerance but is yet to be exploited in finger millet. Recording high H^2 values is promising and indicates the potential for genetic improvement of drought tolerance in finger millet using STG. Significant correlation between STG and drought score was also observed in E2 and E3, suggesting that the observed STG contributes to drought tolerance in these sites. More studies will need to be done to ensure that the stay-green trait being introduced into breeding programs is functional (Jordan et al. 2012) and not cosmetic (Myers et al. 2018).

The NIR-based estimates of Fe and Zn showed a moderate level of H^2 (0.5–0.7). This is a good level of accuracy considering the prohibitive nature of the costs of wet chemistry. Although these results will still need to be validated through wet chemistry, the range of Fe content recorded in the current study was higher than reported for the global finger millet core collection (Upadhyaya et al. 2010) but comparable to Puranik et al. (2020), who analyzed a diverse set of finger millet genotypes that had been evaluated in Kenya. The Zn content reported in this investigation exceeds most of the previous reports and only comparable to previous studies in which plants had been grown in Zn treated soils (Yamunarani et al. 2016; Teklu

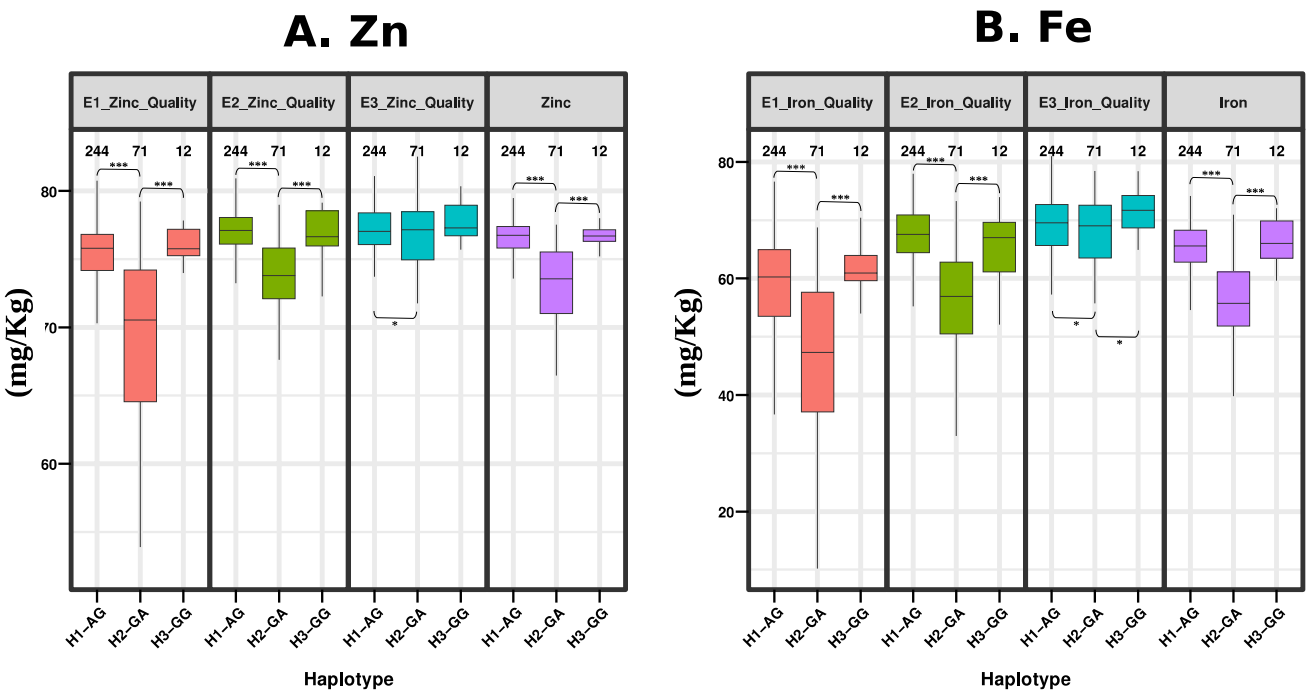

Fig. 9 Boxplots showing differences in performances of individuals harboring different haplotype blocks generated based on a major QTN that was associated with both Fe and Zn. **A.** Differences in haplotypes detected for pleiotropic QTN 2B_35281485 for Zn. **B.** Differences in haplotypes detected for pleiotropic QTN 2B_35281485 for Fe. All haplotypes are significantly different (Supplementary Table S8)

et al. 2024). A recent study on agronomic biofortification of Fe and Zn in Ethiopia reported up to 21.4% increase in grain Fe and Zn content following the application of Zn and Fe fertilizer (Teklu et al. 2023b). Previous studies reported higher Fe grain concentrations than Zn (Upadhyaya et al. 2010; Puranik et al. 2020; Fred et al. 2021; Teklu et al. 2023b) in contrast to our study where Zn concentrations were higher. The relatively higher Zn concentrations were consistent across the 3 locations and therefore reveal a likely unique germplasm set for improving grain Zn content in breeding programs. Our results might also suggest that there could be abundant genetic variation for Zn content in Ethiopia that may not have been captured within the global collections.

Genetic variation, population structure, and linkage disequilibrium

We stringently filtered our markers to retain a highly informative marker set for further characterization of the genotypes. The study also chose to conduct all the genetic analysis using whole genome, sub-genome A and sub-genome B sets of markers, following interesting results from previous studies (Devos et al. 2023; Bančič et al. 2023). The retained markers explained more than 45% genetic variation between the genotypes and successfully defined four sub-populations across the germplasm irrespective of which marker set was used. The marker density of 16 SNPs/Mb was twice higher than reported in Bančič et al. (2023). Retaining an informative set of SNP markers is a pre-requisite for population structure analysis and for determining differences between individuals. The results of the present study successfully showed that most of the genotypes from Tigray and Amhara regions have not been actively used in breeding programs based on the minimal clustering of the released varieties within the Tigray and Amhara clusters. The germplasm used is, therefore, a unique resource that should be introduced into breeding programs to enrich the pool of breeding lines, and will also need to be well maintained and conserved to ensure the allelic richness is available in the future. Being native to Ethiopia, there is likely to be diverse cultivated and wild finger millet accessions that are growing alongside each other with regular gene flow between them (Tsehaye et al. 2006; Tesfaye and Mengistu 2017; Lule et al. 2018; Dida et al. 2021; Pendergast et al. 2022). Gene flow can enhance adaptation and increase genetic variation for traits of interest that should be regularly captured through focused collection missions. Despite being an important center of diversity for finger millet, there are currently only ~30 finger millet genotypes from Ethiopia that are conserved in the ICRISAT global germplasm collection (ICRISAT 2024). This study reveals a need to have better representation of both Tigray

Table 6 Distribution of genotypes harboring favorable pleiotropic haplotypes for DTM/DTF and Fe/Zn

Trait & locus	Haplotype	Genotypes	Region of collection
DTM/DTF 9B_20268455	H3: CACCAACAGGAA ACTT	100065	Oromiya
		208442	Amara
		211029	Tigray
		215850	Amara
		215887	Amara
		234186	Tigray
		237473	Tigray
		243635*	Amara
		AxumARCColl-25*	Tigray
		AxumARCColl-27	Tigray
		AxumARCColl-31*	Tigray
		Tekeze-1	Tigray
Fe & Zn 2B_35281485	H3: GG	219831	Tigray
		234185	Tigray
		235153	Tigray
		238337	Tigray
		243635*	Amara
		AxumARCColl-13	Tigray
		AxumARCColl-25*	Tigray
		AxumARCColl-28	Tigray
		AxumARCColl-29	Tigray
		AxumARCColl-31*	Tigray
		AxumARCColl-35	Tigray
		AxumARCColl-3	Tigray
		AxumARCColl-7	Tigray

*Genotypes having both DTM/DTF and Zn/Fe favorable haplotypes

Table 7 Genes colocalizing with major QTNs

Gene Id	Gene description	Trait	Allele	Pos
ELECO.r07.4 AG0303570	RPM1-interacting protein 4-like protein	PTH	C/G	3,404,460
ELECO.r07.4 AG0304200	Polygalacturonase QRT3	PTH	C/T	3,940,479
ELECO.r07.4 AG0304230	30S ribosomal protein S17	PTH	C/T	3,955,182
ELECO.r07.4 AG0304420	Shikimate/quinate hydroxycinnamoyl transferase	PTH	T/C	4,204,724
ELECO.r07.4 AG0303600*	Multidrug resistance protein ABC transporter family protein	PTH	A/G	3,454,764
			T/C	3,454,863
			A/G	3,454,904
ELECO.r07.5BG0413640	Glycosyltransferase	TSW	G/A	2,841,435
ELECO.r07.8 AG0623600	4'-phosphopantetheinyl transferase	Fe	C/T	14,459,510
ELECO.r07.8 AG0623750	Receptor kinase 1	Fe	T/A	15,284,945
ELECO.r07.9BG0703480*	Ankyrin repeat family protein	DTF and DTM	C/T	19,836,545
			C/T	19,836,552
			G/A	19,836,556
			G/A	19,836,782
			C/T	19,836,804
			G/A	19,836,808
			A/G	19,836,813
			G/A	19,836,895
			G/A	20,349,730

*Highlights genes for which gene-based haplotype analysis was done

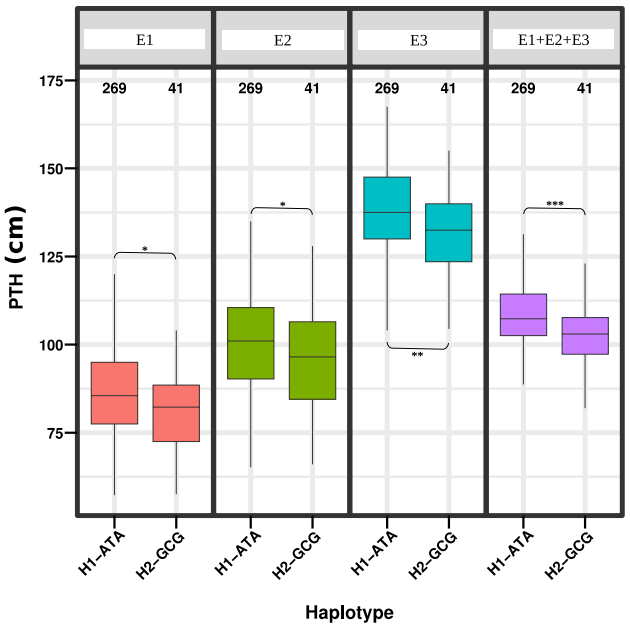

Fig. 10 Boxplots showing differences in performances of individuals harboring the two different haplotypes for PTH within the gene ELECO.r07.4 AG0303600 associated with plant height

and Amhara regions finger millet collection and characterization within the local and global gene banks.

All the three sets of markers revealed four sub-populations using DAPC analysis, although the sub-genome marker sets revealed more distinct clusters. The four clusters are consistent with Bančič et al. (2023) who studied 423 cultivated finger millet accessions from India, Nepal, East Africa, and Southern Africa using whole-genome SNPs. Devos et al. (2023) performed a similar analysis with the addition of wild accessions and reported two distinct clusters irrespective of which marker set was used. A repeat analysis of the same dataset in addition to existing datasets of wild accessions will be interesting in confirming the results reported by Devos et al. (2023). The clear difference in our study was the ability of the sub-genome markers to separate the sub-populations more clearly than the whole-genome SNPs.

We also observed a huge difference in the LD decay points between the two sub-genomes, which would need to be studied further in more diverse germplasm sets. Using naïve models, the study confirmed that LD decayed more rapidly in the sub-genome A (1.2 Mb) in comparison to sub-genome B (3.2 Mb), which is consistent with a previous study in which LD decayed much faster in the sub-genome A with or without genetic structure correction (Bančič et al. 2023). LD is important in association mapping as it determines the level of optimal resolution required for an association study. Although LD has not been studied much in finger millet, the reported slow decay is not surprising given the highly self-pollinating nature of this crop. The difference in LD between the two sub-genomes could also be traced back to the progenitors. It is known that the A sub-genome donor is a wild weedy species, *Eleusine indica* (Zhang et al. 2019), hence a more rapid decay would be expected. LD decays more rapidly in wild populations than in cultivated crops (Sela et al. 2014).

Marker-trait associations, haplotype analysis, and candidate genes

Our stringent filtering of markers and the use of Bonferroni threshold paid off in detecting major QTN regions. Some of the major QTNs detected in the current study were extremely novel and will form the basis for routine breeding and further research on the respective traits. Bančič et al. (2023) and Sood et al. (2023) detected several MTAs while studying agronomic traits in finger millet, but none of the reported associations explained > 20% phenotypic variation. Our study revealed major QTNs for DTM, DTF, TSW, DrtSc, Fe, and Zn alongside several other significant QTNs. The major QTNs were associated with economically important traits in finger millet, all of which are still being improved using traditional methods for most part. We additionally chose a stringent 30% PVE as a cut-off for major QTNs to enhance the immediate application of our results into molecular breeding. There are studies reporting as low as 10% PVE as major QTLs (Pandey et al. 2020). A major QTL/QTN is one that is detected in more than one environment and has a large effect on the phenotype. Most studies recommended at least 15 to 20% PVE for the locus to be considered as major (Du et al. 2015). This study employed a high threshold of 30% PVE as a cut-off to reduce the numbers of false positives when the associated markers are introduced into breeding programs for genomics-assisted selection and introgression. While recent advances in next-generation sequencing (NGS) have led to the development of quality control (QC) markers in finger millet, markers associated with economically important traits are still lacking. These major QTNs will add significant value toward early generation and genomic selection in structured finger millet breeding programs.

The potential reliability of the major QTNs for breeding was further validated using haplotype analysis. Haplotype analysis has been reported to improve the identification of genes of interest in other cereal crops including rice (Yang et al. 2023) and wheat (Taranto et al. 2023) but not yet exploited in finger millet, partly due to the previous lack of high-quality reference genome. Using a chromosome scale reference genome (Devos et al. 2023), we detected significant associations of different haplotype pairs with the target traits across all major QTNs. Associated haplotypes would enhance the precision of selection and introgression and improve breeding efficiency. Our results also provide

opportunity to stack different favorable haplotypes into superior elite backgrounds for better performance. In some cases, we identified individuals harboring more than one favorable haplotypes making such genotypes obvious candidates for immediate release, or introduction into various breeding programs. While a more comprehensive haplotype analysis of all QTNs identified will still be required to demonstrate the potential benefits of the QTNs detected, the results obtained from major QTLs form a baseline from which future similar studies will be anchored. Haplotype based breeding (HBB) promises to enhance genetic gain (Sivabharathi et al. 2024), especially in an under-researched crop such as finger millet.

The highest numbers of major QTNs were reported for TSW, which is one of the important determinants of grain yield in finger millet (Wolie and Dessalegn 2011). Haplotype analysis for seven major QTNs for TSW revealed clear differences in haplotypes demonstrating the readiness of all the markers identified for immediate use in breeding programs. Low yields in finger millet have been reported as one of the limiting factors to commercialization of the crop. Field evaluation, genetic diversity, and haplotype analysis results provide new knowledge and reveal that grain yield in finger millet has not been studied and exploited sufficiently. In rice (*Oryza sativa* L.), where seed weight has been extensively studied, three major genes have been reported for TSW. The first TSW locus in rice, *GS3*, encodes a transmembrane protein (Fan et al. 2006), while the second, *GW5*, encodes a calmodulin binding protein (Liu et al. 2017). A third locus encodes a transcription repressor *OsBZR1*, which represses *FZP* (*Frizzy Panicle*) expression (Bai et al. 2017). This study identified 3 copies of *calmodulin binding proteins* and 4 *transmembrane proteins* in the DTM major QTN. DTM was found to be positively correlated with TSW. A major TSW QTN was also found colocalized with a *glycosyltransferase* gene on chromosome 5B in our study. These interesting candidate genes make a priority list for functional validation toward increasing grain size and yield in finger millet. Finger millet researchers will now have the favorable haplotypes and unique high yielding landraces to deploy in their programs and improve genetic gain as soon as research investments are made available.

Gene-specific haplotype analysis done for two major QTNs reported promising results and suggests that future studies should consider all the significant QTNs for gene-specific haplotype analysis in the future. The two haplotypes identified for *ELECO.r07.4 AG0303600* for PTH are promising for immediate deployment in selection for and/or against the trait. Tall plants in finger millet have been associated with lodging, which is one of the breeding objectives (Lule et al. 2012). Wolie and Dessalegn (2011) reported negative correlation between PTH and grain yield (dwarf plants are better yielding) and a positive correlation with biomass. It has been argued that the introduction of dwarf plants into cereals may have saved more lives than any other scientific development (Hedden 2003). The clear differences in PTH of each of the haplotype pools is proof that this gene-specific haplotype has the potential to play a huge role if introduced into breeding programs, either for grain yield or biomass. Haplotype GCG of *ELECO. r07.4 AG0303600* should be studied further alongside TSW haplotypes for improving yields. Finger millet for fodder also remains an underexploited field that could take advantage of haplotype ATA of the gene *ELECO.r07.4 AG0303600. ELECO.r07.4 AG0303600* codes for ABC transporter gene family that is not well characterized in finger millet. However, there are several reports in many crops demonstrating the involvement of ABC transporter genes in dwarfism (Busov et al. 2008; Wang et al. 2017; Ye et al. 2013; Zhu et al. 2019). The identification of a putative ABC transporter in PTH opens a new chapter for both haplotype-based breeding and functional characterization to confirm its role in finger millet.

Drought tolerance is also an important trait in finger millet that can result in significant yield penalty (Krishnamurthy et al. 2016). This study identified a major QTN for drought tolerance and a relatively minor one for the stay-green trait (21.3% PVE). Haplotype analysis on the major QTNs for DrtSc on chromosomes 5B and 8 A revealed highly significant haplotypes that should be deployed right away in breeding programs. The most applicable would be the negative selection against drought sensitive haplotypes, which were very consistent across environments. The stay-green QTN will also need to be explored further because it revealed significant contribution to the phenotype. Natural variation in the promoter region of a rice stay-green gene has been exploited in breeding and led to delayed senescence and increased grain yield (Shin et al. 2020). Similar studies can be designed in finger millet to further characterize the stay-green QTN alongside the stable DrtSc QTNs and haplotypes.

The pleiotropic QTNs detected between DTM and DTF and between Fe and Zn were not surprising. There is evidence of positive correlation between Fe and Zn in pearl millet (Govindaraj et al. 2022), finger millet (Ojulong et al. 2021), as well as in several other crops (Cichy et al. 2009; Velu et al. 2017; Singh et al. 2018; Sanjana Reddy et al. 2010; Diaz et al. 2022). Our results further confirm a common major QTN for both traits in finger millet for the first time providing more evidence for joint genomics-assisted selection for Fe and Zn. Malnutrition has been identified among the top priority problems that result in high levels of maternal and infant mortality in Ethiopia (Kassebaum et al. 2014). The QTNs identified here can be immediately deployed for the selection of genotypes enriched for the two elements. Once validated, such

markers could also be used in the varietal release process as a standard to ensure only high Fe and Zn varieties are released, or recommended as a biofortification solution to the malnutrition cases, not only in Ethiopia, but in all finger millet producing countries. Both DTF and DTM are traits responsible for earliness, therefore a confirmation of a common QTN is a great validation of our results. Immediate validation and introduction of these markers will make it possible to improve the correlated traits simultaneously.

All the major QTNs detected while using whole-genome SNPs were from the sub-genome B. The major QTNs on chromosome 8 A were only detected when sub-genome A markers alone were used in the QTN analysis. Bančič et al. (2023) also detected significant QTNs only in the sub-genome B when whole-genome SNPs were used in the analysis. This result was not surprising given the differences observed in LD pattern for the different sub-genomes, as well as for the whole-genome SNPs. The appropriate MTA threshold would vary depending on both population- and genome-specific factors such as LD pattern and minor allele frequency (MAF) (Kaler and Purcell 2019). Given the rapid rate of LD decay in sub-genome A and the relatively fewer numbers of markers, the Bonferroni threshold used for the whole-genome SNPs may have resulted in several false negatives for sub-genome A that could be detected when sub-genome A SNPs alone were used for the analysis. Although there are not many GWAS reports for finger millet using cultivated species that distinguish the two sub-genomes, it will be important for future studies to take this point into consideration in order to ensure that no potential associations are missed out.

Conclusion

This study significantly advanced our understanding on the genetic basis of important agronomic and nutrient traits in finger millet. The use of skim-sequencing SNPs alongside a comprehensive and stringent GWAS approach identified major QTNs, favorable haplotypes, and promising candidate genes which are all novel discoveries in finger millet. To ensure the identified MTAs, haplotypes and candidate genes translate into practical applications, allelic and functional characterization will be required for their validation. Once validated, the markers reported in the current study will be great candidates for genomics-assisted selection and haplotype-based breeding in finger millet, in which there are currently none. To enhance conservation efforts and promote utilization in global breeding programs, further characterization of the Ethiopian finger millet germplasm is recommended. Ethiopian breeding program will likely benefit tremendously from integrating exotic germplasm into their conservation and breeding programs. Simultaneous improvement of agronomic and nutritional traits in finger millet breeding programs will pave the way toward making finger millet the desired global food security and climate resilient crop.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00122-025-04892-1.

Acknowledgements This research was financially supported by the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) through the 'Harnessing Opportunities for Productivity Enhancement (HOPE II) for Sorghum and Millets in Sub-Saharan Africa' project (grant number OPP1198373) and CIMMYT Dryland Crops' AVISA projects (OPP1198373), both funded by the Bill and Melinda Gates Foundation (BMGF). Additionally, ICRISAT provided support for the genotyping research work through the project "Safeguarding crop diversity for food security: Pre-breeding complemented with Innovative Finance" (TWCF0400), funded by the Templeton World Charity Foundation, Inc. (TWCF) and managed by the Global Crop Diversity Trust. The authors gratefully acknowledge the Ethiopian Institute of Agricultural Research, Melkassa Agricultural Research Center (Ethiopia) and the University of KwaZulu-Natal for their in-kind support and for facilitating the first author's research and study leave

Author contribution statement AG, HS and DAO conceptualized the study; DAO, DMG conducted genotyping experiments and supervised the data analysis and drafting of the manuscript; AG performed field and laboratory experiments, prepared leaf tissue samples for genotyping, and drafted the original manuscript; AG, HN, PG, SMM, DMG, and VKV analyzed the data; BK, DAO, and COO acquired project funds. All authors have reviewed and approved the final manuscript for publication.

Declarations

Competing interests The corresponding author, DAO, is an Associate Editor of the Theoretical and Applied Genetics Journal in which this manuscript is published. The authors declare that they have no competing interests.

Data availability The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

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

Andri S, Ken A, Andreas A, Nanina A, Tomas A, Chandima A, Antti A, Adrian B, Kamil B, Ben B (2021) DescTools: tools for descriptive statistics. R package version 0.99, 43

Anteneh D, Mekbib F, Tadesse T, Dessalegn Y (2019) Genetic diversity among lowland finger millet *Eleusine coracana* (L.) gaertn accessions. Ethiopian J Agric Sci 29:93–108

Assefa A, Fetene M, Tesfaye K (2013) Agro-morphological, physiological and yield related performances of finger millet [*Eleusine coracana* (L.) gaertn.] accessions evaluated for drought resistance under field conditions. Asian J Agric Rural Dev 3:709–720

Backiyalakshmi C, Babu C, Deshpande S, Govindaraj M, Gupta R, Sudhagar R, Naresh D, Anitha S, Peerzada O, Sajja S, Singh K, Vetriventhan M (2023) Characterization of finger millet global germplasm diversity panel for grain nutrients content for utilization in biofortification breeding. Crop Sci. https://doi.org/10.1002/csc2.21085

Bai X, Huang Y, Hu Y, Liu H, Zhang B, Smaczniak C, Hu G, Han Z, Xing Y (2017) Duplication of an upstream silencer of FZP increases grain yield in rice. Nat Plants 3(11):885–893. https://doi.org/10.1038/s41477-017-0042-4

Bančič J, Odeny DA, Ojulong HF, Josiah SM, Buntjer J, Gaynor RC, Hoad SP, Gorjanc G, Dawson IK (2023) Genomic and phenotypic characterization of finger millet indicates a complex diversification history. The Plant Genome. https://doi.org/10.1002/tpg2.20392

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67(1):1–48. https://doi.org/10.18637/jss.v067.i01

Beres BL, Rahmani E, Clarke JM, Grassini P, Pozniak CJ, Geddes CM, Porker KD, May WE, Ransom JK (2020) A systematic review of durum wheat: enhancing production systems by exploring genotype, environment, and management (g x e x m) synergies. Front Plant Sci 11:568657. https://doi.org/10.3389/fpls.2020.568657

Bolger AM, Lohse M, Trimmomatic UB (2014) A flexible trimmer for Illumina sequence data. Bioinform 30(15):2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Borrell AK, van Oosterom EJ, Mullet JE, George-Jaeggli B, Jordan DR, Klein PE, Hammer GL (2014) Stay-green alleles individually enhance grain yield in sorghum under drought by modifying canopy development and water uptake patterns. New Phytol 203(3):817–830. https://doi.org/10.1111/nph.12869

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23(19):2633–2635

Busov VB, Brunner AM, Strauss SH (2008) Genes for control of plant stature and form. New Phytol 177(3):589

Ceasar SA, Maharajan T, Krishna TPA, Ramakrishnan M, Satish RGV, L, Ignacimuthu S, (2018) Finger millet [*Eleusine coracana* (L.) Gaertn] improvement: current status and future interventions of whole genome sequence. Front Plant Sci 9:1054. https://doi.org/10.3389/fpls.2018.01054

Chandra AK, Pandey D, Sood S, Joshi DC, Tiwari A, Sharma D, Gururani K, Kumar A (2024) Uncovering the genomic regions underlying grain iron and zinc content using genome-wide association mapping in finger millet. Biotech 14:47. https://doi.org/10.1007/s13205-023-03889-1

Cichy KA, Caldas GV, Snapp SS, Blair MW (2009) QTL analysis of seed iron, zinc, and phosphorus levels in an Andean bean population. Crop Sci 49(5):1742–1750

Clevenger JP, Korani W, Ozias-Akins P, Jackson S (2018) Haplotype-based genotyping in polyploids. Front Plant Sci 9:564. https://doi.org/10.3389/fpls.2018.00564

Devi PB, Vijayabharathi R, Sathyabama S, Malleshi NG, Priyadarisini VB (2014) Health benefits of finger millet (*Eleusine coracana* L.) polyphenols and dietary fiber: a review. J Food Sci Techn 51(6):1021–1040

Devos KM, Qi P, Bahri BA, Gimode DM, Jenike K, Manthi SJ, Lule D, Lux T, Martinez-Bello L, Pendergast TH IV, Plott C, Saha D, Sidhu G, Sreedasyam A, Wang X, Wang H, Wright H, Zhao A, Deshpande S, de Villiers S, Dida MM, Grimwood J, Jenkins J, Lovell J, Mayer KFX, Mneney EE, Schatz OHF, Schmutz MC, Song J, Tesfaye B, K, Odeny DA, (2023) Genome analyses reveal population structure and a purple stigma color gene candidate in finger millet. Nat Commun 14:3694. https://doi.org/10.1038/s41467-023-38915-6

Diaz S, Polania J, Ariza-Suarez D, Cajiao C, Grajales M, Raatz B, Beebe SE (2022) Genetic correlation between Fe and Zn biofortification and yield components in a common bean (*Phaseolus vulgaris* L.). Front Plant Sci 12:739033

Dida MM, Oduori CA, Manthi SJ, Avosa MO, Mikwa EO, Ojulong HF, Odeny DA (2021) Novel sources of resistance to blast disease in finger millet: special issue: adapting agriculture to climate change: a walk on the wild side. Crop Sci 61(1):250–262. https://doi.org/10.1002/csc2.20378

Dida MM, Devos KM (2006) Finger millet. In: Kole C (ed) Cereals and millets. Genome mapping and molecular breeding in plants. Springer, Berlin, pp 333–343

Du Z, Che M, Li G, Chen J, Quan W, Guo Y, Wang Z, Ren J, Zhang H, Zhang Z (2015) A QTL with major effect on reducing leaf rust severity on the short arm of chromosome 1A of wheat detected across different genetic backgrounds and diverse environments. TAG 128:1579–1594

Eric MO, Pangirayi T, Paul S, Mwangi G, Abhishek R (2016) Correlations, path coefficient analysis and heritability for quantitative traits in finger millet landraces. Philippine J Sci 145(02):197–208

Estefan G, Sommer R, Ryan J (2013) Methods of soil, plant, and water analysis: a manual for the West Asia and North Africa Region. Third Edition, International center for agricultural research in the dry areas (ICARDA), Beirut, pp 84–105

Fan C, Xing Y, Mao H, Lu T, Han B, Xu C, Li X, Zhang Q (2006) GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. Theor Appl Genet 112:1164–1171. https://doi.org/10.1007/s00122-006-0218-1

FAOSTAT (Food and Agriculture Organization of the United Nations STAT) (2021) Available online: http://www.fao.org/FAOSTAT,/en/#data/QC [Accessed Jul 28 2022]

Fred OH, Sheunda P, Kibuka J, Kumar A, Rathore A, Manyasa E, Ajaku D (2021) Characterization of finger millet germplasm for mineral contents: prospects for breeding. J Cereals and Oilseeds 12(1):33–44

Gebremedhin T, Aschalew AY, Tsehay CT, Dellie E, Atnafu A (2021) Micronutrient intake status and associated factors among children aged 6–23 months in the emerging regions of Ethiopia: a multilevel analysis of the 2016 Ethiopia demographic and health survey. PLoS ONE 16(10):e0258954. https://doi.org/10.1371/journal.pone.0258954

Gebreyohannes A, Shimelis H, Laing M, Mathew I, Odeny DA, Ojulong H (2021) Finger millet production in Ethiopia: opportunities, problem diagnosis, key challenges, and recommendations for breeding. Sustainability 13:13463. https://doi.org/10.3390/su132313463

Gebreyohannes A, Shimelis H, Mashilo J, Odeny DA, Tadesse T, Ojiewo CO (2024) Finger millet (*Eleusine coracana*) improvement: challenges and prospects-a review. Plant Breeding. https://doi.org/10.1111/pbr.13169

Govindaraj M, Kanatti A, Rai KN, Pfeiffer WH, Shivade H (2022) Association of grain iron and zinc content with other nutrients in pearl millet germplasm, breeding lines, and hybrids. Front Nutr 8:746625

Gupta SM, Arora S, Mirza N, Pande A, Lata C, Puranik S, Kumar J, Kumar A (2017) Finger millet: a "certain" crop for an "uncertain" future and a solution to food insecurity and hidden hunger under stressful environments. Front Plant Sci 8:643

Hatti VRBK, Ramachandrappa BK, Sathishand A, Thimmegowda MN (2018) Soil properties and productivity of rain fed finger millet under conservation tillage and nutrient management in Eastern dry zone of Karnataka. J Env Biol 39(5):612–624

Hedden P (2003) Constructing dwarf rice. Nat Biotechnol 21:873–874. https://doi.org/10.1038/nbt0803-873

Hilu KW, de Wet JMJ, Harlan JR (1979) Archaeobotanical studies of *Eleusine coracana* ssp. coracana (finger millet). Amer J Bot 66:330–333. https://doi.org/10.2307/2442610

Hittalmani S, Mahesh HB, Shirke MD, Biradar H, Uday G, Aruna YR, Lohithaswa HC, Mohanrao A (2017) Genome and transcriptome sequence of finger millet (*Eleusine coracana* (L.) Gaertn.) provides insights into drought tolerance and nutraceutical properties. BMC Genom 18:1–16. https://doi.org/10.1186/s12864-017-3850-z

IBPGR (International board for plant genetic resources) (1985) Descriptor for finger millet. IBPGR Secretariat, Rome, pp 1–20

ICRISAT (International Crops Research Institute for Semi-Arid Tropics). (2024). ICRISAT genebank. http://genebank.icrisat.org/ [Accessed Mar 27, 2024]

Indiastat (2020) Socio-economic statistical data and facts about India. https://www.indiastat.com/ [Accessed Jan 20, 2022]

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet 11:94. https://doi.org/10.1186/1471-2156-11-94

Jordan DR, Hunt CH, Cruickshank AW, Borrell AK, Henzell RG (2012) The relationship between the stay-green trait and grain yield in elite sorghum hybrids grown in a range of environments. Crop Sci 52:1153–1161

Kaler AS, Purcell LC (2019) Estimation of a significance threshold for genome-wide association studies. BMC Genomics 20:1–8

Kassebaum NJ, Bertozzi-Villa A, Coggeshall MS, Shackelford KA, Steiner C, Heuton KR, Gonzalez-Medina D, Barber R, Huynh C, Dicker D, Templin T (2014) Global, regional, and national levels and causes of maternal mortality during 1990–2013: a systematic analysis for the global burden of disease study 2013. The Lancet 384(9947):980–1004

Krishnamurthy L, Upadhyaya HD, Purushothaman R, Gowda CL, Kashiwagi J, Dwivedi SL, Singh S, Vadez V (2014) The extent of variation in salinity tolerance of the minicore collection of finger millet (*Eleusine coracana* L. Gaertn.) germplasm. Plant Sci 227:51–59. https://doi.org/10.1016/j.plantsci.2014.07.001

Krishnamurthy L, Upadhyaya HD, Kashiwagi J, Purushothaman R, Dwivedi SL, Vadez V (2016) Variation in drought-tolerance components and their interrelationships in the minicore collection of finger millet germplasm. Crop Sci 56:1914–1926. https://doi.org/10.2135/cropsci2016.03.0191

Krishnan V (2024) Millets: superfood of the century–challenges and opportunities. Front Nut 11:1462567

Kudapa H, Barmukh R, Vemuri H, Gorthy S, Pinnamaneni R, Vetriventhan M, Srivastava RK, Joshi P, Habyarimana E, Gupta SK, Govindaraj M (2023) Genetic and genomic interventions in crop biofortification: examples in millets. Front Plant Sci 14:1123655. https://doi.org/10.3389/fpls.2023.1123655

Ligges U, Mächler M (2003) Scatter plot3d -an R package for visualizing multivariate data. J Stat Softw 8:1–20

Liu J, Chen J, Zheng X, Wu F, Lin Q, Heng Y, Tian P, Cheng Z, Yu X, Zhou K, Zhang X (2017) GW5 acts in the brassinosteroid signaling pathway to regulate grain width and weight in rice. Nat Plants 3(5):17043. https://doi.org/10.1038/nplants.2017.43

Lule D, de Villiers S, Fetene M, Odeny DA, Rathore A, Das RR, Tesfaye K (2018) Genetic diversity and association mapping of Ethiopian and exotic finger millet accessions. Crop and Pasture Sci 69(9):879–891. https://doi.org/10.1071/CP18175

Lule D, Tesfaye K, Fetene M, de Villiers S (2012) Inheritance and association of quantitative traits in finger millet (*Eleusine coracana* subsp. coracana) landraces collected from Eastern and South Eastern Africa. Int J Genet 2(2):12–21

Maqsood M, Ali AA (2007) Effects of drought on growth, development, radiation use efficiency and yield of finger millet (*Eleucine coracana*). Pak J Bot 39:123–134

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA (2010) The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20:1297–1303. https://doi.org/10.1101/gr.107524.110

Myers JR, Aljadi M, Brewer L (2018) The importance of cosmetic stay-green in specialty crops. In: Goldman I (ed) Plant breeding reviews. Wiley, New York, pp 219–256

Odeny DA, Niazi A, Tesfaye K, Lule D, Wanyonyi S, Kunguni JS (2020) Genomic designing for climate smart finger millet. In: Kole C (ed) Genomic designing of climate-smart cereal crops. Springer, Cham, pp 287–307

Ojulong HF, Sheunda P, Kibuka J, Kumar A, Rathore A, Manyasa E, Ajaku D (2021) Characterization of finger millet germplasm for mineral contents: prospects for breeding. J Cereals Oilseeds 12:33–44. https://doi.org/10.5897/JCO2020.0222

Owere L, Tongoona P, Derera J, Wanyera N (2015) Variability and trait relationships among finger millet accessions in Uganda. Uganda J Agric Sci 16(2):161–176

Pandey MK, Gangurde SS, Sharma V, Pattanashetti SK, Naidu GK, Faye I, Hamidou F, Desmae H, Kane NA, Yuan M, Vadez V (2020) Improved genetic map identified major QTLs for drought tolerance-and iron deficiency tolerance-related traits in groundnut. Genes 12(1):37. https://doi.org/10.3390/genes12010037

Pendergast IVTH, Qi P, Odeny DA, Dida MM, Devos KM (2022) A high-density linkage map of finger millet provides QTL for blast resistance and other agronomic traits. Plant Genome 15:e20175. https://doi.org/10.1002/tpg2.20175

Puranik S, Sahu PP, Beynon S, Srivastava RK, Sehgal D, Ojulong H, Yadav R (2020) Genome-wide association mapping and comparative genomics identifies genomic regions governing grain nutritional traits in finger millet (*Eleusine coracana* L. Gaertn). Plants, People, Planet 2(6):649–662. https://doi.org/10.1002/ppp3.10120

Purseglove JW (1985) Tropical crops: monocotyledons (Revised 1985). Longman, London, p 607

R Core Team (2021) R: a language and environment for statistical computing. Vienna: R Core Team. https://www.R-project.org/

Robinson HF, Comstock RE, Harvey PH (1949) Estimates of heritability and the degree of dominance in corn. Agron J 41(8):353–359. https://doi.org/10.2134/agronj1949.00021962004100080005x

Sanjana Reddy P, Reddy BV, Ashok Kumar A, Ramesh S, Sahrawat KL, Venkateswara Rao P (2010) Association of grain Fe and Zn contents with agronomic traits in sorghum. Indian J Plant Genet Resour 23:280–284

SAS (2008) Statistical analysis systems institute. Version 9.1, SAS institute Inc., Cary, North Carolina, USA.

Sela H, Ezrati S, Ben-Yehuda P, Manisterski J, Akhunov E, Dvorak J, Breiman A, Korol A (2014) Linkage disequilibrium and association analysis of stripe rust resistance in wild emmer wheat (*Triticum turgidum* ssp. dicoccoides) population in Israel.

Theor Appl Genet 127:2453–2463. https://doi.org/10.1007/s00122-014-2389-5

Sheoran S, Kumar S, Ramtekey V, Kar P, Meena RS, Jangir CK (2022) Current status and potential of biofortification to enhance crop nutritional quality: an overview. Sustainability 14(6):3301. https://doi.org/10.3390/su14063301

Shin D, Lee S, Kim TH, Lee JH, Park J, Lee J, Lee JY, Cho LH, Choi JY, Lee W, Park JH (2020) Natural variations at the stay-green gene promoter control lifespan and yield in rice cultivars. Nat Commun 11(1):2819. https://doi.org/10.1038/s41467-020-16573-2

Singh R, Van Heusden AW, Kumar R, Visser RG (2018) Genetic variation and correlation studies between micronutrient (Fe and Zn), protein content and yield attributing traits in mungbean (*Vigna. Radiata* L.). Legume Res Int J 41(2):167–174

Sivabharathi RC, Rajagopalan VR, Suresh R, Sudha M, Karthikeyan G, Jayakanthan M, Raveendran M (2024) Haplotype-based breeding: a new insight in crop improvement. Plant Sci 346:112129. https://doi.org/10.1016/j.plantsci.2024.112129

Sood S, Joshi DC, Rajashekara H, Tiwari A, Bhinda MS, Kumar A, Kant L, Pattanayak A (2023) Deciphering the genomic regions governing major agronomic traits and blast resistance using genome wide association mapping in finger millet. Gene 854:147115. https://doi.org/10.1016/j.gene.2022.147115

Tang Y, Liu X, Wang J, Li M, Wang Q, Tian F, Su Z, Pan Y, Liu D, Lipka AE, Buckler ES, Zhang Z (2016) GAPIT version 2: an enhanced integrated tool for genomic association and prediction. Plant Genome. https://doi.org/10.3835/plantgenome2015.11.0120

Taranto F, Esposito S, Fania F, Sica R, Marzario S, Logozzo G, Gioia T, De Vita P (2023) Breeding effects on durum wheat traits detected using GWAS and haplotype block analysis. Front Plant Sci 14:1206517. https://doi.org/10.3389/fpls.2023.1206517

Teklu D, Gashu D, Joy EJM, Amede T, Broadley MR (2023a) Effectiveness of agronomic biofortification strategy in fighting against hidden hunger. Agron 13:2173. https://doi.org/10.3390/agronomy13082173

Teklu D, Gashu D, Joy EJM, Lark RM, Bailey EH, Wilson L, Amede T, Broadley MR (2023b) Impact of zinc and iron agronomic biofortification on grain mineral concentration of finger millet varieties as affected by location and slope. Front Nutr 10:1159833. https://doi.org/10.3389/fnut.2023.1159833

Teklu D, Gashu D, Joy EJM, Bailey EH, Wilson L, Amede T, Broadley MR (2024) Differences in the nutritional quality of improved finger millet genotypes in Ethiopia. Sci Rep 14:460. https://doi.org/10.1038/s41598-023-48749-3

Tesfaye K, Mengistu S (2017) Phenotypic characterization of Ethiopian finger millet accessions (*Eleusine coracana* (L.) Gaertn), for their agronomically important traits. Acta Univ Sapientiae Agric Environ 9:107–118. https://doi.org/10.1515/ausae-2017-0010

Tsehaye Y, Berg T, Tsegaye B, Tanto T (2006) Farmers' management of finger millet (*Eleusine coracana* L.) diversity in Tigray, Ethiopia, and implications for on-farm conservation. Biodivers Conserv 15:4289–4308. https://doi.org/10.1007/s10531-005-3581-3

Turner SD (2018) qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. J Open Source Softw 3(25):731

Upadhyaya HD, Ramesh S, Sharma S, Singh SK, Varshney SK, Sarma NDRK, Ravishankar CR, Narasimhudu Y, Reddy VG, Sahrawat KL, Dhanalakshmi TN (2010) Genetic diversity for grain nutrients contents in a core collection of finger millet (*Eleusine coracana* (L.) Gaertn) germplasm. Field Crops Res 121(1):42–52. https://doi.org/10.1016/j.fcr.2010.11.017

Velu G, Tutus Y, Gomez-Becerra HF, Hao Y, Demir L, Kara R, Crespo-Herrera LA, Orhan S, Yazici A, Singh RP, Cakmak I (2017) QTL mapping for grain zinc and iron concentrations and zinc efficiency in a tetraploid and hexaploid wheat mapping populations. Plant Soil 411:81–99. https://doi.org/10.1007/s11104-016-3025-8

Wang X, Wang C, Sheng H, Wang Y, Zeng J, Kang H, Fan X, Sha L, Zhang H, Zhou Y (2017) Transcriptome-wide identification and expression analyses of ABC transporters in dwarf polish wheat under metal stresses. Biol Plant 61:293–304

Wolie A, Dessalegn T (2011) Correlation and path coefficient analyses of some yield related traits in finger millet (*Eleusine coracana* (L.) Gaertn.) germplasms in Northwest Ethiopia. Afric J Agric Res 6(22):5099–5105

Xiang J, Apea-Bah FB, Ndolo VU, Katundu MC, Beta T (2019) Profile of phenolic compounds and antioxidant activity of finger millet varieties. Food Chem 275:361–368. https://doi.org/10.1016/j.foodchem.2018.09.120

Yamunarani R, Govind G, Ramegowda V, Thammegowda HV, Guligowda SA (2016) Genetic diversity for grain Zn concentration in finger millet genotypes: potential for improving human Zn nutrition. Crop J 4(3):229–234. https://doi.org/10.1016/j.cj.2015.12.001

Yang T, Dong J, Zhao J, Zhang L, Zhou L, Yang W, Ma Y, Wang J, Fu H, Chen J, Li W (2023) Genome-wide association mapping combined with gene-based haplotype analysis identify a novel gene for shoot length in rice (*Oryza sativa* L.). TAAG 136(12):251. https://doi.org/10.1007/s00122-023-04497-6

Ye L, Liu L, Xing A, Kang D (2013) Characterization of a dwarf mutant allele of arabidopsis MDR-like ABC transporter AtPGP1 gene. Biochem Biophys Res Commun 441(4):782–786

Yogeesh LN, Naryanareddy AB, Nanjareddy YA, Gowda MVC (2016) High temperature tolerant genotypes of finger millet (*Eleusine coracana* L.). Nat Environ Pollut Technol 15:1293–1296

Zhang H, Hall N, Goertzen LR, Bi B, Chen CY, Peatman E, Lowe EK, Patel J, McElroy JS (2019) Development of a goose grass (*Eleusine indica*) draft genome and application to weed science research. Pest Manag Sci 75(10):2776–2784. https://doi.org/10.1002/ps.5389

Zhu H, Zhang M, Sun S, Yang S, Li J, Li H, Yang H, Zhang K, Hu J, Liu D, Yang L (2019) A single nucleotide deletion in an ABC transporter gene leads to a dwarf phenotype in watermelon. Front Plant Sci 13(10):1399

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