

Genome-wide association analysis revealed genetic markers linked to grain yield and yield related traits in finger millet [*Eleusine coracana* (L.) Gaertn] grown in acidic soils

Haftom Brhane

`haftom.mesfin@slu.se`

Swedish University of Agricultural Sciences: Sveriges lantbruksuniversitet

Teklehaimanot Haileselassie

Kassahun Tesfaye

Dagnachew Lule

Kibrom B. Abreha

Cecilia Hammenhag

Rodomiro Ortiz

Mulatu Geleta

Research Article

Keywords: Al-tolerance, finger millet, GBS, GWAS, linkage disequilibrium, stress tolerance

Posted Date: July 23rd, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-4607859/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Version of Record: A version of this preprint was published at Plant and Soil on October 8th, 2024. See the published version at <https://doi.org/10.1007/s11104-024-07000-2>.

Abstract

Aim

Finger millet is a nutritionally rich crop cultivated in the arid and semi-arid regions of Africa and Asia. Soil acidity has a major impact on the crop's yield and productivity as tolerant cultivars that perform well in acidic soils are limited. This study aimed at evaluating major finger millet phenotypic traits under acidic soils followed by identifying associated markers.

Method

A total of 288 finger millet genotypes were field evaluated for 8 major phenotypic traits including grain yield under acid soil conditions at Bako and Gute field sites in Ethiopia. In addition, the same genotypes were subjected to genotyping-by-sequencing to generate single nucleotide polymorphism markers to analyze association mapping.

Results

Data analysis revealed significant phenotypic variation in all targeted traits among the studied genotypes. Genotypes Ec-100093 (322.7 g), Ec-215803 (310.2 g), and Ec-203322 (301.2 g) were relatively high-yielding, whereas genotypes Ec-229721 (31 g) and Ec-242110 (31.2 g) had the lowest grain yield across the two locations. The broad-sense heritability of the traits ranged from 0.04 for the number of effective tillers (NET) to 0.78 for days to emergence (GY). The marker-trait association analysis using Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway model revealed 23 SNP markers significantly associated with one or more traits. Among the 273 significant markers, one associated with days to emergence, seven with days to heading, four with days to maturity, three with grain yield (GY), two with ear length, one with plant height, two with number of fingers, and three with the number of effective tillers (NET).

Conclusions

Once validated with further research, the identified novel markers associated with the targeted traits will potentially be useful for genomics-driven finger millet improvement.

Introduction

Finger millet (*Eleusine coracana* L. Gaertn.) is a small seeded cereal crop cultivated in most arid and semi-arid regions of Africa and Asia for food and feed (Goron and Raizada 2015). It is the third most important millet crop after pearl millet and foxtail millet in the world (Singh and Upadhyaya 2015). It is a hardy crop capable of providing reasonable grain yields under biotic and abiotic stresses where most other crops generally fail, and it is often considered a component of food security crop (Kinfe et al. 2017). Finger millet

originated in East Africa, particularly Ethiopia and Uganda, and then extended to India about 3000 years ago (Hilu and De Wet 1976). In Ethiopia, finger millet is widely grown and used by small-scale farmers (Admassu et al. 2009, Kinfu et al. 2017).

Finger millet is a disomic tetraploid ($2n = 4x = 36$; AABB genome) crop and is considered to be originated from hybridization between the diploid species *E. indica* (L.) Gaertn ($2n = 2x = 18$; AA genome) and *E. floccifolia* (Forssk.) Spreng ($2n = 2x = 18$; BB genome) (Bisht and Mukai 2001). It exhibits excellent phenotypic diversity in grain color (dark brown, light brown, radish brown, and white), growth habit (erect, recumbent, and prostrate), panicle shape (open, semi-curved, and curved), and flowering time (Upadhyaya et al. 2007).

Generally, crop improvement involves various complex activities including utilization of existing genetic variation in desired traits and marker-trait associations (MTA) (Mnyenyembe and Gupta 1998). However, research on finger millet has been sparse, mostly focusing on genetic diversity using different morphological (Daba and Haile 2000, Tesfaye and Mengistu 2017), DNA markers (Lule et al. 2018, Brhane et al. 2021, Brhane et al. 2022a), and RNA-seq analysis (Brhane et al. 2022b). One among many factors that have negatively affected advanced research, such as MTA analysis and genetic improvement of finger millet was a lack of reference genome assembled at a chromosome level (Sharma et al. 2018). Recently, a chromosome-level assembly of the finger millet genome has been published, opening up opportunities for advanced research and genomics-driven breeding in this crop (Devos et al. 2023). Single nucleotide polymorphisms (SNPs) have been used to conduct genome-wide association (GWAS) research in various crops including finger millet due to their abundance, co-dominant nature, and high reproducibility (Sharma et al. 2018, Puranik et al. 2020). GWAS relies on next-generation sequencing-based genotyping-by-sequencing (GBS) that allows simultaneous SNP discovery and genotyping of target populations (Elshire et al. 2011). GBS provides a large number of SNP markers that can be used for MTA (Sharma et al. 2018, Kavuluko et al. 2021, Saleem et al. 2022).

Nowadays, GWAS is the most widely used approach for associating phenotypic traits with the underlying genetic background of various crops (Ward et al. 2019). In finger millet, not much research has been published on MTA to identify genomic/scaffold regions governing nutritional traits, blast resistance genes, seed protein governing genes, and genes underlining grain iron (Fe) and zinc (Zn) (Sharma et al. 2018, Puranik et al. 2020, Tiwari et al. 2020, Sood et al. 2023, Chandra et al. 2024). Due to their statistical simplicity and speed of computation, multi-locus mixed model (MLMM) based fixed and random model circulating probability unification (FarmCPU) and Bayesian-information and linkage-disequilibrium Iteratively Nested Keyway (Blink) with sub-setting kinship estimation mostly have been employed in MTA analysis (Liu et al., 2016, Huang et al., 2019). In addition, Blink also eliminates the assumption that causal genes are evenly distributed across the genome, thereby increasing statistical power (Wang and Zhang, 2021). However, no research has been done to uncover the genomic regions and genes regulating yield and yield-related traits of finger millet grown under acidic soil conditions or Al toxicity. Hence, this study aimed at evaluating major finger millet phenotypic traits through replicated field trials at two locations with acidic soils, followed by GWAS utilizing their phenotypic data and GBS-based SNP marker data. This was done to

identify favorable alleles, which could be incorporated into finger millet cultivars through breeding, thereby enabling breeders to improve the crop's grain yield in acidic soils.

Materials and methods

Plant materials

In this study, 288 accessions of finger millet were used; of which 274 accessions were obtained from the gene bank of the Ethiopian Institute of Biodiversity (EIB, Addis Ababa, Ethiopia) while the remaining 14 accessions were cultivars obtained from Bako Agricultural Research Center (BARC) Bako, Ethiopia. Among the gene bank accessions, 228 were landraces originally collected from various regions in Ethiopia, i.e., Amhara (130), Benishangul-Gumuz (2), Oromia (51), SNNPR (4), Tigray (32), and unknown locations (9). The remaining 46 gene bank accessions represent Zimbabwean finger millet genetic resources imported to Ethiopia for various purposes including research. For simplicity, the accessions will be referred to as "genotypes" herein.

Planting, leaf tissue sampling, genomic DNA extraction, and seed multiplication

The seeds of the 288 genotypes were planted in a greenhouse at the Swedish University of Agricultural Sciences (SLU, Alnarp, Sweden) to obtain an adequate number of progeny seeds for field trials. For each genotype, five seeds were planted in a 5 L plastic pot. Upon germination, extra seedlings were removed, and only one seedling was retained per genotype and used for genotyping and seed multiplication.

The leaf tissue of each genotype was collected separately three weeks after planting, DNA was extracted, and GBS-based genotyping was performed as described in Brhane et al. (2022a). To summarize, fresh leaf tissue samples were collected using the BioArk Leaf collection kit and then sent to LGC, Biosearch Technologies (Berlin, Germany) for genomic DNA extraction followed by GBS-based genotyping. High molecular weight genomic DNA was extracted using the Sbeadex plant kit (<https://biosearch-cdn.azureedge.net/assetsv6/sbeadex-plant-data-sheet.pdf>).

The seedlings were further grown to maturity under greenhouse conditions maintained at a light/dark cycle of 11/13 hours day/night temperature of 18°C and 16°C, respectively and 65% relative humidity. Since finger millet is self-pollinating and one plant from each accession was sampled for genotyping, seeds harvested from individual plants were used for the field trials for acidity tolerance evaluation and phenotyping (described below).

Library construction, sequencing and read pre-processing

For the construction of a GBS library, *Pst*I (CTGCA*G, a 6-base cutter) and *Ape*KI (G*CWGC, a 4-base cutter) restriction enzymes were used following the recommendation of the experts at LGC Biosearch Technologies who optimized restriction enzymes for various crops. By combining *Pst*I-*Ape*KI, a fragment size distribution suitable for sequencing on Illumina platforms was produced with a mean insert size of about 220 base pairs. The GBS reads were generated in 2 × 150 bp (paired-end) sequencing mode using

NexSeq 500/550 v2 and NovaSeq SP FC NGS platforms. There were approximately 1.5 million read pairs per sample. After adapter-clipping, N-containing reads and reads that did not match the restriction enzyme site at the 5' end were discarded. To obtain an average Phred quality score of 20 or greater over a window of 10 bases, the 3'-ends of the remaining reads were trimmed using Trimmomatic v. 0.3 (Bolger et al. 2014), and reads with final lengths of less than 20 bases were discarded.

Read alignment, SNP discovery and genotype calling, and filtering

Since the recently released chromosome-level assembly of the finger millet genome (https://www.ncbi.nlm.nih.gov/datasets/genome/GCA_032690845.1/) was not available by the time this study was carried out, the scaffold-level genome assembly (https://www.ncbi.nlm.nih.gov/assembly/GCA_002180455.1/) of the ML-365 cultivar was used as the reference genome (Hittalmani et al. 2017). An alignment of quality-trimmed reads against the reference scaffold-level assembled genome was performed using the BWA-MEM software package v. 0.7.12 (Li and Durbin 2009). Over 99% of the reads were mapped to the reference genome (which contained 525,627 scaffolds spanning 1.3 GB). For variant discovery and genotype calling in diploid format, Freebayes v. 1.2.0 (Garrison and Marth 2012) was used, and 101,889 SNPs were discovered across the 288 finger millet genotypes. GBS-specific criteria were then applied to filter the SNPs (minimum read count > 8, minimum allele frequency (MAF) of 0.05, and percentage of samples with assigned genotype > 66%). The data were further filtered, as described in Brhane et al. (2022a).

A total of 5226 bi-allelic SNPs with below 5% missing data and below 0.05 minor allele frequency (MAF) were used for marker-trait association studies. To generate Manhattan plots, the 5226 SNPs were equally divided into nine scaffold groups (1–9) due to the lack of genome assembly at the chromosome level in finger millet at the time of genotyping. *In silico* analysis of the significantly associated markers was performed on the newly published chromosome-level assembly of finger millet (Devos et al. 2023).

Filed trials and phenotyping

The field trials were carried out during the 2021 main cropping season at two locations with acidic soils in Western Ethiopia where soil acidity is paramount: Bako Agricultural Research Center (BARC) and Gute Research Sub-Center (GRSC). BARC is located at 9°6' N latitude and 37°09' E longitude with an altitude of 1650 m.a.s.l. Gute research sub-center is at 09°01.06' N and 036°38.196' E with an altitude of 1915 m.a.s.l. The dominant soil type in these locations is Nitisol, which is characteristically reddish brown and loamy in texture with pH falling in the range of very strongly acidic to weak alkaline (Keba et al. 2022). They have similar patterns of rain during the rainy season from April to October (Kebede et al. 2019).

The experimental layout for the field trials was an alpha lattice design, comprising 17 blocks with 17 plots each, with two replications at both sites. Seeds of each genotype were sown by hand in two rows of 1 m length on each plot. The spacing between rows within a plot, between plots within a block, and between blocks within a replicate was 40 cm, 60 cm, and 80 cm, respectively. The distance between replicates within a site was 2 m. A seed rate of 15 kg ha⁻¹ and fertilizer rate of 105 kg ha⁻¹ DAP and 65 kg ha⁻¹ urea

were used at both sites. Plant management practices, such as hand-weeding were carried out according to adopted recommendations. The traits studied were days to emergence (DE), days to heading (DH), days to maturity (DM), ear length (EL), plant height (PH), number of fingers (NF), number of effective tillers (NET), and grain yield (GY). Phenotypic data of the traits were collected from five randomly chosen and tagged plants as described by Belete et al. (2020).

Phenotypic data collection

Quantitative data were recorded for days to 50% emergency, days to 50% heading, days to 50% maturity, plant height (cm), finger length (cm), number of fingers per main ear, effective number of tiller per plant, grain yield per plant (g) following the finger millet descriptor (IBPGR, 1985).

Phenotypic data analysis

Data collected for all quantitative characters were subjected to analysis using R-software. Descriptive statistics of the phenotypic data of finger millet genotypes grown under two acidic soil environments in two replications were analyzed in R software using describe By by function. The best linear unbiased estimates (BLUEs), adjusted mean values of all traits from both locations were calculated using the using ASReml package in R and used for downstream analyses. Broad sense heritability was estimated using using the ASReml package in R. In addition, Best Linear Unbiased Predictions (BLUPs) was used to estimate variance components using the ASReml package in R. Correlation analysis was also performed to assess the extent of covariation among them across environments using the R software corrplot package.

Discriminant analysis of principal Components (DAPC)

The population structure and clustering analysis of the genotypes were performed using discriminant analysis of principal components (DAPC) approach using the adgenet package in R software (Jombart and Collins 2015). Bayesian Information Criterion (*BIC*) for $K = 1-10$ (K = number of populations) was used to determine the optimal number of populations based on the minimum observed BIC value.

Linkage disequilibrium

Stringently filtered SNP markers were used to conduct linkage disequilibrium (LD) analysis using Trait Analysis by Association, Evolution, and Linkage (TASSEL v. 5) with the default settings. Pairwise squared allele-frequency correlations (r^2) between SNP markers were generated, and the r^2 values were then plotted against the physical distance between the SNP loci to estimate the extent of LD between pairs of loci using R software.

Genome-wide association study (GWAS)

Genome-wide association study (GWAS) was carried out to determine the association between phenotypic traits collected from field trials and stringently filtered GBS-derived SNP markers. This marker-trait analysis (MTA) was conducted using the Blink models in the Genome Association and Prediction Integrated Tools (GAPIT) package in R software (Lipka et al. 2012, Team 2013). For this purpose, we used 5226 bi-allelic SNPs with MAF above 0.05 and missing values below 5%. MTAs with $p \leq 0.001$ (i.e. $-\log_{10} p \geq 3.00$) and false discovery rate (FDR) < 0.1 were considered significant. The percentage of phenotypic variance

explained (PVE) by each marker was also estimated. The p -values from the model were used to generate Manhattan and quantile-quantile (Q-Q) plots using R software (Turner 2018). The sequences flanking the significantly associated SNP markers (149 kb upstream and downstream of the markers) were extracted and BLAST searched for sequence similarity against the Phytozome 13 *Eleusine coracana* v1.1 database (E.coracana v1.1: Phytozome (doe.gov)) using default parameters to identify putative candidate genes linked to the significant markers.

Results

General phenotypic performance of the genotypes

The analyses of finger millet phenotypic data collected from the two field experimental sites revealed a wide range of variation in the traits analyzed. The genotypes started emerging within 5 to 7 days, with most of them emerging (50%) within five days (Fig. 1A). In addition, the days to heading of the genotypes ranged between 82 to 110 days with an overall mean of 95 days (Fig. 1B). Days to maturity (DM) varied from 152.5 and 176 with an overall mean of 163.3 days. The early maturing genotypes were Ec-211506, Ec-203272, and Ec-203327, reaching maturity in 152.5, 153.5, and 154.5 days, respectively. In contrast, the late maturing genotypes were Ec-243642, Ec-240506, Ec-215888, Ec-215850, Ec-215847, and Ec-208729, reaching maturity in 171.6 days, respectively (Fig. 1C).

The plant height (PH) of the genotypes varied from 59.2 cm to 101.2 cm with an overall mean of 80.01 cm. The top three tallest genotypes were Ec-2215952, Ec-216039, and Ec-203322 with 101.2, 98.4, and 97.4 cm, respectively. Whereas, the shortest genotypes were 203344, 203357, and 237969 with 59.2, 59.9, and 61.1 cm, respectively (Fig. 1D). The number of leaves and the number of fingers were also counted from five individuals per genotype. The number of leaves per plant ranged from 4.5 to 9.5, while the number of fingers per plant was between 3.7 and 7.5 (Fig. 1E and 1F, respectively). The number of effective tillers per plant across genotypes ranged from 1.7 to 4.7 (Fig. 1G).

The highest average grain yield (GY) per plot was recorded for genotype Ec-100093 (322.7 g), followed by Ec-215803 (310.2 g) and Ec-203322 (301.2 g). In contrast, genotypes Ec-229721 and Ec-242110 had the lowest average grain yields of 31 g and 31.2 g, respectively. The overall mean grain yield of the genotypes was 115.6 g. The top three high-yielding genotypes were originally collected from Wollega (Ec-100093) and Gojam (Ec-215803) in Ethiopia, and Zimbabwe (Ec-203322). Among the cultivars Bako-09, Gute, and Wama performed better in the acidic environment than the other cultivars with grain yields of 150.6 g, 151.3 g, and 149.6 g, respectively (Fig. 1H).

Variance components and heritability

The combined variance components of the eight quantitative traits collected across 288 finger millet genotypes, grown at two sites with acidic soils, revealed significant variation among genotypes in all traits. Among the variance components, the genotypic variance was higher in DE, and GY traits compared to the environmental variance. However, it was lower in other traits, such as DH, DM, PH, and NF (Table 1).

Comparatively, the broad sense-heritability (H^2) of the traits was higher in DE (0.86), followed by GY (0.68), DM (0.67), DH (0.49), EL (0.35), PH (0.33), while low in NF (0.09) and NET (0.04) (Table 1).

Table 1
Estimation of the different variance parameters for 8 major quantitative traits of 288 finger millet genotypes evaluated under acidic soil condition

Mean Squared Error									
Source of variation	DF	DE	DH	DM	PH	NF	EL	NET	GY
Environments (ENV)	1	0.01	31764*	26922*	5206*	253.00*	34.48*	0.28	8511*
Genotype (GEN)	284	1.107*	123**	104**	264**	1.52**	9.25**	1.27**	9122**
REP (Location)	2	18.75	405	89	1239	39.52	178.6	106.6	8899
GEN × ENV interaction	284	1.13	151*	72*	525*	5.95*	30.97*	4.62**	5934*
Residuals	811	0.30	52	31	144	1.21	1.21	0.94	2312
Heritability		0.86	0.49	0.67	0.33	0.09	0.35	0.04	0.68
Mean		5.8	95.3	163.2	80.01	5.31	5.31	3.22	115.7
**, * Significant difference at $P < 0.01$ and $P < 0.05$, respectively. DF = Degrees of freedom									

Correlation analysis

Different correlations between the phenotypic traits was obtained with the highest significant positive correlation was observed between DH and DM ($r = 0.78$), followed by between DE and DM ($r = 0.66$), and DE and DH ($r = 0.54$) (Fig. 2). Plant height (PH) showed intermediate positive correlations for all traits except with NET and EL, which were not correlated with PH (Fig. 2). Moreover, GY was significantly positively correlated with PH ($r = 0.23$), EL ($r = 0.32$) and NET ($r = 0.15$), while no correlation was observed with the other phenotypic traits (Fig. 2).

Population structure and linkage disequilibrium

The discriminant analysis of principal components (DAPC) of the 288 finger millet genotypes based on filtered SNP markers, revealed that the genotypes were grouped into six clusters, each representing different genetic populations (Fig. 3a). Using the retained six clusters and 200 PCs, a scatter plot was generated to visualize the placement of each genotype in the scatter plot. Clusters I comprised of 17 genotypes (16 from Pop-3 and 1 from Pop-7), Clusters II comprised of 52 genotypes (45 from Pop-1 and 7 from Pop-2), and Clusters V comprised of 50 genotypes (3 from Pop-1, 1 from Pop-3, 2 from Pop-6, and 44 from Pop-7), respectively, and were dominated by genotypes from a single population. On the other hand, clusters III, IV, and VI comprised 33 genotypes (22 from Pop-1, 10 from Pop-3, and 1 from Pop-6), 48

genotypes (22, 17, 2, 6, and 1 from Pops-1, 2, 3, 5, and 7, respectively), and 88 genotypes (13, 7, 26, 28, 3, and 11 from Pops-1, 2, 3, 4, 5, and 6, respectively), and were more heterogeneous than the other clusters. The analysis revealed that genotypes from Zimbabwe were clustered separately, whereas weak clustering was observed for genotypes from Ethiopia (Fig. 3b).

The 5226 filtered SNP markers located across the finger millet genome scaffolds formed 154,626 SNP pairs. About 53,543 (34.6%) of the SNP pairs had $r^2 \geq 40\%$ while 19.2% were in LD ($r^2 = 0.1$, $P < 0.05$). The LD begins to decay at $r^2 = 0.46$ and drops to its half-decay at $r^2 = 0.1$ at a distance of 149 kb between marker pairs (Fig. 4).

Marker-trait association analysis

Marker-trait association analysis using the BLINK model was employed and significantly associated SNP markers were identified from the finger millet quantitative traits. This analysis resulted in 73 SNP markers significantly associated with the traits ($FDR < 0.05$) after a Bonferroni correction was made. Among the 23 SNP markers significantly associated with the phenotypic traits, one SNP associated with DE, seven with DH, four with DM, one with PH, 11 with NF, two with EL, three with NET and three with GY (Fig. 5 (A-H)). The significantly associated SNP marker (LXGH01314605.1 with DE phenotypic trait was located at the position of 5542 in the finger millet scaffold which explained 0.52% of the total phenotypic variance. Seven markers were associated with DH phenotypic trait which explained phenotypic variance of from 0.56–6.50%. The alleles AA, AT, and TT were recorded at this locus with the TT genotypes significantly early heading than the AA genotypes (Fig. 6A). Four markers were associated with DM trait in the crop explaining phenotypic variance of between 0.68 and 6.80% (Table 2). The significant marker-trait associations for DE, DH, and DM obtained using the Blink model were graphically displayed in a Manhattan plot with their corresponding QQ plot (Fig. 5A, B and E), respectively.

Marker trait association analysis was also performed to find significantly associated SNP markers with PH. This resulted in one significantly ($P < 0.05$) associated markers with MAF varying from 0.06 to 0.39. However, the PVE value of the marker 0.89% (Table 2). The SNP marker that explained the highest PVE (0.94) for PH was located on scaffold LXGH01346160.1 at position 2819 (Table 2). The significant marker-trait associations for PH obtained using the Blink model were graphically displayed in a Manhattan plot with their corresponding QQ plot (Fig. 5F). All three genotypes (AA, AG, and GG) were recorded at this locus with the GG genotypes significantly taller than the AA genotypes (Fig. 6B).

A total of two SNP markers distributed across different scaffolds had a significant association with the NF of which five explained between 0.62 and 0.80% (Table 2). The MAF of the significantly associated markers varied from 0.05 to 0.48. The significant marker-trait associations for NF obtained using the Blink model were graphically displayed in a Manhattan plot with their corresponding QQ plot (Fig. 5C). Two markers were also linked with EL trait. Manhattan plot with their corresponding QQ plot for the EL also provided in Fig. 5D. Three SNPs showed a significant association with the number of effective tillers (NET) each with a PVE > 0.05 (Table 2). The associated markers explained 0.54 to 0.80% of the phenotypic variance. The markers are distributed across different scaffolds with MAF ranging from 0.05 to 0.27. The significant

marker-trait associations for NF obtained using the Blink model were graphically displayed in a Manhattan plot with their corresponding QQ plot (Fig. 5G).

Three markers were significantly associated with GY finger millet trait. The markers explained from 1.38 to 30.8% of the trait (Table 2). These markers were located on different scaffolds. The SNP marker that explained the highest variation in grain yield (30.8%) was located on scaffold LXGH01052700.1 at position 28450. All three genotypes (AA, AG, and GG) were recorded at this locus with the AA genotypes having a significantly higher yield than AG genotypes (Fig. 6C). The Manhattan plot and its corresponding QQ plot for SNP markers significantly associated with GY are presented in Fig. 5H.

Two SNP markers were significantly associated with more than one trait. Table 2 provides details for 23 significant markers with PVE values greater than 0.5 including scaffold ID, SNP alleles, SNP position within a scaffold, P -value, and MAF for each trait. Among the significantly associated markers, those strongly associated with their corresponding traits and explained a relatively high proportion of phenotypic variance were annotated based on finger millet genomic resources in the public databases.

Table 2

List of 23 SNP markers significantly associated ($P < 0.05$) with finger millet traits targeted in this study that explained more than 0.5% of the total variance in their corresponding traits.

Trait	Scaffold ID	alleles	Position	P-value	MAF	PVE (%)
DE	LXGH01314605.1	T/C	5542	0.004	0.498	0.52
DH	LXGH01208436.1	T/A	29293	6.2E-05	0.380	0.677
	LXGH01452763.1	A/T	113304	0.00029	0.053	0.807
	LXGH01485647.1	A/T	1718	0.00050	0.362	9.96
	LXGH01183671.1	T/C	43941	0.00054	0.159	0.56
	LXGH01520281.1	A/G	3954	0.00069	0.156	6.50
	LXGH01099718.1	T/C	3758	0.00083	0.154	2.71
	LXGH01074376.1	G/T	440	0.00095	0.442	1.49
DM	LXGH01139662.1	G/T	19176	0.003	0.179	0.685
	LXGH01369512.1	A/G	5542	0.002	0.080	0.779
	LXGH01384244.1	C/G	9994	0.001	0.061	0.846
	LXGH01337172.1	A/G	20956	0.001	0.138	6.80
GY	LXGH01452763.1	T/G	49927	0.000	0.406	1.387
	LXGH01052700.1	A/G	28450	0.011	0.095	30.84
	LXGH01346143.1	C/G	32418	0.000	0.06	1.40
EL	LXGH01311505.1	A/G	2125	0.003	0.054	0.838
	LXGH01501788.1	T/C	2354	0.007	0.214	0.921
NET	LXGH01386342.1	T/C	28492	0.002	0.068	0.546
	LXGH01038485.1	A/G	14232	0.002	0.068	0.600
	LXGH01368953.1	T/C	39929	0.001	0.229	1.475
NF	LXGH01314605.1	T/C	5542	0.011	0.146	0.624
	LXGH01281367.1	T/C	23526	0.007	0.149	0.888
PH	LXGH01346160.1	A/G	2819	0.008	0.06	0.888

Minor allele frequency = MAF; SNP position in the scaffold = Position; false discovery rate = FDR; phenotypic variance explained = PVE; DE = days to emergence; DH = days to heading; DM = days to maturity; PH = plant height; EL = ear lengths; NF = number of fingers; NET = number of effective tillers; and GY = grain yield.

Discussion

In the present study, a wide range of grain yield variation (31–322.7 g/ m²) with an average yield of 115.6 g/m² was obtained among finger millet genotypes grown in environments with acidic soils. The observed variation in grain yield is in contrast with previously reported grain yield of finger millet grown at Adet and Finoteselam sites in Ethiopia, which yielded 239.7 g/m² and 238.6 g/m², respectively (Marefia et al. 2022). Likely, the difference in grain yields obtained in our study compared to the above report may be due to the acidic soils in Gute and Bako, while the soils at Adet and Finoteselam are not acidic and are therefore expected to produce higher grain yields. Interestingly, the top five high-yielding finger millet genotypes were originally collected from Wollega and Gojam (western Ethiopia), whereas most of the low-yielding genotypes were originally collected from Tigray and Gonder (northern Ethiopia). The genotypes originally collected from areas with acidic soils were high yielding, while those from areas with non-acidic soils were low yielding, and, hence, can be regarded as Al-tolerant and Al-susceptible, respectively. This could be mainly due to enhanced Al-tolerance of genotypes acquired through long-term exposure to acidic soils. Tolerance mechanisms could be by exclusion of Al from the root, sequestration of Al to different plant parts once it enters the plant cell, or both (Barcelo and Poschenrieder 2002, Kochian et al. 2005). Future research based on well-designed experiments will shed light on the tolerance mechanisms of the genotypes investigated in this current study. The consistently high-yielding genotypes across the two environments Ec-215829 (243.5 g), Ec-203322 (228.0 g), and Ec-100093 (213.5 g) should be considered in the finger millet breeding program for developing improved cultivars, particularly targeting acidic soils.

Heritability measures the transmission of characters from parents to their offspring (Falconer 1981). In this study, days to emergence, days to heading, days to maturity, and number of grains per spikelet showed higher broad sense heritability (H^2) compared to plant height, ear length, and number of fingers. A similar heritability pattern among diverse finger millet germplasm was previously reported (Marefia et al. 2022). The moderate and high heritability observed in most of the targeted traits indicates that genotypes have higher effects than environments on their phenotypic variation, unlike traits with low heritability where environmental effects are high.

The correlations between most traits obtained in this study were in line with previous reports (Bharathi 2011, Sharma et al. 2018), where significant positive correlations of DH with EL, DM, and PH were reported. Likewise, the positive correlation between PH and GY obtained in this study agreed with previous findings (Sivagurunathan 2004, Sharma et al. 2018). These investigations showed that finger millet genotypes that grow taller in acidic soils produce higher grain yields. Interestingly, this study revealed genotypes with desirable characteristics in major target traits including grain yield, plant height as well as tolerance to soil acidity. Hence, crossbreeding of genotypes with these desirable characteristics (high grain yield, tall height, and Al-tolerant) will increase the likelihood of recombination of alleles with positive effects on the phenotypes, and eventually developing superior cultivars.

The DAPC and population structure analysis showed that the genotypes originated from six genetic populations ($K = 6$), which is in line with previous findings in finger millet (Brhane et al. 2022a). The detailed

analysis of genetic diversity and population structure of the genotypes used in this study, including their potential significance for finger millet improvement, was previously published (Brhane et al. 2022a). They represent genotypes from diverse agro-ecozones in Ethiopia and Zimbabwe, as well as cultivars released in Ethiopia. Overall, the genotypic and phenotypic variation, and population structure determined in this set of genotypes make it suitable for GWAS to identify genomic regions and candidate genes potentially associated with AI-tolerance.

Determination of LD between markers is an important step in GWAS, as it facilitates the selection of a set of markers suitable for efficient marker-trait association analysis. The LD between markers varies within a genome as well as between genomes of different crops, as it is affected by various factors, such as reproductive mechanism, rate of recombination, selection pressure, genetic drift, physical linkage, and population structure (Puranik et al. 2020). Since the LD analysis in the present study was based on SNP markers across a large number of scaffolds, it was difficult to estimate the markers' rate of LD decay accurately, as was also indicated in previous research (Sharma et al. 2018). In this study, about 19.2% of the SNP marker pairs showed significant LD ($r^2 > 0.1$; $P < 0.05$). A similar finding was reported by Sharma et al. (2018), who reported that 17.9% of finger millet SNP marker pairs were in LD ($r^2 > 0.1$; $p < 0.05$). Another study on finger millet by Puranik et al. (2020) indicated that about 16.8% of finger millet SNP pairs were at LD ($r^2 \geq 0.2$; $p < 0.05$), and the maximum LD was dropped to its half (LD half-decay) at a distance of 28 kbp. Comparatively, genome-wide LD decay in foxtail millet ranged from 100 to 177 kb (Jaiswal et al. 2019).

In this study, 5226 GBS-derived SNP markers and eight phenotypic traits were used to identify significant marker-trait associations (MTA). The analysis revealed 73 novel SNP markers associated with one or more target phenotypic traits at highly significant levels. Sharma et al. (2018) reported 109 SNP markers associated with 14 finger millet agro-morphological traits including grain yield using SLST = 20, MLMM = 36, and MTMM = 53 models. Another GWAS study on finger millet identified 418 SNP markers associated with nutrition-related traits using GLM and mixed MLM models (Puranik et al. 2020). The relatively low number of MTAs identified in this study could be due to factors, such as the number of phenotypic traits studied and the stringency of the GWAS model used. Overall, this study showed that the GBS-derived SNP marker set used is suitable to identify potential markers suitable for marker-aided selection of genotypes for breeding finger millet for cultivation on acidic soils.

Among SNP loci associated with DH, an SNP marker located at position 1528 in scaffold LXGH01485647.1 explained the highest phenotypic variance in this trait. The alignment of this scaffold to the recently released finger millet chromosome-level assembly Devos et al. (2023) showed that this SNP is located on chromosome 5A within the ELEC0.r07.5AG0380680 gene, which codes for Carboxyspermidine synthase /Carboxyspermidine dehydrogenase in finger millet. Carboxyspermidine synthase is an enzyme used for biofilm formation and cell viability in microbes through cationic polyamine spermidine (Ko et al. 2022). Polyamine spermidine promote plant growth in part by the induction of systemic resistance (ISR) to biotic and abiotic stresses (Melnyk et al. 2019). The genotypes at this locus are AA, AT, and TT, which showed significant phenotypic variation in days to heading, and AA genotypes started heading later than TT and AT genotypes.

Among SNP loci significantly associated with GY, LXGH01052700.1_28450 was the most significant, as it explained 30.8% of grain yield variation. The locus is located within the ELECO.r07.1AG0026860 gene on chromosome 1A as per the recently published finger millet genome by (Devos et al. 2023). ELECO.r07.1AG0026860 is an orthologue of the Pgl_GLEAN_10002072 gene in pearl millet. Pgl_GLEAN_10002072 showed significant association with mass features mostly engaged in hydrolase activity, especially acting on ester bonds to promote carotenoid accumulation in pearl millet (Yadav et al. 2021). This suggests the gene plays a key role in pathways that promote the accumulation of antioxidant-related flavonoid compounds. The internal detoxification work in a variety of ways, such as hormone regulation, antioxidant defense system, and vacuole compartmentalization. As an important condition to maintain plant growth and development, the antioxidant defense system (enzyme and non-enzyme system) produces stress response to eliminate the excessive accumulation of reactive oxygen species in the plants (Sharma and Dubey 2007). This ultimately improve grain yield.

The LXGH01346160.1_21363 locus showed a significant association with PH although it explained only 0.94% of the variation in plant height, which is the highest compared to that of other loci associated with this trait. This locus is located within the ELECO.r07.7AG0569730 gene on chromosome 7A. The ELECO.r07.7AG0569730 gene codes for late embryogenesis abundant protein (LEA_2) in finger millet. Late embryogenesis abundant protein, which is also characterized in various other crops, such as *O. sativa* (OsLEAs), *Z. mays* (ZmLEAs), and *S. bicolor* (SbLEAs) was shown to play an important role in plant stress adaptation (Wang et al. 2007, Li and Cao 2016, Nagaraju et al. 2019). When under abiotic stresses such as drought, cold, and salinity, these genes are expressed and play a protective role (Ingram and Bartels 1996). This suggests that the late embryogenesis abundant gene may be induced during AI-stress in finger millet to protect and promote plant growth.

Conclusions

The present study revealed high phenotypic and genetic variations among finger millet genotypes. A total of 23 markers associated with finger millet traits were obtained and could potentially be used for screening finger millet genetic resources for tolerance to acidic soils. However, these markers should be validated with further studies, particularly those that explain substantial phenotypic variations. Genotypes with high yield potential identified in this study can be used as parents in breeding programs aimed at pyramiding desirable alleles through marker-assisted selection. In this study, an insight into the possibility of genomics-driven breeding for finger millet was gained through the identification of an association between markers and major phenotypic traits.

Declarations

Data availability statement

The datasets used in this study are available in online repositories and the link is given under supplementary material section. The GBS quality trimmed reads were submitted to Sequence Reads Archive (SRA), BioProject PRJNA791522.

Author Contributions

Conceptualization and methodology; HB, MG, TH, KT, CH, and RO. Data collection; HB, MG, TH, KT, CH, RO, and DL. Data analysis; HB under the guidance of MG, CH, KBA. Original draft preparation; HB. Review and edit; HB, MG, TH, KT, CH, RO, DL, and KBA. Project funding acquisition and administration; MG, TH, KT, RO, and CH. All authors have read and approved the final version of the manuscript.

Funding

This study was funded by the Swedish International Development Cooperation Agency (Sida) Research and Training grant awarded to Addis Ababa University and the Swedish University of Agricultural Sciences (AAU-SLU Biotech).

Acknowledgments

We thank the Swedish International Development Cooperation Agency (Sida) for financing this study and the Ethiopian Biodiversity Institute (EBI) for providing the finger millet germplasm used in this study. The authors also wish to thank Bako Agricultural Research Center (Ethiopia) for managing the field trials and recording phenotypic data at the two sites.

Conflict of interest

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

References

1. Admassu, S., Teamir, M., and Alemu, D. (2009). Chemical composition of local and improved finger Millet [*Eleusine corocana* (L.) Gaertn.] varieties grown in Ethiopia. *Ethiopian Journal of Health Sciences* 19.
2. Barcelo, J and Poschenrieder, C. (2002). Fast root growth responses, root exudates, and internal detoxification as clues to the mechanisms of aluminium toxicity and resistance: a review. *Environmental and Experimental Botany* 48, 75-92.
3. Belete, T., Tulu, L.A., and Senbetay, T. (2020). Evaluation of finger millet (*Eleusine coracana* (L.) Gaertn.) Varieties at Different Locations of Southwestern Ethiopia. *Journal of Genetic and Environmental Resources Conservation* 8, 9-17.
4. Bharathi, A. (2011). *Phenotypic and genotypic diversity of global finger millet (Eleusine coracana (L.) Gaertn.) composite collection*. Tamil Nadu Agricultural University.
5. Bisht, M. and S.Mukai, Y. (2001). Genomic in situ hybridization identifies genome donor of finger millet (*Eleusine coracana*). *Theoretical and Applied Genetics* 102, 825-832.
6. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114-2120.

7. Brhane, H., Haileselassie, T., Tesfaye, K., Hammenhag, C., Ortiz, R., Abreha, K.B., and Geleta, M. (2021). Novel expressed sequence tag-derived and other genomic simple sequence repeat markers revealed genetic diversity in Ethiopian finger millet landrace populations and cultivars. *Frontiers in Plant science* 12, 735610.
8. Brhane, H., Haileselassie, T., Tesfaye, K., Ortiz, R., Hammenhag, C., Abreha, K.B., and Geleta, M. (2022a). Novel GBS-based SNP markers for finger millet and their use in genetic diversity analyses. *Frontiers in Genetics* 13, 848627.
9. Brhane, H., Haileselassie, T., Tesfaye, K., Ortiz, R., Hammenhag, C., Abreha, K.B., Vetukuri, R.R., and Geleta, M. (2022b). Finger millet RNA-seq reveals differential gene expression associated with tolerance to aluminum toxicity and provides novel genomic resources. *Frontiers in Plant Science* 13, 1068383.
10. Chandra, A.K., Pandey, D., Sood, S., Joshi, D.C., Tiwari, A., Sharma, D., Gururani, K., and Kumar, A. (2024). Uncovering the genomic regions underlying grain iron and zinc content using genome-wide association mapping in finger millet. *3 Biotech* 14, 47.
11. Daba, S.Haile, M. (2000). Effects of rhizobial inoculant and nitrogen fertilizer on yield and nodulation of common bean. *Journal of Plant Nutrition* 23, 581-591.
12. Devos, K.M., Qi, P., Bahri, B.A., Gimode, D.M., Jenike, K., Manthi, S.J., Lule, D., Lux, T., Martinez-Bello, L., and Pendergast Iv, T.H. (2023). Genome analyses reveal population structure and a purple stigma color gene candidate in finger millet. *Nature Communications* 14, 3694.
13. Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S., and Mitchell, S.E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS one* 6, e19379.
14. Falconer, D. (1981). Introduction to quantitative genetics (2d ed.) Longman. *London, UK*.
15. Garrison, E.Marh, G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv preprint arXiv:1207.3907*.
16. Goron, T.L. and Raizada, M.N. (2015). Genetic diversity and genomic resources available for the small millet crops to accelerate a New Green Revolution. *Frontiers in Plant Science* 6.
17. Hilu, K.W. and De Wet, J. (1976). Domestication of Eleusine coracana. *Economic Botany* 30, 199-208.
18. Hittalmani, S., Mahesh, H.B., Shirke, M.D., Biradar, H., Uday, G., Aruna, Y.R., Lohithaswa, H.C., and Mohanrao, A. (2017). Genome and Transcriptome sequence of Finger millet (*Eleusine coracana* (L.) Gaertn.) provides insights into drought tolerance and nutraceutical properties. *Bmc Genomics* 18.
19. Ingram, J. and Bartels, D. (1996). The molecular basis of dehydration tolerance in plants. *Annual review of plant biology* 47, 377-403.
20. Jaiswal, V., Gupta, S., Gahlaut, V., Muthamilarasan, M., Bandyopadhyay, T., Ramchiary, N., and Prasad, M. (2019). "Genome-wide association study of major agronomic traits in foxtail millet (*Setaria italica* L.) using ddRAD sequencing. *Sci Rep*".).
21. Jombart, T. and Collins, C. (2015). A tutorial for discriminant analysis of principal components (DAPC) using adegenet 2.0. 0. *London: Imperial College London, MRC Centre for Outbreak Analysis and Modelling*.

22. Kavuluko, J., Kibe, M., Sugut, I., Kibet, W., Masanga, J., Mutinda, S., Wamalwa, M., Magomere, T., Odeny, D., and Runo, S. (2021). GWAS provides biological insights into mechanisms of the parasitic plant (Striga) resistance in sorghum. *BMC Plant Biology* 21, 1-15.
23. Keba, W., Tolemariam, T., and Mohammed, A. (2022). Evaluation of Grain/Seed Yield and Yield Components of Finger Millet and Three Vetch Species Intercropped at Various Seeding Ratios at Bako, Ethiopia. *Advances in Agriculture* 2022.
24. Kebede, D., Dagnachew, L., Megersa, D., Chemed, B., Girma, M., Geleta, G., and Gudeta, B. (2019). Genotype by environment interaction and grain yield stability of Ethiopian black seeded finger millet genotypes. *African Crop Science Journal* 27.
25. Kinfe, H., Yiergalem, T., Alem, R., Redae, W., Desalegn, Y., and Welegerima, G. (2017). Yield performance and adaptability of finger millet landrace in North Western Tigray, Ethiopia. *World News of Natural Sciences* 15.
26. Ko, K.Y., Park, S.C., Cho, S.Y., and Yoon, S.-I. (2022). Structural analysis of carboxyspermidine dehydrogenase from *Helicobacter pylori*. *Biochemical and Biophysical Research Communications* 635, 210-217.
27. Kochian, L.V., Pineros, M.A., and Hoekenga, O.A. (2005). The physiology, genetics and molecular biology of plant aluminum resistance and toxicity. *Plant and Soil* 274, 175-195.
28. Li, H. and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *bioinformatics* 25, 1754-1760.
29. Li, X. and Cao, J. (2016). Late embryogenesis abundant (LEA) gene family in maize: identification, evolution, and expression profiles. *Plant Molecular Biology Reporter* 34, 15-28.
30. Lipka, A.E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P.J., Gore, M.A., Buckler, E.S., and Zhang, Z. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28, 2397-2399.
31. Lule, D., De Villiers, S., Fetene, M., Odeny, D.A., Rathore, A., Das, R.R., and Tesfaye, K. (2018). Genetic diversity and association mapping of Ethiopian and exotic finger millet accessions. *Crop and Pasture Science* 69, 879-891.
32. Marefia, A., Abate, A., and Bantayehu, M. (2022). Genetic Gain in Yield Potential and Related Traits of Finger millet [*Eleusine coracana* (L.) Gaertn] in Ethiopia. *East African Journal of Sciences* 16, 155-170.
33. Melnyk, R.A., Beskrovnaya, P., Liu, Z., Song, Y., and Haney, C.H. (2019). Bacterially produced spermidine induces plant systemic susceptibility to pathogens. *immunity* 26, 27.
34. Mnyenyembe, P. and Gupta, S. (1998). Variability for grain yield and related traits in finger millet germplasm accessions from Malawi. *African Crop Science Journal* 6, 317-322.
35. Nagaraju, M., Kumar, S.A., Reddy, P.S., Kumar, A., Rao, D.M., and Kavi Kishor, P. (2019). Genome-scale identification, classification, and tissue specific expression analysis of late embryogenesis abundant (LEA) genes under abiotic stress conditions in *Sorghum bicolor* L. *PLoS one* 14, e0209980.
36. Puranik, S., Sahu, P.P., Beynon, S., Srivastava, R.K., Sehgal, D., Ojulong, H., and 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, 649-662.

37. Saleem, K., Shokat, S., Waheed, M.Q., Arshad, H.M.I., and Arif, M.a.R. (2022). A GBS-based GWAS analysis of leaf and stripe rust resistance in diverse pre-breeding germplasm of bread wheat (*Triticum aestivum* L.). *Plants* 11, 2363.
38. Sharma, D., Tiwari, A., Sood, S., Jamra, G., Singh, N., Meher, P.K., and Kumar, A. (2018). Genome wide association mapping of agro-morphological traits among a diverse collection of finger millet (*Eleusine coracana* L.) genotypes using SNP markers. *PloS one* 13, e0199444.
39. Sharma, P. and Dubey, R. (2007). Involvement of oxidative stress and role of antioxidative defense system in growing rice seedlings exposed to toxic concentrations of aluminum. *Plant cell reports* 26, 2027-2038.
40. Singh, M. and Upadhyaya, H.D. (2015). Genetic and genomic resources for grain cereals improvement. Academic Press.
41. Sivagurunathan, M. (2004). Genetic and molecular studies on quantitative and qualitative traits in finger millet (*Eleusine coracana* (L.) Gaertn.). M. Sc.(Ag.). Thesis. TNAU, Coimbatore, India.
42. Sood, S., Joshi, D., Rajashekara, H., Tiwari, A., Bhinda, M., Kumar, A., Kant, L., and 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.
43. Team, R.C. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
44. Tesfaye, K. and Mengistu, S. (2017). Phenotypic characterization of Ethiopian finger millet accessions ((L.) Gaertn), for their agronomically important traits. *Acta Universitatis Sapientiae, Agriculture and Environment* 9, 107-118.
45. Tiwari, A., Sharma, D., Sood, S., Jaiswal, J.P., Pachauri, S.P., Ramteke, P.W., and Kumar, A. (2020). Genome-wide association mapping for seed protein content in finger millet (*Eleusine coracana*) global collection through genotyping by sequencing. *Journal of cereal science* 91, 102888.
46. Turner, S. (2018). "qqman: an R package for visualizing GWAS results using QQ and manhattan plots. J. Open Source Softw. 3: 731".).
47. Upadhyaya, H., Gowda, C., and Reddy, V.G. (2007). Morphological diversity in finger millet germplasm introduced from Southern and Eastern Africa. *Journal of SAT Agricultural Research* 3, 1-3.
48. Wang, X.-S., Zhu, H.-B., Jin, G.-L., Liu, H.-L., Wu, W.-R., and Zhu, J. (2007). Genome-scale identification and analysis of LEA genes in rice (*Oryza sativa* L.). *Plant Science* 172, 414-420.
49. Ward, B.P., Brown-Guedira, G., Kolb, F.L., Van Sanford, D.A., Tyagi, P., Sneller, C.H., and Griffey, C.A. (2019). Genome-wide association studies for yield-related traits in soft red winter wheat grown in Virginia. *PloS one* 14, e0208217.
50. Yadav, C.B., Srivastava, R.K., Gangashetty, P.I., Yadav, R., Mur, L.A., and Yadav, R.S. (2021). Metabolite diversity and metabolic genome-wide marker association studies (MGWAS) for health benefiting nutritional traits in pearl millet grains. *Cells* 10, 3076.

Figures

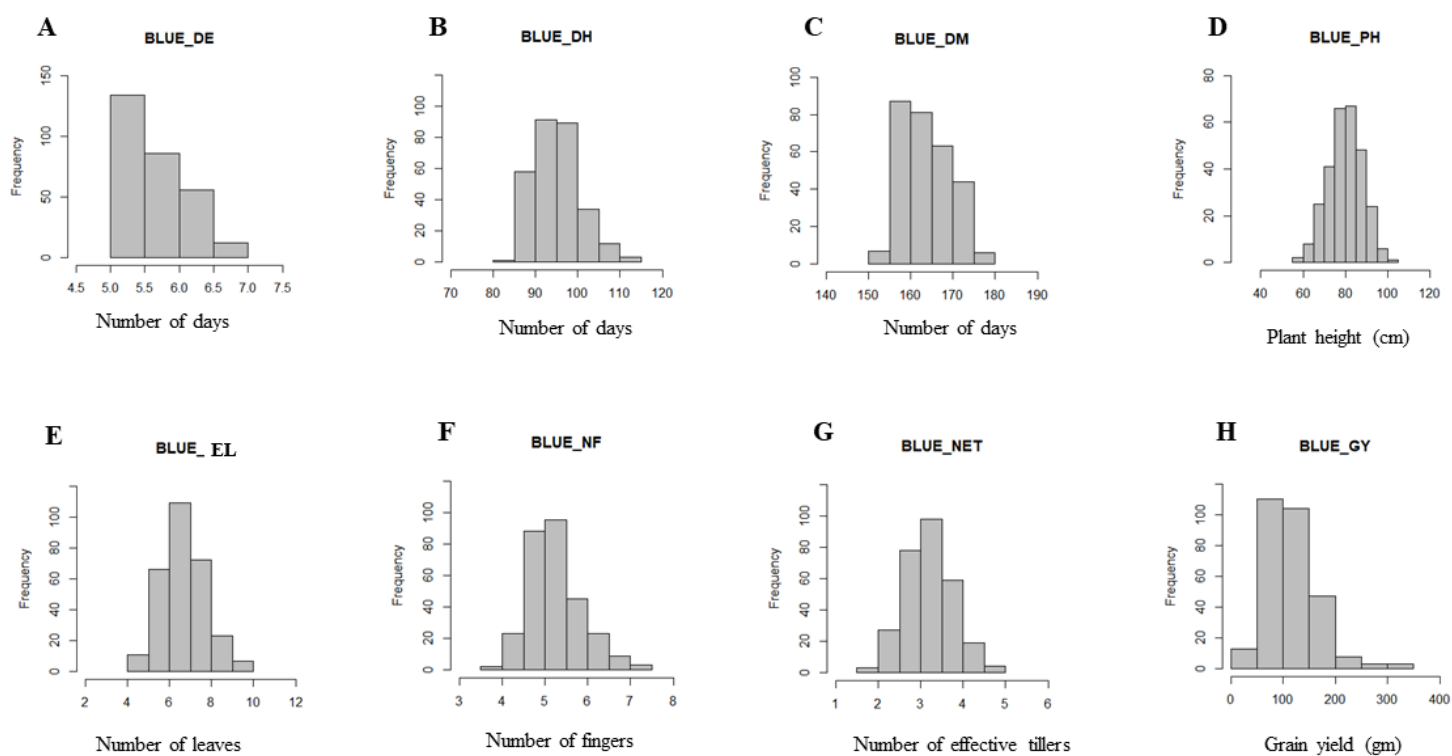


Figure 1

Data distribution of the eight quantitative traits across 288 finger millet genotypes grown under acidic soils in Ethiopia. DE; days to emergence, DH; days to heading, DM; days to maturity, PH; plant height, EL; ear length, NF; number of fingers, NET; number of effective tillers, and GY; grain yield.

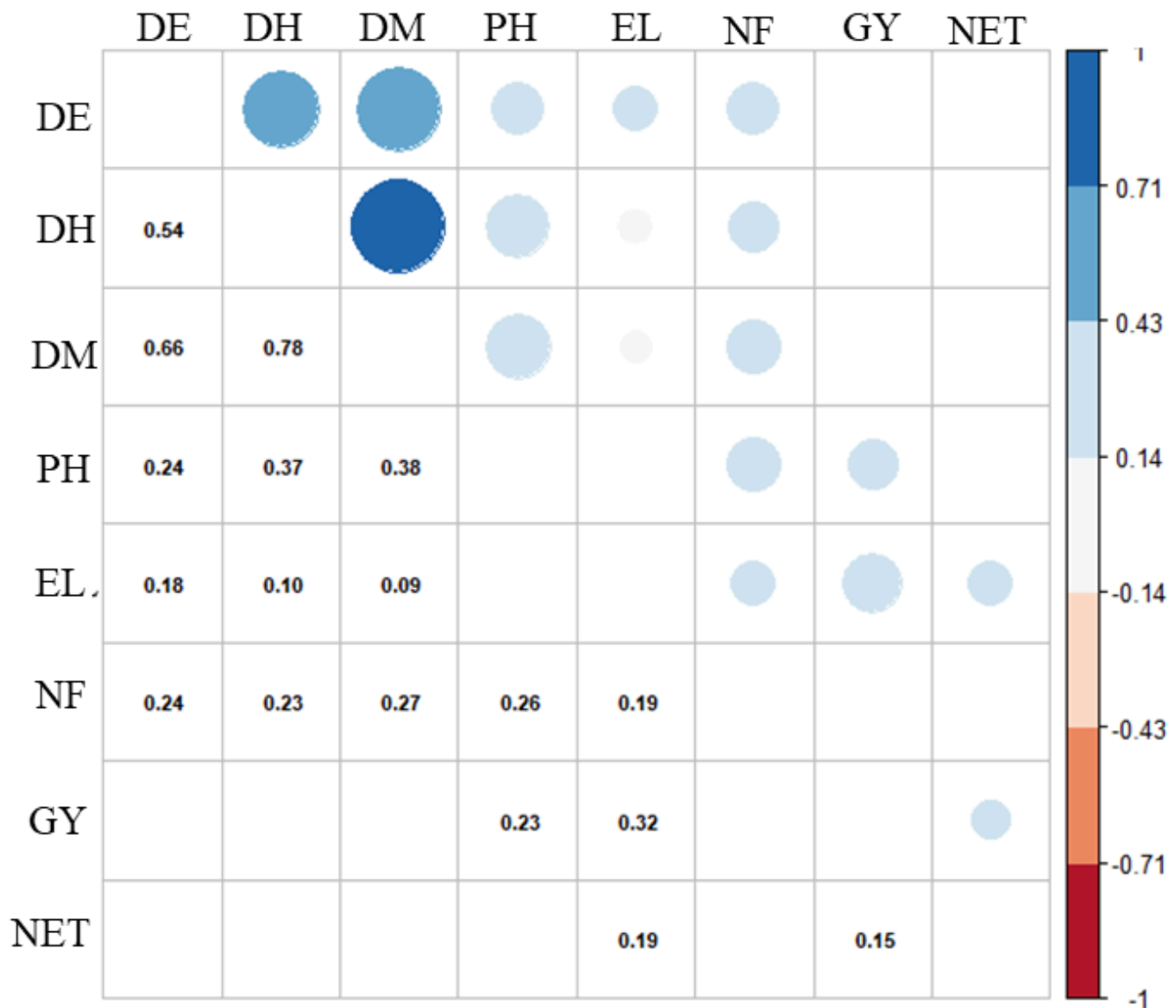


Figure 2

Spearman's phenotypic trait correlation analysis of 288 finger millet genotypes grown evaluated for 8 quantitative traits under acidic soil condition. At two acidic locations. No circle/no value means no significant correlation. Abbreviations: DE; days to emergence, DH; days to heading, DM; days to maturity, PH; plant height, EL; ear lengths, NF; number of fingers, NET; number of effective tillers, and GY; grain yield.

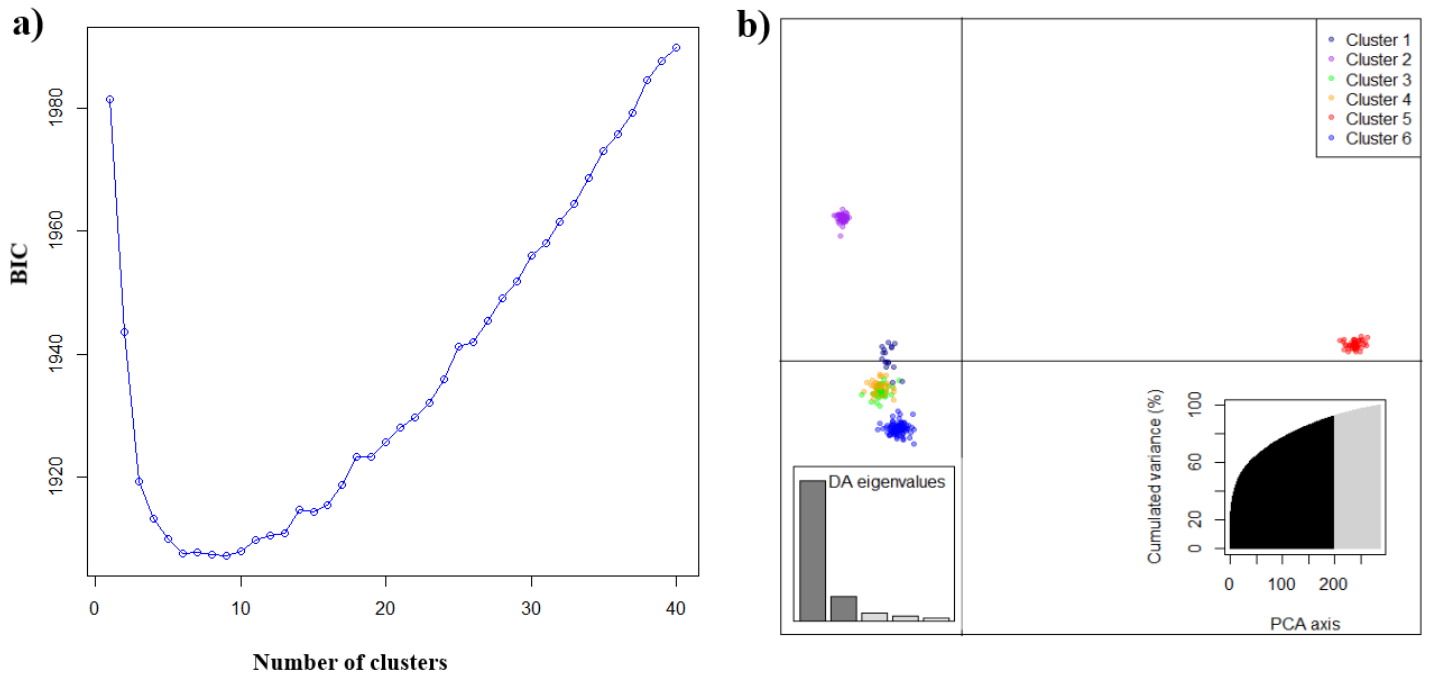


Figure 3

Graphs depicting a discriminant analysis of principal components (DAPC) for 288 finger millet genotypes based on 5226 stringently filtered genotype-by-sequencing (GBS)-based single nucleotide polymorphism (SNP) markers: (a) Bayesian Information Criterion values plotted against number of clusters to select an optimal number of clusters. (b) Scatterplot of the genotypes using the first two principal components distinguishing six genetic clusters. The clusters were distinguished by colors. Discriminant analysis bar plots (bottom left) and principal components (bottom right) were also shown.

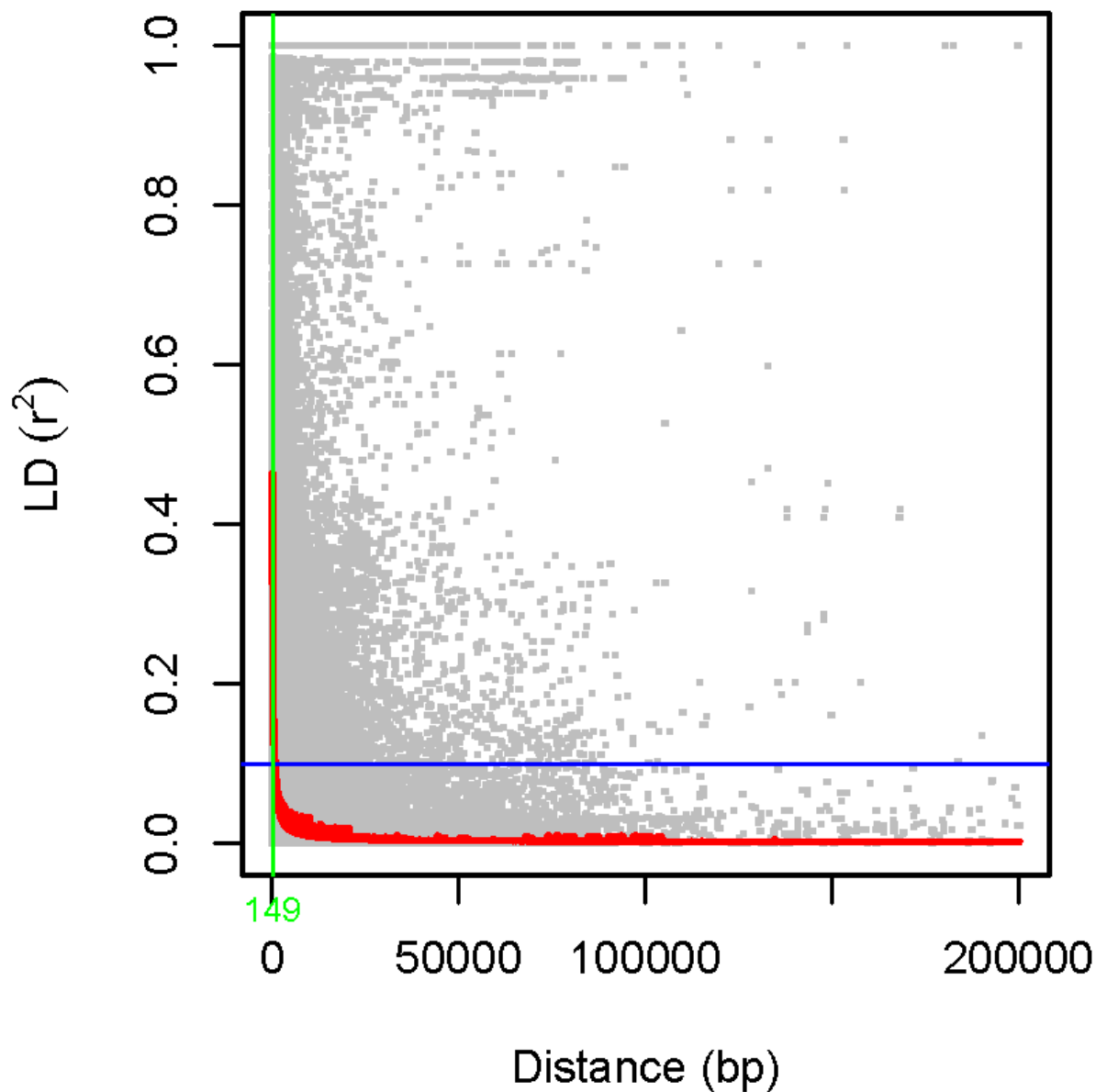


Figure 4

A scatter plot of genome-wide linkage disequilibrium (LD) decay as determined by the r^2 values of the marker pairs. The smoothing spline regression model fitted to LD decay is shown with a red curve line. The horizontal blue line shows the genome's half-decay r^2 value ($r^2 = 0.1$), while the vertical green line shows the genetic distance between markers (149 Kb) at the intersection of the half-decay line and the LD decay curve.

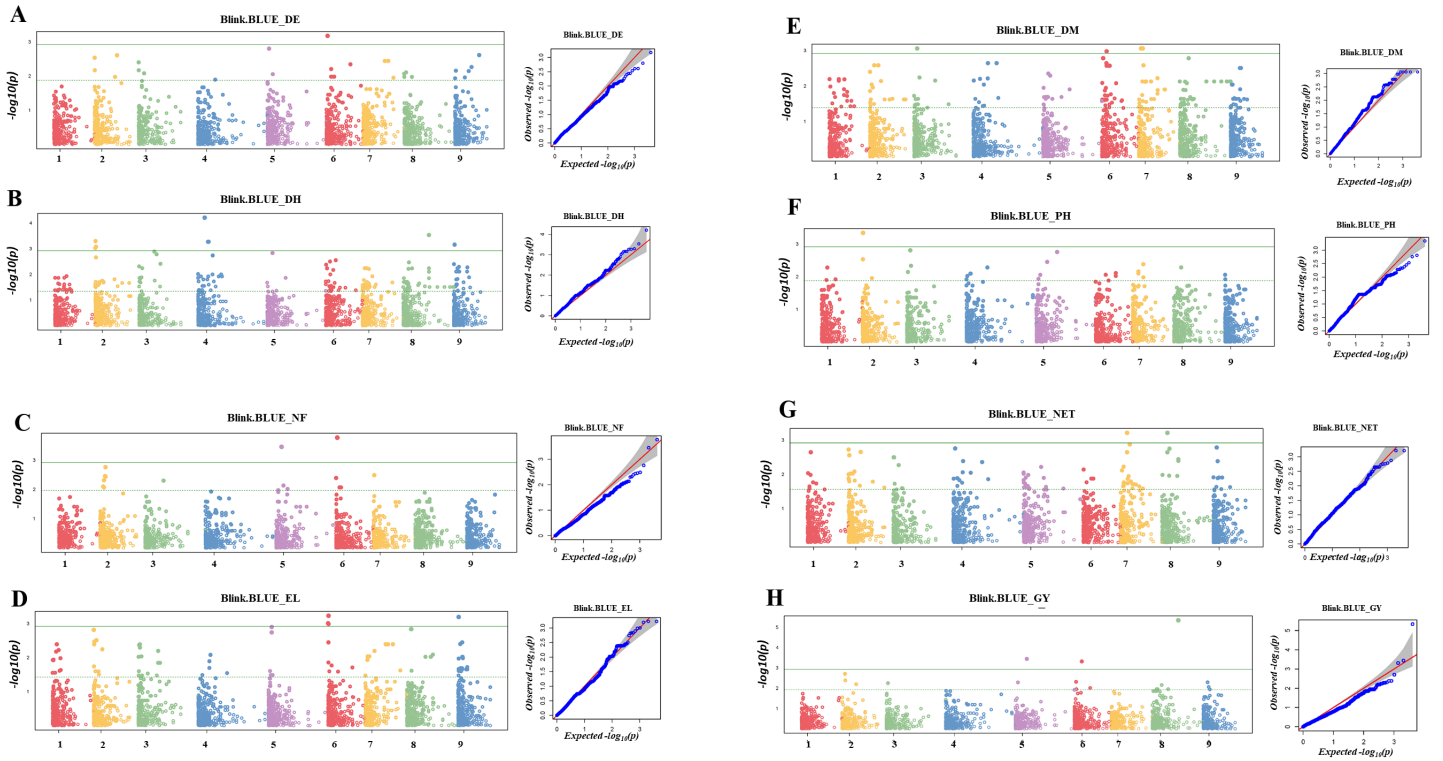


Figure 5

Manhattan's and quantile-quantile plots produced from genome-wide association study (GWAS) in **A**) days to emergence (DE), **B**) days to heading (DH), **C**) days to maturity (DM), **D**) plant height (PH), **E**) number of fingers (NF), **F**) ear length (EL), **G**) number of effective tillers, and **H**) grain yield (GY) traits. The green horizontal line indicates the threshold for significant association between the trait and marker. The X-axis shows the nine scaffolds groups, whereas the Y-axis shows the negative logarithm of p -values.

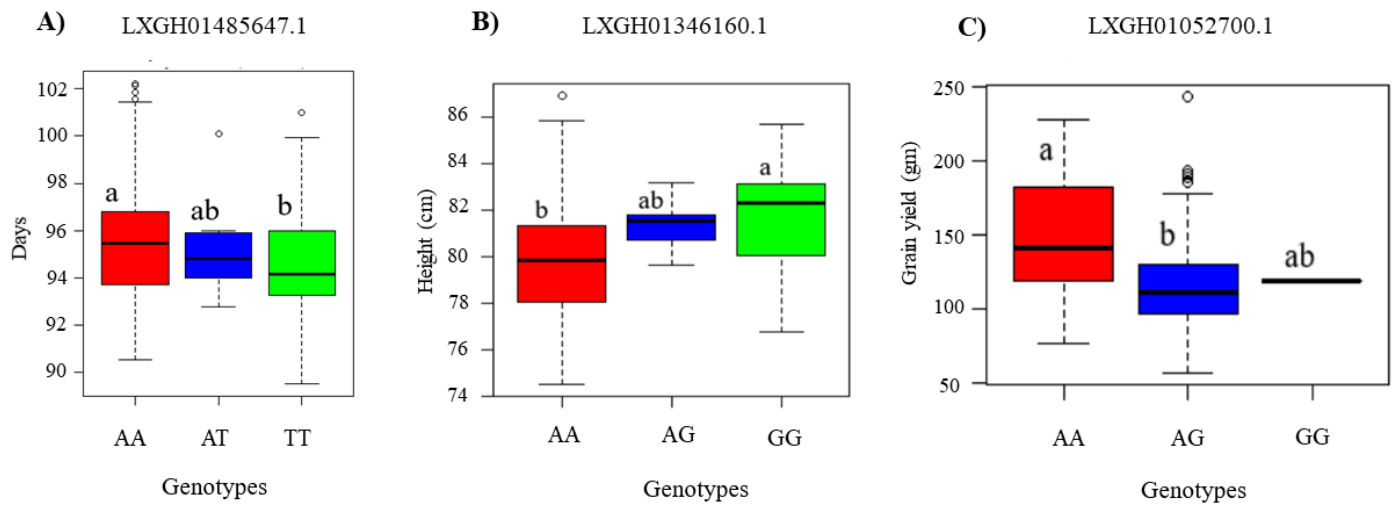


Figure 6

Boxplot showing the allelic effect of significant marker-trait association for A) days to heading (DH), B) plant height (PH), and C) grain yield (GY). The X-axis shows the genotypes of the corresponding locus, while the Y-axis depicts DH in days, PH in cm, and GY in g/m².

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTable1.xlsx](#)
- [SupplementaryTable2.xlsx](#)
- [SupplementaryTable3.xlsx](#)