

RESEARCH

Open Access



Single nucleotide polymorphism (SNP) markers for genetic diversity and population structure study in Ethiopian barley (*Hordeum vulgare* L.) germplasm

Mihret Yirgu^{1,2*}, Mulugeta Kebede³, Tileye Feyissa⁴, Berhane Lakew⁵, Aemiro Bezabih Woldeyohannes⁶ and Mulusew Fikere⁷

Abstract

Background High-density single nucleotide polymorphisms (SNPs) are the most abundant and robust form of genetic variants and hence make highly favorable markers to determine the genetic diversity and relationship, enhancing the selection of breeding materials and the discovery of novel genes associated with economically important traits. In this study, a total of 105 barley genotypes were sampled from various agro-ecologies of Ethiopia and genotyped using 10 K single nucleotide polymorphism (SNP) markers. The refined dataset was used to assess genetic diversity and population structure.

Results The average gene diversity was 0.253, polymorphism information content (PIC) of 0.216, and minor allelic frequency (MAF) of 0.118 this revealed a high genetic variation in barley genotypes. The genetic differentiation also showed the existence of variations, ranging from 0.019 to 0.117, indicating moderate genetic differentiation between barley populations. Analysis of molecular variance (AMOVA) revealed that 46.43% and 52.85% of the total genetic variation occurred within the accessions and populations, respectively. The heat map, principal components and population structure analysis further confirm the presence of four distinct clusters.

Conclusions This study confirmed that there is substantial genetic variation among the different barley genotypes. This information is useful in genomics, genetics and barley breeding.

Keywords Accession, Cluster, Diversity, *Ex-situ* conservation, Genetic differentiation

*Correspondence:

Mihret Yirgu
ymihret@gmail.com

¹ Department of Applied Biology, School of Applied Natural Science, Adama Science and Technology University, P.O.Box 1888, Adama, Ethiopia

² Department of Plant Science, College of Agriculture and Natural Resource, Madda Walabu University, P.O.Box 247, Robe, Oromia, Ethiopia

³ Department of Plant Biology and Biodiversity Management, College of Natural and Computational Sciences, Addis Ababa University, P.O.Box 3434, Addis Ababa, Ethiopia

⁴ Institute of Biotechnology, Addis Ababa University, P.O.Box 1176, Addis Ababa, Ethiopia

⁵ Holetta Agricultural Research Center, Ethiopian Institute of Agricultural Research, Holetta, Ethiopia

⁶ Debre Birhan Agricultural Research Center, Amhara Agricultural Research Institute, Debre Birhan, Ethiopia

⁷ Institute of Molecular Bioscience, University of Queensland, Brisbane, Australia



© The Author(s) 2023. **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/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

Barley (*Hordeum vulgare* L., $2n = 2x = 14$ chromosomes, 5.1 Gb haploid genome size) is one of the main important cereal crops cultivated worldwide in a wide range of environments [1, 2]. Barley has been part of a sustainable food source for humans since pre-historic times. It is mainly used for human food, animal feed, malting and brewing [3, 4]. The crop is a major component of staple food in China, India, Morocco, Ethiopia and Eritrea, the mountainous regions of Bolivia, Ecuador, Colombia and Peru [3].

Barley is the fifth most important cereal crop in Ethiopia after teff, maize, sorghum and wheat both in area of production and amount [5]. Ethiopia is the second-largest barley producer in Africa following Morocco [2], and the average yield of barley in the country is 2.18 tons ha^{-1} [5]. It is grown in various agro-ecologies, ranging from lowland to high altitudes [6, 7] but it performs well at higher altitudes in the northern and central regions of the country [8, 9]. The production of barley in Ethiopia is challenged by abiotic factors (water logging, poor soil fertility, drought and frost) and biotic stress such as net blotch (*Pyrenophora teres*), scald (*Rhynchosporium secalis*) and leaf rust (*Puccinia hordei*) [10, 11]. Understanding, the existence of diverse barley germplasm for yield and yield related traits, resistance against diseases and abiotic stress tolerance, at a molecular level is crucial for barely breeding programs and to enhance its productivity in the country.

Ethiopia is considered as a center of diversity for barley (*Hordeum vulgare* L.), and its landraces are genetically unique and diverse [12, 13]. Earlier studies have shown that the highest genetic diversity within the Ethiopia barley genetic resource for various useful traits can be valorized for barley breeding [8, 14]. This high diversity is due to diverse agro-ecologies, wide ranges of altitude, soil variability, climate and farming systems and topography together with geographical isolation [15, 16]. Barley's population structure also highly depends on the farming system and altitude in Ethiopia [17].

In the past decades, the Ethiopian Biodiversity Institute (EBI) has collected and conserved more than 16,000 accessions sampled from various agro-ecologies across the country. It is a useful resource for genetic diversity and can play an important role in developing new barley varieties with higher yield potential and other desirable agronomic traits [18]. Konopka [19] stated that the Ethiopian barley collection is one of the world's ten largest barley collections and is used as a source of elite breeding material for national as well as global breeding programs. It has useful traits such as the source of disease resistance [20], yellow dwarf virus resistance gene [21, 22], powdery mildew resistance [23], barley leaf scald and net blotch

[24], high lysine content and protein quality [25], malting and brewing quality [26]. However, most of the early studies were based on either collection from distinct geographical regions only or a few samples collected from wider geographical ranges [27].

Analyzing the molecular diversity and its genetic relationship encompassed in crop genetic resources is a prerequisite for designing efficient selection in crop breeding programs and for developing conservation and valorization strategies. Previous studies have used DNA markers to determine the genetic diversity of barley. Markers such as amplified fragment length polymorphism (AFLP) [28], restriction fragment length polymorphisms (RFLPs) [29], and simple sequence repeats (SSR) [27, 30–33] have been employed and provided crucial information for barely breeding programs.

Currently, single nucleotide polymorphism (SNP) markers are the markers of choice for genetic diversity studies [34], genome-wide association mapping [35, 36], genomic selection [37], phylogenetic relationships and population evolutionary history studies [38, 39]. SNPs are the most abundant and robust, feasible for automated high-throughput genotyping [40], highly reproducible and can be used to identify variants [41], replacing the earlier markers due to high throughput, efficiency and cost-effectiveness [12].

A recent study by Teklemariam et al. [42] has observed a weak correlation between geographic distance and genetic differentiation of some Ethiopian barely germplasm collections using SNP markers. However, the genetic diversity of Ethiopia's diverse barley germplasm collection has not been adequately characterized and exploited using advanced tools such as SNP markers. Therefore, this study was designed to assess the genetic diversity and population structure of Ethiopian barley genotypes using SNP markers.

Results

SNP variation and markers distribution

The single nucleotide polymorphism (SNP) distribution of the 10,103 SNPs remaining after imposing a quality control threshold was plotted in 10 Mb (megabase pair) window size across the *H.vulgare* genome (Fig. 1). Variant distribution was not completely uniform across the chromosomes. We detected an average of 97 SNPs per 10 Mb. The highest SNP density (> 129 SNPs/10 Mb) was observed on chromosomes Chr2, Chr3 and Chr6 and Chr7. The lowest average SNP density was found on chromosome Chr1 (< 33 SNPs/10 Mb).

Genetic diversity and relationship

The genetic parameters included gene diversity, heterozygosity (the ratio of observed to expected heterozygosity),

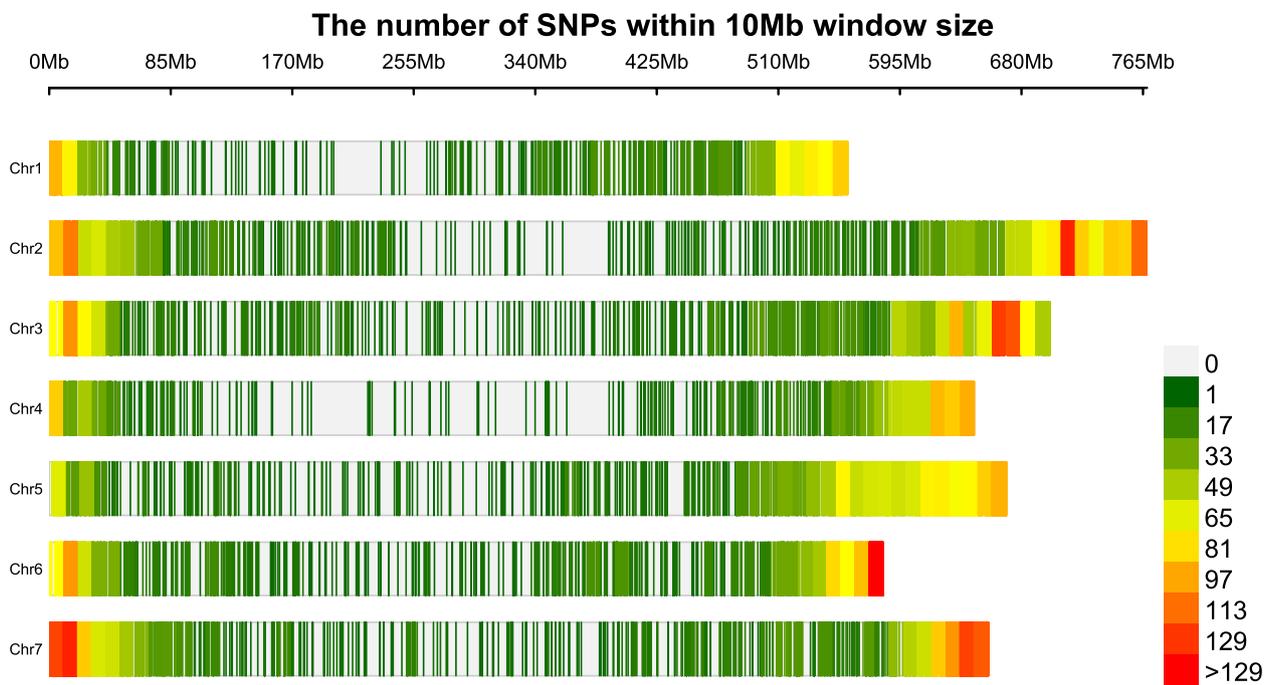


Fig. 1 Distribution of SNP markers within 10 Mb window size across seven chromosomes. Colored bars are SNP counts in 10 Mb interval

minor allelic frequency (MAF) and the polymorphic information content (PIC) of the 10,103 SNP markers of 105 barley genotypes were presented in Table 1. The result showed that the average genetic diversity was 0.253 and ranged from 0.177 to 0.311. The highest gene diversity (0.311) was observed for genotypes collected from the Amhara region followed by the Oromia (0.288), SNNP (0.266), Tigray (0.221) and ICARDA (0.177). The PIC values ranged from 0.151(ICARDA) to 0.271(SNNP) with an average polymorphism value of 0.216, which indicates high genetic diversity between the barley genotypes. The MAF ranged from 0.092 to 0.146, with an average of 0.118. Genotypes from the Oromia region showed the highest MAF (0.146) followed by ICARDA

(0.145) collection and Amhara (0.107) genotypes. However, the lowest MAF was exhibited in the SNNP (0.101) and Tigray (0.092) genotypes. In addition, the observed heterozygosity (H_o) ranged from 0.015 in the Tigray to 0.088 in the Oromia genotypes with an average of 0.045. On the other hand, the expected heterozygosity (H_e) was higher than that observed heterozygosity (H_o), ranging from 0.133 to 0.215 with an average of 0.163.

Analysis of molecular variance (AMOVA)

The analysis of molecular variance (AMOVA) showed that the proportion of variance within the barley populations was significantly higher (52.85%, $P < 0.0001$) than the variation within accessions (46.43%,

Table 1 Gene diversity, observed heterozygosity (H_o), expected heterozygosity (H_e), minor allele frequency (MAF) and the polymorphism information content (PIC) of the 105 barley genotypes

Regions	Genotypes	GD	H_o	H_e	MAF	PIC
Amhara	31	0.311	0.071	0.158	0.107	0.195
Oromia	32	0.288	0.088	0.215	0.146	0.211
SNNP	21	0.266	0.027	0.145	0.101	0.271
Tigray	19	0.221	0.015	0.133	0.092	0.250
ICARDA/Other	2	0.177	0.022	0.164	0.145	0.151
Average values	105	0.253	0.045	0.163	0.118	0.216

Abbreviations: GD Gene diversity, H_o Observed heterozygosity, H_e Expected heterozygosity, MAF Minor allele frequency, PIC Polymorphism information content, SNNP Southern Nations, Nationalities and Peoples, ICARDA International Centre for Agricultural Research in the Dry Areas

Table 2 Analysis of molecular variance (AMOVA) of 105 barley population based on SNP markers

Source of variation	Mean square	Estimated variance	Proportion of Variance (%)	Probability (p) value
Between populations	2349.75	10.8	0.72	***
Within populations	1742.11	820.77	52.85	***
Within accessions	475.33	582.41	46.43	***
Total	1588.53	1413.98	100	

Significant codes: 0 **** 0.001, *** 0.01, ** 0.05, 'ns' 0.1, 'ns' 1, ns-non significant

Table 3 Pairwise genetic differentiation (F_{ST}) among barley populations using SNP markers

Regions	Amhara	SNNP	Oromia
Amhara	-	-	-
SNNP	0.022	-	-
Oromia	0.019	0.038	-
Tigray	0.089	0.117	0.088

Abbreviation: SNNP Southern Nations, Nationalities and Peoples

$P < 0.0001$) (Table 2). Conversely, a significantly lower level of genetic variation (0.72%, $P < 0.0001$) was recorded between the barley populations (Table 2).

Genetic differentiation

Pairwise genetic differentiation (F_{ST}) among barley populations was computed using SNP markers (Table 3). The lower values of 0.019, 0.022, and 0.038 genetic differentiations were recorded between the barley genotypes originating from Amhara and Oromia, Amhara and SNNP, and SNNP and Oromia regions, respectively (Table 3). Although moderate genetic differentiation was recorded between the Tigray and SNNP (0.117) followed by Tigray and Amhara (0.089) and Tigray and Oromia (0.088) (Table 3). This revealed that the barley genotypes obtained from those regions of origin the existence in the highest genetic variations and distant relationships.

Genetic distance and identity

A high genetic distance (0.08) was recorded among barley genotypes obtained from Tigray and ICARDA, which indicates the existence of high genetic differences among the genotypes (Table 4). The genetic distance ranged from 0.002 to 0.08. Besides, the genetic identity ranged from 0.971 to 0.998, and the highest genetic identity was found among Amhara and Tigray, Amhara and ICARDA, and Oromia and SNNP. However, a lower (0.971) genetic identity value was observed between genotypes originating from SNNP and ICARDA (Table 4).

Table 4 Genetic distance (below diagonal) and genetic identity (above diagonal) between 105 barley genotypes based on SNP markers

Regions	Amhara	Oromia	SNNP	Tigray	ICARDA/Other
Amhara		0.991	0.993	0.998	0.998
Oromia	0.009		0.998	0.990	0.990
SNNP	0.007	0.012		0.991	0.971
Tigray	0.012	0.021	0.004		0.992
ICARDA/Other	0.002	0.011	0.031	0.08	

Abbreviations: SNNP Southern Nations, Nationalities and Peoples, ICARDA International Centre for Agricultural Research in the Dry Areas

Principal component analysis

To quantify the genetic variation between genotypes, we perform a principal component analysis (PCA). Four clusters were observed which accounted for 47.7% and 13.5% of the total variation in PC1 and PC2 of the total variance, respectively (Fig. 2). Oromia and Amhara genotypes were loadings in PC1 whereas Tigray and SNNP regions were the major loadings to PC2 (Fig. 2). The lowest total variance was obtained for PC3 (6.7%) and PC4 (5%) (Fig. 3). As shown in Fig. 3, the third and consecutive PCs resulted in a lower percentage of contribution to the total genetic variance. The PCA groups showed slightly unclear-cut separation due to an explicit genetic relatedness between the Amhara and SNNP, and Oromia and Amhara genotypes. The first group assembled Amhara and SNNP genotypes. The second group consisted of Tigray genotypes, and the last group contained the remaining genotypes from Oromia found in all groups (Fig. 2).

Genetic relatedness

We perform genetic relatedness analysis using the refined SNP dataset of 105 genotypes. The heat map was used to visualize the genetic relatedness across the population. The heat map plot depicted four clusters (Fig. 4, Table 5). Cluster I contained the largest number of barley genotypes ($n=60$) followed by cluster III ($n=16$), cluster IV ($n=15$) and cluster II ($n=14$). The number of genotypes belonging to distinctive clusters varies from

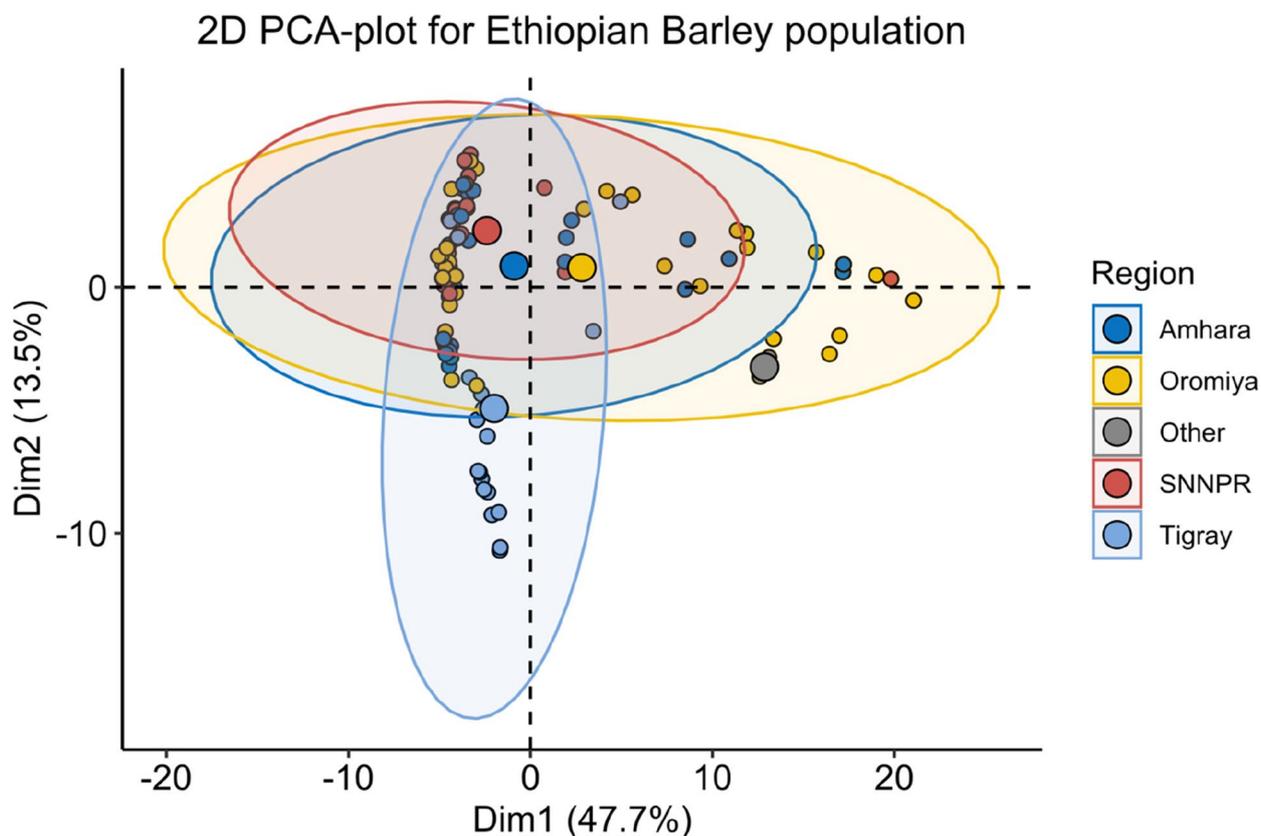


Fig. 2 Principal component analysis showing the clustering among the region of origins of 105 barley genotypes using 10,103 SNP markers. The genotypes are grouped into three main groups. Colored dots are genotypes from various region of origins and grouped by regions, according to the legend. PC1 and PC2 are shown on the X and Y-axis, respectively, aside from their explained variance. Abbreviations: SNNPR = Southern Nations, Nationalities and Peoples Region

14 in Clusters II to 60 in Cluster I (Fig. 4, Table 5). Of all, cluster I was the largest cluster, consisting of 3(2.9%), 16(15.4%), 18(17.2%) and 23(21.9%) from Tigray, Oromia, SNNP and Amhara regions, respectively. In terms of altitude based clustering, the majority of the genotypes were included at an altitude of 2001–3000 masl. However, cluster II consists of a small number of barley genotypes, commonly from Amhara (4.8%) and Oromia (7.6%) regions with originated at medium and high altitude ranges (Fig. 4, Table 5). Cluster III encompassed genotypes only from Oromia (1.9%) and Tigray (13.3%) regions. The greatest number of genotypes included from an altitude of 2001–2500 masl. Cluster IV comprises genotypes from all regions with the highest percentage in the Oromia (5.7%) and the genotypes comprised at the medium altitude (Fig. 4, Table 5).

Population structure

A total of 10,103 SNP markers were used for the population structure analysis of the 105 barley genotypes. The best number of K, which clearly defined the number of

populations as $K=4$, revealed that four subpopulations should include all the 105 barley genotypes with a great probability (Fig. 5). Each K is shown in diverse colors (Fig. 5), the subpopulations described in blue-green and purple observed a high proportion of the variation with the barley genotypes confirming the results of AMOVA and heat map results.

Discussions

The genetic diversity study of germplasm is the greatest way to understand ancestry relations and to achieve proficient management of crop genetic resources to improve breeding programs [43, 44]. Such genetic diversity analysis was essential for plant breeders to perform strategic integration and target selection while preserving important economic traits of individual crops. In this study, the highest gene diversity ($GD=0.311$) was revealed in barley genotypes (Table 1). This high gene diversity may be a natural crossing due to cultivating mixed genotypes in a field [45]. The average gene diversity was 0.253, which is lower than that indicated in worldwide barley genotypes

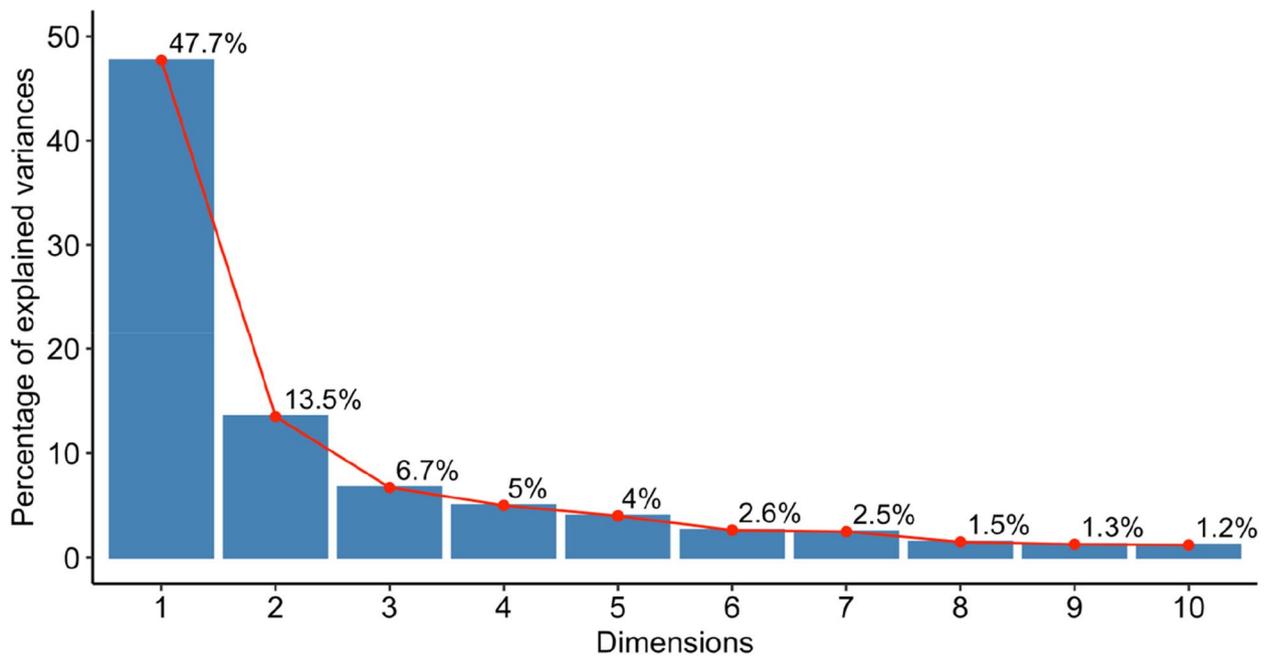


Fig. 3 Percentages of explained variances for 105 barley genotypes using SNP markers

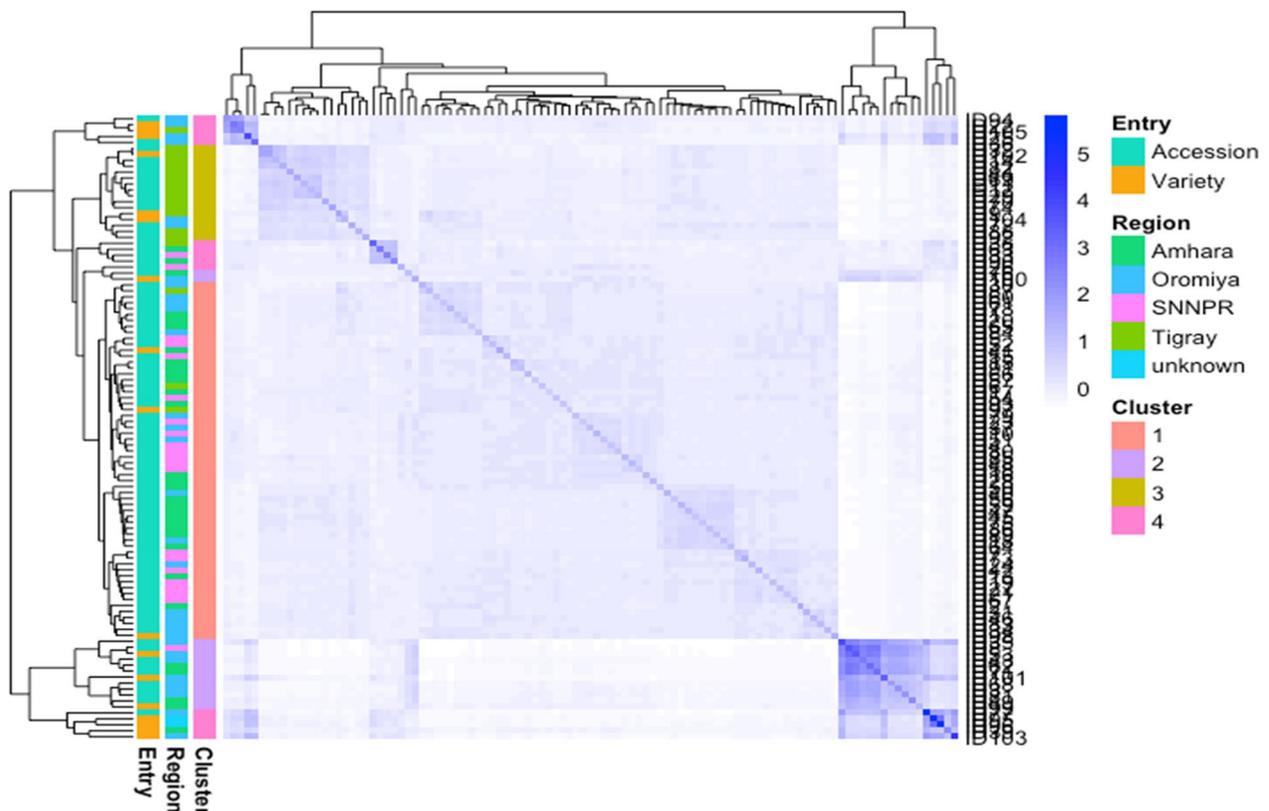


Fig. 4 Heat map analysis shows relationship distance based on the color distribution among 105 barley populations genotyped using 10,103 SNP markers (for genotypes name and collection origins refer to Additional file 1: Table S1). Colored plots on the left side are entry genotypes, regions and clusters at the various stages, according to the legend to the right. On the left of the matrix, barley genotypes are grouped into four distinct clusters. Abbreviations: SNNPR = Southern Nations, Nationalities and Peoples Region, Unknown = unidentified genotype names

Table 5 Clustering of 105 barley genotypes into four groups using SNP markers, the distribution of genotypes across the region of origins and altitudinal ranges

Regions	Cluster I	Cluster II	Cluster III	Cluster IV	Total
Oromia	16	8	2	6	32
Amhara	23	5	0	3	31
Tigray	3	0	14	2	19
SNNP	18	1	0	2	21
ICARDA/Other	0	0	0	2	2
Total	60	14	16	15	105
Altitude Ranges					
≤2000	8	1	2	3	14
2001–2500	22	5	9	8	44
2501–3000	21	3	5	3	32
≥3001	9	5	0	1	15
Total	60	14	16	15	105

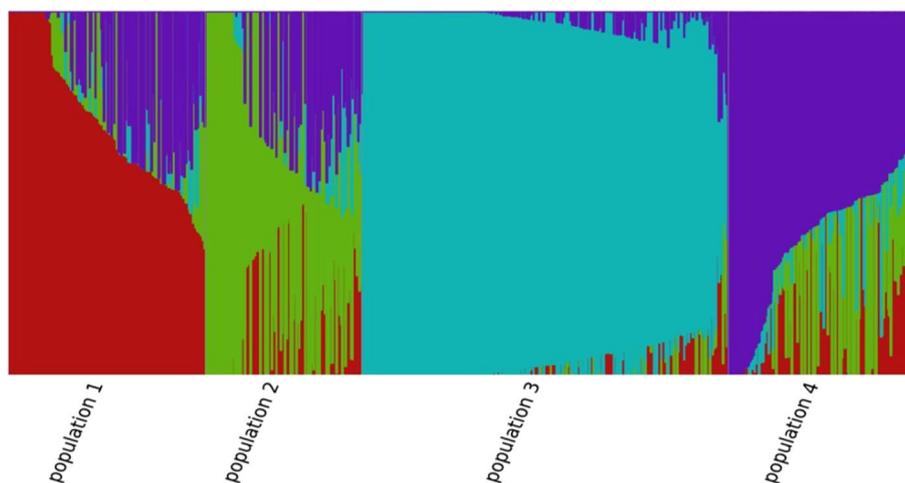
Abbreviations: SNNP Southern Nations, Nationalities and Peoples, ICARDA International Centre for Agricultural Research in the Dry Areas

(GD=0.388) [46], in ICARDA spring barley collections (GD=0.366) [47], in Nordic spring barley collections (GD=0.359) [48] and Egyptian barley collections (GD=0.550) [49]. The PIC ranged was 0.151 to 0.271 with an average polymorphism value of 0.216 (Table 1). A previous study indicated that PIC values >0.5 mean highly informative markers, $0.5 > \text{PIC} > 0.25$ is an informative marker, and $\text{PIC} < 0.25$ is a fairly informative marker [50]. Therefore, our results showed that the SNP markers were informative and polymorphic. The average and range values of polymorphism observed in the present study were lower than the reported range values of 0.474 to 0.652 with an average value of 0.552 in barley landraces [33], 0.34 to 0.83 with an average of 0.57 in barley collections [51]. Another study with Ethiopian durum wheat

by Alemu et al. [52] estimated the ranged value from 0.01 to 0.375 with a mean PIC of 0.375. Higher MAF (MAF=0.146) with an average value of 0.118 was found in barley genotypes (Table 1), indicating further valuable genes can be exploited from those genotypes. The lower MAF obtained may be due to the lower number of genotypes studied from those regions. Our result is lower than that reported for the MAF range was 0.517 to 0.520 in Ethiopian sorghum with a mean value of 0.518 [53].

The observed heterozygosity was higher ($H_o = 0.088$) in the Oromia region, indicating the existence of high genetic variability between the genotypes (Table 1). This may be due to the high rate of outcrossing in barley [54]. In this work, the average observed heterozygosity was smaller ($H_o = 0.045$) than the average expected heterozygosity ($H_e = 0.163$), showing highly related among the barley genotypes. This could be related to the existence of gene flow along with the regions each growing season during the exchange of seeds and interbreeding. Comparatively overall lower heterozygosity is obtained in all the regions, this may be due to the factor explained by the cleistogamy in barley that, the flower sheds its pollen before opening, forcing plants with this habit to be entirely autogamous, which can reduce the heterozygosity [55]. Our result is lower than the reported ranges observed from 0.594 to 0.662, and the expected 0.688 to 0.773 heterozygosity for Ethiopian barley landraces [33]. This is probably due to the source of genotypes and the samples included improved varieties, which are obtained in the lower heterozygosity ranges in our study.

A higher genetic variation of 52.85% was observed within the populations, showing the variations among the barley genotypes in the regions (Table 2). This confirmed that Ethiopia has been recognized as a center

**Fig. 5** Population structure of 105 barley genotypes (K=4). Each vertical bar represents genotypes that are divided into K-colored segments

of diversity for barley [14]. However, in our study, the genetic diversity of barley genotypes was 0.253, which was lower than (60.31%) for the North African germplasm [56]. The existence of variations in our work may be because of the outcrossing that leads to introducing new genetic components into the population. The large genetic diversity within populations could be a result of natural adaptation or extensive exchange of seeds among farmers between environments [57]. Therefore, the presence of genetic variation within the populations obtained in this study can be utilized in breeding programs to enhance barley productivity, the parental line selection from within the populations could be more valuable compared to selection from among the populations. The AMOVA also showed the genetic variation (46.43%) within accessions (Table 2), this could be associated with climatic variability and agro-ecological heterogeneity in the region of origins. The presence of diversity within accessions confirmed the potential of a genetic variant, which is the source material for barley breeding [30, 33]. The utilization of diversity within accessions using pure-bred selection has been confirmed to provide germplasm for desirable traits [8]. The presence of genetic variation within the populations and accessions suggests that there might be a margin for barley improvement in genomic selection for cultivar development.

The genetic variance fixation index (F_{ST}) was an evaluation of population variations because of genetic structure [33, 58]. In the present study, lower values of ($F_{ST}=0.019, 0.022, 0.038$) genetic differentiations were shown in Table 3. This showed that the barley genotypes originating from those regions of origin had very close ancestry, and the close origins may be due to high seed exchange that can help barley breeding to exploit hybrids. This work obtained that the very closest pairwise value ($F_{ST}=0.019$) was observed between the Amhara and Oromia genotypes while distant dissimilarity ($F_{ST}=0.117$) between the SNNP and Tigray genotypes (Table 3). Tigray genotypes were most distantly ($F_{ST}=0.089-0.117$) related to other regions of the populations, which showed geographical origins could be the basis of a genetic variant. The genotypes obtained from the Tigray region may be exploited as a source of breeding materials to improve genetic diversity in barley breeding through hybridization programs. Our study result revealed moderate genetic differentiation ($F_{ST}=0.117$) for the pairwise comparison among the barley population implying a smaller amount of genetic relatedness and this indicated a lesser genetic relationship. This result was confirmed by genetic differentiation largely higher in marginal populations than in the favorable environment. Marginal stands are grouped by geographic and genetic separation because of spatial segregation and restricted

gene flow [59]. A similar finding was reported in this study by Dido et al. [33] indicating moderate genetic differentiation ($F_{ST}=0.082$) was shown in the Ethiopian barley population. On the contrary, the largest genetic differentiation ($F_{ST}=0.257$) was revealed between the barley population by Allel et al. [56].

Genetic distance is the measure of genetic differences that exist among individuals or populations and can be measured by allelic differences [60]. The present study showed the highest genetic distance (0.08) was found between genotypes derived from Tigray and ICARDA (Table 4), which indicated that the barley genotypes had a high genetic difference. This could be due to the restricted gene flow and the influence of eco-geographical differences on the existence of high genetic distance. On the other hand, the least genetic distance (0.002) was recorded between Amhara and ICARDA (Table 4). Research report explained that the lower genetic distance among collection regions probably has high levels of the farmer to farmer seed exchange and gene flow across regions [53].

The genetic identity ranged from 0.971 to 0.998 (Table 4). The least genetic identity was observed between genotypes collected from SNNP and ICARDA. However, the highest genetic identity (0.998) was from Amhara and Tigray, Oromia and SNNP, and Amhara and ICARDA (Table 4). This is possibly due to natural selection for shared genetic components being the key force in determining the high genetic identity between the geographical origins. Genetic divergence as the lower cause is a result of geographical isolation and distinctive agro-ecological conditions shared by genetic materials [61]. Our result is relatively similar to the genetic identity of North African barley germplasm (0.956) reported in an earlier study by Allel et al. [56].

The first two PCA were explaining 61.2% of the total genetic variation (Fig. 2). This indication highlights the potential of highly informative and selective SNP markers for genetic studies in barley, which might underpin conservation and future breeding efforts. Besides, this result showed genetic differences exist among genotypes confirming the result indicated by AMOVA, which revealed a significant genetic variation within population and accessions. There was a comparison to the present study that has been reported in earlier studies on barley germplasm [49, 51], durum wheat [52] and sorghum germplasm [53, 62]. The heat map analysis also grouped the genotypes into four major clusters, reflecting the origin of the genotypes and their genetic relationships (Fig. 4, Table 5). Likewise, the North African region of barley collection from 14 countries was grouped into four clusters [51]. Our result is also parallel with clustering the Ethiopian barley landraces into three clusters [33,

42] and the Egyptian barley genotypes into three main groups using SNP markers [49].

The present result has shown that the distribution pattern of the genotypes into different groups indicated the existence of significant genetic variations among the barley genotypes. For example, among the clusters, cluster I was the largest cluster containing 60 genotypes (57.2%), and most of the genotypes included an altitude of 2001–3000 masl (Fig. 4, Table 5). These stated that structure analysis grouped the barley genotypes with greater genotypic similarity and this may be used as a source of breeding material to enhance genetic variants in barley breeding. Clustering genotypes into genotypically similar clusters of diverse collections are significant for barley improvements such as selecting parents for hybrid [49, 63] and the development of modern breeding lines [64]. In the genotypes, groups were admixed into the varied clusters irrespective of their collections origins (Fig. 4, Table 5). For instance, genotypes collected from the Oromia region (32) were grouped into four clusters (sixteen genotypes in cluster I, eight genotypes in cluster II, two genotypes in cluster III and six genotypes in cluster IV). Genotypes (31) were collected from the Amhara region also clustered into three groups; 23 genotypes in cluster I, 5 genotypes in cluster II and 3 genotypes in cluster IV (Fig. 4, Table 5). This finding agrees with the admixture that could be associated with gene flow facilitated by the continuous exchange of seeds among smallholder farmers in various agro-ecologies in the shared market and the contentious introduction of new seeds into the respective growing regions [33, 52, 65, 66]. The analysis of population structure also confirms the barley genotypes clustered into four subgroups (Fig. 5), while they showed no clear clustering pattern of grouping. These unstructured grouping and the mixed genetic background suggested that the genotype shared a similar lineage, this is possibly due to the high exchange of planting material between the region of origins. A similar result was reported by the absence of consistent population structure among the Ethiopian barley landraces [31, 42] and Jordan barley germplasm [67], which was related to high seed-mediated gene flow.

Conclusions

In the current study, the SNP data generated using advanced molecular tools provide useful information that can be utilized in breeding and genetic research in barley. Based on SNP data, the barley genotypes were genetically divergent. The AMOVA indicated high genetic variations within the populations and genotypes. This high diversity could be the foundation for developing and generating desirable new barley varieties with superior grain yield potential and wide adaptability, enhanced with abiotic

and biotic resistant traits. This study identified four subpopulations, considered as four independent subpopulations in the improvement program, but was not grouped the genotypes according to collection origin and adaptation zone. This information on genetic diversity and population structure of Ethiopian barley genotypes will be applied for current and future research using genome-wide association and genomic selection for economically useful traits in barley.

Methods

Plant materials

One hundred five barely panels comprised of eighty-five barley accessions and sixteen improved varieties including two landraces and two wild crosses were used in this study. These lines were chosen based on cultivars' regional passport information and the breeding merit they have for subsequent germplasm enhancement. The eighty-five barley accessions were obtained from the *ex-situ* collection of the Ethiopian Biodiversity Institute (EBI) along with their passport data (Additional file 1: Table S1). The random sampling procedure was modified to allow the equal representation of barley accessions from the 1976 to 2018 collection periods in four regions of Ethiopia (Oromia, Amhara, Tigray, and Southern Nations, Nationalities and Peoples) (Additional file 2: Fig. S1). To get detailed information on genetic diversity, genotypes were collected from a wide range of altitudes (≤ 2000 , 2001–2500, 2501–3000 and ≥ 3001 masl) (Additional file 2: Fig. S1). The improved varieties including landraces were obtained from Universities, national and regional agricultural research centers, and two wild crosses of barley from Debre Zeit Agricultural Research Center (DZARC) primarily introduced from the International Centre for Agricultural Research in the Dry Areas (ICARDA) (Additional file 1: Table S1).

Planting, leaf sampling and genomic DNA extraction

Each barley genotype was sowed on 02 August 2021 in a seedling tray, five seedlings each, in a greenhouse at National Agricultural Biotechnology Research Center (NABRC), Holetta. For two weeks old barley leaf samples were collected and pooled from the same genotype in equal amounts. Then samples were placed into 96 collection plates (96-well plate holds 12×8 -strip tubes). During the sample preparation, the leafcutters (scissors) were sterilized with 70% alcohol before cutting the next genotype to prevent cross-contaminate. The collected leaf samples were first freeze-dried at -20 °C for 24 h. Then the freeze-dried leaf samples were shipped for genotyping to the Integrated Genotyping Service and Support (IGSS) platform located at Biosciences Eastern and Central Africa- International Livestock Research Institute

Hub (BecA-ILRI Hub) based in Nairobi. Genomic DNA from 105 barley genotypes was extracted using the Nucleomag Plant Genomic DNA extraction kit following the manufacturer's instructions. The genomic DNA concentration was used in the range of 50–100 ng/μl. The quality and quantity of extracted DNA in each sample were determined by electrophoresis on 0.8% agarose gel.

SNP genotyping

Genomic libraries were constructed according to Kilian et al. [68], complexity reduction method through digestion of genomic DNA using a combination of PstI and HpaII enzymes and ligation of bar-coded adapters followed by PCR amplification of adapter-ligated fragments. Libraries were sequenced using single-read sequencing runs for 77 bases. Next-generation sequencing technology was carried out using HiSeq2500 (Illumina). The markers scoring was achieved using DArTsoft14, which is an in-house marker scoring pipeline based on algorithms. Two types of markers were scored, SilicoDArT markers and SNP markers which were both scored '1' for presence, and '0' for absence and '-' for calls with non-zero count however too low counts to scored confidently as "1" for the SilicoDArT while the sequences SNPs were scored '0' for reference allele homozygote, '1' for SNP allele homozygote and '2' for heterozygote. Totalities of 31,646 silicoDArTs and SNP markers were employed to genotype the materials. Both SilicoDArT and SNP markers were aligned to the reference genome of *Hordeum vulgare_v2.0* [69], to identify chromosome positions. Our study which included only SNP markers data was used for this analysis after SNP calling and imputation.

SNP calling and data filtering

An initial set of 14,454 single nucleotide polymorphisms (SNPs) containing monomorphic and undefined chromosomes or positions were removed from the raw dataset. Then further quality control was performed, and SNP markers with > 95% call rate, minor allele frequency (> 5%), and missing rate per sample (< 10%) and per SNP (< 30%) were retained for downstream analysis using R software version 4.1.3 [70]. The diversity analysis was performed using complete (non-missing) data; therefore, SNP with missing loci was imputed using the R package snpReady [71]. Finally, 10,103 (69.90%) SNP markers were retained for further analysis.

Data analysis

Genetic parameters such as polymorphism information content (PIC), genetic diversity, heterozygosity (observed and expected), minor allele frequency (MAF), genetic differentiation and genetic distance

were computed using the R package snpReady [71]. The molecular variance analysis (AMOVA) was performed at different hierarchical levels (among and within the populations and accessions) in the stats R package using the *aov* function.

To explore the genetic structure of the 105 barley accessions we first undertook a cluster analysis using the agglomerative hierarchical algorithm (wards method, Euclidean distance) (Anderberg [72] of the locus-by-entry reference allele count table for the SNP markers that passed the filtration process) and in the dendrogram was plotted across the testing location or regions. The elbow method (implemented in the R package factextra) was used to determine the optimum number of clusters (k) [73, 74]. The best *K*-value for estimating an optimum subpopulation size for the dataset was determined based on peak ΔK values following the Evanno method [75]. A principal component analysis (PCA) of the SNP markers was conducted using the *prcomp* function to summarise the contributions of each part to the variation that existed in the population. The genetic diversity pattern among the population was visualized using the *heatmap* R package [71].

Abbreviations

EBI	Ethiopian Biodiversity Institute
EIAR	Ethiopian Institute of Agricultural Research
DZARC	Debre Zeit Agricultural Research Center
ICARDA	International Centre for Agricultural Research in the Dry Areas
SNNP	Southern Nations, Nationalities and Peoples
NABRC	National Agricultural Biotechnology Research
AMOVA	Analysis of molecular variance
IGSS	Integrated Genotyping Service and Support
BecA-ILRI Hub	Biosciences Eastern and Central Africa- International Livestock Research Institute Hub

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12863-023-01109-6>.

Additional file 1: Table S1. List of 105 barley genotypes obtained from the *ex-situ* collection of the Ethiopian Biodiversity Institute (EBI), Ethiopian Institute of Agricultural Research (EIAR) and Universities with the region of origins and altitude ranges.

Additional file 2: Figure S1. The administrative map of Ethiopia indicates the collection points of the barley genotypes. Dots represent the barley genotypes in different colors depending on various region of origins and altitude ranges, according to the legend.

Acknowledgements

The authors would like to acknowledge the EBI, Ethiopian research centers and Universities for the provision of plant materials for this study, and NABRC for the providing of initial laboratory and greenhouse facilities. The authors are also acknowledging research fellows and experts who provide valuable input to this work. A special thanks to Girma Mengistu (PhD), Tilahun Mekonnen (PhD) and Mr. Sisay Kidane (PhD candidate) for their support.

Authors' contributions

MY, MK, TF, BL and ABW: conceived and designed the study, visualization and interpreted the data. MY: analyzed and wrote the manuscript. MK, TF, BL and ABW: revised and edited the manuscript. MF: contributed to data quality control and analysis, visualization of the data, revised and edited the manuscript. All authors had read and agreed to the published version of the manuscript.

Funding

The Ethiopian Ministry of Science and Higher Education financially supported this work, and genotyping was made possible with the support/low-priced services through the Integrated Genotyping Service and Support (IGSS) of BecA-ILRI Hub based in Nairobi.

Availability of data and materials

The datasets generated and/or analyzed during the current study are available in the European Variation Archive (EVA) at EMBL-EBI under accession number PRJEB59022 (<https://www.ebi.ac.uk/eva/?eva-study=PRJEB59022>).

Declarations

Ethics approval and consent to participate

All experimental plant materials (barley genotypes) were obtained from the *ex-situ* collection of the Ethiopian Biodiversity Institute (EBI), the Ethiopian Institute of Agricultural Research (EIAR) and Universities. The materials used in this study were owned by governmental institutions established for research, and freely available for non-commercial purposes. Data exchange was under institutional, national and international plant import/export guidelines.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 28 August 2022 Accepted: 31 January 2023

Published online: 14 February 2023

References

- Sato K. History and future perspectives of barley genomics. *DNA Res*. 2020;27:1–8. <https://doi.org/10.1093/dnares/dsaa023>.
- FAOSTAT. Production and Trade. Food and Agriculture Organization of the United Nations. Rome: FAO. 2020. <http://www.fao.org/faostat/en/#data/QC/visualize>. Accessed 10 Mar 2022.
- Newman CW, Newman RK. A brief history of barley foods. *Cereal Foods World*. 2006;51:4–7.
- Baik BK, Ullrich SE. Barley for Food: Characteristics, Improvement, and Renewed Interest. *J Cereal Sci*. 2008;48:233–42. <https://doi.org/10.1016/j.jcs.2008.02.002>.
- CSA(Central Statistical Agency). Federal democratic republic of Ethiopia. Central statistical agency. Agricultural sample survey, Volume III, Report on Crop and Livestock Product Utilization. Statistical. Cent Stat Agency (CSA), Addis Ababa, Ethiopia. 2018/2019.
- TantoHadad T, Rau D, Bitocchi E, Papa R. Genetic diversity of barley (*Hordeum vulgare* L.) landraces from the central highlands of Ethiopia: Comparison between the Belg and Meher growing seasons using morphological traits. *Genet Resour Crop Evol*. 2009;56:1131–48. <https://doi.org/10.1007/s10722-009-9437-z>.
- Abebe TD, Bauer AM, On JL. Morphological diversity of Ethiopian barley (*Hordeum vulgare* L.) in relation to geographic regions and altitudes. *Hereditas*. 2010;147:154–64. <https://doi.org/10.1111/j.1601-5223.2010.02173>.
- Lakew B, Semeane T, Alemayehu F, et al. Exploiting the diversity of barley landraces in Ethiopia. *Genet Resour Crop Evol*. 1997;44:109–16.
- Kaso T, Guben G. Review of barley value chain management in Ethiopia. *J Biol Agric Healthcare*. 2015;5:84e97.
- Yitbarek S, Bekele H, Getaneh W, Dereje T. Disease surveys and loss assessment studies on barley, In Hailu Gebre and Joop van Leur(eds.). Barley Research in Ethiopia. Past work and prospects. Proceedings of the 1st Barley Review Work Shop, Addis Ababa: IAR/ICARDA. 1996.p.105–115.
- Mashilla DW. Importance, Biology, Epidemiology, and Management of Loose Smut (*Ustilago nuda*) of Barley (*Hordeum vulgare*). *East Afr J Sci*. 2019;13:89–108.
- Muñoz-Amatriaín M, Cuesta-Marcos A, Endelman JB, Comadran J, Bonman JM, et al. The USDA barley core collection: Genetic diversity, population structure, and potential for genome-wide association studies. *PLoS ONE*. 2014;9:1–13. <https://doi.org/10.1371/journal.pone.0094688>.
- Milner SG, Jost M, Taketa S, Mazón ER, Himmelbach A, Oppermann M, et al. Genebank genomics highlights the diversity of a global barley collection. *Nat Genet*. 2019;51:319–26. <https://doi.org/10.1038/s41588-018-0266-x>.
- Harlan JR. Ethiopia: A center of diversity. *Econ Bot*. 1969;23:309–14. <https://doi.org/10.1007/BF02860676>.
- Eticha F, Gausgruber H, Berghoffer E. Multivariate analysis of agronomic and quality traits of hull-less spring barley (*Hordeum vulgare* L.). *J Plant Breed Crop Sci*. 2010;2:81–95.
- Mekonnen B, Lakew B, Dessalegn T. Morphological diversity and association of traits in Ethiopian food barley (*Hordeum vulgare* L.) landraces in relation to regions of origin and altitudes. *J Plant Breed Crop Sci*. 2015;7:44–54. <https://doi.org/10.5897/JPBCS2014.0480>.
- Samberg LH, Fishman L, Allendorf FW. Population genetic structure in a social landscape: barley in a traditional Ethiopian agricultural system. *Evol Appl*. 2013;6:1133–45. <https://doi.org/10.1111/eva.12091>.
- EBI. Ethiopian Biodiversity Institute. 2018. <http://www.ebi.gov.et/>. Accessed 20 Jan 2021.
- Konopka J. Global strategy for the *ex-situ* conservation and use of barley germplasm. 4–6 June 2007, Tunis, Tunisia. 2007.
- Bonman JM, Bockelman HE, Jackson LF, Steffenson BJ. Disease and insect resistance in cultivated barley accession from the USDA National Small Grains Collection. *Crop Sci*. 2005;45:1271–80.
- Qualse CO, Mcguire PE, Vogt HE, Topcu MA. Ethiopia as a Source of Resistance to the Barley Yellow Dwarf Virus in Tetraploid Wheat. *Crop Sci*. 1977;17:527–9. <https://doi.org/10.2135/cropsci1977.0011183X001700040011x>.
- Beoni E, Chrpová J, Jarošová J, Kundu JK. Survey of Barley yellow dwarf virus incidence in winter cereal crops, and assessment of wheat and barley resistance to the virus. *Crop Pasture Sci*. 2016;67:1054–63. <https://doi.org/10.1071/CP16167>.
- Piffanelli P, Ramsay L, Waugh R, Benabdelmoun A, D'Hont A, et al. A barley cultivation-associated polymorphism conveys resistance to powdery mildew. *Nature*. 2004;430:887–91.
- Yitbarek S, Berhane L, Fikadu A, Van Leur JA, Grando S, Ceccarelli S. Variation in Ethiopian barley landrace populations for resistance to barley leaf scald and net blotch. *Plant Breeding*. 1998;117:419–23. <https://doi.org/10.1111/j.1439-0523.1998.tb01966.x>.
- Munck L, Karlsson KE, Hagberg A, Eggum BO. Gene for improved nutritional value in barley seed protein. *Science*. 1970;168:985–7.
- Lance RC, Nilan RA. Screening for low acid-soluble β -glucan barleys. *Barley Genet Newsl*. 1980;10:41.
- Dido AA, Krishna MSR, Assefa E, Degefu DT, Singh BJK, Tesfaye K. Genetic diversity, population structure and relationship of Ethiopian barley (*Hordeum vulgare* L.) landraces as revealed by SSR markers. *J Genet*. 2022;101:1–20. <https://doi.org/10.1007/s12041-021-01346-7>.
- Assefa A, Labuschagne MT, Viljoen CD. AFLP analysis of genetic relationships between barley (*Hordeum vulgare* L.) landraces from north Shewa in Ethiopia. *Conserv Genet*. 2007;8:273–80.
- Demissie A, Bjørnstad Å, Kleinhofs A. Restriction Fragment Length Polymorphisms in Landrace Barleys from Ethiopia in Relation to Geographic, Altitude, and Agro-Ecological Factors. *Crop Science*. 1998;38:237–43. <https://doi.org/10.2135/cropsci1998.0011183X0038000100>.
- Tanto Hadado T, Rau D, Bitocchi E, Papa R. Adaptation and diversity along an altitudinal gradient in Ethiopian barley (*Hordeum vulgare* L.) landraces revealed by molecular analysis. *BMC Plant Biology*. 2010;10. <https://doi.org/10.1186/1471-2229-10-121>.
- Abebe TD, Léon J. Spatial and temporal genetic analyses of Ethiopian barley (*Hordeum vulgare* L.) landraces reveal the absence of a distinct population structure. *Genet Resour Crop Evol*. 2013;60:1547–58. <https://doi.org/10.1007/s10722-012-9941-4>.

32. Misganaw A, Kidane S, Tesfu K. Assessment of genetic diversity among released and elite Ethiopian barley genotypes using simple sequence repeat (SSR) markers. *Afr J Plant Sci.* 2017;11:114–22. <https://doi.org/10.5897/AJPS2017.1541>.
33. Dido AA, Degefu DT, Assefa E, Krishna MSR, Singh BJK, Tesfaye K. Spatial and temporal genetic variation in Ethiopian barley (*Hordeum vulgare* L.) landraces as revealed by simple sequence repeat (SSR) markers. *Agri Food Secur.* 2021;10:1–14. <https://doi.org/10.1186/s40066-021-00336-3>.
34. Close TJ, Bhat PR, Lonardi S, Wu YH, Rostoks N, et al. Development and implementation of high-throughput SNP genotyping in barley. *BMC Genomics.* 2009;10:582.
35. Rostoks N, Mudie S, Cardle L, Russell J, Ramsay L, Booth A, Svensson JT, et al. Genome-wide SNP discovery and linkage analysis in barley based on genes responsive to abiotic stress. *Mol Genet Genomics.* 2005;274:515–27. <https://doi.org/10.1007/s00438-005-0046-z>.
36. Li Z, Lhundrup N, Guo G, Dol K, Chen P, Gao L, et al. Characterization of Genetic Diversity and Genome-Wide Association Mapping of Three Agronomic Traits in Qingke Barley (*Hordeum Vulgare* L.) in the Qinghai-Tibet Plateau. *Front Genet.* 2020;11:638. <https://doi.org/10.3389/fgene.2020.00638>.
37. Morgil H, Gercek YC, Tulum I. Single nucleotide polymorphisms (SNPs) in plant genetics and breeding. *The Recent Topics in Genetic Polymorphisms: IntechOpen.* 2020. <https://doi.org/10.5772/intechopen.91886>.
38. Nordborg M, Hu TT, Ishino Y, Jhaveri J, Toomajian C, Zheng H, et al. The pattern of polymorphism in *Arabidopsis thaliana*. *PLoS Biol.* 2005;3:1289–99.
39. Amdadul H, Shahina A, Sup N, Hoy T, Yu J, Kwon K. Identification of functional SNPs in genes and their effects on plant phenotypes. *J Plant Biotechnology.* 2016;43:1–11.
40. Alkan C, Coe BP, Eichler EE. Genome structural variation discovery and genotyping. *Nat Rev Genet.* 2011;12:363–76.
41. Qiu X, Gong R, Tan Y, Yu S. Mapping and characterization of the major quantitative trait locus qSS7 associated with increased length and decreased width of rice seeds. *Theor Appl Genet.* 2012;125:1717–26.
42. Teklemariam SS, Bayissa KN, Matros A, Pillen K, Ordon F, Wehner G. The genetic diversity of Ethiopian barley genotypes in relation to their geographical origin. *PLoS ONE.* 2022;17:e0260422. <https://doi.org/10.1371/journal.pone.0260422>.
43. Al-Abdallat AM, Karadshah A, Hadadd NI, Akash MW, Ceccarelli S, Baum M, et al. Assessment of genetic diversity and yield performance in Jordanian barley (*Hordeum vulgare* L.) landraces grown under Rainfed conditions. *BMC Plant Biology.* 2017;17:1–13. <https://doi.org/10.1186/s12870-017-1140-1>.
44. Ketema S, Tesfaye B, Keneni G, et al. DArTSeq SNP-based markers revealed high genetic diversity and structured population in Ethiopian cowpea [*Vigna unguiculata* (L) Walp] germplasm. *PloS one.* 2020;15:e0239122.
45. Mengistu DK, Kiros AY, Pè ME. Phenotypic diversity in Ethiopian durum wheat (*Triticum turgidum* var. durum) landraces. *Crop J.* 2015;3:190–9. <https://doi.org/10.1016/j.cj.2015.04.003>.
46. Sun D, Ren W, Sun G, Peng J. Molecular diversity and association mapping of quantitative traits in Tibetan wild and worldwide originated barley (*Hordeum vulgare* L.) germplasm. *Euphytica.* 2011;178:31–43. <https://doi.org/10.1007/s10681-010-0260-6>.
47. Amezrou R, Gyawali S, Belqadi L, Chao S, Arbaoui M, Mamidi S, et al. Molecular and phenotypic diversity of ICARDA spring barley (*Hordeum vulgare* L.) collection. *Genet Resour Crop Evol.* 2018;65:255–69. <https://doi.org/10.1007/s10722-017-0527-z>.
48. Bengtsson T, Åhman I, Bengtsson T, Manninen O, Veteläinen M, et al. Genetic diversity, population structure and linkage disequilibrium in Nordic spring barley (*Hordeum vulgare* L. subsp. *vulgare*). *Genet Resour Crop Evol.* 2017;64:2021–33. <https://doi.org/10.1007/s10722-017-0493-5>.
49. Elakhdar A, Kumamaru T, Qualset CO, Brueggeman RS, Amer K, Capochichi L. Assessment of genetic diversity in Egyptian barley (*Hordeum vulgare* L.) genotypes using SSR and SNP markers. *Genet Resour Crop Evol.* 2018;65:1937–51. <https://doi.org/10.1007/s10722-018-0666-x>.
50. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet.* 1980;32:314–31.
51. Nyiraguhirwa S, Grana Z, Henkrar F, Ouabbou H, Mohammed I, Udupa SM. Genetic diversity and structure of a barley collection predominantly from the North African region. *Cereal Res Commun.* 2021. <https://doi.org/10.1007/s42976-021-00209-2>.
52. Alemu A, Feyissa T, Letta T, Abeyo B. Genetic diversity and population structure analysis based on the high-density SNP markers in Ethiopian durum wheat (*Triticum turgidum* ssp. durum). *BMC Genetics.* 2020;21:1–12. <https://doi.org/10.1186/s12863-020-0825-x>.
53. Girma M, Hussein S, Mark L, Dagnachew L, Ermias A, Isack M. Genetic diversity assessment of sorghum (*Sorghum bicolor* (L.) Moench) landraces using SNP markers. *South Afr J Plant Soil.* 2020. <https://doi.org/10.1080/02571862.2020.1736346>.
54. Parzies HK, Spoor W, Ennos RA. Genetic diversity of barley landrace accessions (*Hordeum vulgare* ssp. *vulgare*) conserved for different lengths of time in ex-situ gene banks. *Heredity.* 2000;84:476–86.
55. Wang N, Ning S, Pourkheirandish M, Honda I, Komatsuda T. An alternative mechanism for cleistogamy in barley. *Theor Appl Genet.* 2013;126:2753–62.
56. Allel D, Ben-Amar A, Lamine M, Abdely C. Relationships and genetic structure of North African barley (*Hordeum vulgare* L.) germplasm revealed by morphological and molecular markers: Biogeographical considerations. *South Afr J Botany.* 2017;112:1–10. <https://doi.org/10.1016/j.sajb.2017.05.005>.
57. Uba CU, Oselebe HO, Tesfaye AA, Abteu WG. Genetic diversity and population structure analysis of Bambara groundnut (*Vigna subterranean* L.) landraces using DArT SNP markers. *PLoS ONE.* 2021;16:e0253600. <https://doi.org/10.1371/journal.pone.0253600>.
58. Tehseen MM, Istiqlil D, Kehel Z, Sansaloni CP, da Silva Lopes M, Kurtulus E, et al. Genetic Diversity and Population Structure Analysis of *Triticum aestivum* L. Landrace Panel from Afghanistan. *Genes.* 2021;12:340. <https://doi.org/10.3390/genes12030340>.
59. Tóth EG, Tremblay F, Housset JM, et al. Geographic isolation and climatic variability contribute to genetic differentiation in fragmented populations of the long-lived subalpine conifer *Pinus cembra* L. in the western Alps. *BMC Evol Biol.* 2019;19:190.
60. Nei M. *Molecular Evolutionary Genetics.* New York: Columbia University Press; 1987.
61. Guillot G, Rousset F. Dismantling the Mantel tests. *Methods Ecol Evol.* 2013;4:336–44.
62. Enyew M, Feyissa T, Carlsson AS, Tesfaye K, Hammenhag C, Geleta M. Genetic Diversity and Population Structure of Sorghum [*Sorghum Bicolor* (L.) Moench] Accessions as Revealed by Single Nucleotide Polymorphism Markers. *Front Plant Sci.* 2022;12. <https://doi.org/10.3389/fpls.2021.799482>.
63. Kumar Y, Sehrawat KD, Singh J, Shehrawat S. Identification of promising barley genotypes based on morphological genetic diversity. *J Cereal Res.* 2021;13:79–88. <https://doi.org/10.25174/2582-2675/2021/108051>.
64. Sansaloni C, Franco J, Santos B, Percival-Alwyn L, Singh S, Petroli C, et al. Diversity analysis of 80,000 wheat accessions reveals consequences and opportunities of selection footprints. *Nat Commun.* 2020;11:1–12. <https://doi.org/10.1038/s41467-020-18404-w>.
65. Desmae H, Jordan D, Godwin I. DNA markers reveal genetic structure and localized diversity of Ethiopian sorghum landraces. *Afr J Biotechnol.* 2016;15:2301–11. <https://doi.org/10.5897/AJB2016.1540>.
66. Woldeyohannes AB, Accotto C, Desta EA, Kidane YG, Fadda C, Pè ME, Dell'Acqua M. Current and projected eco-geographic adaptation and phenotypic diversity of Ethiopian teff (*Eragrostis teff*) across its cultivation range. *Agr Ecosyst Environ.* 2020;300:1–10. <https://doi.org/10.1016/j.agee.2020.107020>.
67. Thormann I, Reeves P, Thumm S, Reilley A, Engels JMM, Biradar CM, et al. Changes in barley (*Hordeum vulgare* L subsp *vulgare*) genetic diversity and structure in Jordan over a period of 31 years. *Plant Genet Resour.* 2018;16:112–26. <https://doi.org/10.1017/S1479262117000028>.
68. Kilian A, Wenzl P, Huttner E, Carling J, Xia L, Blois H, et al. Diversity arrays technology: a generic genome profiling technology on open platforms. *Methods Mol Biol.* 2012;888:67–89. https://doi.org/10.1007/978-1-61779-870-2_5.
69. Mayer KFX, Waugh R, Langridge P, Close TJ, Wise RP, Graner A, et al. A physical, genetic and functional sequence assembly of the barley genome. *Nature.* 2012;491:711–6. <https://doi.org/10.1038/nature11543>.
70. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2022. <https://www.R-project.org/>.

71. Granato IS, Galli G, de Oliveira Couto EG, e Souza MB, Mendonça LF, Fritsche-Neto R. *snpReady*: a tool to assist breeders in genomic analysis. *Mol Breeding*. 2018;38:102. <https://doi.org/10.1007/s11032-018-0844-8>.
72. Anderberg MR. *Cluster Analysis for Applications*. New York: Academic Press; 1972.
73. Lengyel A, Botta-Dukát Z. Silhouette width using a generalized mean-A flexible method for assessing clustering efficiency. *Ecol Evol*. 2019;9:13231–43. <https://doi.org/10.1002/ece3.5774>.
74. Kassambara A, Mundt F. *Factoextra: Extract and Visualize the Results of Multivariate Data Analyses*. R Package Version 1.0.7. 2020. <https://CRAN.Rproject.org/package=factoextra>. Accessed 15 May 2022.
75. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol*. 2005;14:2611–20.

Publisher's Note

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

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

