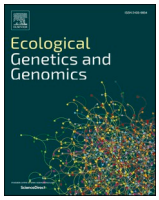
Ecological Genetics and Genomics 30 (2024) 100218

Contents lists available at ScienceDirect

Ecological Genetics and Genomics

journal homepage: www.elsevier.com/locate/egg



Exploring the extents of genetic diversity and population structure of enset [*Ensete ventricosum* (Welw.) cheesman] from southern Ethiopia using simple sequence repeat markers: Implications for crop improvement and conservation

Tesfaye Dilebo [a,b,*], Tileye Feyissa [a,c], Zemede Asfaw [d], Fekadu Gadissa [e]

[a] Department of Microbial, Cellular and Molecular Biology, Addis Ababa University, P. O. Box: 1176, Addis Ababa, Ethiopia
[b] Department of Natural Science, Hossana College of Education, P. O. Box: 94, Hossana, Ethiopia
[c] Institute of Biotechnology, Addis Ababa University, P. O. Box: 1176, Addis Ababa, Ethiopia
[d] Department of Plant Biology and Biodiversity Management, Addis Ababa University, P.O. Box 1176, Addis Ababa, Ethiopia
[e] Department of Biology, College of Natural Sciences, Madda Walabu University, P.O. Box 247, Bale-Robe, Ethiopia

ARTICLE INFO

Keywords:
Enset
Enset conservation
Gene flow
Genetic variation
SSR markers polymorphism

ABSTRACT

Enset (*Ensete ventricosum*) is a multi-use perennial herbaceous crop used as a staple food for over 20 million people in Ethiopia. Despite its high use values, very few studies have been conducted to improve this crop, particularly using molecular marker systems. In this context, the study aimed at evaluating the magnitude of genetic diversity and population structure of enset germplasm collections from four major enset growing zones in southern Ethiopia using 12 simple sequence repeat (SSR) markers. A total of 147 individual leaf samples were collected from the entire enset populations and gave 289 alleles, ranging from 12 to 41 alleles per locus, with a mean of 24.5. The polymorphism information content for each locus varied from 0.86 to 0.95, with a mean of 0.91. The number of effective alleles ranged from 5.13 to 11.79 with a mean of 8.27. The expected and observed heterozygosity showed average values of 0.85 and 0.84, respectively. The greatest genetic distance (1.16) was between Gurage and wild populations, while the shortest (0.37) was between Gurage and Silte. Among the six populations, the wild had the highest percentage of polymorphic loci (100%). AMOVA attributed 89% of the genetic variation to intra-population and only 11% to among populations. The whole set of germplasm indicates low genetic differentiation and high gene flow (Nm). The UPGMA and principal coordinates largely correspond to each other and indicate three major groups. Overall, the information gained from this study would be useful for enset improvements and conservation strategies.

1. Introduction

Ethiopia is recognized as both the center of origin and the center of diversity for many crops, including enset [1,2]. Enset [*Ensete ventricosum* (Welw.) Cheesman] is a large herbaceous monocarpic perennial plant belonging to the order Zingiberales, family Musaceae, and genus *Ensete* [3]. It is sometimes known as a false banana or Ethiopian banana because of its similar form to the banana (*Musa* x *paradisiaca*) [4]. The huge size, single pseudostem structure, dilated base, and upright leaves of enset, among other characteristics, set it apart from the banana [5]. Moreover, enset is a diploid plant with a haploid chromosome number of n = 9, whereas *Musa* species, including edible bananas, have different ploidy levels and chromosome numbers (diploid, triploid, or tetraploid), with n = 10 or n = 11 [4,6,7].

Enset domestication, cultivation, and human consumption as food are restricted to Ethiopia [2,4]. In most of the areas where enset is used as human food, it is usually a staple or a co-staple crop. It is a multi-purpose crop that is selected mainly for assurances of food security and stability for over 20 million of the Ethiopian population [8,9] and also produces fodder, fiber, and other materials. In addition, some enset landraces are used in human and livestock medicine [10–12]. Besides these products and services, enset plays important roles in the

* Corresponding author. Department of Microbial, Cellular and Molecular Biology, Addis Ababa University, P. O. Box: 1176, Addis Ababa, Ethiopia.
E-mail addresses: tesfayedilebo1@gmail.com (T. Dilebo), tileye.feyissa@aau.edu.et (T. Feyissa), zemede.asfaw@aau.edu.et (Z. Asfaw), fikega2000@gmail.com (F. Gadissa).

https://doi.org/10.1016/j.egg.2024.100218
Received 7 August 2023; Received in revised form 24 December 2023; Accepted 2 January 2024
Available online 5 January 2024

sociocultural and environmental relations of the central, southern, and southwestern administrative regions of Ethiopia [13,14]. Furthermore, it is comparatively drought-resilient, as it can withstand a shortage of rainfall for a certain period and is highly productive with minimum labor input [15,16].

Enset-producing farmers have maintained and enriched the diversity of the crop through identification and selection using their accumulated indigenous and local knowledge, and for them, the use of enset diversity is directly linked to the diverse methods of generating products that are used as food as well as non-food values [17,18]. Furthermore, they are the primary guardians of enset diversity, and previous authors have documented hundreds of types [12,13,15,17,19–21]. Enset shows noticeable variations in terms of morphological characteristics (e.g., the color of pseudostem, petiole, and midrib, size (width and length), and angle of leaf orientation), agronomic features (maturity rate, reaction to different diseases and pests, and agroecological adaptability), use values (quantity and quality of *kocho* (*qocho)* and *bulla*, use of corms, and fiber quality), and other attributes [13,15,17,22,23]. However, such features may change under the influence of ecological conditions [24], and thus, their use as markers is not fully effective in determining the existing diversity, and which is crucial in looking for the relevant markers.

Maintaining plant diversity and using it sustainably is vital to addressing concerns related to food security systems as well as other issues created by human population growth [25,26]. Genetic diversity is the degree of genetic variability found among individuals of a variety or a population within a species [27], and the existence of variation among or within groups is crucial to improving any plant species either by a classical or molecular approach [28]. Genetic diversity has partly resulted from genetic recombination, mutations, gene flow, and genetic drift [29,30]. In addition, migration, hybridization, and polyploidization are responsible for creating variation in plants [31]. These result in variations in DNA sequence, protein structure or isozymes, physiological properties, and morphological traits [32].

Sustainable utilization of the genetic diversity of the existing enset genotypes is expected to play an important role in selecting desirable genotypes with better yield and quality traits, along with resistance/tolerance to biotic and abiotic stresses through the breeding system to increase food security. So far, the extent of enset genetic diversity has been detected using varied molecular methods such as random amplified polymorphic DNA (RAPD) [33], amplified fragment length polymorphism (AFLP) [34], inter simple sequence repeat (ISSR) [35], and simple sequence repeats (SSRs) from banana primers [36]. In this regard, the enset genotypes examined by SSR markers in earlier studies were collected either from a limited geographical and agroecological range and/or were very limited in number. For example, Olango et al. [37] worked only on 66 samples maintained at the Areka agricultural research center that had been originally collected from Wolaita, Gamo-Gofa, Dawro, and Ari areas; Gerura et al. [38], used 83 samples from the Gurage Zone, and Nuraga et al. [39] used 38 medicinal and other 13 enset landraces from southern Ethiopia. A few other markers have also been conducted recently, such as those using AFLP and SNPs (single nucleotide polymorphisms) from Dawro, Gurage, Keffa, South Omo, Sheka, and Sidama [40] and SNPs from the Sidama region, Gurage, and South Omo Zones [41], which studied the genetic diversity among and within wild and cultivated enset landraces of Ethiopia.

Thus, it is essential to assess the genetic diversity of available enset populations that were not included in the previous studies and that grow in the diverse environments and farming practices of the Hadiya, Kembata-Tembaro, Silte, and Gurage zones of southern Ethiopia. Therefore, the objectives of this study were to evaluate the level of genetic diversity and population structure in germplasm collections from the above-listed zones and the Areka Agricultural Research Center by using the SSR molecular marker method, which would back-up the ethnobotanical and other studies by Dilebo et al. [12] to ultimately provide the foundation for improvement and sustained conservation of the enset crop.

2. Materials and methods

2.1. Plant material and sample collection

For this study, a total of 147 individuals (Table 1, Fig. 1) from newly emerged central cigar leaf samples (132 cultivated landraces, 6 released varieties, and 9 wild-growing types) were collected from the four administrative zones (Gurage (29), Silte (29), Hadiya (42) and Kembata-Tembaro (32) in the former Southern Nations, Nationalities, and Peoples Region (SNNPR) and the Areka Agricultural Research Center (AARC) enset germplasm conservation site, located in Wolaita Zone of southern Ethiopia. The cultivated landraces were known by farmers with different local names and were accessed from farmers' homegardens. The collected individual samples were grouped into six populations depending on the administrative zones of collection, each of which is largely inhabited by a unique ethnic group with its own cultural practices. In addition, the wild and released samples of enset varieties were included to represent their own populations (Appendix 1).

Enset-producing farmers recognize each landrace by its unique local name, which refers to different features such as morphological, agronomic, or quality of end products [12]. Therefore, each landrace was considered a separate sample, and the identification and description for sample collection were carefully applied with the help of senior knowledgeable enset farmers in each zone, and the collection was performed from farmers' homegardens with the informed consent/permission and support of each man or woman knowledgeable former. Moreover, landraces having the same local names but originating from different zones were collected and labeled as separate samples (Appendix 1) because each farming community considered them distinct farmers' varieties or landraces. The collected clean inner cigar-shaped enset leaf tissue, cut into smaller pieces were kept in zip-locked plastic bags containing silica gel (JHD, Guangdong Guanghua Sci-Tech Co, Ltd.) until the extraction of genomic DNA, and to improve proper desiccation, the silica gel (weight ratio of silica gel to enset leaf tissue: 10:1) was changed twice. Prior to the excision of the leaf samples, ethanol (75%) was used for cleaning and to avoid cross-contamination within the samples. The samples were brought to The Plant Biotechnology Research Laboratory, Addis Ababa University, Ethiopia, and kept there at room temperature until DNA extraction was carried out.

Table 1
List of enset populations used in this study and their altitudes, geographic positions and collection districts along with their sample sizes.

Population name	Collection districts	Altitude ranges (m) of collection sites	Latitude ranges (N)	Longitude ranges (E)	Sample size
Gurage	Endegegn, Gumer, Enemore Ener	1850–2790	7°76′- 8° 45′	37,°46' −38°71′	29
Silte	Mirab A, Misraq A, Alicho Weriro	2218–3195	7° 49′- 8^{0} 01	37,°96' −38°17′	29
Kembata	Doyogena, Angacha, Damboya	2145–2822	7^{0} 08′- 7^{0}30′	37,°22′- 38,^{0}04′	32
Hadiya	Misha, Lemmo, Dunna	2132–2915	7^{0} 07′- 7°92′	37,°29′- 38,°13′	42
Released	AARC	1774	7^{0} 04′	37,°42′	6
Wild	AARC	1774	7^{0} 04′	37,°42′	9
Total					**147**

Note: A = Azernet; AARC= Areka Agricultural Research Center; Original provenance of the wild populations: Ari District, South Omo Zone (A1 and A2), Dawro Zone (W1–W4), and Keffa Zone (W5–W7).

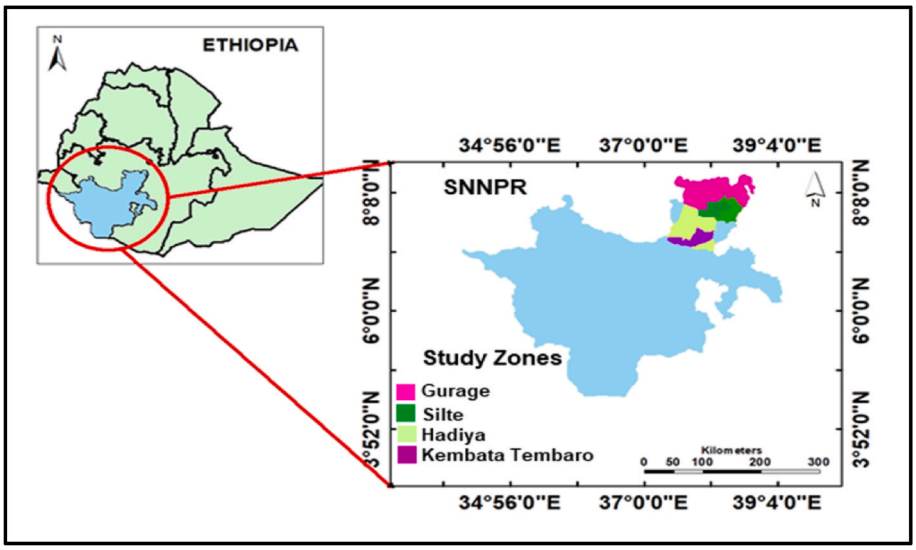


Fig. 1. Map of Ethiopia showing the location of the former Southern Nations, Nationalities and Peoples Region (SNNPR) and the four enset populations' collection zones.

2.2. DNA extraction

The silica-dried leaf samples of about 1 g were ground using a Mixer and Miller (RETSCH Mixer - Mill MM400) for 5 min. The total genomic DNA was extracted from each powdered leaf sample following the mini-prep DNA double extraction protocol of Cetyl Trimethyl Ammonium Bromide (CTAB) with little modification [42]. The quality of the extracted genomic DNA was checked using 1% agarose gel electrophoresis in 1 × TBE buffer, which was run for 1 h at 100 V. Similarly, the quantity (concentration) of the DNA was measured using a NanoDrop Spectrophotometer (ND-2000, USA). Finally, the DNA was adjusted to a concentration of 50 ng/μL using molecular-grade water and stored at −20 °C for further SSR analysis.

2.3. Primer screening and PCR amplifications

In the present study, a total of 15 SSR markers, 11 of which were developed by Olango et al. [37] and four from the Enset Molecular Data Base [43] (Table 2), were effectively screened, optimized and used to reveal the extents of genetic diversity in the populations considered.

The PCR amplifications were conducted in a 96-well thermo-cycler (Eppendorf AG Mastercycler 22,331, Hamburg, Germany) using a total

Table 2
List and Characteristics of the 15 SSR markers used in the analyses of Genetic Diversity and Population Structure of Enset (*E. ventricosum*) Populations in the Present Study.

Marker code	Primer sequence of forward (F) and reverse (R) (5′–3′)	Repeat motif	References	Size (bp)	Ta (^{0}C)
Evg-01	F: AGTCATTGTGCGCAGTTTCC R: CGGAGGACTCCATGTGGATGAG	(CTT)8	Olango et al. [37]	103–135	56
Evg-02	F: GGAGAAGCATTTGAAGGTTCTTG R: TTCGCATTTATCCCTGGCAC	(AG)12	Olango et al. [37]	108–158	56
Evg-03	F: ACAGCATAAGCGAAATAGCAG R: ACAGCATAAGCGAAATAGCAG	(AG)12	Olango et al. [37]	109–125	58
Evg-04	F: GCCATCGAGAGCTAAGGGG R: GGCAAGGCCGTAAGATCAAC	(AG)21	Olango et al. [37]	106–163	58
Evg-06	F: CCGAAGTGCAACACCAGAG R: TCGCTTTGCTCAACATCACC	(GAA)9	Olango et al. [37]	200–219	57
Evg-08	F: CCATCGACGCCTTAACAGAG R: TGAACCTCGGGAGTGACATAAG	(GA)21	Olango et al. [37]	152–196	58
Evg-09	F: GCCTTTCGTATGCTTGGTGG R: ACGTTGTTGCCGACATTCTG	(GA)13	Olango et al. [37]	139–179	56
Evg-10	F: CAGCCTGTGCAGCTAATCAC R: CAGCAGTTGCAGATCGTGTC	(AG)21	Olango et al. [37]	187–217	57
Evg-11	F: GGCCTAGTGACATGATGGTG R: TGATGCTAGATTCAAAGTCAAGG	(AC)13	Olango et al. [37]	130–168	54
Evg-12	F: TGCAACCCTTTGCTGCATTC R: AGCATCATTCGCCATGGTTG	(TG)14	Olango et al. [37]	117–161	58
Evg-17	F: GCGTCTGGTATGCTCAACTG R: TCGGGAATGATACAGAGGCG	(TCA)8	Olango et al. [37]	110–164	58
EnO-04	F: ATCTGCATGCACCCTAGCTT R: AAACCCTAACGTCCCTCCTC	(GT)10	Biswas et al. [43]	120–150	58
EnO-06	F: TGCCCAAAAACTTTGATGTG R; CCACACATCTCAGAGCCTCA	(AGA)11	Biswas et al. [43]	130–152	54
EnB-02	F: ATCAAGGTCATGTGCTGTGC R: ATCAAGGTCATGTGCTGTGC	(CT)11	Biswas et al. [43]	–	56
EnB-07	F: AGATCAACCGCATCCATCAT R: GCACGTGTGTCACATTGCAT	(GAA)15	Biswas et al. [43]	255–272	56

EnO-04 = EnOnjSSR049028, EnO-06 = EnOnjSSR061390, EnB-02 = EnBedSSR020585, EnB-07 = EnBedSSR071399.

reaction volume of 15 μl containing 1.5 μl of genomic DNA (50 ng/μl), 0.25 μl of each primer (10 pmol/μl) (Sigma-Aldrich), 3 μl of 5xFIRPol@ Master Mix ready (5× reaction buffer B: 0.4 M Tris-Hcl, 0.1 M (NH4)2, and 0.1% w/v Tween-20; 7.5 mM MgCl2; 1 mM dNTPs of each, blue and yellow dyes, Solis BioDyne, Tartu, Estonia), and 10.0 μl of molecular grade water. The amplification condition consisted of an initial denaturing step of 94 °C for 5 min, followed by 35 cycles of 94 °C for 30 s, optimum annealing temperature (specific to each primer) for 1 min, primer elongation at 72 °C for 1 min, and a final elongation step of 72 °C for 15 min. The amplified products were separated by electrophoresis in a 3.5% (w/v) agarose gel (Sigma) in 1xTBE buffer containing ethidium bromide ($EtBr_2$) at 100V for 3 h. The generated fragments were visualized using a gel documentation system (BIORAD Gel DocTM EZ System Imager) and scored against a 100 base-pair DNA ladder (Thermo Fisher Scientific, Massachusetts, USA).

2.4. Data analysis

Genetic diversity indices such as the total number of alleles (Na), the number of effective alleles (Ne), observed heterozygosity (Ho), expected heterozygosity (He), unbiased expected heterozygosity (uHe), Shannon's Information Index (I), percentage of polymorphic loci (PPL), and Nei's gene diversity (GD) over the entire loci were computed using GenAlEx version 6.501 [44]. Other locus-based genetic diversity parameters, including major allele frequency (MAF) and polymorphic information content (PIC) were estimated using PowerMarker version 3.25 [45].

To estimate the correlation between observed allelic diversity and sample size of populations, rarified allelic richness (Ar) and private rarified allelic richness (Ap) were computed using the rarefaction procedure applied in the HP-Rare 1.1 software [46]. An analysis of molecular variation (AMOVA) was computed to analyze the distribution pattern of genetic variances between groups and individuals within a group using Arlequin version 3.5.2.2 [47]. Wright's fixation index (F_{ST}) of the total populations and pairwise F_{ST} among all pairings of populations were computed using GenAlEx version 6.501 [44] to estimate population differentiation and the significance was evaluated using 1000 bootstraps. Gene flow (Nm), which measures the average number of individuals in each generation migrating among populations, was computed using the formula, $Nm = 0.25\,(1-Fst)/Fst$ [48] Nei's standard genetic distance (DST, corrected) [49] based on the Unweighted Pair Group Method with Arithmetic Mean (UPGMA) [50] tree was constructed using DARwin var. 6.0.14 (Perrier and Jacquemoud-Collet) and POPTREE2 [51], respectively, and significance was tested based on 1000 bootstraps [52]. TreeView (Win32) 1.6.6 program [53] and FigTree var. 1.4.3 [54] were used to display the produced trees.

To assess the pattern of population structure of the studied enset populations and to detect the extents of admixture, a Bayesian model-based clustering algorithm was inferred using STRUCTURE version 2.3.4 [55,56]. A burn-in period of 50,000 was employed in each run, and data were gathered over 500,000 Markov Chain Monte Carlo (MCMC) replications for $K = 1$ to $K = 6$ using 20 iterations for each K. This was done to estimate the most likely number of populations (K). The web-based STRUCTURE HARVESTER version. 0.6.92 [57], was used to estimate the ideal K value using the simulation method developed by Evanno et al. [58]. The bar plot for the most probable K was calculated by employing the Clumpak beta version [59].

3. Results

To develop any conservation strategy for sustainable utilization and crop improvement programs, knowledge of the genetic diversity and population structure of any species is indispensable. In the present study, a total of 15 SSR markers have been used to estimate the extent of genetic diversity and population structure of enset populations from different administrative zones of the former SNNPR of Ethiopia. Out of which, twelve SSR markers (Table 2) exhibited reproducible polymorphic amplification during the screening process and were employed to analyze the 147 individual genotypes of the enset populations. The remaining two markers (Evg-03 and EnB-07) were observed to be monomorphic and hence, omitted from further analysis. The SSR fragment size ranged from 103 bp (Evg-01) to 217 bp (Evg-10) (Table 2).

3.1. Genetic diversity

In the 147 enset genotypes considered, the selected 12 polymorphic SSR loci yielded a total of 289 alleles, with a mean of 24.5 alleles per locus. The number of alleles identified per locus ranged from 12 (in EnO-06) to 41 (in Evg-08) (Table 3). The major allele frequency (MAF) ranged from 0.09 (in Evg-12) to 0.25 (in Evg-10), with an average value of 0.154. The polymorphic information content (PIC) values ranged from 0.86 (in EnO-06) to 0.95 (in Evg-09), with an average of 0.91. Relatively, the highest effective number of alleles (Ne) (11.79), Nei's gene diversity (GD) (0.95), Shannon information index (I) (2.57), expected heterozygosity (He) (0.91), and unbiased expected heterozygosity (uHe) (0.94) were recorded for Evg-09, while the least parameters such as Ne, I, He, and uHe were observed in EnO-06. Similarly, Evg-09 revealed the highest value of gene flow ($Nm = 7.02$) and the lowest inbreeding coefficient based on genetic differentiation ($Fst = 0.03$) (Table 3).

3.2. Extents of genetic diversity in the populations considered

Estimates of genetic diversity in the six enset populations considered are presented in Table 4. Accordingly, the Hadiya population scored the highest values with regard to the number of different alleles (Na), number of effective alleles (Ne), Shannon diversity index (I), and expected heterozygosity (He). However, the wild population scored the least in all these diversity indices. Similarly, the highest estimate of allelic richness (Ar) (7.80) was also recorded in the Hadiya population, followed by the Kambata-Tembaro (7.73) population. On the other hand, the highest number of private allelic richness (Arp) (1.29) was observed in the wild population, whereas, released population contained a lower number of private allelic richness (0.74). Relatively higher observed heterozygosity (Ho) was recorded in both the wild and released ($Ho = 0.88$) populations, while it was least in Gurage and Hadiya ($I = 0.81$) populations. The percentage polymorphic loci (PPL) ranged from 89% (in the Hadiya population) to 100% (in the wild population), with a mean value of 94.83%.

3.3. Population genetic differentiation and gene flow

Analysis of molecular variance (AMOVA) revealed that 89% of the total variation is accounted for the individuals within a group (population); whereas variation among the six populations and individual genotypes (samples) in each population contributed to only 5% and 6% of the total variation, respectively (Table 5). Thus, the heterozygosity of the individuals within each population accounts for most of the within-population variation. The overall genetic differentiation among the six enset populations was limited and significantly low ($F_{ST} = 0.05$; $p < 0.001$). On the other hand, the average gene flow among the analyzed enset populations was high ($Nm = 5.14$).

3.4. Extents of genetic distance between the populations

The extent of pairwise genetic distances between the six enset populations is presented is Table 6. The highest genetic distance (1.16) was observed between populations from Gurage and wild, followed by populations from Silte and wild (1.02). Conversely, lower pairwise genetic distance was identified between populations from Gurage and Silte (0.37) and between Hadiya and Kembata (0.47).

Table 3
Informativeness and levels of the different diversity indices of the SSR loci across the enset populations considered.

Marker	MAF	NA	GD	PIC	Ne	I	Ho	He	uHe	Fst	Nm	PHWE
Evg-01	0.12	22.00	0.92	0.92	6.45	2.04	0.73	0.82	0.84	0.10	2.22	0.00**
Evg-02	0.12	26.00	0.93	0.93	9.39	2.37	0.88	0.87	0.90	0.08	2.91	0.00**
Evg-04	0.15	32.00	0.93	0.92	9.03	2.32	0.85	0.87	0.90	0.06	3.71	0.00**
Evg-06	0.19	18.00	0.87	0.86	6.03	1.95	0.77	0.82	0.85	0.06	4.10	0.00**
Evg-08	0.13	41.00	0.94	0.94	10.53	2.44	0.88	0.90	0.93	0.04	5.54	0.03*
Evg-09	0.10	31.00	0.95	0.95	11.79	2.57	0.86	0.91	0.94	0.03	7.02	0.03*
Evg-10	0.25	18.00	0.90	0.89	7.43	2.15	0.80	0.85	0.88	0.05	4.87	0.00**
Evg-11	0.11	21.00	0.93	0.93	9.05	2.23	0.92	0.85	0.88	0.07	3.19	0.21^{ns}
Evg-12	0.09	26.00	0.94	0.94	9.67	2.40	0.88	0.88	0.91	0.06	3.67	0.04*
Evg-17	0.18	27.00	0.92	0.92	8.25	2.17	0.85	0.85	0.87	0.09	2.41	0.04*
EnO-04	0.20	15.00	0.90	0.89	6.54	1.94	0.89	0.81	0.84	0.10	2.29	0.02*
EnO-06	0.21	12.00	0.87	0.86	5.13	1.78	0.79	0.78	0.80	0.10	2.30	0.00**

MAF = Major allele frequency; NA=Number of alleles; GD = Gene diversity; Ne = Effective number of alleles; PIC=Polymorphic information content; I=Shannon's Information Index; Ho=Observed heterozygosity; He = Expected heterozygosity; uHe = unbiased heterozygosity; Fst = inbreeding coefficient within subpopulations relative to total (genetic differentiation among subpopulations); Nm = gene flow estimated from Fst 0.25(1- Fst)/Fst; PHWE P-value for deviation from Hardy Weinberg equilibrium, ns not significant, * = $P < 0.05$, ** = $P < 0.0001$ and hence highly significant.

Table 4
Summary of the different diversity indices scored for each population over the 12 loci considered.

Population	Na	Ne	Ar	Arp	I	Ho	He	uHe	PPL%	F
Gurage	14.25	8.56	7.09	1.14	2.24	0.81	0.84	0.86	95.00	0.014
Silte	13.25	8.42	7.24	0.89	2.28	0.84	0.87	0.89	99.00	0.032
Kembata	16.00	9.76	7.73	0.91	2.44	0.83	0.88	0.90	96.00	0.060
Hadiya	17.25	10.64	7.80	1.09	2.50	0.81	0.89	0.90	89.00	0.085
Wild	6.83	5.44	6.83	1.29	1.72	0.88	0.78	0.85	100.00	−0.116
Released	8.50	6.83	7.06	0.74	1.98	0.88	0.84	0.89	90.00	−0.048
Mean	**12.68**	**8.27**	**7.29**	**1.01**	**2.20**	**0.84**	**0.85**	**0.88**	**94.83**	**0.005**

NA=Number of alleles; Ne = Effective number of alleles; Ar=Allelic richness; Arp = Private allelic richness; I=Shannon's Information Index; H_O=Observed heterozygosity; H_e = Expected heterozygosity; uHe = unbiased heterozygosity; PPL = percentage polymorphic loci; F=Fixation index.

Table 5
AMOVA among and within enset populations based on the 12 SSR markers.

Source	Df	SS	MS	Est. Var.	%	Nm	F_{st}
Among Populations	5	88.36	17.67	0.26	5.00%		
Among Individuals	141	799.69	5.67	0.35	6.00%		
Within Individuals	147	731.00	4.97	4.97	89.00%		
Total	293	1619.05		5.58	100.00%	5.14	0.05

Df = Degrees of freedom, SS=Sum of squares, MS = mean squares, Est.Var = Estimated variation, Nm = gene flow.

Table 6
Pairwise population matrix of Nei's genetic distance scored over the 12 SSR markers.

Populations	Gurage	Silte	Kembata	Hadiya	Wild	Released
Gurage	[a]					
Silte	0.37	[a]				
Kembata	0.79	0.59	[a]			
Hadiya	0.74	0.59	0.47	[a]		
Wild	1.16	1.02	1.01	0.71	[a]	
Released	0.83	0.63	0.73	0.47	0.74	[a]

[a] = Not applicable.

3.5. Cluster analysis

The Unweighted Pair-group Method with Arithmetic Mean (UPGMA) based cluster analysis using the 12 polymorphic SSR markers grouped the 147 enset genotypes into three major clusters (I, II, and III) (Fig. 2). Cluster I had the largest number of genotypes, which consisted of 81 (55.1%) of the total enset genotypes, that formed two separate sub-clusters (IA and IB). Among these sub-clusters, IA contained 16 landraces, which represent exclusively Kembata populations. On the other hand, sub-cluster IB consisted of 65 (44.2%) enset genotypes, of which 35 (53.85%) were from Hadiya, 13 (20.0%) from Kembata, and seven (10.77%) each from the Released and Wild populations. Cluster II comprised 60 genotypes, which were further divided into two distinct sub-clusters (IIA and IIB). Sub-cluster IIB contained 27 (75%) and 9 (25%) enset genotypes representing Gurage and Silte populations, respectively. In the same pattern, sub-cluster IIA counted 19 (79.17%) from the Silte population. Cluster III encompassed only six enset landraces, four from Hadiya and the rest two from Gurage populations. The grouping pattern revealed some degree of parallelism with the geographic or administrative region of collections though intermixes from different geographic or administrative zones of collections have been evidenced (Fig. 2).

3.6. Principal coordinate analysis

The principal coordinate analysis (PCoA) created using the first two components showed a similar pattern to UPGMA and followed a weak pattern of grouping following their geographic location of collections. Collections from different but adjacent collection zones were lumped together with their distinctive small group. The majority of the enset samples collected from Gurage zone are grouped together on the upper left, and those from Silte administrative zone get clustered at the center. The predominant enset population obtained from Kembata is clustered around the lower left and right sides of the axis (Fig. 3). Enset populations collected from the Hadiya administrative zone are distributed on the upper right and lower right of the axis. The wild and released samples are more or less distributed on the upper right and upper left sides of Axis 1 (Fig. 3).

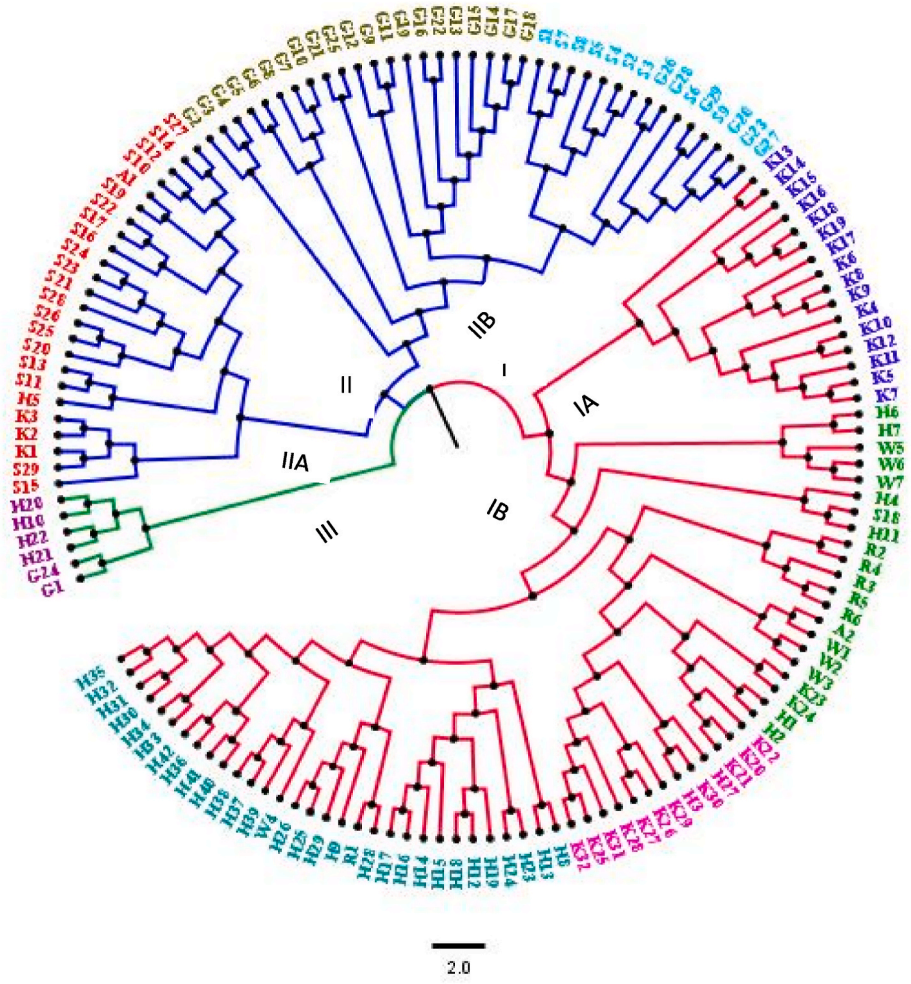


Fig. 2. A dendrogram based on the unweighted pair-group method with the arithmetic mean (UPGMA) of the 147 *E. ventricosum* individuals representing the six populations (Gurage, Silte, Kembata, Hadiya, Released, and Wild varieties) grouped in three major clusters (I, II, and III) and sub-clusters (IA, IB, IIA, and IIB).

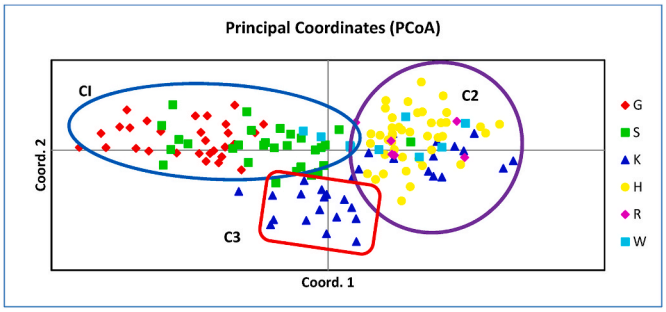


Fig. 3. Two-dimensional (2D) PCoA representation of the genetic relationships among the 147 individuals of *E. ventricosum* representing six populations (G, Gurage = red; S, Silte = deep green; K, Kembata = blue; H, Hadiya = yellow; R, released = pink; W, wild = light green) from Ethiopia.

3.7. Structure analysis

To further comprehend the genetic structure of the entire 147 enset samples (genotypes), a Bayesian-based population structure following the Evanno et al. [58] analysis was determined using STRUCTURE program outputs. Accordingly, the predicted highest K value was two (K = 2) (Fig. 4). Based on this value, the Clumpak result (bar plot) revealed a limited admixtures, and as a result, the populations were largely structured together according to the closeness of their collection sites or geographic origin.

4. Discussion

To further enhance enset crop conservation and improvement, the use of molecular markers, such as SSR markers, is a widely recognized approach. As enset is a diploid species [4,6,7], an aggregate of two alleles per individual sample is anticipated at the single-copy microsatellite locus [60]. The detection of two hundred and eighty-nine alleles (an average of 24.5 alleles per locus) using the twelve SSR markers, the present study is indicative of the presence of high genetic diversity in the populations considered. The number of alleles we detected is high compared to several of the reports so far. For example; Getachew et al. [36] reported 61 alleles from 220 enset genotypes by applying 11 banana SSR markers; Olango et al. [37] reported 202 alleles from 66 enset genotypes employing 34 enset SSR loci, Gerura et al. [38] reported 77 alleles from 83 enset genotypes using 12 enset SSR microsatellites, and Nuraga et al. [39] detected 38 alleles from 52 genotypes using 15 SSR loci. Such a large number of alleles detected in the present study suggest that the microsatellite markers we selected are appropriate and provide an excellent chance to assess enset genetic resources for conservation and breeding.

4.1. Estimates of genetic diversity in the populations

Genetic diversity offers opportunities to understand the genetic

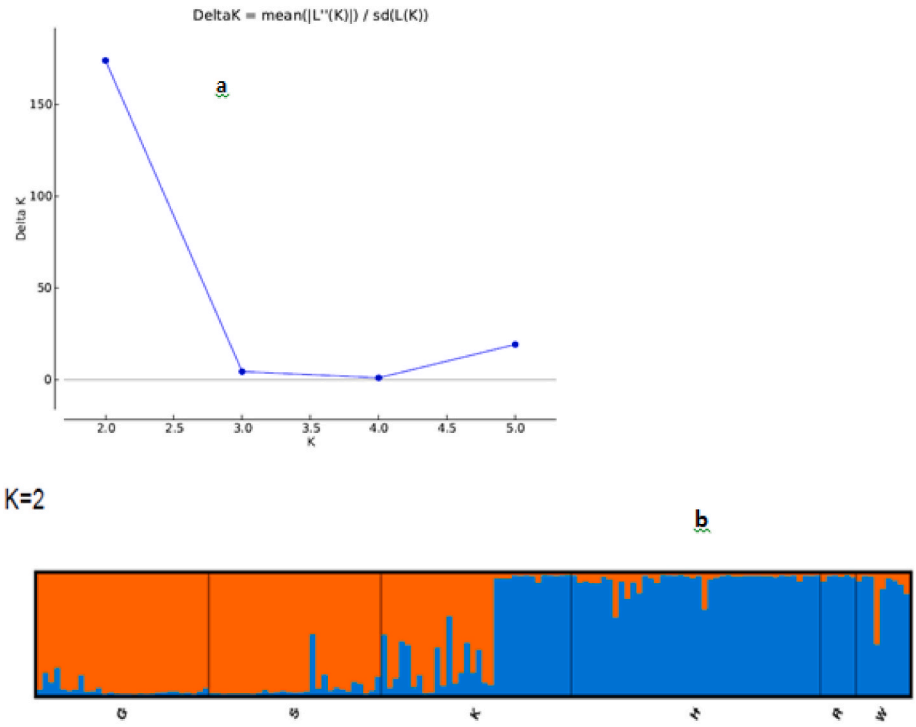


Fig. 4. Results of the population structure of the 147 *E. ventricosum* individuals in six pre-determined populations using Evano et al. [58] method and Bayesian model-based estimation shows Delta K plot for maximum value at K = 2 (a), CLUMPAK plot for K = 2 using the CLUMPP software, *G* Gurage, *S* Silte, *K* Kembata, *H* Hadiya, *R* Released, and *W* Wild (b).

structure and level of diversity within and among genotypes, populations, and species [26]. In comparison to smaller, recently evolved populations, larger and older populations are predicted to have higher levels of accumulated and maintained genetic variation [61], which is crucial for enhancing fitness and consequently decreases the possibility of local extinction [62]. In the present study, the polymorphic information content (PIC) of the 12 SSR markers applied showed larger values suggesting the high discriminatory power of the SSR markers employed following recommendations by Prabakaran et al. [63] and Tiago et al. [64] that suggested a PIC value of greater than 0.5 as very informative and helpful for determining the degree of polymorphism at a given locus. Similarly, Dutta et al. [65] had also described a high PIC value as one of the most critical indications to effectively assess genetic diversity in a given species. Previous studies had shown slightly inconsistent PIC values using different numbers of SSR markers on enset. For instance, Olango et al. [37] observed a PIC value of 0.41–0.77 by evaluating the genetic diversity of 66 enset landrace collections; Gerura et al. [38] described a PIC value of 0.62–0.77 in 83 enset landraces from the Gurage Zone using 12 SSR markers; and Nuraga et al. [39] obtained by far a lower PIC value ranging from 0.26 to 0.72 in 52 cultivated enset landraces using 15 SSR markers. The use of various types and numbers of genotypes and primers in each of those studies could be the reason for the disparity. It has further been explained by Biswas et al. [66] who suggest that the number of genotypes and the genetic background of the genotypes have a significant impact on PIC values.

Of the total enset populations considered, the wild varieties retained the highest private allelic richness, which suggests it is a source of novel alleles for breeding and conservation. Such high private alleles could be attributed to a relatively high rate of mutation, particularly at SSR loci [67]. The result corroborates with the reports by Gerura et al. [38].

The estimate of Shannon's information index (I) in this study also revealed a high extent of genetic variation within the examined enset populations, which is an indication of the population diversity in a specific environment, according to Dido et al. [68]. Overall, the values we detected were higher as compared to the previous findings by Olango et al. [37] (1.16), (Gerura et al. [38] (1.17), and Nuraga et al. [39] (0.74).

The expected (He) and observed (Ho) heterozygosities in the present study again revealed high extents of genetic diversity in the enset populations considered, with the observed and expected heterozygotes closely matching the same frequency (He = 0.85; Ho = 0.84). Likewise, the detected mean expected heterozygosity (He = 0.85) is higher than the values reported by Olango et al. [37] (0.59), Gerura et al. [38] (0.59), and Nuraga et al. [39] (0.47).

The mean expected heterozygosity (He) within a population is the most reliable way to quantify genetic variation [69]. Similarly, Jump et al. [70] stated that heterozygosity is one of the key parameters to study genetic variation in natural populations in addition to revealing the structure and even history of a population. In this regard, the present study indicated that the loci considered exhibited the existence of high heterozygosity that caused a considerable deviation from Hardy-Weinberg Equilibrium (HWE). Gadissa et al. [71] described high heterozygosity as being expected in historically outcrossing plant species that preserve their heterozygosity through vegetative propagation. This was expected considering that the majority of vegetatively propagated crops are highly heterozygous [36,40,41]. Entirely all cultivated enset crops in general and the samples we collected from the farmers' homegardens are vegetatively propagating from young mother plants, but the wild relatives are originally regenerating through sexual reproduction and maintained by vegetative propagation for a number of years, *ex-situ* conservation sites for the enset crop. The same situations have been noted in *Plectranthus edulis* by Gadissa et al. [71], cassava (*Manihot esculenta*) by Tiago et al. [64], and banana by Biswas et al. [66].

The enset populations we considered revealed a low fixation index (F = 0.005), suggesting accumulation and maintenance of the existing heterozygosity either through somatic mutation over time or through other mechanisms [41,72]. On the other hand, the result suggests a minimized role of inbreeding that could facilitate homozygosity and hence a larger coefficient [64,73].

The present study showed a relatively high extent of genetic differences within populations in terms of gene diversity, Shannon's information index, and heterozygosity. This may be attributed to the very diverse landrace resources cultivated and maintained for a long period by local farmers for their livelihood and other different socio-cultural purposes. Consequently, the regions could be viewed as genetic diversity hotspots and one of the potential in-situ conservation sites for the enset crop. This result also supports the existence of high on-farm or farmer diversity in the studied regions that have been sustainably conserved for generations [12,74,75].

In general, the computed genetic diversity of enset in the present study is greater than that determined by earlier studies using different DNA markers such as RAPD [33], AFLP [34], and ISSR [35] which is more likely as SSRs are more variable markers than RAPD, AFLP, and ISSR [76]. Such a large allelic variability, as other SSR marker studies in enset may be attributed to the nature of SSR markers since they have a high mutation rate both forward and backward. However, drawing general conclusions from a direct comparison of these studies is challenging due to the differences in the number as well as types of populations and DNA markers employed. In the same manner, Hamza et al. [77] described that comparisons of comprehensive diversity estimates from marker systems with various properties and origins of difference do not permit relevant conclusions.

4.2. Implications of population genetic differentiation and gene flow

The computed analysis of molecular variance (AMOVA) demonstrated a higher within-population variance than among the populations, suggesting less impact of the region or zone of origin. The result is in line with previous reports by Olango et al. [37], Gerura et al. [38], and Nuraga et al. [39] using SSR marker systems. In the same manner, Negash et al. [34] and Tesfamicael et al. [40], Birmeta et al. [33], and Tobiaw and Bekele [35] have reported that the among-population variance is lower than the within-population variance for enset genotypes using AFLP, RAPD, and ISSR markers, respectively. Overall, the result indicates the importance of within-population diversity to initiate conservation and utilization of the current enset diversity in the country at large.

Similarly, a low F_{ST} value (0.046) observed among the analyzed genotypes once again shows low genetic differentiation among the studied enset populations. Nkhata et al. [24] indicated that genetic divergence among populations can be low if the value of F_{ST} is less than 0.05. This is most likely related to the high rate of gene flow between enset farming communities in different zones due to the close interaction between farming households and the deep-rooted tradition of sharing and exchanging enset germplasm by raising suckers of desired/preferred genotypes. Local cultures and relatedness of families in the enset-growing areas of Ethiopia as revealed in their languages and practices, including collaborations in enset farming and harvesting activities of men and women. In general, such commonality in farming is attributed to the sociological characteristics and psychological make-up of the enset culture people of Ethiopia. A similar trend and observation have been documented by other researchers [12,13,15]. Such relatively high gene flow (mean $Nm = 5.14$) in the present study, contributes to the low level of population differentiation. In the same line, Slatkin and Barton [48] had stated that if the average number of migrants (gene flow) per generation (Nm) is less than 1, then no random differentiation across populations may be anticipated. Wright [78] had also observed an inverse correlation between gene flow (Nm) and population divergence (F_{ST}). Extensive vegetative propagation practices of enset by farmers may further restrict the extent of gene flow and consequently limit its population differentiation.

4.3. Patterns of genetic distance between the populations

In terms of pair-wise population genetic distances, the present study showed a relatively higher Nei's standard genetic distance (1.16) between Gurage and wild populations. On the other hand, the lowest genetic distance (0.37) was recorded between Silte and Gurage, followed by Kembata and Hadiya populations. The result suggests that the extent is parallel to the closeness of the collection administrative zones and vice versa, which could be attributed to the high sharing of enset genotypes as a planting source among the adjacent administrative or geographic regions. A similar pattern was reported in populations from Gamo Gofa *vs.* Wolaita and Wild *vs.* Ari [37] and the enset populations from the Gurage zone of Ethiopia [38].

4.4. Patterns of grouping and population structure

Understanding the genetic relationships and population structures between the diverse enset populations collected from different locations could provide a substantial advantage for improvement programs. In addition, it offers valuable knowledge regarding the extent of gene flow or evolutionary linkages in crop species at large. Accordingly, the UPGMA based cluster analysis confirmed the absence of pronounced intermixes of the enset samples from different collection sites or sources. Similar results have been reported by Olango et al. [37] and Gerura et al. [38]. Similarly, the principal coordinate analysis (PCoA), a multivariate dataset that allows us to explore the spatial distribution of the genetic distances between populations using the two-dimensional plan, according to Özkan et al. [79], also revealed the same patterns of grouping, suggesting low genetic divergence among the populations. As reported by Tesfamicael et al. [40], PCoA employing AFLP markers revealed that wild and cultivated enset landraces formed clusters with a significant number of overlapping individuals from the two groups.

The population structure analysis also showed that the enset collections from the four collection zones and AARC were closely related to one another and overall, two inferred groups ($K = 2$) with little admixtures have been detected. It is remarkable to note that each enset genotype analyzed has alleles that come from the two groupings, the cultivated landraces, and wild collections. This signifies the existence of a higher level of admixture among the different populations, which confirms the high gene flow and eventually minimal population differentiation. This could result from the exchange of genetic materials between the local and regional enset farming communities. The result is in agreement with the previous reports by Birmeta et al. [33], Olango et al. [37], Gerura et al. [38], Nuraga et al. [39], Tesfamicael et al. [40], and Haile et al. [41].

4.5. SSR markers for conservation and utilization of enset populations

One of the main uses of the SSR marker is offering a room for evaluation of genetic diversity within or among natural populations, which is essential for effective conservation and utilization of genetic resources [66]. Moreover, it is an essential tool in maintaining indigenous plant genetic resources for use in food security, a very important issue in reducing vulnerability since they have been adapted for centuries to the local climate, soil, landscape, and cultural heritage of each region [80] and thus, ensures the future needs of subsistence farmers, breeders, researchers, and the entire community [81]. The SSR marker system also generates baseline information on local crop diversity that could benefit farmers in terms of local adaptation to ecological diversity, pests, pathogens, risk management, rituals, and food culture [82]. In this regard, the present study has established baseline information from the points of view of genetic diversity estimate in the enset population, one of the most economically important but neglected indigenous food crops, to help facilitate its conservation and utilization. In addition to the SSR markers, other more reliable and evenly distributed markers may be important, even if the allelic richness may be lower for inferring long-term evolutionary relationships or ancestry.

According to report, a significant number of the enset populations and their genetic diversity have decreased in different parts of the

country over the past 50 years [83]. As a result, maintenance of the remaining diversity is being undertaken by combining both *in situ* (on-farm) and *ex-situ* (*in vitro*, or in-field gene banks) conservation initiatives. Negash [17] and Tsegaye and Struik [15] suggested that on-farm conservation is an unquestionably important component to be supported along with the formal (*ex-situ*) conservation initiatives as it enables continued maintenance and evolution of landraces using conventional cultivating practices. To supposrt community initiatives to conserve agro-biodiversity in many indigenous communities, strengthening local institutions and farmer leadership has been shown to be effective. Olango et al. [23] suggested that enset bio-cultural resource continuity must be based on participatory community approaches and the mobilization of both the young and the elder groups.

Nowadays, the Areka Agricultural Research Center (AARC) has maintained over 600 enset germplasm collections from 12 enset cultivating areas of Ethiopia in field gene banks. Olango et al. [23] reported that only 40% of the landraces that were known to the Wolaita agricultural community were represented by the AARC collection, indicating that the actual diversity is still not fully sampled. Similarly, Yemataw [84] asserted that the Areka collections have not undergone a thorough assessment. This indicated that there is a need for systematic collection from the major enset growing regions and other locations of interest, including wild populations, to apply appropriate conservation strategies (both *in situ* (on-farm) and *ex-situ* (*in vitro*, or in-field gene banks), and to undertake full characterization of enset germplasm for its role in prospective breeding initiatives, and as a source of better-performing landraces. In this regard, the present study in general and the marker systems used in particular are essential and could offer tips to initiate and facilitate maintenance and enhancement of the crop.

5. Conclusions

Given the significance of the enset crop in the central, southern, and southwestern regions of the Ethiopian agriculture and food production system, the present study provides additional information on the extent of genetic diversity and population structure of enset genetic resources to help implement suitable conservation plans, crop improvement and breeding programs. The study revealed high genetic diversity among the enset samples with very low fixation indices among the populations. The computed AMOVA also showed the occurrence of higher diversity among individuals within populations than among populations. This can be explained by a significant gene flow and maintained low genetic divergence among the populations. The patterns of genetic variability in cultivated enset landraces showed a weak link with cultivation regions. The findings of this study have demonstrated that the SSR marker system can be applied effectively to identify genetic diversity among enset genotypes and it also suggests including more germplasm collections from the remaining enset growing areas for conservation and to develop the best-performing enset varieties. Furthermore, this study inspires researchers to create genomic tools for this crucial crop, to enable faster delivery of improvements to Ethiopia's food and nutrition security through the use of modern breeding techniques like marker-assisted selection (MAS) and genomic selection (GS).

Funding statement

This research was financially supported by Addis Ababa University, Ethiopia as part of a PhD dissertation on consumables and during data collection.

CRediT authorship contribution statement

Tesfaye Dilebo: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Visualization, Writing – original draft, Writing – review & editing. **Tileye Feyissa:** Conceptualization, Investigation, Methodology, Resources, Supervision, Validation, Visualization, Writing – review & editing. **Zemede Asfaw:** Supervision, Validation, Writing – review & editing. **Fekadu Gadissa:** Software, Validation, Visualization, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

Data availability

Data will be made available on request.

Acknowledgements

The authors would like to thank the Department of Microbial, Cellular, and Molecular Biology, Addis Ababa University (AAU), Ethiopia for financial and technical supports. The authors also acknowledge the Institute of Biotechnology and Plant Genetic Research Laboratory of Addis Ababa University for providing us laboratory facilities. We are also grateful to enset-producing farmers and Areka Agricultural Research Center for their kind support during sample collection.

Appendix: 1

Indicates the local names of enset landraces or varieties with their code.

Agade	G1	*Agade*	S1	*Aganche*	K1	*Alabite*	H1	*Endale*	R1
Ahiro	G2	*Ahiro*	S2	*Ashura*	K2	*Anchire*	H2	Gewada	R2
Astara	G3	*Ankfuye*	S3	*Banko*	K3	*Awunada*	H3	*Kelisa*	R3
Awunad	G4	*Ashaqite*	S4	*Badede*	K4	*Bamba*	H4	*Messena*	R4
Aywogna	G5	*Astara*	S5	*Cherquwa*	K5	*Beneje*	H5	*Yambule*	R5
Bededet red	G6	*Beneje*	S6	*Degoblack*	K6	*Bequcho*	H6	*Zereta*	R6
Bezeria	G7	*Darye*	S7	*Derqeta*	K7	*Boicho*	H7	Sucker 1	A1
Bitena	G8	*Dem-worad*	S8	*Dirbo*	K8	*BuchTor*	H8	Sucker 2	A2
Dem-wored	G9	*Fengo*	S9	*Etine*	K9	*Dirbo*	H9	Dawro 1	W1
Feraziya	G10	*Gariye*	S10	*Fecheche*	K10	*Fello*	H10	Dawro 2	W2
Gegered	G11	*Hanzana*	S11	*Gimbo*	K11	*Gimbo*	H11	Dawro 3	W3
Gimbuwa	G12	*Hiniba*	S12	*Gishira*	K12	*Ginjowona*	H12	Dawro 4	W4
Ginnad	G13	*Kembat*	S13	*Godorete*	K13	*Gogoricho*	H13	Kaffa-old-Red	W5
Gozoda	G14	*Megribe*	S14	*Gomorsa*	K14	*Gomorsa*	H14	Kaffa-old-white	W6

(*continued on next page*)

(*continued*)

Agade	G1	*Agade*	S1	*Aganche*	K1	*Alabite*	H1	*Endale*	R1
Gumbura	G15	*Merza Kembt*	S15	*Hella*	K15	*Gozoda*	H15	Kaffa-new-Red	W7
Hanzana	G16	*Orad*	S16	*Leqeqe*	K16	*Gudere*	H16		
Hiniwa	G17	*Qeshqeshe*	S17	*Merza-black*	K17	*Hella*	H17		
Kemele	G18	*Qiniware*	S18	*Mesmesa*	K18	*Kaseta mo*	H18		
Lemat	G19	*Separa*	S19	*Qeqille*	K19	*Kekera-red*	H19		
Nechewa	G20	*Sherafire*	S20	*Qorate*	K20	*Korina*	H20		
Oniya	G21	*Shigez*	S21	*Quina*	K21	*Laddare*	H21		
Qesewa	G22	*Shirtye*	S22	*Sebera*	K22	*Mariye*	H22		
Separa	G23	*Sino*	S23	*Shelleqe*	K23	*Meqelwesa*	H23		
Shewatia	G24	*Sisqella*	S24	*Sinera*	K24	*Moche*	H24		
Shewora	G25	*Tegeded*	S25	*Sorpie*	K25	*Oniya*	H25		
Shirafere	G26	*Wonade*	S26	*Torore*	K26	*Qeteqeta*	H26		
Wonadia	G27	*Woshemaja*	S27	*Unjame*	K27	*Qombotira*	H27		
Zobir-red	G28	*Zegizig*	S28	*Wachiso*	K28	*Quiena*	H28		
Zogirad	G29	*Zobir*	S29	*Wohe*	K29	*Separa*	H29		
				Xebere	K30	*Shate*	H30		
				Xessa	K31	*Sinera*	H31		
				Zobiro	K32	*Sisqella mo*	H32		
						Sisasura	H33		
						Soqido-Le	H34		
						Suwandiya	H35		
						Tegaded	H36		
						Xiggo	H37		
						Zobira	H38		
						Gariya	H39		
						Wonade	H40		
						Kaseta-Le	H41		
						Sisqella Du	H42		

G, Gurage; S, Silte; K, Kembata and H, Hadiya administrative zones of farmers homegardens; R, released; A, sucker-wild and W, wild varieties, which are maintained in the Areka Agricultural Research Center (AARC) (*ex-situ* conservation site of enset crop) of Ethiopia.

References

[1] J.R. Harlan, Ethiopia: a center of diversity, Econ. Bot. 23 (1969) 309–314, https://doi.org/10.1007/BF02860676.

[2] J. Engels, J. Hawkes, The Ethiopian gene center and its genetic diversity, in: J. Engels, J. Hawkes, M. Worede (Eds.), Plant Genetic Resources of Ethiopia, Cambridge University Press, Cambridge, 1991.

[3] P. Tomilson, Anatomy of the Monocotyledons III. Commelinales Zingiberales, Clarendon Press, Oxford, UK, 1969.

[4] J.S. Borrell, M.K. Biswas, M. Goodwin, G. Blomme, T. Schwarzacher, J.S. Heslop-Harrison, A.M. Wendawek, A. Berhanu, S. Kallow, S. Janssens, E.L. Molla, Enset in Ethiopia: a poorly characterized but resilient starch staple, Ann. bot. 123 (5) (2019) 747–766, https://doi.org/10.1093/aob/mcy214.

[5] T. Abebe, K.F. Wiersum, F. Bongers, Spatial and temporal variation in crop diversity in agroforestry homegardens of southern Ethiopia, Agroforest syst 78 (2010) 309–322, https://doi.org/10.1007/s10457-009-9246-6.

[6] E. Cheesman, Classification of the bananas: the genus ensete horan, Kew Bull. 2 (1947) 97–106.

[7] E. Bekele, M. Shigeta, Phylogenetic relationships between Ensete and Musa species as revealed by the trnT trnF region of cpDNA, Genet. Resour. Crop Evol. 58 (2011) 259–269, https://doi.org/10.1007/s10722-010-9568-2.

[8] S.A. Brandt, A. Spring, C. Hiebsch, J.T. McCabe, E. Tabogie, M. Diro, G. Wolde-Michael, G. Yntiso, M. Shiegeta, S. Tesfaye, The "Tree against Hunger": Enset-Based Agricultural Systems in Ethiopia, American Association for the Advancement of Science, Washington D.C., 1997, p. 56. http://users.clas.ufl.edu/aspring/publications/enset.pdf.

[9] J.S. Borrell, M. Goodwin, G. Blomme, K. Jacobsen, A. Wendawek, D. Gashu, E. Lulekal, Z. Asfaw, S. Demissew, P. Wilkin, Enset-based agricultural systems in Ethiopia: a systematic review of production trends, agronomy, processing and the wider food security applications of a neglected banana relative, Plants people planet 2 (3) (2020) 1–17, https://doi.org/10.1002/ppp3.10084.

[10] Y. Tsehay, F. Kebebew, Diversity and cultural use of (*Ensete ventricosum* (welw.) cheesman) in bonga *in- situ* conservation site Ethiopia, Ethnobot. Res. Appl. 4 (2006) 147–157, https://doi.org/10.17348/era.4.0.147-158.

[11] G. Nuraga, T. Feyissa, K. Tesfaye, S. Demissew, Z. Tadele, Phenotypic diversity of enset (*Ensete ventricosum* (Welw.) Cheesman) landraces used in traditional medicine, Genet. Resour. Crop Evol. 66 (2019) 1761–1772, https://doi.org/10.1007/s10722-019-00832-3.

[12] T. Dilebo, T. Feyissa, Z. Asfaw, A. Zewdu, On-farm diversity, use pattern, and conservation of enset (*Ensete ventricosum*) genetic resources in southern Ethiopia, J. Ethnobiol. Ethnomed. 19 (1) (2023) 1–18, https://doi.org/10.1186/s13002-022-00569-x.

[13] Z. Yemataw, K. Tesfaye, A. Zeberga, G. Blomme, Exploiting indigenous knowledge of subsistence farmers for the management and conservation of enset (*Ensete ventricosum* (Welw.) Cheesman) (*musaceae* family) diversity on-farm, J. Ethnobiol. Ethnomed. 12 (2016) 34, https://doi.org/10.1186/s13002-016-0109-8.

[14] B. Haile, B. Tesfaye, T.M. Olango, Indigenous biosystematics of Enset (*Ensete ventricosum* [Welw.] Cheesman) in its center of origin and diversity, Southwest Ethiopia: folk nomenclature, classification, and descriptors, Econ. Bot. 76 (4) (2022) 382–413, https://doi.org/10.1007/s12231-022-09553-8.

[15] A. Tsegaye, P.C. Struik, Analysis of enset (*Ensete ventricosum*) indigenous production methods and farm-based biodiversity in major enset growing regions of Southern Ethiopia, Exp. Agric. 38 (2002) 292–315, https://doi.org/10.1017/S0014479702003046.

[16] R.R. Chase, L. Buchi, J. Rodenburg, N. Roux, A. Wendawek, J.S. Borrell, Smallholder farmers expand production area of the perennial crop enset as a climate coping strategy in a drought-prone indigenous agrisystem, Agri (2022) 20220283369, https://doi.org/10.1002/ppp3.10339.

[17] A. Negash, Diversity and Conservation of Enset (*Ensete Ventricosum* Welw. Cheesman) and its Relation to House- Hold Food and Livelihood Security in South-Western Ethiopia, PhD thesis Wageningen University and Research Centre, The Netherlands, 2001, p. 247p.

[18] Z. Asfaw, Enset (*Ensete ventricosum*, musaceae) ethnobotany: research status, gaps and key messages, Ethiop. J. Biol. Sci. 17 (2018) 51–62.

[19] M. Shigeta, Folk in–situ conservation of ensete. Towards the interpretation of indig- enous agricultural science of the Ari, south– western Ethiopia, Afr. Stud. Monogr. 10 (3) (1990) 93–107, 10008234879.

[20] K. Alemu, S. Sandford, Enset in North Omo Region, 1991, p. 49p. FRP Technical pampHlet number 1.Farm Africa, Addid Abeba, Ethiopia.

[21] G. Birmeta, H. Nybom, E. Bekele, Distinction between wild and cultivated Enset (*Ensete ventricosum*) gene pools in Ethiopia using RAPD markers, Hered 140 (2004) 139–148, https://doi.org/10.1111/j.1601-5223.2004.01792.x.

[22] B. Tesfaye, On Sidama folk identification, naming, and classification of cultivated enset (*Ensete ventricosum*) varieties, Genet. Resour. Crop Evol. 55 (2008) 1359–1370, https://doi.org/10.1007/s10722-008-9334-x.

[23] T. Olango, B. Tesfaye, C. Marcello, E. Mario, Indigenous knowledge, use and on-farm management of enset (*Ensete ventricosum* (Welw.) Cheesman) diversity in Wolaita, Southern Ethiopia, J. Ethnobiol. Ethnomed. 10 (2014) 41, https://doi.org/10.1186/1746-4269-10-41.

[24] W. Nkhata, H. Shimelis, R. Melis, R. Chirwa, T. Mzengeza, I. Mathew, A. Shayanowako, Population structure and genetic diversity analyses of common bean germplasm collections of East and Southern Africa using morphological traits and high-density SNP markers, PLoS One 15 (12) (2020) e0243238, https://doi.org/10.1371/journal.pone.0243238.

[25] M.J. Chappell, L.A. LaValle, Food security and biodiversity: can we have both? An agroecological analysis, Agric. Hum. Val. 28 (2011) 3–26, https://doi.org/10.1007/s10460-009-9251-4.

[26] J. Delfini, V. Moda-Cirino, J. dos Santos Neto, P.M. Ruas, G.C. Sant'Ana, P. Gepts, L.S. Gonçalves, Population structure, genetic diversity and genomic selection signatures among a Brazilian common bean germplasm, Sci. Rep. 11 (1) (2021) 2964, https://doi.org/10.1038/s41598-021-82437-4.

[27] R.K. Salgotra, B.S. Chauhan, Genetic diversity, conservation, and utilization of plant genetic resources, Genes 14 (1) (2023) 174, 2023.

[28] T. Feyissa, *Coccinia abyssinica* (Lam.) Cogn.(anchote) biology, productivity, and prospects of genetic improvement using biotechnological tools, J. Hortic. Res. 28 (2) (2020) 1–10, https://doi.org/10.2478/johr-2020-0023.
[29] P.W. Hedrick, Sex: differences in mutation, recombination, selection, gene flow, and genetic drift, Evolution 61 (12) (2007) 2750–2771, https://doi.org/10.1111/j.1558-5646.2007.00250.x.
[30] D. Ortiz-Barrientos, J. Engelstädter, L.H. Rieseberg, Recombination rate evolution and the origin of species, Trends Ecol. Evol. 31 (3) (2016) 226–236, https://doi.org/10.1016/j.tree.2015.12.016.
[31] M. Vallejo-Marín, S.J. Hiscock, Hybridization and hybrid speciation under global Change, New Phytol. 211 (4) (2016) 1170–1187, https://doi.org/10.1111/nph.14004.
[32] P. Kumar, V.K. Gupta, A.K. Misra, D.R. Modi, B.K. Pandey, Potential of molecular markers in plant biotechnology, Plant Omics 2 (4) (2009) 141–162, https://doi.org/10.3316/090706285698938.
[33] G. Birmeta, H. Nybom, E. Bekele, RAPD analysis of genetic diversity among clones of the Ethiopian crop plant *Ensete ventricosum*, Euphytica 124 (3) (2002) 315–325, https://doi.org/10.1023/A:1015733723349.
[34] A. Negash, A. Tsegaye, R. Van Treuren, B. Visser, AFLP analysis of enset clonal diversity in south and Southwestern Ethiopia for conservation, Crop Sci. 42 (2002) 1105–1111, https://doi.org/10.2135/cropsci2002.1105.
[35] D.C. Tobiaw, E. Bekele, Analysis of genetic diversity among cultivated enset (*Ensete ventricosum*) populations from Essera and Kefficho, southwestern part of Ethiopia using inter simple sequence repeats (ISSRs) marker, Afr. J. Biotechnol. 10 (70) (2011) 15697–15709, https://doi.org/10.5897/AJB11.885.
[36] S. Getachew, F. Mekbib, B. Admassu, S. Kelemu, S. Kidane, K. Negisho, A. Djikeng, I. Nzuki, Look into genetic diversity of Enset (*Ensete ventricosum* (Welw.) Cheesman) using transferable microsatellite sequences of banana in Ethiopia, J. Crop Improv. 28 (2014) 159–183, https://doi.org/10.1080/15427528.2013.861889.
[37] T.M. Olango, B. Tesfaye, M.A. Pagnotta, M.E. Pè, M. Catellani, Development of SSR markers and genetic diversity analysis in enset (Ensete ventricosum (Welw.) Cheesman), an orphan food security crop from Southern Ethiopia, BMC Genet. 16 (2015) 98, https://doi.org/10.1186/s12863-12015-10250-12868.
[38] F.N. Gerura, B.H. Meressa, K. Martina, A. Tesfaye, T.M. Olango, Y. Nasser, Genetic diversity and population structure of enset (*Ensete ventricosum* Welw Cheesman) landraces of Gurage zone, Ethiopia, Genet. Resour. Crop Evol. 66 (2019) 1813–1824, https://doi.org/10.1007/s10722-019-00825-2.
[39] G.W. Nuraga, T. Feyissa, K. Tesfaye, M.K. Biswas, T. Schwarzacher, J.S. Borrell, P. Wilkin, S. Demissew, Z. Tadele, J.S. Heslop-Harrison, The genetic diversity of enset (*Ensete ventricosum*) landraces used in traditional medicine is similar to the diversity found in non-medicinal landraces, Front. Plant Sci. 12 (2022) 756182, https://doi.org/10.3389/fpls.2021.756182.
[40] K.G. Tesfamicael, E. Gebre, T.J. March, B. Sznajder, D.E. Mather, C.M. Rodríguez López, Accumulation of mutations in genes associated with sexual reproduction contributed to the domestication of a vegetatively propagated staple crop, enset, Hortic. Res. 7 (2020) 185, https://doi.org/10.1038/s41438-020-00409-7.
[41] A.T. Haile, M.R. Kovi, S.S. Johnsen, B. Tesfaye, T. Hvoslef-Eide, O.A. Rognli, Genetic diversity, population structure and selection signatures in enset (*Ensete ventricosum*,(Welw.) cheesman), an underutilized and key food security crop in Ethiopia, Genet. Resour. Crop Evol. 70 (2023), https://doi.org/10.1007/s10722-023-01683-9.
[42] T. Borsch, K.W. Hilu, D. Quandt, V. Wilde, C. Neinhuis, W. Barthlott, Noncoding plastid trnT-trnF sequences reveal a well resolved phylogeny of basal angiosperms, J. Evol. Biol. 16 (2003) 558–576, https://doi.org/10.1046/j.1420-Ne9101.2003.00577.x.
[43] M.K. Biswas, J.N. Darbar, J.S. Borrell, M. Bagchi, D. Biswas, G.W. Nuraga, S. Demissew, P. Wilkin, T. Schwarzacher, J.S. Heslop-Harrison, The landscape of microsatellites in the enset (*Ensete ventricosum*) genome and web-based marker resource development, Sci. Rep. 10 (1) (2020) 1–11, https://doi.org/10.1038/s41598-020-71984-x.
[44] P.R. Peakall, P.P. Smouse, GenAlEx 6.5: genetic analysis in excel. Population genetics software for teaching and research– an update, Bioinformatics 28 (2012) 2537–2539.
[45] K. Liu, S.V. Muse, PowerMarker: an integrated analysis environment for genetic marker analysis, Bioinformatics 21 (9) (2005) 2128–2129, https://doi.org/10.1093/bioinformatics/bti282.
[46] S.T. Kalinowski, HP-RARE 1.0: a computer program for performing rarefaction on measures of allelic richness, Mol. Ecol. Notes 5 (2005) 187–189, https://doi.org/10.1111/j.1471-8286.2004.00845.x.
[47] L. Excoffier, H.E. Lischer, Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows, Mol. Ecol. Resour 10 (3) (2010) 564–567, https://doi.org/10.1111/j.1755-0998.2010.02847.x.
[48] M. Slatkin, N.H. Barton, A comparison of three indirect methods for estimating average levels of gene flow, Evolution 43 (7) (1989) 1349–1368, https://doi.org/10.1111/j.1558-5646.1989.tb02587.x.
[49] M. Nei, Genetic distance between populations, Am. Nat. 106 (1972) 283–292.
[50] P.H.A. Sneath, R.R. Sokal, Numerical Taxonomy, WH Freeman and Company, San Francisco, 1973, pp. 1–573.
[51] N. Takezaki, M. Nei, K. Tamura, POPTREE2: software for constructing population trees from allele frequency data and computing some other population statistics with windows interface, Mol. Biol. Evol. 27 (2010) 747–752, https://doi.org/10.1093/molbev/msp312.
[52] J. Felsenstein, Phylogenies and the comparative methods, Am. Nat. 125 (1985) 1–15, https://doi.org/10.1086/284325.
[53] R.D. Page, Tree View: an application to display phylogenetic trees on personal computers, Bioinformatics 12 (4) (1996) 357–358.
[54] R. Andrew, FigTree: Tree Figure Drawing Tool, Version 1.4.3. Institute of Evolutionary Biology, University of Edinburgh, United Kingdom, 2016. http://tree.bio.ed.ac.uk/software/figtree/.
[55] J.K. Pritchard, M. Stephens, P. Donnelly, Inference of population structure using multi-locus genotype data, For. Genet. 155 (2000) 945–959, https://doi.org/10.1093/genetics/155.2.945.
[56] D. Falush, M. Stephens, J.K. Pritchard, Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies, For. Genet. 164 (4) (2003) 1567–1587, https://doi.org/10.1093/genetics/164.4.1567.
[57] D.A. Earl, B.M. VonHoldt, Structure HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method, Cons. Genet. Resour. 4 (2012) 359–361, https://doi.org/10.1007/s12686-011-9548-7.
[58] G. Evanno, S. Regnaut, J. Goudet, Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study, Mol. Ecol. 14 (8) (2005) 2611–2620, https://doi.org/10.1111/j.1365-294X.2005.02553.x.
[59] N.M. Kopelman, J. Mayzel, M. Jakobsson, N.M. Rosenberg, I. Mayrose, Clumpak: a program for identifying clustering modes and packaging population structure inferences across K, Mol. Ecol. Res. 15 (5) (2015) 1179–1191, https://doi.org/10.1111/1755-0998.12387.
[60] E.J. Howells, B.L. Willis, L.K. Bay, M.J.H. Van Oppen, Microsatellite allele sizes alone are insufficient to delineate species boundaries in Symbiodinium, Mol. Ecol. 25 (12) (2016) 2719–2723, https://doi.org/10.1111/mec.13631.
[61] S.N. Rampersad, D. Perez-Brito, C. Torres-Calzada, R. Tapia-Tussell, V. F. Carrington, Genetic structure and demographic history of Colletotrichum gloeosporioides sensu lato and C truncatum isolates from Trinidad and Mexico, BMC Evol. Biol. 13 (2013) 130, https://doi.org/10.1186/1471-2148-13-130.
[62] D.J. Futuyma, Sympatric speciation: norm or exception? in: K.J. Tilmon (Ed.), Specialization, Speciation, and Radiation: the Evolutionary Biology of Herbivorous Insects Univ. of California Press, Berkeley, 2008, pp. 136–148.
[63] A. Prabakaran, K. Paramasivam, T. Rajesh, D. Rajarajan, Molecular characterization of rice landraces using SSR markers, J. Plant Breed. 1 (4) (2010) 512–516.
[64] A.V. Tiago, E.S.S. Hoogerheide, E.C.M. Pedri, F.S. Rossi, E.S. Cardoso, J.M.A. Pinto, G. F Pena, A.A.B. Rossi, Genetic diversity and population structure of cassava ethno-varieties grown in six municipalities in the state of Mato Grosso, Brazil, Genet. Mol. Res. 18 (4) (2019), https://doi.org/10.4238/gmr18357 gmr18357.
[65] S.K. Dutta, S. Singh, D. Chatterjee, T. Boopathi, A. Singh, S. Saha, Morphological and genetic diversity of pole type common bean (Phaseolus vulgaris L.) landraces of Mizoram (India), Indian J. Biotechnol. 15 (4) (2016) 550–559. http://nopr.niscpr.res.in/handle/123456789/41026.
[66] M.K. Biswas, M. Bagchi, D. Biswas, J.A. Harikrishna, Y. Liu, C. Li, O. Sheng, C. Mayer, G. Yi, G. Deng, Genome-wide novel genic microsatellite marker resource development and validation for genetic diversity and population structure analysis of banana, Genes 11 (12) (2020) 1479, https://doi.org/10.3390/genes11121479.
[67] I.A. Matus, P.M. Hayes, Genetic diversity in three groups of barley germplasm assessed by simple sequence repeats, Genome 45 (2002) 1095–1106, https://doi.org/10.1139/g02-071.
[68] A.A. Dido, M.S. Krishna, E. Assefa, D.T. Degefu, B.J. Singh, K. Tesfaye, Genetic diversity, population structure and relationship of Ethiopian barley (*Hordeum vulgare* L.) landraces as revealed by SSR markers, J. Genet. 101 (1) (2022) 9, https://doi.org/10.1007/s12041-021-01346-7.
[69] W.A. Mengesha, S. Demissew, M.F. Fay, R.J. Smith, I. Nordal, P. Wilkin, Genetic diversity and population structure of Guinea yams and their wild relatives in South and South West Ethiopia as revealed by microsatellite markers, Genet. Resour. Crop Evol. 60 (2013) 529–541, https://doi.org/10.1007/s10722-012-9856-0.
[70] A.S. Jump, R. Marchant, J. Peñuelas, Environmental change and the option value of genetic diversity, Trends Plant Sci. 14 (1) (2009) 51–58, https://doi.org/10.1016/j.tplants.2008.10.002.
[71] F. Gadissa, K. Tesfaye, K. Dagne, M. Geleta, Genetic diversity and population structure analyses of *Plectranthus edulis* (Vatke) Agnew collections from diverse agro-ecologies in Ethiopia using newly developed EST-SSRs marker system, BMC Genet. 19 (1) (2018) 1–15, https://doi.org/10.1186/s12863-018-0682-z.
[72] D. McKey, M. Elias, B. Pujol, A. Duputié, The evolutionary ecology of clonally propagated domesticated plants, New Phytol. 186 (2) (2010) 318–332, https://doi.org/10.1111/j.1469-8137.2010.03210.x.
[73] Z. Teshome, M.T. Terfa, B. Tesfaye, E. Shiferaw, T.M. Olango, Genetic diversity in anchote (*Coccinia abyssinica* (Lam.) Cogn) using microsatellite markers, Curr. Plant Biol. 24 (2020) 100167, https://doi.org/10.1016/j.cpb.2020.100167.
[74] Z. Yemataw, H. Mohamed, M. Diro, T. Addis, G. Blomme, Enset (*Ensete ventricosum*) landrace selection by farmers and their cultural practices in southern Ethiopia, Genet. Resour. Crop Evol. 61 (6) (2014) 1091–1104, https://doi.org/10.1007/s10722-014-0093-6.
[75] A. Zeberga, Z. Yemataw, W. Sinebo, S. Musemil, D. Ambachew, On farm cultivar diversity of Enset (*Ensete ventricosum* W.) in Southern Ethiopia, J. Agric. Dev. 4 (1) (2014) 63–85. https://opendocs.ids.ac.uk/opendocs/handle/20.500.12413/8780.
[76] S. Abdelhamid, C.L. Lê, M. Conedera, P. Kuepfer, The assessment of genetic diversity of Castanea species by RAPD, AFLP, ISSR, and SSR markers, Turk. J. Bot. 38 (5) (2014) 835–850, https://doi.org/10.3906/bot-1303-30.
[77] H. Hamza, M.A. Abederrahim, M. Elbekkay, A. Ferchichi, Comparison of the effectiveness of ISSR and SSR markers in determination of date palm (*Phoenix dactylifera* L.) agronomic traits, Aust. J. Crop. Sci. 7 (2013) 763–769.
[78] S. Wright, The genetical structure of populations, Ann. Eugen. 15 (1951) 323–354.
[79] G. Özkan, K. Haliloğlu, A. Türkoğlu, H.I. Özturk, E. Elkoca, P. Poczai, Determining genetic diversity and population structure of common bean (*Phaseolus vulgaris* L.)

landraces from Türkiye using SSR markers, Genes 13 (8) (2022) 1410, https://doi.org/10.3390/genes13081410.
[80] P.M. Berry, V. Fabók, M. Blicharska, Y.K. Bredin, M.G. Llorente, E. Kovács, N. Geamana, A. Stanciu, M. Termansen, T. Jääskeläinen, J.R. Haslett, Why conserve biodiversity? A multi-national exploration of stakeholders' views on the arguments for biodiversity conservation, Biodivers. Conserv. 27 (2018) 1741–1762, https://doi.org/10.1007/s10531-016-1173-z.
[81] C. Khoury, B. Laliberté, L. Guarino, Trends in ex situ conservation of plant genetic resources: a review of global crop and regional conservation strategies, Genet. Resour. Crop Evol. 57 (2010) 625–639, https://doi.org/10.1007/s10722-010-9534-z.
[82] R. Rao, B. Sthapit, B, Conservation of tropical plant genetic resources: in situ approach, in: Conservation of Tropical Plant Species, Springer New York, New York, NY, 2012, pp. 3–26, https://doi.org/10.1007/978-1-4614-3776-5_1.
[83] K. Zippel, P. Ludders, The Global Food and Product Chain Dynamics, Innovations, Conflicts, Strategies: *Ensete ventricosum* in Ethiopia: the Need to Grow More than One Landraces, Deutscher Tropentag, 2005, pp. 11–13.
[84] Z. Yemataw, Agro-morphological and molecular characterization of enset (*Ensete ventricosum* (welw.) cheesman) landraces from Ethiopia, in: PhD Dissertation, Addis Ababa University, 2018.