

SSR Based Assessment of Genetic Diversity in Hot Pepper (*Capsicum annuum* L.) Genotypes from Central Ethiopia Regional State, Ethiopia

Henok Ayelign

ayelignhenok2005@gmail.com

Werabe University

Eleni Shiferaw

Ethiopian Biodiversity Institute

Ebrahim Sirgaga

Werabe University

Bahiru Wabela

Werabe University

Article

Keywords: *Capsicum annuum*, Genetic diversity, SSR marker

Posted Date: November 12th, 2025

DOI: <https://doi.org/10.21203/rs.3.rs-7929864/v1>

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

Additional Declarations: No competing interests reported.

Abstract

Genetic diversity assessment and estimation of relationship among genotypes at DNA level are very helpful for germplasm conservation and improvement in breeding programs. This study was carried out to assess the genetic variability of 57 hot pepper genotypes using SSR markers. Thirteen SSR markers generated a total of 106 scorable alleles; ranging from 5 to 14 alleles per locus. The polymorphism information content indices were highly variable and ranged from 0.21 to 0.81, with an average of 0.68. The gene diversity was wider in range, 0.21 to 0.83, which implies an immense variability in the population. The overall percentage of polymorphic bands (73.4%), genetic diversity ($H=0.53$) and Shannon's diversity information index ($I=0.60$) showed high level of variation in the analysed genotypes. The Nei's genetic distance ranging from 0.108 to 0.925 indicated presence of nearly identical genotypes to a diverse relationship. Clustering, principal coordinate and STRUCTURE analysis results indicated that most genotypes clustered together with no clear separation based on zones, this underlines the importance of considering genetic variation across genotypes rather than focusing on collection sites. The finding confirmed the presence of genetic diversity which provides baseline information for conservation, utilization and breeding.

Introduction

Pepper (*Capsicums* spp.) which is believed to have its ancestral origins in the tropical South American region centered in what is now Bolivia (Patricia *et al.* 2003) and are the most important spice and vegetable crop grown at an altitude ranging from 1400–1900 m a.s.l.. Hot pepper (*C. annum*) is the main species in Ethiopia grown under rainfed and irrigated conditions [1].

Hot pepper is a high value crop due to its high pungency and serve as food consumption and source of cash earning for smallholder producers in both green and dry form as well as for medicines and lachrymatory agents [1, 2, 3, 4]. It is rich in proteins, lipids, carbohydrates, fibers, mineral, salts and vitamins. Fruits are an excellent source of health-related photochemical compounds that are very important in preventing chronic diseases such as cancer, asthma, coughs, sore throats, toothache, diabetes and cardiovascular diseases [5]. It is also a rich source of vitamin A, E and contains five to six times as much vitamin C as an orange or a lemon, making it an ideal vegetable to prevent flu colds more than any other vegetable crop [6]. The average daily consumption of hot pepper by Ethiopian adults is about 15 g, which is higher than most other vegetables [7]. Ethiopia achieved a record harvest of 5.1 million quintals of hot pepper from 274,000 ha in 2023/24 production year [8]. The Southern Nations Nationalities and Peoples Regional State (SNNPRs) had 1,999,668 farmers that cultivate hot pepper in an area of 174,463 ha and produce 3,131,154 quintal and the average productivity was 17.95 quintal per hectare [9]. Similarly, the Gurage, Silte and Alaba zones in the Central Ethiopia Regional state, have suitable agro-ecologies for hot pepper production and are the major producing areas [10]. In these zones it is a key cash crop and contributes to a significant portion of smallholder farmers' income, though the overall market system faces challenges like declining output quality crop and susceptibility to diseases [11].

In spite of its use and a wide range of genetic resources, little effort has been made to improve the quality and productivity of hot pepper in Ethiopia [12]. Genetic diversity in plant species can be determined using morphological, biochemical, and molecular markers. However, morphological markers are strongly affected by environmental factors since plants are dependent on environmental conditions during cultivation. Biochemical markers are less affected by environmental influences but they are unable to detect low level of variation because they screen only a very small expressed genome section [13]. Molecular techniques have been widely used for estimating the genetic diversity, identifying cultivars and plant populations [14], as they are not influenced by many factors, including environment, plant growth stage and level of variation [15]. SR markers became one of the main molecular markers for genetic studies, especially as a result of the high level of polymorphism, multiallelism and high reproducibility.

Despite enormous uses of hot pepper as vegetable, spice, medicine, ornamental and source of income, extensive production practiced and its productivity in Central Ethiopia is low. This might be attributed to the use of low yielding varieties, using unknown seed sources, poor irrigation system, lack of information on soil fertility, drought, insect, pest, diseases and lack of research outputs on production techniques [16]. In the region, farmers usually save their own seed and transfer it from one generation to the next. However, proper seed production methods including isolation techniques are not in practice within and among farmers, giving chance to outcross and introgression among different genotypes.

Genetic diversity assessment and estimation of relationship among genotypes are very helpful for efficient preservation and proper utilization of the crop. There are few studies on Ethiopian hot pepper using morphological [2, 3], ISSR [17] and SSR [1] markers. However, information on the magnitude and pattern of genetic diversity of the hot pepper genotypes using molecular markers are so

far limited and no study carried out on the area based on SSR marker. So, this study was aimed to assess the existing genetic diversity of hot pepper genotypes from Gurage, Silte and Alaba zones using SSR marker. This could provide information on the genetic variability, which may assist in the identification and selection of the genetic materials for improvement, distribution, conservation and breeding.

Materials and Methods

Plant materials

A total of 57 genotypes (47 accessions from Ethiopian Biodiversity Institute and 10 Farmers' varieties) collected from Gurage, Silte and Alaba zones of Central Ethiopia Regional State were used to study the genetic variation of hot pepper genotypes using SSR markers. Samples were sown on the plastic pot to obtain young leaves for DNA extraction.

DNA extraction

Ten young leaves from each genotype was bulked in to two bulked samples (containing five genotypes each) and grounded into fine powder using a mortar and pestle and DNA was extracted using a modified CTAB protocol [18]. Two types of buffers – buffer 1 (0.1 M Tris-HCl (pH 7.6), 0.005 M EDTA, 0.35 M sorbitol) and buffer 2 (0.2 M Tris-HCl (pH 8.0), 0.05 M EDTA, 2M NaCl, 2% CTAB), were mixed in equal amounts and 5% sodium Laurylsarcosine (0.8 final volume) and 0.38% sodium bisulfate were added just before extraction. The mixture was incubated at 65°C for 60 min, then two extractions were carried out with chloroform/isoamyl alcohol (24:1). After which the nucleic acids precipitated with Cold isopropanol, and the pellet was dissolved in 1 mL TE 0.1X (Tris-EDTA) buffer. The quality and presence of isolated DNA was checked by gel electrophoresis using 1 % agarose at 100 V constant for 30 minutes.

PCR Amplification and Electrophoresis

Fifteen SSR primers, publicly available from [19], were screened and thirteen primers that produced clear and reproducible band patterns were selected for further study using the 57 genotypes. Primer names, forward and reverse primer sequences, annealing temperature and expected fragment sizes are shown in Table 1. Polymerase chain reactions (PCR) was performed in a PCR express thermocycler (HYBAID, UK), in a final reaction volume of 10 µl containing 20 ng (1 µl) genomic DNA, 0.4 µM forward primer, 0.4 µM reverse primer, 0.2 µl dNTPs, 1µl PCR buffer (10x), 0.6 mM MgCl₂, 5.1 µl PCR water and 1U Taq DNA polymerase. Amplification was carried out using the following program: initial denaturation step of 3 min at 94°C followed by 35 cycles with denaturation for 30 sec at 94°C, annealing for 30 sec at 50-60°C (depending on the individual primer), and extension for 45 sec at 72°C. A final extension step at 72°C was performed for 10 minutes. The amplification products were differentiated by electrophoresis using 6% polyacrylamide gel in 0.5x TBE buffer with a 6x DNA loading dye. DNA marker 100 bp (Solis Bio Dyne, Estonia) was used to estimate molecular weight and size of the fragments. Finally, the gel was stained using silver nitrate following the procedure developed by [20]. Gels were documented using digital camera and the allele sizes were determined using UVITEC software (UVITEC, Cambridge, UK).

Table1. Amplified primers, their sequences and annealing temperatures used for the analysis of *Capsicum annum*

Marker	Repeat	Primer sequence	Ta0C	Fragment size (bp) ¹
HPMS1-5	(AT)11(GT)17	F:CCAAACGAACCGATGAACACTC R:GACAATGTTGAAAAAGGTGGAAGAC	55	300-340
GPMS29	(GT)15GGT7(GTT)2	F:CAGGCAATACGGAGCATC R:TGTGTTGCTTCTTGACGAC	50	238-288
HPMS1-111	(AAT)12	F:AAGCTTATCCCTTTCAAATATAA R:ATATCTCACGTATTGCGGATTCTT	50	159-190
GPMS161	(AAT)25	F:CGAAATCCAATAAACGAGTGAAG R:CCTGTGTGAACAAGTTTTTCAGG	53	269-310
HPMS1-214	(GTTT)2(TTG)9	F:TGCGAGTACCGAGTTCTTTCTAG R:GGCAGTCCTGGGACAACCTCG	55	78-104
HPMS2-21	AT)11 (AC)9 (ATAC)10	F:TTTTTCAATTGATGCATGACCGATA R:CATGTCATTTTGTGATTGTTGG	52	307-350
HPMS2-24	(CT)17(CA)5A21	F:TCGTATTGGCTTGTGATTTACCG R:TTGAATCGAATACCCGACAGGAG	52	196-221
EPMS310	(CAT)13	F:TGGGAAGAGAAATTGTGAAAGC R:AGGAAACATGGTTCAATGCC	53	159-175
EPMS342	(CTT)7	F:CTGGTAGTTGCAAGAGTAGATCG R:ATGATCTTTGACGACGAGGG	55	309-347
EPMS386	(CA)15	F:ACGCCAAGAAAATCATCTCC R:CCATTGCTGAAGAAAATGGG	50	139-183
EPMS426	(AT)15	F:GAGGAAACACTCTCTCTCTCTCTC R:TCAAGAGACCCCAAATAGGG	55	106-145
GPMS169	(ATT)5T(TTA)7	F:TCGAACAAATGGGTCATGTG R:GATGAGGGTCCTGTGCTACC	53	179-200
GPMS197	(GA)3(TAT)16	F:GCAGAGAAAATAAAATTCTCGG R:CAATGGAAATTTTCATCGACG	52	290-360

Genetic diversity analysis

The number of alleles, gene diversity, heterozygosity and the polymorphic information content (PIC) was calculated using Power Marker version 3.25 [21]. The dendrogram tree was constructed by the unweighted pair-group method with arithmetic average (UPGMA) based on Euclidian distance using Power marker 3.25 and the tree was visualized using Molecular Evolutionary Genetics Analysis (MEGA 6) [22]. The number of effective alleles, analysis of the molecular variation (AMOVA) and Principal Coordinate Analysis (PCA) was performed using GenALEX software version 6.5 [23]. The Bayesian model-based software STRUCTURE 2.3.4 Pritchard *et al.* (2000) was used to infer the population structure. The analysis of Population structure was performed 10 times for each K value (K=1 to 10) using a burn-in period of 50,000 and 100,000 Markov chain Monte Carlo (MCMC) iterations. The optimal K value among K groups was determined using STRUCTURE HARVESTER [24]. The result files were aligned and displayed in each K drawn as colored box plots by CLUMPP software [25].

Result

Amplification and polymorphism of SSR markers

Table 2. Characteristics of the 13 analyzed microsatellite loci in 57 hot pepper (*capsicum annuum*) genotypes

No.	Marker	No. of effective alleles	Allele	No.	Gene diversity	Heterozygosity	PIC
1	HPMS1-5	1.690	8		0.830024623	0.166666667	0.808185797
2	GPMS29	1.561	6		0.796398892	0.210526316	0.766531966
3	HPMS1-111	2.506	9		0.756694367	0.877192982	0.721200613
4	GPMS161	2.429	11		0.778508772	0.929824561	0.751608535
5	HPMS1-214	1.977	5		0.645544783	0.859649123	0.578652595
6	HPMS2-21	1.799	7		0.792166821	0.464912281	0.761591246
7	HPMS2-24	1.607	6		0.620575562	0.245614035	0.579591529
8	EPMS310	1.052	6		0.211911357	0	0.206952839
9	EPMS342	1.997	7		0.731494306	0.605263157	0.694321421
10	EPMS386	2.844	12		0.799938443	0.912280702	0.774874478
11	EPMS426	2.072	8		0.694252078	0.789473684	0.64609312
12	GPMS169	1.525	7		0.797129886	0.175438596	0.772545789
13	GPMS197	2.245	14		0.778431825	0.649122807	0.762502842
	Average	1.946	8.1538		0.710236286	0.529689609	0.67957869

A total of 106 clear and scorable bands were recorded by the polymorphic primers for the tasted 57 pepper populations having SSR fragment sizes ranging from 78 - 360 base pairs. The number of different alleles produced per primer varied from five for primer HPMS1-214 to fourteen for primer GPMS197, with an average of 7.385 (Table 2). The average numbers of effective alleles of SSR markers were 1.946. Higher and lower number of effective alleles was recorded on primer GPMS386 and EPMS310 with a value of 2.844 and 1.052 respectively.

The polymorphism information content (PIC) indices, which represent their distribution, were highly variable ranged from 0.21 to 0.81, with an average of 0.68. Among the 13 SSR used, the highest PIC value (0.81) was recorded by primer EPMS1-5 followed by HPMS386 (0.775), whereas the lowest PIC value (0.21) was detected by marker EPMS310.

Gene diversity of hot pepper populations

The gene diversity indices were highly variable across SSR loci and ranged from 0.21 to 0.83, with mean value of 0.71. Primer HPMS1-5 exhibited highest gene diversity (0.83) followed by EPMS386 with a value of 0.80, whereas lowest gene diversity value (0.21) was detected for marker EPMS310 (Table 2). Out of the thirteen, nine SSR primers showed higher gene diversity than the average value (0.71). Observed heterozygosity was also highly variable between loci and ranged from 0 to 0.93 with mean value of 0.53. Marker GPMS161 had the maximum heterozygosity (0.93), followed by EPMS386 (0.91); while the minimum was recorded by primer EPMS310 with a value of 0.

Population genetic diversity

The percentage of polymorphic bands ranged from 38.5% to 100% and averaged 73.4%. High percentages of polymorphic band were observed in both pop1 and pop5 which was collected from Gurage zone (Abshege wereda) with a value of 100% (Table 3). The least band polymorphism was found in pop37 followed by pop32 which was collected from Alaba zone (Kembata wereda) with a value of 38.5% and 46.2% respectively. The mean number of different alleles per locus (N_a) and effective alleles of all populations were 2.0 and 1.9 respectively, with pop5 having the maximum value ($N_a=2.5$ and $N_e=2.4$). The highest gene diversity was shown by pop5 ($H=0.72$, $I=0.85$) followed by pop1 ($H=0.73$, $I=0.84$) both from Gurage zone (Abshege wereda) and pop10 from Gurage zone (Meskan wereda) ($H=0.67$, $I=0.76$). The least gene diversity was observed in pop37 collected from Alaba zone (Kembata wereda) ($H=0.28$, $I=0.32$).

The 57 tested hot pepper genotypes were grouped into three populations based on origin of collection. Within populations, Shannon's informative index ranged from 1.254 (Alaba zone) to 1.548 (Gurage zone) with average of 1.402. The maximum and the minimum expected heterozygosity was revealed from Gurage zone (0.725) and Alaba zone (0.628) respectively with mean value of 0.672. Similarly, the observed heterozygosity was varied from 0.508 (Alaba zone) to 0.580 (Silte zone) with mean value 0.540. The number of effective alleles ranged from 3.117 to 4.062 with an average value of 3.591. High percentages of polymorphic band were observed in both Gurage and Silte zones with a value of 100%, while the least was from Alaba zone (92.31%) (Table 4).

Table 3. Summary of parameters for genetic diversity in hot pepper genotypes: No. of different alleles (Na), No. of effective alleles (Ne), Percentage of polymorphic bands (PPB), Heterozygosity (H) and Diversity index (I).

Code	Accen.	CS	Na	Ne	PPB	H	I	Code	Accen.	CS	Na	Ne	PPB	H	I
Pop1	25809	GZ	2.5	2.3	100	0.73	0.84	Pop30	223642	GZ	1.8	1.8	69.2	0.46	0.51
Pop2	25810	GZ	2.4	2.1	92.3	0.63	0.75	Pop31	223654	AZ	1.6	1.6	69.2	0.54	0.64
Pop3	25811	GZ	2.1	2.0	61.5	0.49	0.58	Pop32	223651	AZ	1.8	1.8	46.2	0.35	0.36
Pop4	25812	GZ	2.2	2.1	69.2	0.54	0.66	Pop33	223648	AZ	1.8	1.8	61.5	0.45	0.51
Pop5	25813	GZ	2.5	2.4	100	0.72	0.85	Pop34	223655	AZ	1.8	1.7	69.2	0.47	0.51
Pop6	25815	GZ	2.0	1.9	76.9	0.58	0.61	Pop35	223653	AZ	2.0	1.9	84.6	0.59	0.64
Pop7	25816	GZ	2.2	2.0	84.6	0.59	0.67	Pop36	223652	AZ	1.8	1.8	53.9	0.40	0.47
Pop8	25821	GZ	2.2	2.0	92.3	0.64	0.71	Pop37	223656	AZ	1.5	1.5	38.5	0.28	0.32
Pop9	25822	GZ	1.9	1.9	69.2	0.49	0.55	Pop38	223657	AZ	1.9	1.8	69.2	0.47	0.54
Pop10	25823	GZ	2.3	2.2	92.3	0.67	0.76	Pop39	223650	AZ	2.2	2.1	76.9	0.58	0.68
Pop11	25824	GZ	2.1	2.1	76.9	0.56	0.64	Pop40	223658	AZ	2.0	1.9	69.2	0.50	0.58
Pop12	25826	GZ	2.1	2.0	84.6	0.59	0.66	Pop41	223659	AZ	2.3	2.2	76.9	0.59	0.71
Pop13	25827	GZ	2.0	1.9	84.6	0.60	0.60	Pop42	223649	AZ	2.4	2.2	76.9	0.62	0.75
Pop14	223638	GZ	2.2	2.1	92.3	0.63	0.70	Pop43	212912	AZ	2.3	2.2	84.6	0.65	0.74
Pop15	223647	GZ	2.2	2.1	84.6	0.65	0.72	Pop44	Zala	AZ	1.8	1.8	76.9	0.55	0.56
Pop16	223640	GZ	2.3	2.1	76.9	0.59	0.71	Pop45	Zala	AZ	1.7	1.6	61.5	0.41	0.44
Pop17	223635	GZ	2.2	2.1	76.9	0.58	0.67	Pop46	25817	SZ	2.2	2.0	92.3	0.64	0.69
Pop18	223633	GZ	1.9	1.9	53.9	0.42	0.50	Pop47	25818	SZ	2.1	1.9	76.9	0.55	0.61
Pop19	223639	GZ	2.3	2.1	84.6	0.62	0.73	Pop48	25819	SZ	2.2	2.0	76.9	0.56	0.66
Pop20	223645	GZ	2.2	2.1	69.2	0.55	0.67	Pop49	223646	SZ	1.8	1.7	69.2	0.46	0.50
Pop21	Karya	GZ	1.8	1.7	61.5	0.42	0.50	Pop50	25820	SZ	2.2	2.0	76.9	0.56	0.66
Pop22	Marekofana	GZ	2.4	2.3	76.9	0.58	0.72	Pop51	25828	SZ	1.8	1.7	69.2	0.46	0.51
Pop23	Torbe	GZ	2.1	2.0	61.5	0.47	0.58	Pop52	25829	SZ	1.7	1.7	61.5	0.42	0.45
Pop25	223634	GZ	1.7	1.6	61.5	0.41	0.44	Pop53	25830	SZ	2.3	2.2	84.6	0.60	0.72
Pop24	Marekofana	GZ	1.9	1.8	69.2	0.47	0.54	Pop54	Melka awaze	SZ	2.0	1.9	61.5	0.47	0.56
Pop26	223637	GZ	1.6	1.5	53.9	0.35	0.38	Pop55	Zala	SZ	1.8	1.7	69.2	0.45	0.49
Pop27	223643	GZ	1.8	1.7	53.9	0.37	0.45	Pop56	Kondo	SZ	1.8	1.8	84.6	0.54	0.57
Pop28	223641	GZ	2.0	1.9	69.2	0.50	0.58	Pop57	Tenkeso	SZ	2.2	2.1	76.9	0.56	0.67
Pop29	223644	GZ	2.1	2.0	69.2	0.51	0.60								
Mean											2.0	1.9	73.4	0.53	0.63

Accen = Accessions, CS= Collected Site, GZ = Gurage Zone, AZ = Alaba Zone, SZ = Silte Zone

Genetic distance

From Nei unbiased genetic distance matrix which compares pairwise genetic distances between all combinations of any two populations, the most distantly related populations were pop20 and pop45 which were collected from Gurage zone, Mareko wereda and Alaba zone, Weyradijo wereda, with a value of 0.925. Pop49 from Silte zone (Lanfro wereda) and pop50 from Silte zone (Dalucha

woreda); followed by pop30 from Gurage Zone (Mareko woreda) and pop 34 from Alaba zone (Kembata woreda) had the smallest genetic distance with a value of 0.108 and 0.111 respectively.

Analysis of genetic distance among accessions pooled by origin of collection were investigated and the most distantly related accessions found between Alaba zone and Silte zone with a value of 0.098. The lowest genetic distance was revealed between Gurage zone and Silte zone (0.075) (Table 5).

Table 4. Summary of genetic diversity in hot pepper genotypes from three geographic origins

Population	N	Na	Ne	I	Ho	He	%P
Gurage Zone	60	7.231	4.062	1.548	0.534	0.725	100
Alaba Zone	30	5.692	3.117	1.254	0.508	0.628	92.31
Silte Zone	24	6.00	3.591	1.405	0.580	0.694	100
Mean	38	6.308	3.591	1.402	0.540	0.672	97.44

Mean number of different alleles (Na), the mean number of effective alleles (Ne), Shannon’s index (I), observed heterozygosity (Ho), and expected heterozygosity (He), Percentage of polymorphic bands (%P)

Table 5. Pairwise matrix of Nei’s unbiased genetic distance by geographic origins

	Gurage Zone	Alaba Zone	Silte Zone	
0.000				Gurage Zone
0.076		0.000		Alaba Zone
0.075		0.098	0.000	Silte Zone

Analysis of molecular variance

Analysis of molecular variance (AMOVA) indicated a total of 97% within zones and 3% among zones variation with gene flow of Nm=7.768 (Table 6). Total SSR marker variation in the analysed populations was also partitioned by AMOVA into among population; among individuals and within individual components which showed 22% into among population; 4% among individuals and 74% within individuals. Wright’s F-statistics for the tested markers were Fis (0.054) and Fit (0.260) with the rate gene flow (0.9). The SSR markers had a mean fixation index (Fst) of 0.217, indicating about 21.7% genetic differentiation among populations independent of the marker used (P< 0.01) (Table 7).

Table 6. Analysis of molecular variance among the 3 Zones of *Capsicum annum*

Source	Df	SS	MS	Est. Var.	% of variance	Nm	P-Value
Among Zones	2	48.576	24.288	0.369	3%	7.768	0.010
Within Zones	111	1271.565	11.456	11.456	97%		
Total	113	1320.140		11.824	100%		

DF, degree of freedom; SS, sum of squared observations; MS, mean of squared observations; Est. Var., estimated variance; Nm, gene flow

Table 7. Analysis of molecular variance among the 57 populations of *C. annum*

Source	Df	SS	MS	Est. Var.	% of variance	F _{st}	P-Value
Among Populations	56	441.320	7.881	1.011	22%	0.217	0.010
Among Individuals	57	218.750	3.838	0.197	4%		
Within region							
Within Individuals	114	392.500	3.443	3.443	74%		
Total	227	1052.570		4.651	100%		

DF, degree of freedom; SS, sum of squared observations; MS, mean of squared observations; Est. Var., estimated variance, F_{st}, genetic differentiation.; P-value, probability for F_{st} based on standard permutation across the full data set.

Clustering

The 57 genotypes collected from the three zones of Central Ethiopia were divided into three clusters, with cluster I, II, and III consisting of 6, 19, and 89 genotypes, respectively. Cluster II was divided into three subgroups of which subgroup I was constructed by five genotypes (G3, G4, G5, G6 and G102) out of these four genotypes were collected from Gurage zone, Abshege woreda and the remaining one (G102) was from Silte zone, Silti woreda. Subgroup II of cluster II showed admixture of six genotypes from the three zones. Subgroup III contains eight genotypes of which seven of them were collected from Gurage zone, Meskan woreda. Cluster III was constructed through the intermixing of 89 genotypes from the three zones (Fig1).

Principal Coordinates and population structure Analysis

The PCA showed that all genotypes were distributed across the plot, with 59.59% of the total variation explained in the first three axes, with 23.18%, 19.13% and 17.28% contributions respectively. Populations (27, 42, and 57) which were collected from different zones tended to form their own grouping. Similarly, pop12 from Gurage zone, Meskan woreda, pop15 from Gurage zone, Mareko woreda, pop40 from Alaba zone, Kembata woreda and pop47 from Silte zone, Dalucha woreda tended to form their own grouping coordinate. Five populations (49, 51, 55, 56 and 57) collected from Silte zone showed different grouping coordinate (Fig 2A).

The population structure of 57 pepper genotypes was determined using structure 2.3.4, and the result of the structure harvester revealed that the largest peak was found at K=3, indicating the presence of three populations (Fig 2B). The populations of Gurage zone showed admixture with populations of Alaba and Silte zones. Similarly, there was also some overlap between the population of Alaba and those of Silte Zone (Fig 2C).

Discussion

Knowledge of level of genetic diversity of plants provides a great chance for selecting heterogeneity patterns to develop heterosis effect for breeding programs and future conservation of the germplasm. The methodology for the estimation of genetic diversity can be different: morphological, biochemical and molecular markers. In this work, fifty-seven hot pepper accessions from Silte, Gurage and Alaba zones were assessed by SSR markers for determining the level of genetic variation. SSR markers were used successfully for genetic diversity assessment in capsicum species in previous studies [1, 26]. The 13 microsatellite markers used in this study generated a significant number of different alleles with total 106 scorable alleles. Slightly lower average allelic number per locus (8.15) was recorded in this study than an earlier study by [1], who reported 8.54 average alleles per locus for 75 hot pepper accessions. The difference in allele numbers may be due to the different sample sizes and genotypes used. Polymorphism information content (PIC) values above 0.5 are consistent with the usefulness of the marker [27]. Here, PIC values ranged from 0.21 to 0.81, with a mean value of 0.68, which implied the presence of higher genetic diversity among the germplasms. The PIC results of this study are almost similar to values reported in genetic diversity of pepper germplasm resources in China [28].

Gene diversity index which represent the number of alleles and their distribution, were highly variable across the SSR loci ranged from 0.21 to 0.83 with average value of 0.71, represents that the genotypes were not fixed to homozygosity. The level of diversity observed in this study is nearly similar to the results of [1] with average gene diversity of 0.62 among Ethiopian hot pepper accessions. [2, 17]

also reported the presence of high genetic variation in Ethiopian pepper using morphological and ISSR markers respectively. In contrast our study showed high heterozygosity (0.393) than those reported on the germplasm resources in China which reported average heterozygosity of 0.07 [28].

Percentage polymorphic bands which ranged from 38.5 to 100% with average value of 73.4% revealed a wide and diverse genetic base in *C. annuum*. Similarly, high percentage polymorphism was found in the Capsicum species by earlier studies [17, 29, 30, 31, 32] using different types of markers. The least percentage of the polymorphic band (38.5%), gene diversity (0.28) and Shannon information index (0.32) were exhibited by pop37 collected from Alaba zone. The genotypes at population level, the diversity estimate of this study showed high level of variation within each population (Mean $H = 0.53$, $I = 0.60$). Diversity estimates from Shannon's informative index (1.548), expected heterozygosity (0.725) and percentages of polymorphic bands (100%) obtained by grouping accessions based on their origin of collection showed that genotypes from Gurage zone was highly diverse. The highest diversity obtained from Gurage zone might be due to the largest number of samples used for this investigation. The lowest mean number of effective alleles, diversity index (1.254), expected heterozygosity (0.628) and percentages of polymorphic bands (92.31) were shown by Alaba zone. The presence of high genetic variation confirms the finding by [2, 33, 34] who reported high genetic variation in Ethiopian hot pepper. Using microsatellite markers, [1] also reported higher level of diversity among accessions collected from different regions of Ethiopia. This showed that the SSR markers were strong enough in revealing the level of variation within and among populations.

The Nei's genetic distance ranging from 0.108 to 0.925 indicated a range from nearly identical genotypes to a diverse relationship. The highest and lowest genetic diversity observed could be due to variation in the collection site, agronomic practice and seed exchange practice. Similar studies on the accessions of hot pepper resulted wide range of similarity coefficients [1]. The highest genetic distance (0.925) was reported between populations from Gurage zone (Mareko wereda) and Alaba zone (Weyradijo wereda), whereas the smallest (0.108) was found from Silte zone (Lanfro and Dalucha weredas). The list genetic distance for the same site may be due to the presence of few seed sources during its introduction. The Nei's genetic distance among accessions pooled by origin of collection were also estimated and the most distantly related populations were between Alaba and Silte zone (0.098), while lowest was revealed between Gurage and Silte zone (0.075). In general, analysis of genetic distance among accessions pooled by origin of collection showed low genetic distance than the previous findings on Ethiopian pepper using SSR marker [1] and ISSR marker [17]. Genetic diversity is strongly associated with life form, geographic range, breeding system, seed dispersal mechanism, and succession status [35]. This finding may imply there is multiple introductions of some very close genotypes, which is likely due to the fact that their parental breeding lines were the same or were very close to each other, bearing in mind that pepper is a self-pollinated crop. Genetic distance is an important parameter for germplasm improvement and conservation of genetic resources. The greater the genetic diversity of germplasm, the greater is the chance of success in breeding desirable traits. Germplasms from the study area can be exploited to do crossing between distantly related populations which may result in vigorous varieties that combine the traits of the distantly related parents.

Analysis of molecular variance (AMOVA) indicated a total of 97% within zones and 3% among zones variation. High within population variation (86.87%) was also reported using ISSR markers [17]. Lower molecular variation among population could be caused by the presence of higher level of gene flow and higher seed exchange rate among the study areas which could lead to the intermix of genotypes [36]. This result contrasts with the finding of [37, 38, 39] who reported presence of genetic variations of *Capsicum* spp. mainly among population. The estimate of population differentiation of hot pepper genotypes using F_{ST} was 0.217 indicating 21.7% of the genetic variation was due to differences between populations; which was nearly similar with the finding by [1], F_{ST} 0.214, and slightly higher than the result of the study by [17], F_{ST} 0.13.

Cluster analysis indicates the genetic relationships between the various breeding materials and provides enough information to allow the selection of genetically diverse and superior parental lines for breeding programs. The dendrogram of the current study revealed three major clusters and there was no clear separation between accessions based on their origin. This underlines the importance of considering genetic variation across genotypes rather than focusing on geographical differences in breeding programs. The present result confirmed that genetic distance is unrelated to geographic distance. Earlier study on hot pepper revealed incomparable result using SSR [1] and ISSR [17], who reported geographical proximity in accession that were clustered together.

Principal coordinate analysis (PCA) as a complementary technique for cluster analysis is one of the important multivariate statistical approaches to group individual samples based on similarity/dissimilarity coefficients. As revealed by the result of the cluster analysis, the PCA provided grouping of the populations regardless of their geographical locations. This could be attributed to the characters

acquired during the domestication process and the presence of related hot pepper productions practice in the region. In both cluster and PCA analysis, genotypes from different collection sites clustered together also indicating there were a significant exchange of plant materials in different hot pepper growing zones of Central Ethiopia. Their similarities may either be caused by close kinship or indicate that they are derivatives from the same cultivar [28]. PCA analysis is useful not only for estimating the genetic diversity of the genotypes but also for selecting crossing parental lines. To produce progeny with a heterosis, varieties with a greater genetic distance are generally recommended as parents since crossing between two closely related individuals results the expression of unexpected recessive alleles [40].

Model-based clustering techniques were applied in STRUCTURE software to identify the admixed and separated genotypes from the existence of population structure [41]. The intermixing distribution and admixture structure of populations on the PCA and population structure were supported by the result of the clustering analysis. The results of the UPGMA dendrogram, PCA and STRUCTURE analysis revealed the presence of high gene flow and intermixing of genotypes regardless of their collection area, with the exception of few genotypes clustered with respect to their origin. The finding contrasted the finding by [17] who reported individual based UPGMA analysis showing clustering of individuals with respect to their populations except few intermixed individuals from other population.

Conclusion and Recommendation

The result of this study has provided that SSR markers detected a considerable genetic diversity in the studied accessions, which could be used in conservation of genetic resources of hot pepper germplasms and crop breeding programs. This study is the first attempt to assess the genetic diversity of *C. annuum* L. from Central Ethiopia Regional state, Ethiopia. However, the current study has a limitation in the number of accessions analysed and agro-ecologies considered. Therefore, genetic diversity of Ethiopian populations from different agro-ecologies should be studied in detail.

Declarations

Conflict of Interest

The authors declare no conflict of interest

Funding

This research was funded by Werabe University

Author Contribution

Conceptualization : Henok Ayelign, Eleni Shiferaw, Ebrahim Sirgaga, Bahiru Wabela
Data curation : Henok Ayelign, Eleni Shiferaw
Investigation : Henok Ayelign, Eleni Shiferaw
Supervision : Eleni Shiferaw
Data analysis and interpretation : Henok Ayelign, Eleni Shiferaw
Writing- original draft : Henok Ayelign
Critical revision : Henok Ayelign, Eleni Shiferaw, Ebrahim Sirgaga, Bahiru Wabela
Final approval : Henok Ayelign, Eleni Shiferaw, Ebrahim Sirgaga, Bahiru Wabela

Acknowledgement

The authors thank Ethiopian Biodiversity Institute, Addis Ababa, Ethiopia, for kindly providing laboratory chemicals and consumables.

Data Availability

The datasets generated during and/or analysed during the current study are available from the corresponding author Henok Ayelign on reasonable request via ayelignhenok2005@gmail.com

References

1. Woubit, D. B., Eleni, S., Fitsum, S., Mekonnen, A., & Shimeles, A.. Assessment of genetic diversity and wilt disease resistance in hot pepper (*Capsicum annuum*) germplasm from Ethiopia. 31–40. <https://doi.org/10.2478/biorc-2021-0006> (2021).
2. Belay, F., Abate, B., & Tsehaye, Y. Genetic diversity studies for morphological traits of hot pepper (*Capsicum annuum* L .) genotypes in Central Zone of Tigray Region , Northern Ethiopia. <https://doi.org/10.5797/AJAR2019.14256> (2019).
3. Habtie, B., Dejen, T., & Dessalegn, Y. Morphological Characterization of Hot Pepper (*Capsicum annuum* . L) Land Races of Ethiopia for Qualitative Characters. **4**(9), 4–9(2017).
4. Getahun, D., & Habtie, B. Growth and Yielding Potential of Hot Pepper Varieties under Rain-Fed Production at Woreta , Northwestern Ethiopia. 2017; **3**(3), 11–18.
5. El-Ghoraba AH, Javedb Q, Anjumb FM, Hamedc SF, Shaabana HA. Pakistani bell pepper (*Capsicum annum* L.): Chemical compositions and its antioxidant activity. *International Journal of Food Properties*. **16**(1),18-32 (2013).
6. Bosland. P. W. and E. J. Votava. Peppers, Vegetables and Spices *Capsicum*. *CABI Publishing*. New York. 198p (2000).
7. EEPA (Ethiopian Export Promotion Agency). Export performance of Agricultural products. Addis Ababa, Ethiopia. (2014).
8. CSA. Agricultural sample survey. Report on Area and Production of Major Crops. Central Statistics Agency, Federal Democratic Republic Ethiopia, *Statistical Bulletin*. Addis Ababa. (2024).
9. CSA. Agricultural Statistics. Addis Ababa. 2022;128. (2022).
10. Rehima Mussema. "Analysis of red pepper marketing: the case of Alaba and Siltie in SNNPRS of Ethiopia." PhD diss., Haramaya University. (2006).
11. Mekdes Dessie, Mufti Jailan, Habtamu Mosi and GetachewMekonnen. Value Chain Analysis of Red Pepper: The Case of Mareko District, Guragie Zone, Southern Ethiopia. *Agri Res & Tech :Open Access J*. **17**(2), 556015. DOI: 10.19080/ARTOAJ.2018.17.556015 (2018).
12. Shimeles Aklilu, Getachew Ayana, Bekele Abebie and Tesfaye Abdissa. Screening for Resistance Sources in Local and Exotic Hot Pepper Genotypes to Fusarium Wilt (*Fusarium oxysporium*) and Associated Quality Traits in Ethiopia. <https://doi.org/10.4172/2329-8863.1000367> (2018).
13. Rao, V. R., and Hodgkin, T..Genetic diversity and conservation and utilization of plant genetic resources. *Plant Cell Tissue Organ Cult*. **68**(1), 1-19 (2002b).
14. Debener, T., and Mattiesch, L. Effective pairwise combination of long primers for RAPD analyses in roses. *Plant Breeding*. **117**(2), 147-151 (1998).
15. Tanksley, S., Young, N., Paterson, A., and Bonierbale, M. RFLP mapping in plant breeding: new tools for an old science. *Nat Biotechnol*. **7**(3), 257-264 (1989).
16. Ahmed Beyan, Haji Jema and Endrias Geta. Analysis of farm households' technical efficiency in production of smallholder farmers: the case of Girawa District, Ethiopia. *Journal of Agriculture and Environmental Science*. **13** no. 12, 1615-1621 (2013).
17. Alayachew, S. A., Atnafu, D. M., & Gedefa, S. A. Research Article Genetic Diversity Study of Ethiopian Hot Pepper Cultivars (*Capsicum* species) Using Inter Simple Sequence Repeat (ISSR) Marker. <https://doi.org/10.3923/ijar.65.73> (2018).
18. Doyle J. J. & Doyle J. L. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*. **19**, 11-15 (1987).
19. Nicolai M., Cantet M., Lefebvre V.M., Sage-Palloix A. M. & Palloix A. Genotyping a large collection of pepper (*Capsicum* spp.) with SSR loci brings new evidence for the wild origin of cultivated *C. annuum* and the structuring of genetic diversity by human selection of cultivar types. *Genet Resour Crop Ev*. **60**(8), 2375-2390. <https://doi.org/10.1007/s10722-013-0006-0> (2013).
20. Huang L, Deng X, Li R, Xia Y, Bai G, Siddique KHM, *et al*. A fast silver staining protocol enabling simple and efficient detection of ssr markers using a non-denaturing polyacrylamide gel. *Journal of Visualized Experiments* [Internet]. **2018**(134), 1–7. Available from: <https://doi.org/10.3791/57192> PMID: 29733306 (2019).
21. Liu K. & Muse S. V. Power Marker: an integrated analysis environment for genetic Marker analysis. *Bioinfo*. **21**, 2128-2129. <http://dx.doi.org/10.1093/bioinformatics/bti282> (2005).
22. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution* [Internet]. (2016).
23. Peakall R. & Smouse P. E. GenAEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*. **28**, 2537-2579. DOI: 10.1093/bioinformatics/bts460 (2012).

24. Earl DA, VonHoldt BM. STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* [Internet]. **4**(2), 359–61. Available from: <https://doi.org/10.1007/s12686-011-9548-7> (2012).
25. Jakobsson M, Rosenberg NA. CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* [Internet]. **23**(14), 1801–6. Available from :<https://doi.org/10.1093/bioinformatics/btm233> PMID: 17485429 (2007).
26. Guzmab, F.A., Moore S., Vicente M,C, & Jahn M,M. Microsatellites to enhance characterization, conservation and breeding value of Capsicum germplasm. *Genet Resour Crop Ev.* **67**, 569-585 (2019).
27. Botstein D., White R. L., Skolnick M. and Davis R. W. Construction of a genetic linkage map in man using restriction fragment length polymorphism. *Am J Hum Genet.* **2**, 314-331 (1980).
28. Xiao-min, Z., Zheng-hai, Z., Xiao-zhen, G. U., Sheng-li, M. A. O., Xi-xiang, L. I., & Chadœuf, J. Genetic diversity of pepper (Capsicum spp .) germplasm resources in China reflects selection for cultivar types and spatial distribution. **15**(9), 1991–2001. [https://doi.org/10.1016/S2095-3119\(16\)61364-3](https://doi.org/10.1016/S2095-3119(16)61364-3) (2016).
29. Ballina-Gómez, H., Latournerie-Moreno, L., Ruiz Sánchez, E., Pérez-Gutiérrez, A., Rosado-Lugo, G. Morphological characterization of *Capsicum annuum* L. accessions from southern Mexico and their response to the Bemisia tabaci-Begomovirus complex. *Chilean Journal of Agricultural Research.* **73**(4), 330-338 (2013).
30. Dias, G.B., Gomes, V.M., Moraes, T.M.S., Zottich, U.P., Rabelo, G.R., Carvalho, A.O., Moulin, M., Gonçalves, L.S.A., Rodrigues, R., Cunha, M.D. Characterization of Capsicum species using anatomical and molecular data. *Genetics and Molecular Research.* **12** (4), 6488-6501 (2013).
31. Thul, S.T., Darokar, M.P., Shaseny, A.K., Khanuja, S.P.S. Molecular profiling of genetic variability in Capsicum species based on ISSR and RAPD markers. *Mol. Biotechnol.* **10**, 1007/s12033-011-9446 (2011).
32. Tanksley, S.D. High rates of cross-pollination in chill pepper. *Hortscience.* **19**, 580–582 (1984).
33. Aklilu S., Abebie B., Wegarya D. and Tekleword A. Analysis of morphological diversity among hot pepper (*Capsicum annuum* L.) collections in the Rift Valley area of Ethiopia. *Trop Agr.* **93** (3), 152-164 (2016).
34. Solomon A. M., Han K., Lee J, H., Jang S. and Kang B., C. Genetic diversity and population structure of Ethiopian Capsicum germplasms. *PLoS ONE.* **14**(5), 1-20. <https://doi.org/10.1371/journal.pone.0216886> (2019).
35. Nybom H. Comparison of different nuclear DNA markers for estimating intra-specific genetic diversity in plants. *Mol Ecol.* **13**, 1143–1155 (2004).
36. Kapital, B., Feyissa, T., Petros, Y., & Mohammed, S. Molecular diversity study of black cumin (*Nigella sativa* L .) from Ethiopia as revealed by inter simple sequence repeat (ISSR) markers. **14**(18), 1543–1551. <https://doi.org/10.5797/AJB2015.14567> (2015).
37. Ilbi, H.. RAPD markers assisted varietal identification and genetic purity test in pepper, *Capsicum annuum*. *Scientia Hort.* **97**, 211-218 (2003).
38. Yang, R., Kong, J., Wu, X., Deng, Z., Chen, Q., Liu, W. Application of ISSR markers in genetic polymorphism of *Capsicum frutescens* L. *Journal of Shanghai University* (Natural Science Edition). pp. 4-20 (2005).
39. Lijun, O., Xuexiao, Z. Inter simple sequence repeat analysis of genetic diversity of five cultivated pepper species. *Afr. J. Biotechnol.* **11**, 752-757 (2012).
40. Nugroho, K., Cimanggu, I., Nugroho, K., Terryana, R. T., Manzila, I., Priyatno, T. P., & Lestari, P. The Use of Molecular Markers to Analyze the Genetic Diversity of Indonesian Pepper (*Capsicum* spp.) Varieties Based on Anthracnose Resistance. **23**(3). <https://doi.org/10.7454/mss.v23i3.11261> (2019).
41. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics* [Internet]. **155**(2), 945–59. Available from: <https://doi.org/10.1093/genetics/155.2.945> PMID: 10835412 (2000).

Figures



Page 14/15

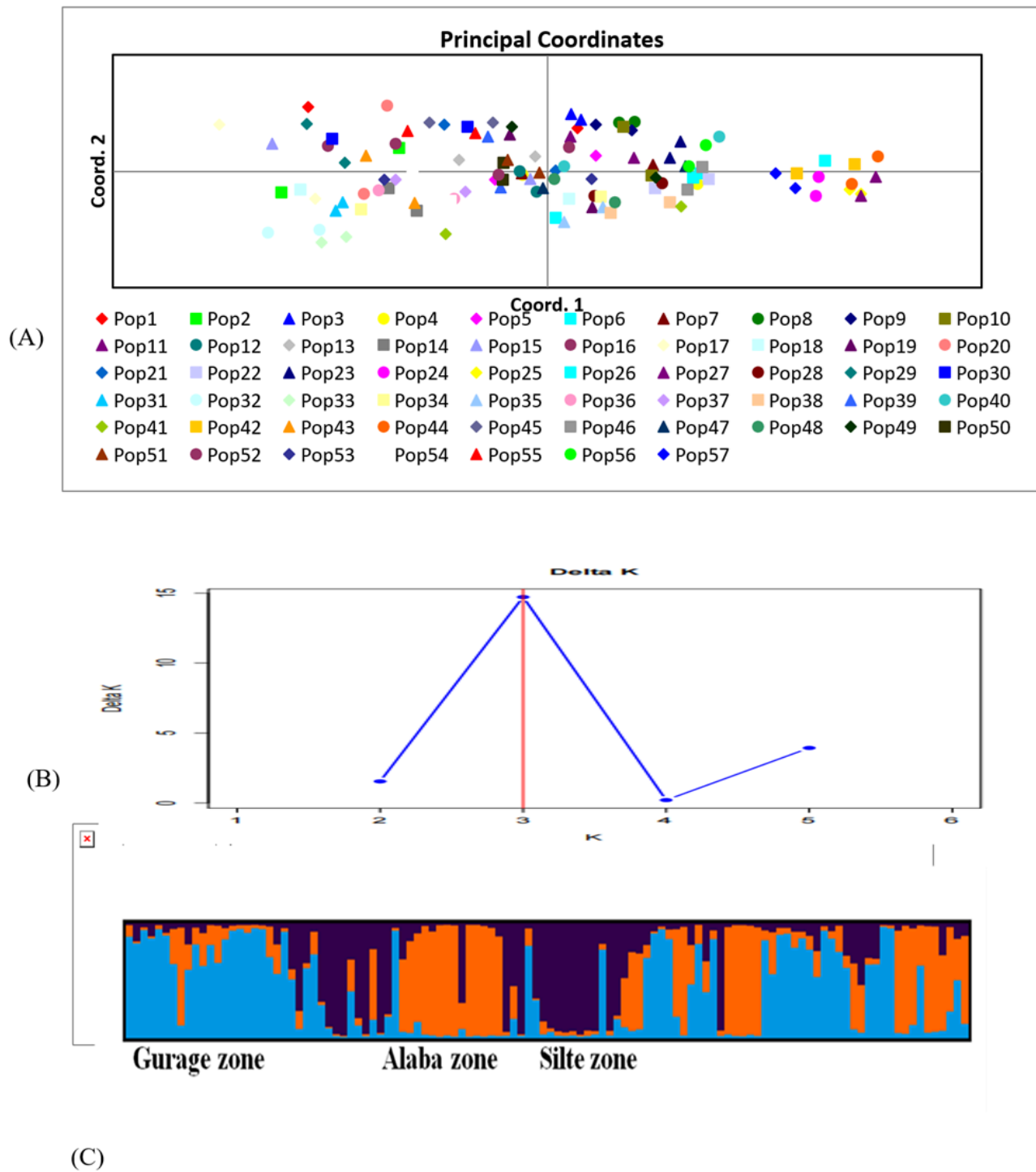


Figure 2

(A) Principal coordinate analysis of the 57 *Capsicum annuum* genotypes; (B) Delta K value for determining the optimal number of subpopulations; (C) Population structure of 57 *Capsicum annuum* genotypes representing three zones at K=3.