Vol. 16(9), pp. 258-269, October 2022
DOI: 10.5897/AJPS2022.2268
Article Number: 81D98F269863
ISSN 1996-0824

http://www.academicjournals.org/AJPS

ACADEMIC JOURNALS
expand your knowledge

African Journal of Plant Science

Full Length Research Paper

Molecular characterization of banana genotypes by SSR markers

Solomon Tamiru Workneh[1*], Sisay Kidane Alemu[1], Gamachu Olani[2], Abel Debebe[2], Biruktait Berhanu[1], Asmare Dagnew[2] and Wegayehu Assefa[2]

[1]Ethiopian Institute of Agricultural Research (EIAR), National Agricultural Biotechnology Research Center, Holeta, Ethiopia.
[2]Ethiopian Institute of Agricultural Research (EIAR), Melkasa Agricultural Research Center, Adama, Ethiopia.

Received 20 June, 2022; Accepted 13 October, 2022

Banana and plantain are monocotyledonous herbs in the genus Musa (Musa spp.) and grown globally due to their diverse uses as export income, food and fiber. Banana is also grown widely in Ethiopia; however, the crop attracted very limited research attention and has little genetic information available. Therefore, the main objective of this study was to assess the genetic diversity and population structure of banana genotypes using 14 SSR markers. The genetic diversity of 96 banana genotypes obtained from Melkasa Agricultural Research center was analyzed using 14 SSR markers. A total of 187 alleles were identified, and the number of alleles per marker (Na) ranged from 6 to 21 with an average of 13.36. The range of polymorphic information content per marker was 0.52 to 0.93, with an average of 0.82. The results of phylogenetic analysis, principal coordinate analysis, and structure analysis showed an admixture of the populations indicating that the genotype grouping pattern did not exactly correspond to the genotypes' breeding history and genome composition. However, Clusters I to III from the phylogenetic analysis and K=3 from the population structure analysis confirmed the existence of 3 major groups among the genotypes as a whole. Analysis of molecular variance (AMOVA) revealed the presence of higher genetic variation within the population than between the population. Generally, genetic diversity and population structure obtained from this study provide inputs for the improvement of the crop.

Key words: Alleles, banana, genetic diversity, *Musa* spp., polymorphic information content, SSR markers.

INTRODUCTION

Banana and plantains are monocotyledonous herbs in the genus Musa (*Musa spp*.) which originated in Southeast Asia and the western Pacific (Langhe et al., 2009). They are year-round, perennial fruit crops with a rapid rate of growth that are commonly grown throughout the world's tropics and subtropics. After cassava, maize, and yams, bananas and plantains are rated the fourth most important crop in sub-Saharan Africa (SSA) (Igwe et al., 2022). The crop makes a considerable contribution to the export revenue and food security of these regions,

*Corresponding author. E-mail: soltamiru24@gmail.com.

and other banana plant parts are used locally for food and fiber. The annual global banana production was 116million tons in 2017 to 2019, with an estimated value of USD 31 billion (FAO, 2020).

Different genotypes, including diploids (AA, AB and BB), triploids (AAA, AAB and ABB), and tetraploids (AAAA, AAAB, AABB and ABBB), were generated from *Musa acuminata* (AA) and *Musa balbisiana* (BB). The different ploidy levels and chromosome numbers of bananas have led to complexity in taxonomy and are associated with parthenocarpy, leading to female sterility, seedless fruit, and non-viable seeds (De Jesus et al., 2013). The majority of edible genotypes grown by farmers worldwide are triploids (2n = 3x = 33 chromosomes), and they are categorized into three main groups: AAA (Cavendish or dessert bananas), ABB (cooking bananas), and AAB (plantain). The plantain subgroup (AAB), is primarily cultivated by small farmers for local use and is crucial to the agriculture of tropical humid forest regions in Africa, Central and South America, and Asia (Noyer et al., 2005; Okeh et al., 2022).

Ethiopia, which is located in the tropics, has great potential for banana production (Gebre et al., 2022). The most widely grown and consumed fruit in the country is the Cavendish or dessert banana. It has a significant socio-economic role in the general well-being of rural communities, including food security and income generation, especially in the south and southwest of the country (Gebre et al., 2022). Banana production contributes around 47.83% of producers' consumption, specifically, about 49.19% for income generation, 0.47% for animal feed and 2.52% for other purposes (CSA, 2015; Gebre et al., 2022). It covers about 59.64% (53,956.16 ha) of the total fruit area, about 68.00% (478,251.04 tones) of the total fruits produced, and about 38.30% (2,574,035) of the total fruit-producing farmers in the country (CSA, 2015; Gebre et al., 2022). However, the production of this crop has become susceptible to both abiotic and biotic factors depending on their genomic constitutions. It has been discovered that genotypes with the "B" genome (specifically the ABB type) are more resistant to abiotic and biotic stress than those with only the "A" genome (Igwe et al., 2022). This imposes the need for genotype testing and crop improvement supported by the conservation of characterized banana genotypes. The global population is rapidly increasing, with over 9 billion people predicted by 2050 (Ehrlich and Harte, 2015). Feeding this excessively growing population is generating a lot of pressure on agricultural crop production. Therefore, knowing the plant's genetic background is crucial for developing high-yielding and pest resistance genotypes. Assessing variability among available genotypes is a crucial task for identifying variation with useful traits for crop improvement and conservation programs. These characterizations can be achieved using conventional methods based on morphological traits, and the use of molecular markers. However, molecular methods provide precise genetic information based on the advancement and use of technologies used to identify genetic variability present in the genotypes, offering information about the variability found at the DNA level and enabling the genetic differentiation among individuals (Christelová et al., 2017). Due to their multi-allelic nature, high reproducibility, and co-dominant inheritance, microsatellites or simple sequence repeats (SSRs) are among the most useful PCR-based DNA markers in population genetics and germplasm characterization studies (Powell et al., 1996).

So far, several scientific research have been conducted on genetic diversity and genotype selections of bananas in various countries (Opara et al., 2010; De Jesus et al., 2013; Christelová et al., 2017; Paofa et al., 2018; Marian et al., 2018; Biswas et al., 2020). However, in Ethiopia, the crop attracted very limited research attention and there is little genetic information available. Therefore, the purpose of this study was to assess the genetic diversity and population structure analysis of banana genotypes using SSR markers and identify diverse genotypes potentially useful for future breeding programs.

MATERIALS AND METHODS

Plant materials and DNA extraction

A total of 96 genotypes obtained from Melkasa Agricultural Research Center (MARC) including improved variety, advanced clones, and landrace were used in this experiment (Table 1). Young Leaves were taken from vegetatively propagated plants of each 96 individuals maintained at MARC. The collected fresh leaves were placed in 2 ml autoclaved and labeled Eppendorf tubes and freezed for 24 h at -20°C. After 24 h the leaves were further dried in a freeze dryer (John Morris group) for 24 h and then grounded using Geno Grinder (MM-200, Retsch) for 3 min. Genomic DNA was extracted in two replications using plant DNA extraction protocol based on the method of Diversity Array Technology (DArT, 2019) with some minor modifications. The qualification and quantification of genomic DNA were done using a Nanodrop spectrophotometer (ND-8000, Thermo Scientific). The level of quality was further assessed by running the genomic DNA on 1% agarose gel in TAE buffer using a standard lambda DNA as a reference band. A gel documentation system (Bio Doc-IT Imaging system) was used to visualize the DNA bands under UV light (Cambridge, UK). Samples with high band intensity, purity, and lesser smear were selected and normalized to 50 ng/µl for further PCR analysis.

Primer selection and PCR optimization

A total of 20 lyophilized primer pairs were reconstituted using nuclease-free water to obtain 100 µM stock solutions. All primers were diluted to a working concentration of 10 µM and finally stored at 20°C. Applicability of each primer was optimized using the "gradient pcr" methodology. PCR based on reliability, polymorphism, and specificity to the target region, 14 SSR markers were selected (Table 2) for final analysis out of the 20 tested primers.

Table 1. List of 96 banana genotypes used in this study.

No.	Genotype	Level of advancement	Genome group	Banana type
1	Americani	Improved variety	AAA	Dessert banana
2	Dinke-2	Landrace accession	AAA	Dessert banana
3	Dwarf Cavendish	Improved variety	AAA	Dessert banana
4	Green Red	Improved variety	AAA	Dessert banana
5	Poyo	Improved variety	AAA	Dessert banana
6	Williams-1	Improved variety	AAA	Dessert banana
7	Ducasse Hybrid	Improved variety	ABB	Cooking banana
8	Butuza	Improved variety	AAA	Dessert banana
9	Robusta	Improved variety	AAA	Dessert banana
10	Grande Naine	Improved variety	AAA	Dessert banana
11	Giant Cavendish	Improved variety	AAA	Dessert banana
12	Red	Improved variety	AAA	Dessert banana
13	Silk	Improved variety	AAB	Plantain banana
14	Wondo genet-2	Landrace accession	AAA	Dessert banana
15	Matoke	Improved variety	AAA	Dessert banana
16	Wondo genet-1	Landrace accession	AAA	Dessert banana
17	Kitawira	Improved variety	AAA	Dessert banana
18	Kibungo-1	Improved variety	----	Unknown
19	Cardaba	Improved variety	ABB	Cooking banana
20	Saba	Improved variety	ABB	Cooking banana
21	Bluggoe	Improved variety	ABB	Cooking banana
22	Pelipita	Improved variety	ABB	Cooking banana
23	Prata	Improved variety	AAB	Plantain banana
24	Cubian Red	Advanced clones	AAA	Dessert banana
25	Gros Michel	Advanced clones	AAA	Dessert banana
26	Giant Parfitt (ITC1246)	Advanced clones	------	Unknown
27	NamO ITC1303	Advanced clones	-------	Unknown
28	NtebwaO ITC1461	Advanced clones	------	Unknown
29	Chinese Cavendish ITC0547	Advanced clones	AAA	Dessert banana
30	Ice cream	Advanced clones	ABB	Cooking banana
31	Thai (Aka kluay Khay)	Advanced clones	------	Unknown
32	Pisan Umbuk ITC30686	Advanced clones	------	Unknown
33	FHIA-25 ITC1418	Advanced clones	AAB	Plantain banana
34	SuuO ITC1462	Advanced clones	-----	Unknown
35	Nante ITC1353	Advanced clones	-----	Unknown
36	Veimama ITC0576	Advanced clones	-----	Unknown
37	FHIA#18 hybrids	Advanced clones	AAAB	Unknown
38	Lakika	Advanced clones	-----	Unknown
39	Ibwi ITC1465	Advanced clones	-----	Unknown
40	Plantain or Cooking made size	Advanced clones	AAB	Plantain banana
41	FHIA-23 ITC1265	Advanced clones	AAAA	Unknown
42	Fai palagi ITC1059	Advanced clones	----	Unknown
43	Naine de China ITC0178	Advanced clones	----	Unknown
44	Suce (French for Sugar)	Advanced clones	AA	Unknown
45	Champa Nasik	Improved variety	----	Unknown
46	Meraro	Improved variety	AAA	Dessert banana
47	Ikimaga	Improved variety	AAA	Dessert banana
48	Borocemsa	Improved variety	ABB	Cooking banana
49	Ginir-1	Landrace accession	AAA	Dessert banana
50	Nijuru	Improved variety	AAA	Dessert banana
51	Kitawira	Improved variety	AAA	Dessert banana

Table 1. Contd.

52	Chibul Angombe	Improved variety	AAA	Dessert banana
53	Matoke	Improved variety	AAA	Dessert banana
54	Ghana Cooking	Advanced clones	AAB	Plantain banana
55	Pisang Raja	Improved variety	AAB	Plantain banana
56	Ambowoha Selle-1	Landrace accession	----	Unknown
57	Giant Cavendish	Improved variety	AAA	Dessert banana
58	Grande Naine	Improved variety	AAA	Dessert banana
59	Robusta	Improved variety	AAA	Dessert banana
60	Ambowoha Selle-2	Landrace accession	AAA	Dessert banana
61	Williams-2	Improved variety	AAA	Dessert banana
62	Dinke-1	Improved variety	AAA	Dessert banana
63	Poyo	Improved variety	AAA	Dessert banana
64	Kamara Masenge	Improved variety	AAB	Plantain banana
65	Dwarf Cavendish	Improved variety	AAA	Dessert banana
66	Chinese Dwarf	Improved variety	AAA	Dessert banana
67	Horn	Improved variety	AAB	Plantain banana
68	Ambo-3	Landrace accession	AAA	Dessert banana
69	Lady Finger	Improved variety	AAB	Plantain banana
70	Paracido Alrey	Improved variety	----	Unknown
71	Ambo-2	Landrace accession	AAA	Dessert banana
72	Ambowoha Selle-3	Landrace accession	AAA	Dessert banana
73	Williams-1	Improved variety	AAA	Dessert banana
74	Ducasse Hybrid	Improved variety	ABB	Cooking banana
75	Butuza	Improved variety	AAA	Dessert banana
76	Williams Hybrid	Improved variety	AAA	Dessert banana
77	Giner-2	Landrace accession	AAA	Dessert banana
78	Lacatan	Improved variety	AAA	Dessert banana
79	Uganda Red	Improved variety	----	Unknown
80	Pisang sri	Improved variety	----	Unknown
81	Gittity	Improved variety	----	Unknown
82	Wondo gent-3	Landrace accession	ABB	Cooking banana
83	Wondo gent-4	Landrace accession	AAA	Dessert banana
84	Nijuru	Improved variety	AAA	Dessert banana
85	Imbogo	Improved variety	----	Unknown
86	Cardaba	Improved variety	ABB	Cooking banana
87	Kenya-1	Landrace accession	AAA	Dessert banana
88	Cachaco	Improved variety	ABB	Dessert banana
89	Bodles Altafort	Improved variety	AAAA	Unknown
90	Suce Sugar	Advanced clones	AA	Unknown
91	Dwarf parfitt ITC0548	Advanced clones	----	Unknown
92	FHIA-17 ITC1264	Advanced clones	AAAA	Unknown
93	Ntindii ITC1464	Advanced clones	-----	Unknown
94	Cuban Yellow	Advanced clones	-----	Unknown
95	Cocos ITC0451	Advanced clones	-----	Unknown
96	Kitarasa ITC1451	Advanced clones	------	Unknown

Source: Authors

Genotyping (whole sample amplification)

The GeneAmp®PCR System 9700 thermal cycler was used to conduct the PCR reaction in a total of 12.5 µl volume. Briefly, the recipe contains 6.25 µl one Taq 2x Master Mix (M04821) Biolabs England, with standard buffer (which contains all PCR reaction components, $MgCl_2$, PCR buffer, dNTPs and Taq DNA polymerase), 0.5 µl forward primer, 0.5 µl reverse primer, 0.25 µl DMSO, 3 µl nuclease-free water and 2 µl genomic DNA. The PCR was programmed with an initial denaturation phase of 3 min at

Table 2. List of SSR markers, primer sequences, Repeat motives and their annealing temperatures.

Marker	Primer's sequence (5' - 3')	Repeat motives	Annealing temp. (°C)	Expected size	Reference
AGMI101/102	F: TGCAGTTGACAAACCCCACACA R: TTGGGAAGGAAAATAAGAAGATAGA	(GA)3	54.6	190-250*	Kaemmer et al., 1997
Ma1/24	F: GAGCCCATTAAGCTGAACA R: CCGACAGTCAACATACAATACA	(CT)13	56.7	170-250*	Crouch et al., 1998
Ma1/27	F: TGAATCCCAAGTTTGGTCAAG R: CAAAACACTGTCCCCATCTC	(GA)9	51.7	120-185*	Crouch et al., 1998
Ma3/103	F: TCGCCTCTCTTTAGCTCTG R: TGTTGGAGGATCTGAGATTG	(CT)10	56.7	135-165*	Crouch et al., 1998
AGMI103/104	F: ACAGAATCGCTAACCCTAATCCTCA R: CCCTTTGCGTGCCCCTAA	(GAGAAA)3GATGA(GAA)2	57.9	150-260	Kaemmer et al., 1997
AGMI95/96	F: ACTTATTCCCCCGCACTCAA R: ACTCTCGCCCATCTTCATCC	(TC)6N24 (TC)7	58.8	263-275	Kaemmer et al., 1997
AGMI93/94	F: AACAACTAGGATGGTAATGTGTGGAA R: GATCTGAGGATGGTTCTGTTGGAGTG	(GA)13	58	152-176	Kaemmer et al., 1997
AGMI187/188	F: GCAACTTTGGCAGCATTTT R: TGATGGACTCATGTGTACCTACTAT	(CT)12	50.8	303-306	Kaemmer et al., 1997
AGMI105/108	F: TCCCAACCCCTGCAACCACT R: ATGACCTGTCGAACATCCTTT	(GA)16N76	54.6	285-299	Kaemmer et al., 1997
Ma1/17	F: AGGCGGGGAATCGGTAGA R: GGCGGGAGACAGATGGAGT	(GA)14	58	280-375*	Crouch et al., 1998
MaOCEN13	F: GCTGCTATTTTGTCCTTGGTG R: CTTGATGCTGGGATTCTGG	(TC)16	50.2	141-200	Creste et al., 2006
MaOCEN1	F: TCTCAGGAAGGGCAACAATC R: GGACCAAAGGGAAAGAAACC	(CT)17	58	210-250	Creste et al., 2006
MaOCEN3	F: GGAGGAAATGGAGGTCAACA R: TTCGGGATAGGAGGAGGAG	(GA)10	60	180-250	Creste et al., 2006
AGMI125/126	F: TTAAAGGTGGGTTAGCATTAGG R: TTTGATGTCACAATGGTGTTCC	(GA)20	57.2	360-372	Kaemmer et al., 1997

*Those fragment size information are obtained from the present study while the rest are from literature sources.
Source: Authors

94°C, followed by 35 cycles of denaturation at 94°C for 1 min, annealing at 50.2-60°C depending on the primers for 1 min, and elongation at 72°C for 2 min, with a final elongation at 72°C for 10 min, and a holding step at 4°C.

PCR products were loaded on a 3% percent agarose gel (w/v) using a 6x loading dye that contained gel red. Electrophoresis was carried out in 1 TAE buffer, for three h at 100 constant volts. The gel was stained with gel red and a BioDoc-it imaging device (Cambridge, UK) was used to

Table 3. Summary of various diversity parameters for 96 banana genotypes using 14 SSR markers.

Marker	MAF	GD	PIC	Na	I	Ne	Ho	He	PHWE	F
AGMI101/102	0.37	0.77	0.74	6.00	1.471	4.003	0.000	0.749	0.000***	1.000
Ma1/24	0.30	0.86	0.85	14.00	1.841	5.622	0.022	0.791	0.000***	0.970
Ma1/27	0.27	0.84	0.82	11.00	1.722	4.926	0.067	0.783	0.000***	0.919
Ma3/Ma3	0.23	0.83	0.80	7.00	1.453	3.811	0.000	0.727	0.000***	1.000
AGMI103/104	0.55	0.59	0.52	6.00	0.995	2.304	0.006	0.561	0.000***	0.990
AGMI95/96	0.15	0.91	0.90	15.00	2.202	8.053	0.000	0.872	0.000***	1.000
AGMI93/94	0.16	0.89	0.88	13.00	1.995	6.607	0.000	0.837	0.000***	1.000
AGMI187/188	0.21	0.91	0.90	19.00	2.237	8.134	0.000	0.870	0.000***	1.000
AGMI105/108	0.33	0.84	0.83	17.00	2.022	5.906	0.000	0.826	0.000***	1.000
Ma1/17	0.35	0.79	0.77	9.00	1.563	4.255	0.000	0.739	0.000***	1.000
MaOCEN13	0.36	0.81	0.80	13.00	1.844	5.144	0.000	0.800	0.000***	1.000
MaOCEN1F	0.19	0.91	0.90	19.00	2.161	7.392	0.000	0.856	0.000***	1.000
MaOCEN3	0.34	0.84	0.82	17.00	1.843	5.122	0.035	0.801	0.000***	0.956
AGMI125/126	0.12	0.93	0.93	21.00	2.341	9.236	0.035	0.887	0.000***	0.960
Mean	0.28	0.84	0.82	13.36	1.835	5.751	0.012	0.793	-	0.985

Where MAF = Major allele frequency, GD= Gene diversity, PIC = Polymorphic information content, Na= Number of alleles, I = Shannon's Information Index, Ne = Effective number of alleles, Ho = Observed heterozygosity, He = Expected heterozygosity, F = Fixation Index, PHWE = *P*-value for deviation from Hardy Weinberg Equilibrium, ns = not significant, *** = $P < 0.0001$ and hence highly significant.
Source: Authors

view the gel image under UV light. A 100 base pair SMO BIO DNA ladder was used as a molecular ruler to compare DNA bands and estimate the sizes of the DNA fragments.

Data scoring and analysis

The PyElph 1.4 image analyzer software tool was used to score the molecular weights of clearly resolved bands amplified by the SSR markers (Pavel and Vasile, 2012). The software power marker ver. 3.5 was used to calculate the number of alleles (Na), major allele frequency (MAF), and the polymorphic information content (PIC) (Liu and Muse, 2005). The number of effective alleles (Ne), observed (Ho), and expected (He) heterozygosity, Shannon's information index (I), gene flow (Nm), Wright's fixation index (FST, FIS, FIT) were computed with GenAlEx ver. 6.502 software (White and Peakall, 2015). Cluster analysis was performed using DARwin software version 6.0 (Perrier and Jacque moud-Collet, 2006).

A dendrogram was constructed using the dissimilarity matrix as input data to visualize the pattern of a cluster both within and between genotypes.

Patterns of genotype clustering were examined using the Unweighted Neighbor-Joining (NJ) method (Saitou and Nei, 1987). Principal coordinate analysis (PCoA) was carried out using DARwin software version 6.0 to examine the distribution of variance among samples and determine the power of ordination. (Perrier and Jean-Claude Colllet, 2006)

The population genetic structure was computed using the Bayesian statistical model calculated using the Structure software version 2.3.4 (Pritchard et al., 2000). A burn-in period of 100,000 and a run length of 200,000 for the Monte Carlo Markov chain (MCMC) with 20 iterations for each K were used to calculate the best likely number of population clusters (K-value). The web-based structure harvester version 0.6.92 (Earl and Von Holdt, 2012) was used to identify best probable K value according to Evanno et al. (2005). The labeling of each population cluster based on K value was determined using Clumpak beta (Kopelman et al., 2015). GenAlEx version 6.503 software (Peakall and Smouse, 2006) was used to calculate the analysis of molecular variance (AMOVA) after the result of population structure analysis based on the value of K obtained.

RESULTS AND DISCUSSION

Gene diversity indices and marker polymorphism

The 14 SSR markers used to assess the genetic diversity of banana genotypes detected a total of 187 alleles with a mean of 13.36 alleles per locus (Table 3). The number of alleles per marker ranged from 6 to 21 with an average of 13.36 alleles. The lowest and the highest numbers of effective alleles were observed for markers AGMI103/104 (2.304) and AGMI125/126 (9.236) in the same order. The lowest value of Shannon's information (I) index was recorded for marker AGMI103/104 (0.99) while AGMI125/126 exhibited the highest value (2.341). The mean expected heterozygosity was 0.793 with a minimum of 0.561(AGMI103/104) and a maximum of 0.887 (AGMI125/126). The observed heterozygosity has a range of 0.0 to 0.067, with a mean of 0.012. The current study found more alleles than the report by Jesus et al. (2013), which found 182 alleles in total, ranging in number from 7 to 15. On the other hand, Christelová et al. (2017) reported a higher mean number of alleles (21.5) for 695 Musa accessions using 19 SSR markers. A relatively higher number of alleles were observed in the present study compared to the previous reports indicating the existence of high genetic variation among the genotypes.

The markers in the current study showed higher

polymorphic information content (PIC), with values ranging from 0.52 (AGMI103/104) to 0.93 (AGMI125/126) and an average value of 0.82. The 14 SSR markers used in this study revealed an average PIC value greater than 0.5 indicating a high discriminating power of the analyzed loci. According to Botstein et al. (1980), a PIC value greater than 0.5 implies a locus with a high degree of polymorphism, whereas a value less than 0.25 indicates a locus with a lower degree of polymorphism. Christelová et al. (2017) found a similar result for 695 accessions using 19 SSR markers, with a PIC value ranging from 0.561 to 0.933 and a mean value of 0.789. Similarly, Changadeya et al. (2012), reported an average PIC value of 0.74 which is also lower than the result obtained in this study. The mean expected and observed heterozygosity, and effective alleles obtained in this study differed from the previous reports using SSR markers on bananas (Changadeya et al., 2012; De Jesus et al., 2013; Marian et al., 2018).

The difference in the number and type of genotypes used as well as the number and type of markers may be contributing factors to the variation in the level of allele number and heterozygosity. The study's findings regarding the number of alleles, the effective number of alleles, polymorphic information content (PIC), and expected heterozygosity showed that there was significant genetic diversity present as well as the high potential of the markers to be used in analyzing the genetic diversity of banana genotypes. For all markers, the observed heterozygosity showed low values to the expected heterozygosity, indicating a high level of homozygosity. The observed lower heterozygosity in banana genotypes might possibly be a result of minimal outcrossing.

The average gene diversity (0.84) and expected heterozygosity (0.79) detected among the 96 banana genotypes revealed high levels of variation within the studied banana genotypes.

Phylogenetic relationship

Phylogenetic analysis based on Unweighted Neighbor-joining (NJ) methods grouped the 96 banana genotypes into three major groups, by forming different hierarchical sub-groups (Figure 1). Clusters I, II, and III each included 54% (52 genotypes), 24% (23 genotypes), and 22% (21 genotypes) respectively, out of the total genotypes. The majority of genotypes from improved variety were found in the first cluster (37 improved out of 52 genotypes). In Cluster II, majority of the genotypes were from advanced clones. Cluster III mainly consisted of advanced clones and landrace accessions. The current study revealed that genotype grouping patterns are unrelated to breeding history (improved variety, advanced clones, and landrace). The genotypes were mixed up in all the clusters except for a few cases where clusters were dominated by genotypes of a particular genome composition. The lack of clear clustering among banana genotypes based on breeding history could be attributable to a similar genetic basis that unites all genotypes. A high level of genetic similarity is expected among genotypes from the same sub-cluster because of common ancestry or mutations (Creste et al., 2003). Variations within each cluster in vegetatively propagated crops such as bananas are mainly determined by genotype and genome change caused by mutations whose frequency is determined by how many times a clone has been multiplied and planted (Changadeya et al., 2012).

Further analysis of the 96 genotypes of bananas resolved them into various distinct coordinates based on dessert, cooking, and plantains bananas. Most dessert bananas (AAA genome composition), almost all cooking bananas (ABB genome composition), and almost all plantain bananas (AAB genome composition) formed distinct sub-clusters based on genomic composition within the main cluster (Figure 2). Eight of the 9 coocking type genotypes (cardaba, Saba, Bluggoe, pelipita, Ice cream, Borocemsa, Wondo gent-3, and Cachaco), co-clustered together based on their genomic constitutions, while Ducasse Hybrid fell in a separate sub-cluster with other genotypes of AAA genome. Similar sub-clustering was noted in plantain genotypes such as silk, FHI A-25 ITC1418, plantain or cooking made size, Ghana cooking, Pisang Raja, and Horn that have AAB genomes. Three plantain banana genotypes such as Parata, Kamara Masenge and Lady Finger which have AAB genomes did not cluster with other known AAB genotypes. Similarly, most dessert banana genotypes in which the AAA genome is grouped based on their genomic constitutions. Some genotypes with unknown genomic constitutions grouped closely with the AAA genome, implying that they are members of the AAA genomic group. Furthermore, some genomic groups were successfully resolved, while others, such as those with mixed ploidy groups, were clustered together based on their genetic closeness to their progenitors, *M. acumminata* (A genome), and *M. balbisiana* (B genome). The ABB genomes for example Parata, Kamara Masenge, and Lady Finger were closely grouped, with the same relatedness found between AAA genome groups. The current findings agrees with those of Okeh et al. (2022) who used ISSR and SCoT markers, and Igwe et al. (2021) who used CDDP markers and found genotype clustering mainly based on the genomic constitution.

Principal coordinate analysis (PCoA)

The principal coordinate analysis (PCoA) generated by the genetic dissimilarity matrix explained 22.23% of the total genetic variation. The first, second, and third axis explained 8.08, 7.61 and 6.54% of the genetic variation, respectively. The two-dimensional plot of PCoA analysis

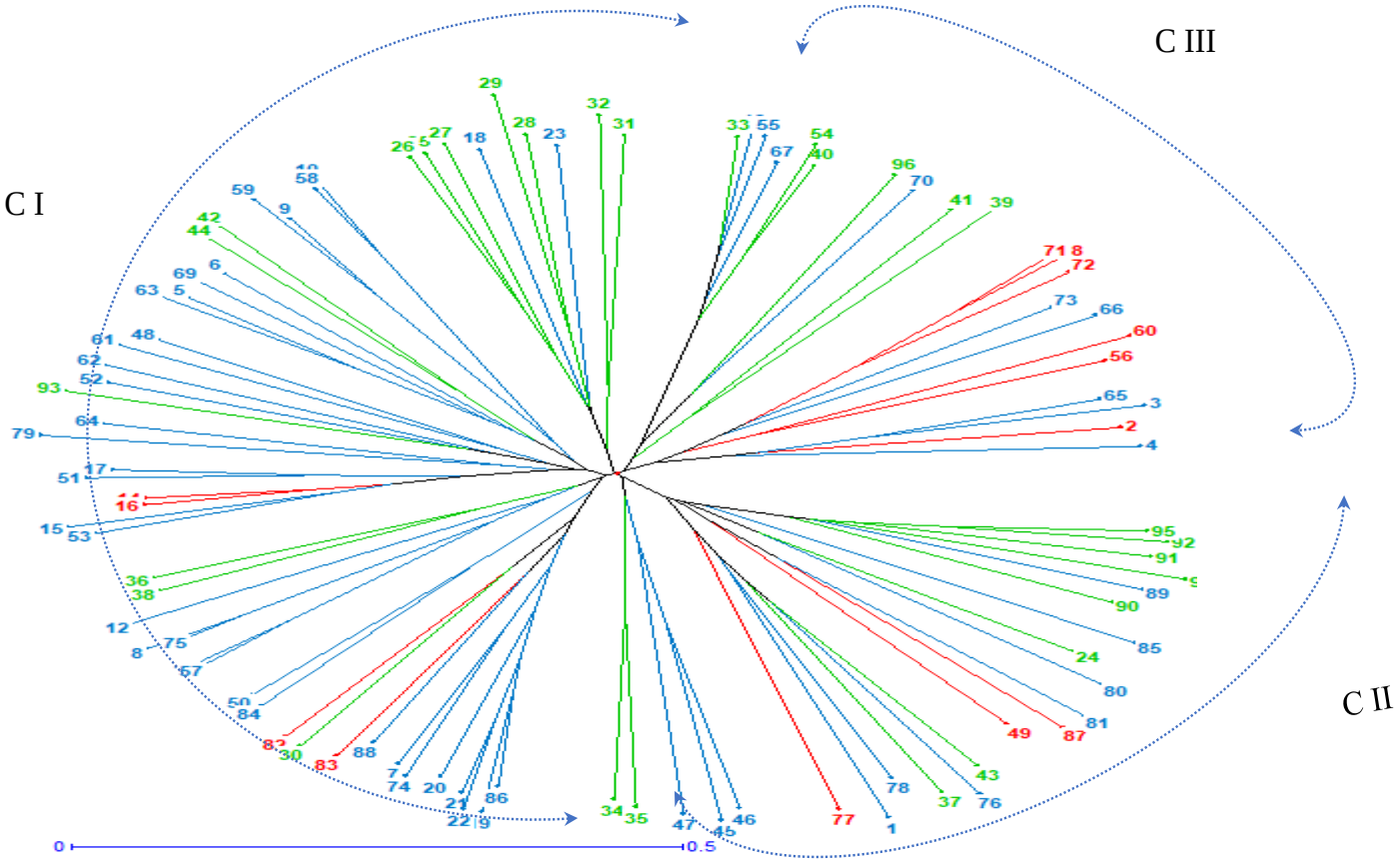

Figure 1. Genetic relationship of 96 banana genotypes by 14 SSR markers as resulted from unweighted neighbor-joining (NJ) analysis (Blue= Improved variety, Green= Advanced clones, and Red= Landrace accessions).
Source: Authors

displayed in Figure 3 showed that the genotypes spread all over the plot in such a way that none of the populations (improved variety, advanced clones, and landraces) formed a distinct group. The overall PCoA analysis grouping pattern is consistent with the NJ cluster. In both analyses, there was no unique clustering of populations from priory information with the grouping patterns obtained in the results (Figures 1 to 3) confirming an admixed population. Hence pattern of genetic relationship and variation among the genotypes is not based on the anticipated priory information rather it is seems the result of actual difference in the genetic background of the genotypes studied.

Population structure

The model-based Bayesian algorithm allowed three clusters (k = 3) to be identified as the best fit for capturing the major structure in the entire data set. Based on the K value, the clumpak population structure bar plot revealed no clear grouping of populations based on breeding history and genomic background. The observed high admixture structuring also agreed with the pattern of clustering in PCoA and phylogenetic analysis as far as the priory population information is concerned. This could be possible because of two speculative reasons: one is if genotypes in different populations are derived from the same parents, or if genotypes in a single population evolved from a different line of ancestry.

Analysis of molecular variance

Analysis of molecular variance (AMOVA) was conducted on the basis of the most likely number of grouping (K=3) as resulted from population structure analysis. The results revealed that variation within a cluster was accounting for higher variation (97%) than the variation among clusters (2%) (Figure 4) (Table 4), signifying variability of the genotypes within each groups. On the other hand, variability within individuals contributed a smaller proportion (1%) to the total variation both at priory population information level and after the result of population structure analysis (K= 3).

The study also revealed that there was little genetic

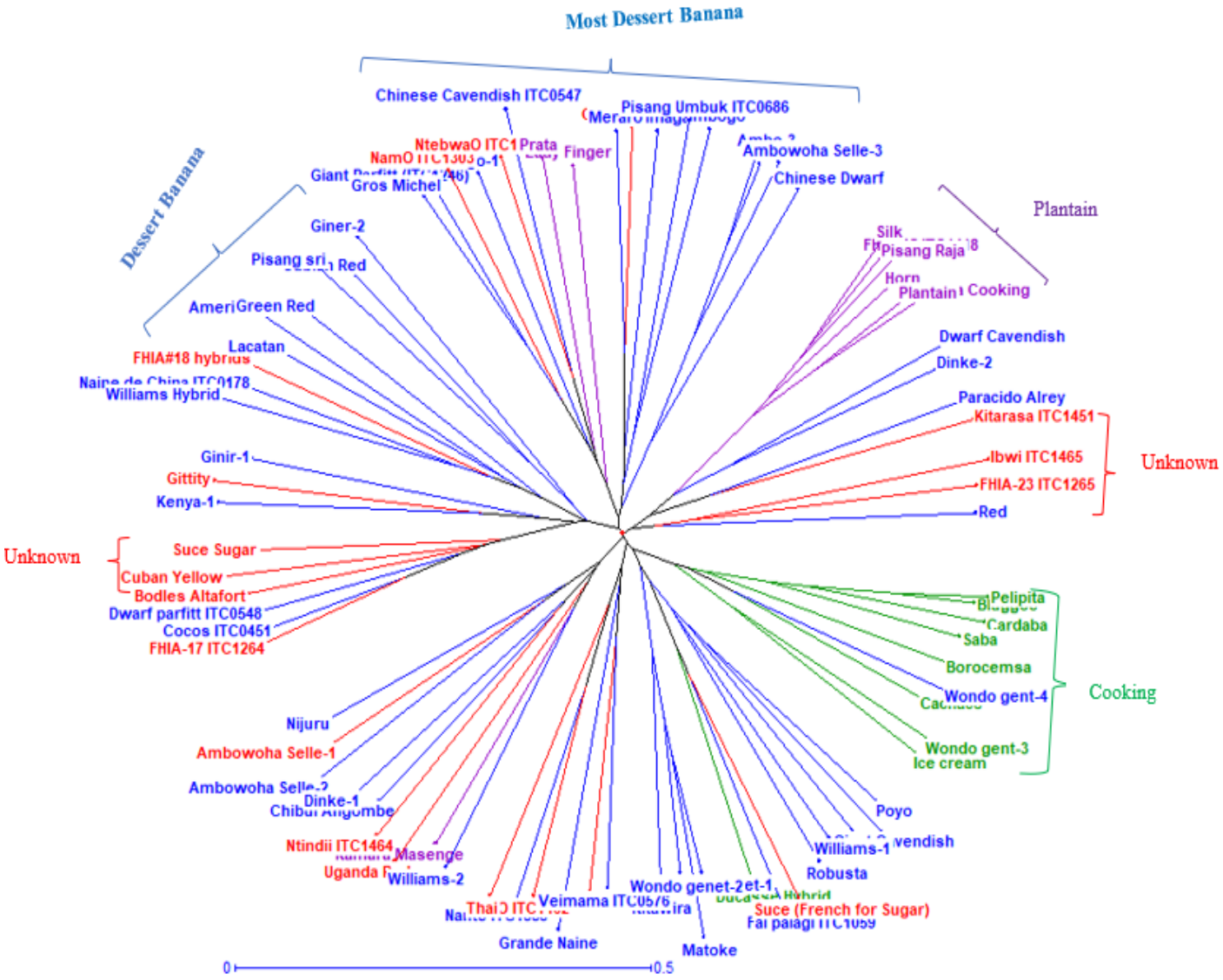

Figure 2. Genetic relationships of 96 banana genotypes by 14 SSR markers as resulted from Unweighted Neighbor-Joining (NJ) analysis and patterns grouping based on use type (Blue=Dessert banana; Green=Cooking banana; pink =plantain).
Source: Authors

variation among populations of bananas (FST = 0.023). The magnitude among and within population differentiation was quantified using F-statistics (Fit, Fis and Fst) (Wright, 1951). Population differentiation due to genetic structure is quantified by the fixation index (Fst). Wright (1951) defined the Fst value range as follows: 0 to 0.05 = low, 0.05 to 0.15 = moderate, 0.15 to 0.25 = high, and greater than 0.25 = very large genetic differentiations.

Conclusion

Knowledge of distribution of genetic diversity, and relatedness among genotypes contributes significantly towards crop improvement. In this study, the genetic diversity of 96 banana genotypes was assessed using 14 SSR markers. The various diversity indices generated showed that the SSR markers utilized in the present study were effective and informative for the banana diversity study. A high level of genetic variation was observed within genotypes, implying that the genotypes within clusters have huge usefulness in the improvement program. Furthermore, the limited sub-grouping patterns observed in relation to the various use types (Dessert, Cooking and Plantain) also confirms the association of the genome based classification with use type grouping. It is crucial to conduct more genetic variability studies using high-resolution markers and using a wider range of

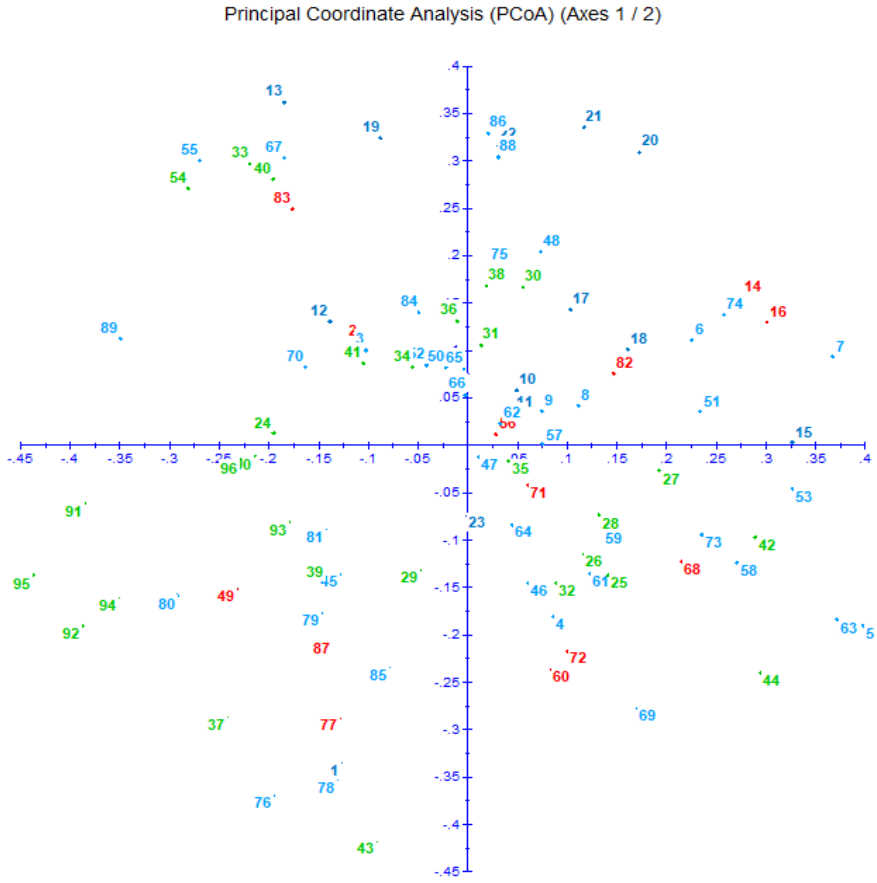

Figure 3. Principal coordinates analysis (PCoA) bi-plot showing the grouping pattern of 96 banana genotypes from three populations (Light blue = Improved varieties; Green=Advanced Clones; Red=Landraces).
Source: Authors

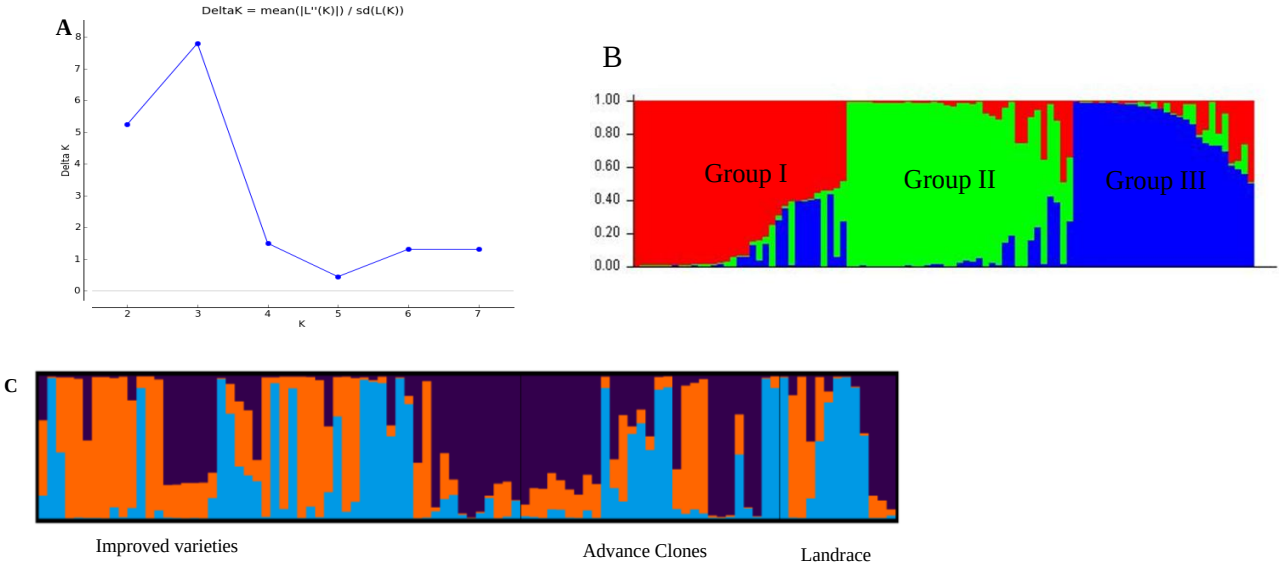

Figure 4. Population structure of 96 banana genotypes. (A) The highest peak at K=3 based on Evano et al. (2005); (B) The three major clusters of 96 genotypes; (C) Estimated population structure for K = 3 based on improved variety, advanced clone, and landrace.
Source: Authors

Table 4. Analysis of molecular variance of 96 banana genotypes after the result of population structure analysis (k= 3).

Source	DF	SS	MS	Est. Var.	% of variation	F-statistics	P value
Among population	2	40.597	20.298	0.140	2	0.023	0.001
Among individuals	93	1099.674	11.824	5.876	97	0.988	0.001
Within individuals	96	7.000	0.073	0.073	1	0.988	0.001
Total	191	1147.271		6.089	100		

DF = degree of freedom, SS=sum of squares, MS=mean squares.
Source: Authors

genotypes in the future to screen further variability and reveal any significant co-clustering patterns with some traits.

CONFLICT OF INTERESTS

The authors declare that they have no conflict of interests.

ACKNOWLEDGMENTS

The authors acknowledge the Ethiopian Institute of Agricultural Research (EIAR), National Agricultural Biotechnology Research Center (NABRC) for the financial support and the lab facility. The authors also thank Melkassa Agricultural Research Center (MARC) for providing the study materials.

REFERENCES

Biswa M K, Bagchi M, Biswas D, Harikrishna JA, Liu Y, Li C, Sheng O, Mayer C, Yi G, Deng G (2020). Genome-wide novel genic microsatellite marker resource development and validation for genetic diversity and population structure analysis of banana. Genes 11(12): 1-21.

Botstein D, White RL, Skalnick MH, Davies RW (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphism. American Journal of Human Genetics 32:314-331.

Changadey W, Kaunda E, Ambali AJ (2012). Molecular characterisation of *Musa L*. cultivars cultivated in Malawi using microsatellite markers. African Journal of Biotechnology 11(18):4140-4157.

Christelová P, De Langhe E, Hřibová E, Čížková J, Sardos J, Hušáková M, Van den houwe I, Sutanto A, Kepler A K, Swennen R, Roux N, Doležel J (2017). Molecular and cytological characterization of the global Musa germplasm collection provides insights into the treasure of banana diversity. Biodiversity and Conservation 26(4):801-824.

Creste S, Benatti TR, Orsi MR, Risterucci AM, Figueira A (2006). Isolation and characterization of microsatellite loci from a commercial cultivar of Musa acuminata. Molecular Ecology Notes 6(2):303-306.

Creste S, Tulmann Neto A, De Oliveira Silva S, Figueira A (2003). Genetic characterization of banana cultivars (*Musa spp*.) from Brazil using microsatellite markers. Euphytica 132(3):259-268.

Crouch HK, Crouch JH, Jarret RL, Cregan PB, Ortiz R (1998). Segregation at microsatellite loci in haploid and diploid gametes of *Musa Spp*. Crop Science 38(1):211-217.

Central Statistical Agency of Ethiopia (CSA) (2015). Agricultural Sample Survey. Time Series Data for National & Regional Level (From 1995/96 (1988 E.C) - 2014/15 (2007 E.C). Report On Area And Production of Crops (Private Peasant Holdings, Meher Season.

Diversity Arrays Technology (DArT) (2019). Plant DNA Extraction Protocol for DArT. https://www.diversityarrays.com/orderinstructions/plant-dna-extraction-protocol-for-dart/

De Jesus ON, Amorim EP, Ferreira CF, de Campos JMS, Silva GDG, Figueira A (2013). Genetic diversity and population structure of Musa accessions in ex situconservation. BMC Plant Biology 13(1):1-22

Earl DA, vonHoldt BM (2012). Structure Harvester: A website and program for visualizing structure output and implementing the Evanno method. Conservation Genetics Resources 4(2):359-361.

Ehrlich PR, Harte J (2015). Opinion: To feed the world in 2050 will require a global revolution. Proceedings of the National Academy of Sciences of the United States of America 112(48):14743-14744.

Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software structure: A simulation study. Molecular Ecology 14(8):2611-2620.

Food and Agriculture Organization (FAO) (2021). Banana Market Review 2020. Rome. https://agfstorage.blob.core.windows.net/misc/FP_com/2021/10/11/Ban.pdf

Gebre GG, Fikadu AA, Gebeyehu TK (2022). Is banana value chain in East Africa sustainable? Evidence from Ethiopia. Resources, Environment and Sustainability, 8(October 2021), 100060. https://doi.org/10.1016/j.resenv.2022.100060

Igwe DO, Constance O, Anne I, Osano A, Acquaah G, Ude GN (2021). Genetic diversity and population assessment of *Musa L*.(Musaceae) employing CDDP markers. Plant Molecular Biology Reporter 39(4):801-820.

Igwe DO, Ihearahu OC, Osano AA, Acquaah G, Ude GN (2022). Assessment of genetic diversity of Musa species accessions with variable genomes using ISSR and SCoT markers. Genetic Resources and Crop Evolution 69(1):49-70.

Jesus ON De, Oliveira S De, Amorim EP, Ferreira CF, Marcello J, Campos S De, Silva DG, Figueira A (2013). Genetic diversity and population structure of Musa accessions in ex situ conservation. BMC Plant Biology 13: 41.

Kaemmer D, Fischer D, Jarret RL, Baurens FC, Grapin A, Dambier D, Noyer JL, Lanaud C, Kahl G, Lagoda PJL (1997). Molecular breeding in the genus Musa: A strong case for STMS marker technology. Euphytica 96(1):49-63.

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. Molecular Ecology Resources 15(5):1179-1191.

Langhe ED, Vrydaghs L, De P, Perrier X, Denham T (2009). 356-1778-1-Pb. A Journal of Plants, People and Appllied Research 7:165-177.

Liu K, Muse SV (2005). PowerMaker: An integrated analysis environment for genetic maker analysis. Bioinformatics 21(9):2128-2129.

Marian DQ, Adelaide A, Beloved MD (2018). Assessment of plantain (*Musa sapientum* L.) accessions genotypic groups relatedness using simple sequence repeats markers. African Journal of Biotechnology 17(16):541-551.

Noyer JL, Causse S, Tomekpe K, Bouet A, Baurens FC (2005). A new image of plantain diversity assessed by SSR, AFLP and MSAP markers. Genetica 124(1):61-69.

Okeh D, Onyinye I, Ihearahu C, Adhiambo A, Acquaah G, Nkem G

(2022). Assessment of genetic diversity of Musa species accessions with variable genomes using ISSR and SCoT markers. Genetic Resources and Crop Evolution 69(1):49-70.
Opara UL, Jacobson D, Al-Saady NA (2010). Analysis of genetic diversity in banana cultivars (*Musa cvs*.) from the South of Oman using AFLP markers and classification by phylogenetic, hierarchical clustering and principal component analyses. Journal of Zhejiang University: Science B 11(5):332-341.
Paofa J, Sardos J, Christelova P, Dolez J, Roux N (2018). Collection of new diversity of wild and cultivated bananas (*Musa spp*.) in the Autonomous Region of Bougainville, Papua New Guinea. 0123456789, 2267-2286.
Pavel AB, Vasile CI (2012). PyElph - a software tool for gel images analysis and phylogenetics. BMC Bioinformatics 13(1).
Peakall ROD, Smouse PE (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6(1):288-295.
Perrier X, Jacquemoud-Collet JP (2006). DARwin software: Dissimilarity analysis and representation for windows. Website http://darwin.cirad.fr/darwin [accessed 1 March 2013].
Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A (1996). The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Molecular Breeding 2(3):225-238.
Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. Genetics 155(2):945-959.
Saitou N, Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 4(4):406-425.
White H, Peakall R (2015). GENALEX 6.502: Genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6:288-295
Wright S (1951). The genetic structure of populations. Annal Eugenics 15:323-354.