Peer

Comparative Assessment of SSR and RAPD markers for genetic diversity in some Mango cultivars

Mohammed A. A. Hussein¹, Manal Eid¹, Mehdi Rahimi², Faten Zubair Filimban³ and Diaa Abd El-Moneim⁴

¹ Department of Botany (Genetic Branch), Faculty of Agriculture, Suez Canal University, Ismailia, Egypt

² Department of Biotechnology, Institute of Science and High Technology and Environmental Sciences, Graduate University of Advanced Technology, Kerman, Iran

³ Division of Botany, Department of Biology, Faculty of Sciences, King Abdulaziz University, Jeddah, Saudi Arabia

⁴ Faculty of Environmental Agricultural Sciences, Plant Production Department - Genetic Branch, Arish University, El Arish, Egypt

ABSTRACT

Genetic improvement mainly depends on the level of genetic variability present in the population, and the degree of genetic diversity in a population largely determines the rate of genetic advancement. For analyzing genetic diversity and determining cultivar identities, a molecular marker is a useful tool. Using 30 SSR (simple sequence repeat) and 30 RAPD (randomly amplified polymorphic DNA) markers, this study evaluated the genetic divergence of 17 mango cultivars. The effectiveness of the two marker systems was evaluated using their genetic diversity characteristics. Additionally, the effects of SM (simple matching) and Dice similarity coefficients and their effects on mango clustering were evaluated. The findings showed that SSR markers generated 192 alleles, all of which were polymorphic (100%). With RAPD markers, 434 bands were obtained, 361 of which were polymorphic (83%). The average polymorphic information content (PIC) for RAPD and SSR was 0.378 and 0.735, respectively. Using SSR markers resulted in much higher values for other genetic diversity parameters compared to RAPD markers. Furthermore, grouping the genotypes according to the two similarity coefficients without detailed consideration of these coefficients could not influence the study results. The RAPD markers OPA 01, OPM 12 followed by OPO 12 and SSR markers MIAC_4, MIAC_5 followed by mMiCIR_21 were the most informative in terms of describing genetic variability among the cultivars under study; they can be used in further investigations such as genetic mapping or marker-assisted selection. Overall, 'Zebda' cultivar was the most diverse of the studied cultivars.

Subjects Agricultural Science, Biotechnology, Genetics, Molecular Biology, Plant Science **Keywords** Comparative assessment, Dice and Simple matching coefficients, Genetic diversity, Mango, Population structure, RAPD, SSR markers

INTRODUCTION

The mango fruit (*Mangifera indica* L.) is one of the most nutritious and expensive edible fruits globally. It has a diploid genome (2n = 2x = 40 chromosomes) and belongs to the order Sapindales, family Anacardiaceae, and genus *Mangifera*. Mangoes are grown on

Submitted 12 April 2023 Accepted 24 August 2023 Published 28 September 2023

Corresponding author Diaa Abd El-Moneim, dabdelmoniem@aru.edu.eg

Academic editor Imren Kutlu

Additional Information and Declarations can be found on page 18

DOI 10.7717/peerj.15722

Copyright 2023 A. A. Hussein et al.

Distributed under Creative Commons CC-BY 4.0

OPEN ACCESS

2.5 million ha in tropical and subtropical regions of the world, with an annual production of approximately 46.6 million tons (*Wang, Luo & Huang, 2020*). The mango is the fifth most essential fruit crop, followed by bananas, grapes, apples, and oranges (*Deshpande et al., 2017*).

The mango is the most exquisite fruit in Egypt and one of the world's best fruits. Mangoes have been planted in 289,020 feddans (1 feddans = 0.42 hectare) in Egypt, according to statistics from the Ministry of Agriculture and Land Reclamation (2020), with a maximum annual yield of two million and 800 tons.

Several varieties with different origins grow in Egypt. From India and Sri Lanka come the 'Hindi Besennara', long, 'Banarasi Langra', and 'Mabrouka' varieties, and from Florida and South Africa come the 'Carrie', 'Glenn', 'Keitt', and 'Kent' varieties. The local varieties include the 'Zebda', 'Taymour', 'Mesk', 'Senarry', 'Mstikawy', and 'Dabsha' mangoes (*Abdelsalam et al., 2018*).

The traditional methods of identifying or distinguishing cultivars have used the morphological traits of the plant's leaves, flowers, and fruit (*Avramidou et al., 2023*). This approach has been unsuccessful since closely related cultivars often cannot be differentiated by these traits and since environmental factors can affect the expression of the traits (*e.g.*, climate conditions or cultivation procedures may have impacts).

Genetic variation in the germplasm of plants is a crucial factor in new cultivar creation because inbreeding in cultivated plants leads to a quick loss in vigor, yield, and fruit size (*Spangelo et al., 1971*). For these reasons, modern breeders need more efficient methods of breeding that are rapid, informative, and unaffected by environmental factors. Molecular markers have been replacing or complementing traditional morphological and agronomic characterizations since they are extremely plentiful, cover the genome, and are not influenced by the environment. In the case of fruit trees with a long juvenile period, these markers can make the characterization of new cultivars less time-consuming (*Ramadasappa, Mallaiah & Ramakrishnaiah, 2021*).

Previous studies have found that molecular markers are unlimited in number, remain unaffected by the environment and growth conditions, and are simply inherited (*Karihaloo et al., 2003*). Various molecular markers such as the randomly amplified polymorphic DNA (RAPDs) of *Adhikari & Sinchan (2015)*, amplified fragment length polymorphisms (AFLPs) of *Eiadthong et al. (2000)*, and simple sequence repeats (SSRs) of *Bukhari et al. (2022)* have been tested for genetic diversity assessment in mangoes. There has been increasing development and generalized use of many methodologies during recent years, and now comparative studies are needed to choose the best DNA marker technology for fingerprinting and diversity studies in terms of reproducibility, cost, sensibility, and level of polymorphism detection. One technique may be more appropriate than another in a study, and different techniques may be informative at different taxonomic levels. Similarities between the different molecular techniques have begun to be debated, but the results have conflicted among authors. Several works have reported comparable results among different markers (*Bally et al., 2021; Luo et al., 2011; Powell et al., 1996; Srivastava et al., 2012*).

Although molecular markers only reflect a small portion of a plant's genome, they are used to infer links between the complete genome among a number of populations. The precision of resulting estimations of genetic distance will depend on how the loci discovered by individual marker analysis approaches are distributed (*Nei*, 1987). Therefore, it is ideal for assessing genetic diversity because the loci discovered are scattered randomly throughout and must be a sample to reflect the entire genome. Genetic maps are required to compare the distribution of various individual markers, but this takes much time. As a result, an alternate strategy has been put out to calculate the relationships between the germplasm accessions produced from various marker approaches utilizing statistics analysis and various similarity coefficients (Liu, 1998). The molecular marker data have been compared using different similarity coefficients and clustering methods (Meyer et al., 2004). Using different similarity coefficients such as Dice and simple matching could affect the results of the unweighted pair method of groups with arithmetic (UPGMA) and other methods (*Reif, Melchinger & Frisch, 2005*), whereas Egyptian mango cultivars differ in fruit quality and disease tolerance, but the extent of genetic differences between each of them is not well known. This study aims to evaluate the impact of two DNA markers (RAPD and SSR markers) on the genetic diversity levels of 17 cultivars of mango, as well as to compare two alternative similarity coefficients and clustering techniques to identify the most dissimilar cultivars in genetic diversity assessments that will serve as parents in mango breeding and improvement programs in order to maintain the unique genetic resources.

MATERIALS AND METHODS

Plant materials and processing of leaves

All plant materials used and collected in the study met Egypt's guidelines and legislative regulations. Fifty-one mature mango trees (*Mangifera indica* L.) encompassing 17 cultivars were included in the present investigation (Table 1). The experimental trees were grown in private orchards in Abou Swear City, Ismailia Governorate, Egypt. It was chosen for the quality of the fruit produced there. Three trees per cultivar were selected; all of them were vegetatively propagated. The trees were labeled in March and April of 2021 at the time of blooming, and leaf material for DNA extraction was gathered at that time. Care was taken in selecting samples to gather only young, tender, healthy leaves. Samples were wrapped in aluminum foil, appropriately labeled, fixed by submersion briefly in liquid nitrogen, and stored at -80 °C until DNA extraction.

DNA extraction

The genomic DNA from leaf samples was extracted by a modified CTAB method (*Porebski, Bailey & Baum, 1997*). DNA was quantified by gel electrophoresis, and its quality was verified by the Nano Drop ND-1000 spectrophotometer (GMI, Ramsey, MN, USA). DNA samples were then stored at 4 °C. The stock DNA samples of each cultivar were diluted with tris-EDTA and analyzed individually to detect intra-cultivar variations and bulked to detect inter-cultivar variations.

1 able 1 List of 17 Mango cultivars used in the study.									
No.	Name	Origin	No.	Name	Origin				
1	Banarasi Langra	India	10	Sukkary white	Selected seedy clone in Egypt				
2	Mabrouka	India	11	Mstikawy	Selected seedy clone in Egypt				
3	Ewais	Selected seedy clone in Egypt	12	Fajri kalan	India				
4	Taymour	Selected seedy clone in Egypt	13	Elwazza neck	Selected seedy clone in Egypt				
5	Hindi Besennara	India	14	Keitt	USA				
6	Mesk	Selected seedy clone in Egypt	15	Kent	USA				
7	Zebda	Selected seedy clone in Egypt	16	Senarry	Selected seedy clone in Egypt				
8	Hindi mlooky	India	17	Nabiel	Selected seedy clone in Egypt				
9	Companeit Elsowwa	Selected seedy clone in Egypt							

SSR amplification

Thirty of the microsatellite markers utilized in this investigation were previously described by *Duval et al.* (2005), *Honsho et al.* (2005), *Schnell et al.* (2005) and *Viruel et al.* (2005) (Table 2) and these primers were synthesized by the Oligosystem (Macrogen, Seoul, Korea). The polymerase chain reactions were performed in 10 μ l volume containing: 0.2 μ M of each forward and reverse primers, 0.2 mM of dNTPs, 1.8 mM MgCl₂, 0.05 U Taq polymerase, 1X PCR buffer, and 10 ng of template DNA, according to *Schnell et al.* (2005) with slight modifications. The amplification was done in a thermocycler (Eppendorf MasterCycler Gradient; Eppendorf, Hamburg, Germany). After a first denaturation step at 94 °C for 2 min, thereafter 30 cycles at 94 °C for 1 min, 51 °C for 30 s, 72 °C for 1 min, and finally extension at 72 °C for 5 min.

Every reaction was repeated twice to ensure the reproducibility of the results. The PCR mixer and cycling PCR products were separated on agarose gel (2%), and ethidium bromide was utilized for staining to ensure the PCR amplification and determine the approximate size of the amplified fragments. Then, products were separated on polyacrylamide gels (7%) to confirm the allele sizing of the SSR loci and stained with ethidium bromide solution and visualized using the gel documentation model (Gel-Doc 2000 with Diversity Database software Ver. 2.1; Bio-Rad Laboratories, Hercules, CA, USA) for gel analysis. Quantity One software was used to estimate the sizes of the products by comparison with the size marker.

RAPD amplification

Thirty 10-mer RAPD primers were screened for the RAPD-PCR reactions that resulted in distinct and well-separated bands on a polyacrylamide gel. These primers (Table 3) were synthesized by Oligo (Macrogen, Seoul, Korea) and utilized to detect polymorphisms in the 17 mango cultivars. The RAPD-PCR reaction was achieved, according to *Williams et al.* (1990). PCR amplification was carried out in a total volume of 10 µl, containing 10 ng genomic DNA, 0.3 pmol 10 mer random primer, 1X reaction buffer, 5 µl of Go-Taq (ready mastermix) including (1.5 mM Mg Cl₂, 0.2 mM dNTPs, 0.5 U Taq polymerase) PCR amplification was performed in a primus 384 well thermocycler (MWG Biotech AG), Ebensburg, Germany; http://www.mwg-biotech.com), programmed to include a pre-denaturation step at 94 °C for 60 s and followed by 34 cycles of denaturation at 94 °C

Table 2 List of polymorphic SSR markers used in this study.									
No.	Locus	SSR primers sequence 5—>>3	No.	Locus	SSR primers sequence 5—>>3				
1 MiSHRS	MISHDS 1	F: TAACAGCTTTGCTTGCCTCC	16	mMiCID 5	F: GCCCTTGCATAAGTTG				
	WII5HK5_1	R: TCCGCCGATAAACATCAGAC	10	IIIWIICIK_3	R: TAAGTGATGCTGCTGGT				
C	MISHDS A	F: CCACGAATATCAACTGCTGCC	17	mMiCID 8	F: GACCCAACAAATCCAA				
2	WII511K5_4	R: TCTGACACTGCTCTTCCACC	17	IIIWIICIK_0	R: ACTGTGCAAACCAAAAG				
2	MICHDS 18	F: AAACGAGGAAACAGAGCAC	10	mMiCIP 0	F: AAAGATAAGATTGGGAAGAG				
5	WII511K5_16	R: CAAGTACCTGCTGCAACTAG	10	IIIWIICIK_9	R: CGTAAGAAGAGCAAAGGT				
4	MISHDS 32	R: AGAAACATGATGTGAACC	10	mMiCID 18	F: CCTCAATCTCACTCAACA				
4	WII511K5_52	F: TTGATGCAACTTTCTGCC	19	IIIWIICIK_10	R: ACCCCACAATCAAACTAC				
F	mMiCID 26	F: GTTTTCATTCTCAAAATGTGTG	20	mMiCID 21	F: CCATTCTCCATCCAAA				
5	IIIMICIK_50	R: CTTTCATGTTCATAGATGCAA	20	IIIWIICIK_21	R: TGCATAGCAGAAAGAAGA				
6	MISHDS 18	F: TTTACCAAGCTAGGGTCA	21	mMiCIP 22	F: TGTCTACCATCAAGTTCG				
0	WII3FIK3_40	R: CACTCTTAAACTATTCAACCA	21	IIIWIICIK_22	R: GCTGTTGTTGCTTTACTG				
7		F:ATGGAGACTAGAATGTACAGAG	22	mMiCIP 25	F: ATCCCCAGTAGCTTTGT				
/	LIVIIVIA_1	R: ATTAAATCTCGTCCACAAGT	22	IIIWIICIK_23	R: TGAGAGTTGGCAGTGTT				
0	IMMA 7	F: ATTTAACTCTTCAACTTTCAAC	22	MISHRS 20	F: CAACTTGGCAACATAGAC				
0	LIVINA_/	R: AGATTTAGTTTTGATTATGGAG	25	WII311K3_29	R: ATACAGGAATCCAGCTTC				
0	LMMA_8	F: CATGGAGTTGTGATACCTAC	24	mMiCIP 20	F: GCGTGTCAATCTAGTGG				
9		R: CAGAGTTAGCCATATAGAGTG	24	IIIWIICIK_29	R: GCTTTGGTAAAAGGATAAG				
10	ι ΜΜΑ Ο	F:TTGCAACTGATAACAAATATAG	25	mMiCID 20	F: GCTCTTTCCTTGACCTT				
10	LIVIIVIA_9	R: TTCACATGACAGATATACACTT	23	IIIWIICIK_50	R: TCAAAATCGTGTCATTTC				
11	I MMA 10	F: TTCTTTAGACTAAGAGCACATT	26	MIAC 2	F: GCTTTATCCACATCAATATCC				
11	LIVINA_10	R: AGTTACAGATCTTCTCCAATT	20	WIAC_2	R: TCCTACAATAACTTGCC				
10	I MMA 11	F: ATTATTTACCCTACAGAGTGC	27	MIAC 3	F: TAAGCTAAAAAGGTTATAG				
12	LIVINA_11	R: GTATTATCGGTAATGTCTTCAT	27	MIAC_5	R: CCATAGGTGAATGTAGAGAG				
12		F: CACAGCTCAATAAACTCTATG	20	MIAC 4	F: CGTCATCCTTTACAGCGAACT				
13	LIVINIA_15	R: CATTATCCCTAATCTAATCATC	20	MIAC_4	R: CATCTTTGATCATCCGAAAC				
1.4	I MMA 14	F: ATTATCCCTATAATGCCCTAT	20	MIAC 5	F: AATTATCCTATCCCTCGTATC				
14	LIMINIA_14	R: CTCGGTTAACCTTTGACTAC	29	MIAC_3	R: AGAAACATGATGTGAACC				
15	LMMA_15	F: AACTACTGTGGCTGACATAT	20	MIAC 6	F: CGCTCTGTGAGAATCAAATGGT				
15		R: CTGATTAACATAATGACCATCT	50	MIAC_0	R: GGACTCTTATTAGCCAATGGGATG				

for 60 s, annealing at 36 °C for 45 s and extension at 72 °C for 30 s, finished with a final extension step of 5 min at 72 °C.

Electrophoresis of the fragments was done on 7% polyacrylamide gels in a vertical CBS Scientific (San Diego, CA, USA; http://www.cbsscientific.com) electrophoresis unit in 1x TBE buffer at a voltage of 480 V and 80 mA for 90 min (until the lower band of dye escaped from the gel). Gel staining was done using silver nitrate according to a modified protocol of *Bassam & Caetano-Anollés (1993)*. Dried gels were scanned, and the sizes of the amplified products were visually examined and estimated by Quantity One software.

Data scoring and analysis

Bands were scored from the images. The presence of a band was scored as 1, and the absence of a band as 0. Cluster analysis was used to identify the studied populations and to

No. Name Sequence No. Name Sequence 1 OPM_20 5'-AGG TCT TGG G-3' 16 OPH_18 5'-GAA TCG GCC A-3' 2 OPQ_05 5'-CCG CGT CTT G-3' 17 OPB_17 5'-AGG GAA CGA G-3' 3 OPS_17 5'-TGG GGA CCA C-3' 18 OPA_04 5'-AAT CGG GCT G-3' 4 OPD_20 5'-ACC CGG TCA C-3' 19 OPE_12 5'-TTA TCG CCC C-3'			r			
1 OPM_20 5'-AGG TCT TGG G-3' 16 OPH_18 5'-GAA TCG GCC A-3' 2 OPQ_05 5'-CCG CGT CTT G-3' 17 OPB_17 5'-AGG GAA CGA G-3' 3 OPS_17 5'-TGG GGA CCA C-3' 18 OPA_04 5'-AAT CGG GCT G-3' 4 OPD_20 5'-ACC CGG TCA C-3' 19 OPE_12 5'-TTA TCG CCC C-3'	No.	Name	Sequence	No.	Name	Sequence
2 OPQ_05 5'-CCG CGT CTT G-3' 17 OPB_17 5'-AGG GAA CGA G-3' 3 OPS_17 5'-TGG GGA CCA C-3' 18 OPA_04 5'-AAT CGG GCT G-3' 4 OPD_20 5'-ACC CGG TCA C-3' 19 OPE_12 5'-TTA TCG CCC C-3'	1	OPM_20	5'-AGG TCT TGG G-3'	16	OPH_18	5'-GAA TCG GCC A-3'
3 OPS_17 5'-TGG GGA CCA C-3' 18 OPA_04 5'-AAT CGG GCT G-3' 4 OPD_20 5'-ACC CGG TCA C-3' 19 OPE_12 5'-TTA TCG CCC C-3'	2	OPQ_05	5'-CCG CGT CTT G-3'	17	OPB_17	5'-AGG GAA CGA G-3'
4 OPD_20 5'-ACC CGG TCA C-3' 19 OPE_12 5'-TTA TCG CCC C-3'	3	OPS_17	5'-TGG GGA CCA C-3'	18	OPA_04	5'-AAT CGG GCT G-3'
	4	OPD_20	5'-ACC CGG TCA C-3'	19	OPE_12	5'-TTA TCG CCC C-3'
5 OPO_12 5'-CAG TGC TGT G-3' 20 OPA_01 5'-CAG GCC CTT C-3'	5	OPO_12	5'-CAG TGC TGT G-3'	20	OPA_01	5'-CAG GCC CTT C-3'
6 OPA_13 5'-CAG CAC CCA C-3' 21 OPX_08 5'-CAG GGG TGG A-3'	6	OPA_13	5'-CAG CAC CCA C-3'	21	OPX_08	5'-CAG GGG TGG A-3'
7 OPB_11 5'-GTA GAC CCG T-3' 22 OPX_17 5'-GAC ACG GAC C-3'	7	OPB_11	5'-GTA GAC CCG T-3'	22	OPX_17	5'-GAC ACG GAC C-3'
8 OPJ_20 5'-AAG CGG CCT C-3' 23 OPX_02 5'-TTC CGC CAC C-3'	8	OPJ_20	5'-AAG CGG CCT C-3'	23	OPX_02	5'-TTC CGC CAC C-3'
9 OPM_06 5'-CTG GGC AAC T-3' 24 OPX_01 5'-CTG GGC ACG A-3'	9	OPM_06	5'-CTG GGC AAC T-3'	24	OPX_01	5'-CTG GGC ACG A-3'
10 OPA_09 5'-GGG TAA CGC C-3' 25 OPA_18 5'-AGG TGA CCG T-3'	10	OPA_09	5'-GGG TAA CGC C-3'	25	OPA_18	5'-AGG TGA CCG T-3'
11 OPX-18 5'-GAC TAG GTG G-3' 26 OPA_07 5'-GAA ACG GGT G-3'	11	OPX-18	5'-GAC TAG GTG G-3'	26	OPA_07	5'-GAA ACG GGT G-3'
12 OPM-12 5'-GGG ACG TTG G-3' 27 OPA_11 5'-CAA TCG CCG T-3'	12	OPM-12	5'-GGG ACG TTG G-3'	27	OPA_11	5'-CAA TCG CCG T-3'
13 OPD_13 5'-GGG GTG ACG A-3' 28 OPW_11 5'-CTG ATG CGT G-3'	13	OPD_13	5'-GGG GTG ACG A-3'	28	OPW_11	5'-CTG ATG CGT G-3'
14 OPM_15 5'-GAC CTA CCA C-3' 29 OPC_05 5'-GAT GAC CGC C-3'	14	OPM_15	5'-GAC CTA CCA C-3'	29	OPC_05	5'-GAT GAC CGC C-3'
15 OPG_11 5'-TGC CCG TCG T-3' 30 OPK_10 5'-GTG CAA CGT G-3'	15	OPG_11	5'-TGC CCG TCG T-3'	30	OPK_10	5'-GTG CAA CGT G-3'

 Table 3
 List of polymorphic RAPD markers used in this study.

group similar populations in terms of all measured traits. Various cluster analysis methods and different distance criteria were evaluated, and the method with the highest cophenetic correlation coefficient was selected. The number of amplified bands, polymorphic bands, and polymorphisms percentage for each primer were obtained based on the band patterns and used for grouping the cultivars and estimating the indices.

The polymorphism information content (PIC) (*Botstein et al.*, 1980), effective multiplex ratio, mean heterozygosity, marker index (*Powell et al.*, 1996), expected heterozygosity (*Liu*, 1998), and discriminating power (*Tessier et al.*, 1999) were estimated using an Excel program based on the relevant formulas. The number of effective alleles, Shannon's information index, and Nei's measure of gene diversity were calculated using POPGEN software version 1.31 (*Yeh*, 1999).

The probability of identity (PI) was calculated for each marker according to *Peakall & Smouse (2012)*. Differences and similarities were first calculated based on different similarity criteria for grouping the genotypes cultivars. Then the cultivars were grouped using different cluster analysis methods (*e.g.*, unweighted pair-group method using arithmetic average (UPGMA) and Ward). The cluster analysis method that had the highest cophenetic correlation coefficient was selected. The dendrogram and cluster analysis was performed using NTSYS-PC (numerical taxonomy system) version 2.11 (*Rohlf, 2000*). Structure software (*Pritchard, Stephens & Donnelly, 2000*) was used to identify the population structure in the studied mango cultivar with RAPD and SSR markers data. To choose the optimal level of K, for each K, the first 10,000 repetitions were followed by 10,000 Marcov chain Monte Carlo calculations (MCMCs) based on the mixed model from K equals 2 to K equals 10 with three (3) repetitions. Then the best K utilizing structure harvester software (*Earl & von Holdt, 2012*) was identified based on the Δk method (*Evanno, Regnaut & Goudet, 2005*).



Figure 1 An example of SSR markers polymorphism. Miac_4 (A), Miac_5 (B) and RAPD markers OPK_10 (C) and OPA_18 (D) among seventeen mango cultivars. Full-size DOI: 10.7717/peerj.15722/fig-1

RESULTS

Molecular characterization of SSR and RAPD markers

Thirty RAPD markers (all with multiple loci) produced 434 bands ranging from 6 to 26, with an average of 14.5 bands. OPA_18 had the most bands, with 26, while OPK_10, OPX_02, and OPO_12 had the fewest, with 6 bands, followed by OPG_11 with 8 bands (Figs. 1C and 1D). Thirty RAPD markers generated 361 polymorphic bands (83% polymorphisms), and 73 monomorphic bands, which were typical of all cultivars, existed. The 30 SSR markers (all of them having a single locus) showed high levels of polymorphisms (~100%) and produced 192 polymorphic alleles averaging 6.4/locus. The number of total alleles/loci ranged from 4 (MIAC_4, MiSHRS_18, and mMiCIR_36) to 10 (MIAC_5) (Figs. 1A and 1B).

According to this study's RAPD markers, the expected heterozygosity (He) ranged from 0.415 (OPA_01) to 0.500 (OPQ_05, OPM_06, and OPM_12), with a mean of 0.480. For SSR markers, expected heterozygosity (He) values varied from 0.524 (mMiCIR_21) to 0.0.860 (MIAC_5), with a mean of 0.752.

The PIC value was calculated separately for each studied marker, and the results are shown for the RAPD markers in Table 4 and SSR markers in Table 5. The PIC values, which reflect allele diversity and frequency among the cultivars, were not uniformly higher for

Table 4 Characteristics and studied indices of RAPD primers used in the study of genetic diversity of mango genotypes.											
Primer	Na	Р	He	PIC	Havp	EMR	MI	D	NEA	Н	S
OPM_20	21	15	0.477	0.380	0.002	5.882	0.011	0.847	1.488	0.295	0.454
OPQ_05	24	22	0.500	0.368	0.001	11.059	0.015	0.748	1.634	0.359	0.532
OPS_17	19	15	0.484	0.376	0.002	6.176	0.012	0.831	1.473	0.290	0.450
OPD_20	19	13	0.477	0.379	0.002	5.118	0.011	0.846	1.494	0.281	0.428
OPO_12	6	5	0.494	0.371	0.006	2.235	0.013	0.803	1.603	0.368	0.553
OPA_13	11	7	0.470	0.383	0.004	2.647	0.010	0.859	1.403	0.264	0.420
OPB_11	10	10	0.496	0.371	0.003	4.529	0.013	0.796	1.534	0.304	0.513
OPJ_20	14	12	0.500	0.369	0.002	5.824	0.014	0.766	1.583	0.347	0.524
OPA_09	13	12	0.481	0.378	0.002	4.824	0.011	0.840	1.484	0.303	0.470
OPX_18	14	14	0.487	0.375	0.002	8.118	0.017	0.665	1.527	0.328	0.501
OPM_06	16	12	0.500	0.369	0.002	6.176	0.015	0.736	1.578	0.341	0.512
OPM_12	13	9	0.500	0.368	0.003	4.471	0.015	0.755	1.643	0.377	0.560
OPD_13	18	13	0.486	0.375	0.002	7.000	0.017	0.661	1.605	0.356	0.532
OPM_15	14	14	0.496	0.370	0.002	6.412	0.013	0.791	1.600	0.347	0.518
OPG_11	8	7	0.484	0.376	0.004	4.118	0.017	0.656	1.598	0.338	0.501
OPH_18	19	18	0.490	0.373	0.002	7.706	0.012	0.818	1.517	0.312	0.478
OPB_17	9	6	0.457	0.389	0.004	2.118	0.009	0.878	1.455	0.285	0.445
OPA_04	16	12	0.477	0.380	0.002	4.706	0.011	0.847	1.479	0.287	0.441
OPE_12	17	14	0.489	0.374	0.002	5.941	0.012	0.821	1.469	0.297	0.460
OPA_01	10	8	0.415	0.407	0.003	5.647	0.017	0.503	1.693	0.386	0.562
OPX_08	17	12	0.499	0.369	0.002	5.765	0.014	0.770	1.637	0.369	0.547
OPX_17	16	13	0.491	0.373	0.002	5.647	0.013	0.812	1.472	0.287	0.447
OPX_02	6	2	0.484	0.376	0.014	0.824	0.012	0.838	1.567	0.325	0.492
OPX_01	10	10	0.457	0.389	0.003	3.529	0.009	0.877	1.359	0.232	0.373
OPA_18	26	22	0.420	0.405	0.001	6.588	0.007	0.911	1.392	0.248	0.393
OPA_07	14	14	0.480	0.378	0.002	5.588	0.011	0.842	1.456	0.283	0.441
OPA_11	17	16	0.457	0.389	0.002	5.647	0.009	0.876	1.439	0.275	0.431
OPW_11	20	20	0.497	0.370	0.001	9.176	0.013	0.790	1.588	0.352	0.529
OPC_05	11	9	0.479	0.378	0.003	3.588	0.011	0.843	1.451	0.272	0.423
OPK_10	6	6	0.491	0.373	0.005	2.588	0.012	0.816	1.518	0.318	0.488
Average	14.5	12.03	0.480	0.378	0.003	5.322	0.013	0.795	1.525	0.314	0.481
polymorphis	m 83%										

Notes.

NA, Total bands; P, polymorphic bands; He, Expected heterozygosity; PIC, Polymorphism information content; EMR, Effective multiplex ratio; Havp, Mean heterozygosity; MI, Marker index; D, Discriminating power; NEA, Number of effective alleles; H, Nei's gene diversity; S, Shannon's Information index.

> all the RAPD loci tested. The PIC value ranged from 0.368 (OPQ_05 & OPM_12) to 0.405 (OPA_18), with a mean of 0.378. On the other hand, PIC values for SSR markers ranged from 0.510 (mMiCIR_21) to 0.852 (MIAC_5), with a mean of 0.735.

> The mean heterozygosity (Havp) of the RAPD markers ranged from 0.001 to 0.014 with an average of 0.003; for the SSR markers, it ranged from 0.014 to 0.095 with an average of 0.043. The OPX_02 (RAPD) and the mMiCIR_36 (SSR) had the highest mean heterozygosity value.

Table 5 Characteristics and studied indices of 55K primers used in the study of genetic diversity of mango cultivars.											
Primer	NA	Р	He	PIC	Havp	EMR	MI	D	NEA	Н	S
MiSHRS_1	5	5	0.767	0.748	0.047	1.412	0.066	0.815	1.361	0.254	0.414
MiSHRS_4	5	5	0.720	0.700	0.056	1.765	0.099	0.765	1.448	0.282	0.446
MiSHRS_18	4	4	0.681	0.649	0.080	1.647	0.131	0.724	1.580	0.334	0.495
MiSHRS_32	7	7	0.749	0.733	0.036	1.706	0.061	0.796	1.311	0.204	0.338
mMiCIR_36	4	4	0.618	0.590	0.095	1.529	0.146	0.657	1.460	0.282	0.437
MiSHRS_48	5	5	0.745	0.724	0.051	1.176	0.060	0.792	1.295	0.214	0.361
LMMA_1	9	9	0.829	0.818	0.019	1.765	0.034	0.881	1.246	0.178	0.305
LMMA_7	7	7	0.785	0.770	0.031	1.882	0.058	0.834	1.351	0.231	0.374
LMMA_8	8	8	0.770	0.753	0.029	1.941	0.056	0.818	1.322	0.202	0.332
LMMA_9	5	5	0.749	0.729	0.050	1.471	0.074	0.796	1.384	0.258	0.418
LMMA_10	8	8	0.782	0.768	0.027	2.000	0.054	0.831	1.324	0.209	0.344
LMMA_11	6	6	0.807	0.793	0.032	1.765	0.057	0.857	1.379	0.263	0.426
LMMA_13	5	5	0.749	0.727	0.050	1.529	0.077	0.795	1.404	0.268	0.425
LMMA_14	7	7	0.793	0.779	0.030	1.765	0.052	0.843	1.328	0.222	0.366
LMMA_15	6	6	0.719	0.696	0.047	1.882	0.088	0.764	1.410	0.247	0.386
mMiCIR_5	6	6	0.751	0.730	0.041	1.824	0.076	0.798	1.415	0.256	0.400
mMiCIR_8	8	8	0.839	0.829	0.020	1.941	0.039	0.892	1.306	0.219	0.366
mMiCIR_9	6	6	0.793	0.777	0.021	3.039	0.063	0.952	1.399	0.267	0.371
mMiCIR_18	9	9	0.838	0.829	0.018	1.471	0.026	0.891	1.196	0.152	0.274
mMiCIR_21	5	5	0.524	0.510	0.095	0.882	0.084	0.557	1.233	0.150	0.257
mMiCIR_22	7	7	0.814	0.801	0.027	1.647	0.044	0.865	1.296	0.213	0.358
mMiCIR_25	5	5	0.675	0.655	0.065	1.176	0.076	0.717	1.305	0.204	0.340
MiSHRS _29	5	5	0.584	0.561	0.083	1.588	0.132	0.621	1.273	0.193	0.325
mMiCIR_29	9	9	0.827	0.817	0.019	1.765	0.034	0.878	1.245	0.177	0.305
mMiCIR_30	6	6	0.782	0.766	0.036	1.765	0.064	0.831	1.385	0.256	0.411
MIAC_2	5	5	0.692	0.666	0.062	1.353	0.083	0.735	1.360	0.232	0.371
MIAC_3	7	7	0.752	0.733	0.035	1.294	0.046	0.799	1.235	0.166	0.285
MIAC_4	4	4	0.737	0.713	0.053	1.529	0.081	0.783	1.511	0.332	0.513
MIAC_5	10	10	0.860	0.852	0.014	1.647	0.023	0.913	1.197	0.154	0.276
MIAC_6	9	9	0.831	0.821	0.019	1.765	0.033	0.883	1.245	0.178	0.306
Average	6.4	6.4	0.752	0.735	0.043	1.664	0.067	0.803	1.340	0.227	0.369
Polymorphism 10	0%										

Table 5 Characteristics and studied indices of SSR primers used in the study of genetic diversity of mango cultivars.

Notes.

NA, Total bands; P, polymorphic bands; He, Expected heterozygosity; PIC, Polymorphism information content; EMR, Effective multiplex ratio; Havp, Mean heterozygosity; MI, Marker index; D, Discriminating power; NEA, Number of effective alleles; H, Nei's gene diversity; S, Shannon's Information index.

The effective multiple ratio (EMR), which indicates the number of polymorphic gene loci in germplasm, ranged from 0.824 for OPX_02 to 11.059 for OPQ_05 of the RAPD markers with an average of 5.322, while for the SSR markers, it ranged from 0.882 to 3.039 for the MiCIR_21 and mMiCIR_9, respectively, with an average of 1.664 (Tables 4 and 5).

The marker index (MI) was developed to evaluate how well the markers could detect polymorphisms. The highest value of the RAPD marker index was 0.017 for the following four markers: OPX_18, OPD_13, OPG_11, and OPA_01, while the lowest value was 0.007 for OPA_18 and an average of 0.013 (Table 4). The SSR markers had an average marker

index of 0.067, with the MIAC_5 having the lowest value (0.023) and the MiCIR_36 having the highest value (0.146) (Table 5).

Discrimination power varied from 0.503 for OPA_01 to 0.911 for OPA_18 with an average of 0.795 for the RAPD markers, and for the SSR markers, it varied from 0.557 to 0.952 for MiCIR_21 and mMiCIR_9, respectively with an average of 0.

The number of effective alleles for the RAPD markers ranged from 1.359 (OPX_01) to 1.693(OPA_01), and the mean in the study population was 1.525 Table 4). On the other hand, it ranged from 1.196 (mMiCiR_18) to 1.58 (MiSHRS_18), and its mean in the study population was 1.34 for the SSR markers (Table 5).

The Nei gene diversity index (H) is one of the most crucial metrics for assessing gene diversity among populations and cultivars. The RAPD markers' H values range from 0.232 to 0.386 and an average of 0.314. The OPX_01 displayed the lowest amount of H, and the OPA_01 displayed the greatest value of H (Table 4). SSR markers had an average value of H (0.226), with mMiCIR_21 recording the lowest value (0.150) and MiSHRS_18 recording the highest value (0.334). The Shannon index shows the level of variation between cultivars. The mean Shannon index of RAPD markers in this study was 0.481.

OPA_01 had the highest value (0.562), followed by OPM_12 and OPO_12 (0.560– 0.553). The OPX_01 value (0.373) was the lowest. The mean Shannon index for SSR markers was 0.369. The MIAC_4 and MiSHRS_18 showed the highest values (0.513 and 0.495. respectively), while the mMiCIR_21 had the lowest value (0.257).

Power of the probability of identity

According to *Mirimin et al. (2015)*, the probability of identify (PI) represents the possibility of discovering two people with the same genotype at particular loci in the population. The SSR probability value ranged from 0.475 (mMiCIR_21) to 0.14 (MIAC_5), and was typically low, with an average of 0.25. The RAPD markers' probability value was very low and ranged from 0.224 (OPK_10) to 0.053 (OPQ_05), with a mean of 0.109 (Table 6).

Genetic similarity and cluster analysis

The pairwise comparison of cultivars based on simple matching similarity coefficients indicated likely high genetic similarity between the 17 mango cultivars, ranging from a maximum of 0.83 for 'Hindi Besennara' and 'Hindi mlooky' to a minimum of 0.58 for 'Zebda' and 'Elwazza neck' for the SSR markers and from a maximum of 0.80 for the 'Taymour', 'Elwazza neck', and 'Sukkary white' to a minimum of 0.65 for the 'Keitt' and 'Ewais' for the RAPD markers. A dendrogram was generated from the binary data of the SSR marker score results based on the simple matching similarity coefficients, as shown in Fig. 2A. The dendrogram showed that the genetic similarity coefficient of the 17 mango cultivars ranged from 0.58 to 0.83. It can be seen that the 17 mango cultivars were divided into three clusters, with a mean similarity of 0.66 for cluster 1 (11 cultivars) and cluster 2 (5 cultivars). The third cluster contained only the 'Zebda' cultivar. The first cluster was divided into two groups with a mean of 0.67. Group 1 had two cultivars, 'Banarasi Langra' and 'Fajri kalan', with a similarity coefficient 0.72. Group 2 was divided into two subgroups, A and B, with a similarity coefficient of 0.68. Subgroup A had seven cultivars, and Subgroup

Table 6	Table 6 values of probability of identify for SSR and RAPD markers.							
No.	SSR markers	Probability of identify	No.	RAPD markers	Probability of identify			
1	MiSHRS_1	0.233	1	OPM_20	0.066			
2	MiSHRS_4	0.280	2	OPQ_05	0.053			
3	MiSHRS_18	0.318	3	OPS_17	0.08			
4	MiSHRS_32	0.250	4	OPD_20	0.075			
5	mMiCIR_36	0.381	5	OPO_12	0.199			
6	MiSHRS_48	0.255	6	OPA_13	0.126			
7	LMMA_1	0.171	7	OPB_11	0.149			
8	LMMA_7	0.214	8	OPJ_20	0.091			
9	LMMA_8	0.230	9	OPM_06	0.079			
10	LMMA_9	0.251	10	OPA_09	0.107			
11	LMMA_10	0.217	11	OPX_18	0.093			
12	LMMA_11	0.193	12	OPM_12	0.092			
13	LMMA_13	0.251	13	OPD_13	0.066			
14	LMMA_14	0.206	14	OPM_15	0.091			
15	LMMA_15	0.281	15	OPG_11	0.161			
16	mMiCIR_5	0.248	16	OPH_18	0.073			
17	mMiCIR_8	0.161	17	OPB_17	0.154			
18	mMiCIR_9	0.207	18	OPA_04	0.089			
19	mMiCIR_18	0.162	19	OPE_12	0.082			
20	mMiCIR_21	0.475	20	OPA_01	0.113			
21	mMiCIR_22	0.186	21	OPX_08	0.072			
22	mMiCIR_25	0.325	22	OPX_17	0.091			
23	MiSHRS _29	0.416	23	OPX_02	0.191			
24	mMiCIR_29	0.173	24	OPX_01	0.182			
25	mMiCIR_30	0.217	25	OPA_18	0.063			
26	MIAC_2	0.308	26	OPA_07	0.107			
27	MIAC_3	0.248	27	OPA_11	0.089			
28	MIAC_4	0.263	28	OPW_11	0.060			
29	MIAC_5	0.140	29	OPC_05	0.139			
30	MIAC_6	0.168	30	OPK_10	0.224			
Sum		7.428	Sum		3.26			
Average		0.25	Average	2	0.109			

 Table 6
 Values of probability of identify for SSR and RAPD markers.

B had two cultivars, Ewais and Nabiel. The second cluster contained five cultivars, and two of the cultivars, 'Hindi Besennara' and 'Hindi mlooky', achieved a similarity coefficient of 0.83.

The dendrogram was generated from the binary data of the RAPD marker scoring results based on similarity coefficients for simple matches, as shown in Fig. 2B. From the dendrogram, it can be seen that the genetic similarity coefficient of the 17 mango cultivars varied from 0.68 to 0.80. In this dendrogram, the 17 mango cultivars were divided into three clusters with a similarity coefficient of 0.69, namely, cluster 1 (12 cultivars), cluster 2 (four cultivars), and cluster 3 (one cultivar). The first cluster was separated into two





Full-size DOI: 10.7717/peerj.15722/fig-2

groups with a similarity coefficient of 0.71. Group1 had seven cultivars, 'Banarasi Langra', 'Mabrouka', 'Ewais', 'Companeit Elsowwa', 'Fajri kalan', 'Mstikawy' and 'Nabiel', with a mean similarity of 0.72.

Group 2 had five cultivars, 'Taymour', 'Elwazza neck', 'Hindi mlooky', 'Sukkary white', and 'Hindi Besennara'. The 'Tayimour' and 'Elwazza neck' cultivars revealed a mean similarity of 0.80. In the second cluster, there were two groups. Group 1 contained 'Mesk' and 'Senarry' and had a coefficient of 0.73. Group 2 consisted of 'Keitt' and 'Kent' and had a similarity coefficient of 0.74. Zebda, the sole cultivar in the third cluster, was the most diverse of the 17 cultivars and appeared as an outlier for both the SSR and RAPD markers.

Dice similarity values were generated for the 30 SSR markers in the 17 mango cultivars (not shown). The Dice pairwise similarity coefficients and similarity values varied from 0.22 between 'Mstikawy' and 'Banarasi Langra' to 0.67 between 'Hindi mlooky' and 'Hindi Besennara', and from a minimum of 0.66 for 'Ewais', 'Sennary', and 'Keitt' to a maximum of 0.81 for 'Elwazza neck' and 'Taymour' for the RAPD markers.

A UPGMA dendrogram was generated using Dice similarity coefficients and applied for SSR markers, as shown in Fig. 3A. The 17 mango cultivars were divided into three main clusters. Cluster 1 contained 'Banarasi Langra' and 'Fajri kalan', with a similarity coefficient of 0.45. Cluster 2 was divided into two subgroups. 'Mabrouka' and 'Mesk' were in group 1. The second group had three subgroups. Subgroup A contained 'Ewais' and 'Nabiel'; subgroup B had 'Companeit Elsowwa', 'Kent', and 'Mstikawy'; and subgroup C had 'Keitt' and 'Senarry'. The 'Zebda' cultivar was the only cultivar in the second group. Cluster 3 consisted of two groups. Group one contained 'Taymour', 'Elwazza neck', and 'Sukkary white'. 'Taymour' and Elwazza neck had similar coefficients of 0.61. 'Hindi mlooky' and 'Hindi Besennara' were tied in the second group, with a coefficient of 0.67.

The dendrogram was generated from the binary data of the RAPD marker scoring results based on the Dice similarity coefficients, as shown in Figs. 3B. The dendrogram shows that the value of the genetic similarity coefficient of the 17 mango cultivars varied from 0.70 to 0.81. The 17 mango cultivars were separated into three clusters with a coefficient of 0.70; they were Cluster 1 (12 cultivars), cluster 2 (four cultivars), and Cluster 3 (one cultivar). The first cluster was divided into three groups with a similarity coefficient of 0.71. Group 1 consisted of six cultivars, 'Banarasi Langra', 'Mabrouka', 'Ewais', 'Companeit Elsowwa', 'Mstikawy', and 'Nabiel', with a similarity coefficient of 0.73. Group 2 had five cultivars, 'Taymour', 'Elwazza neck', 'Hindi mlooky', 'Sukkary white', and 'Hindi Besennara'. The cultivar 'Fajri kalan' was the only one in the third group.

The 'Taymour' and 'Elwazza neck' cultivars had a coefficient of 0.81. The second cluster was divided into two groups. Group 1 contained 'Mesk' and 'Sennary', with a coefficient of 0.73. Group 2 consisted of 'Keitt' and 'Kent', with a mean similarity of 0.75. The third cluster contained 'Zebda'. Notably, 'Zebda' was the most diverse of the 17 cultivars, appearing as an outlier in the UPGMA dendrogram for the SSR and RAPD data.

Population structure

The population structure was determined based on the model presented in the structure software. Based on the data, the suitable number of groups for the population was two to 10. Using the structure harvester website for the best k, four subgroups were selected for the population (Fig. 4). The cluster analysis was based on the Bayesian statistical model to understand the populations' distance structure. Assuming that the lineage model was of a mixed type and the allelic abundance model was of a continuous type, the results showed that there were four populations of germplasms based on cultivar and genome. These cultivars were not completely separate (Fig. 5), and each cultivar was assigned to each group with almost equal probabilities.



Figure 3 Genetic similarity dendrogram (Dice coefficient) based on unweighted pair-group method using arithmetic average (UPGMA) analysis and constructed of SSR marker (A) and RAPD marker (B) shows the relationships among the 17 mango cultivars.

0.76

0.78

Full-size DOI: 10.7717/peerj.15722/fig-3

7 Zebda

0.81

DISCUSSION

Cluster 3

0.73

0.70

The genetic diversity of 17 mango cultivars was investigated in this work using two PCR-based systems: SSR and RAPD. The kind and quantity of polymorphisms identified vary depending on the system in principle. Despite the number of polymorphic alleles generated by 30 SSR markers being different compared to the number of polymorphic bands generated by RAPD markers, they recognized 100% of the polymorphisms. The differences in polymorphisms of the DNA markers generated by each primer indicated the complexity of the plant genome, as the DNA bands resulted from the binding of primer nucleotides in plant chromosomes. However, the number of polymorphic DNA bands was



Figure 4 Determining subgroups using Structure Harvester.

Full-size DOI: 10.7717/peerj.15722/fig-4





Full-size DOI: 10.7717/peerj.15722/fig-5

able to indicate the genomic profile of the mango species due to the distribution of the primer binding sites (*Fitmawati Hartana & Purwoko, 2010*).

Nevertheless, a substantial percentage of polymorphism from 30 RAPD markers was still present, and this was deemed sufficient and particularly instructive in calculating genetic diversity. Previous studies of mangoes have reported similar levels of polymorphisms associated with RAPD markers (*Islam et al., 2018*). According to the banding patterns obtained from RAPD and SSR markers, the 17 tested mango cultivars could be distinguished from each other. These results confirmed previous findings that the mango is highly heterozygous (*Sherman, Rubinstein & Eshed, 2015*). This could be explained by the fact that the mating system in mangoes typically involves outcross pollination with some self-pollination, and as a result, the cultivars employed acquired enough free external gene flow (*Kumari et al., 2020*).

The PIC was determined by the quantity of alleles found, their frequency of distribution, the placement of the study's primers in the genome, and the sensitivity of the genotype to the technique employed (*Pachauri et al., 2013; Abd El-Moneim et al., 2021*). The 30 RAPD markers utilized in the 17 mango cultivars had a moderate polymorphism, as indicated by the mean of PIC value. This result was greater than the mean showed by *Srivastava et al. (2012)*. However, the average PIC value of RAPD markers studied by various studies differs depending on how many RAPD markers were utilized and how many cultivars were examined. The average PIC value of the present analysis supports previously published mean PIC values for SSR markers in mango (*Ravishankar et al., 2015*; and *Padmakar, Dinesh & Ravishankar, 2017*). Allelic variation may be revealed more readily with a high PIC value. In contrast to RAPD markers, the PIC values for SSR markers were obviously greater.

It is important to note that the high mutational rate at SSR loci, which is impacted by the structure, number of repeated nucleotides, and type of the locus (g-SSR or est-SSR), may potentially be a contributing factor to the level of genetic variability. While most of the amplified SSR loci in the current investigation are based on dinucleotide repeats, loci with a small number of repeated units display a high mutational rate, as *Gadaleta et al. (2007)* described. Because of this, a heterozygous state could be produced by any of the mutated alleles. Consequently, based on the PIC value in this investigation, the SSR markers were more discriminative than the RAPD markers.

Probability of identity (PI) is the probability that two individuals in a population or sample will have the same genotype by chance rather than by relationship (*Mirimin et al., 2015*). The average PI values resulting from SSR markers were small compared to the RAPD markers, which were very small. This indicated that a small number of identical alleles were discovered in the mango cultivars analyzed. These results matched those of *Ravishankar et al. (2015*), who identified 387 mango accessions using six microsatellite markers. They suggested that the probability of finding two individuals with the same genotype was very close to zero, considering both the marker set and the sample size utilized. However, the PI is considered the most widely used theoretical estimator for accessing supporting statistics for individual identification and quantifying the level of genetic variability in populations or samples (*Dokupilová et al., 2014*; *Mollet et al., 2015*).

Compared to RAPD data, the SSR data in this investigation produced lower similarity values. This was primarily due to the co-dominant character of the SSR markers. Which made it possible to identify many alleles per locus and led to larger levels of anticipated heterozygosity than would be conceivable with RAPD markers. More dinucleotide-type

SSRs were employed than other types, and their mutation rates were considerably higher (*Islam et al., 2018* and *Padmakar, Dinesh & Ravishankar, 2017*).

The SSR data exhibited that the higher the value of the similarity coefficient between the cultivars 'Hindi Besennara' and 'Hindi mlooky' (and they are in the same place in the dendrogram clusters), the more similar the DNA banding patterns would become between the cultivars, meaning that the cultivars were becoming more and more similar. In contrast, the RAPD data showed that 'Taymour' and 'Elwazza neck' were in the same position in the dendrogram clusters. This could not be related to the nature of the used similarity coefficients set but somewhat to a limited number of genomic regions where varieties differed. Such differences could not be assessed with a few markers (*Denčić et al., 2016*).

The selection of the similarity coefficient must be based on several criteria because even a few structural changes in more differentiated groups can change the relationship between varieties with high genetic similarity. Considering the genetic basis of the RAPD markers (*Williams et al.*, 1990), the lack of amplification of a particular band in two cultivars did not necessarily represent a genetic similarity between them, so coefficients that exclude these common negative occurrences in their expressions of similarity (*e.g.*, Dice, simple matching) are better suited for use with this type of marker.

Remarkably, all the coefficients showed that the 'Zebda' cultivar was clearly separate from the other mango cultivars in the SSR and RAPD groups. There is no doubt that this cultivar is genetically far from other cultivars, which is essential in breeding and improving mangoes. The reason for this is probably its segregation from the base population in the first selection phase. The dendrograms for the SSR and RAPD markers were not affected by the coefficient type, even by simple matching or Dice. The UPGMA method can provide consistent results in terms of clustering regardless of the similarity coefficient.

Furthermore, the dendrograms generated from the examined coefficients all showed the same general structure (Figs. 2 and 3), so it is evident that the different coefficients caused few changes depending on the type of markers. Nevertheless, different groups were formed for the SSR markers than for the RAPD markers, which is illustrated by the several properties of these markers (*Garcia et al., 2004*). Consequently, the dendrogram for either the SSR markers or RAPD markers in Figs. 2 and 3 obtained by the simple matching similarity coefficient with a single hit was identical to that obtained by Dice.

The 17 cultivars in each of the four subpopulations were almost identical to the other subgroups in other parts of their genome, which is why all mango cultivars were placed in a mixed structure (the probability of each cultivar belonging to each subgroup was less than 0.7). No cultivar was definitely in a specific group. Due to gene flow over time and the possibility of shared genetic ancestry amongst cultivars, this mixing may have occurred. The type of markers used or the fact that the chosen markers were not evenly dispersed across the genome could also be to blame. Additionally, some of these cultivars originated in America and India, while the remainder had been chosen and developed in Egypt. As a result, these chosen cultivars were genetically diverse, most of which were crossbred.

CONCLUSIONS

The extent of mango polymorphism has been usefully shown by both the RAPD and SSR marker approaches. They are more helpful in evaluating the genetic diversity of the studied cultivars. Although reliability and transferability are two drawbacks of RAPD-based analysis, RAPD data can become highly reliable, provided a set methodology is followed. However, the findings showed that RAPD and SSR markers systems effectively classify the 17 mango cultivars based on where they were first cultivated. Interestingly, it is essential to emphasize that SSRs demonstrated superior performance by displaying higher values for most of the parameters that determine the potential of markers in diversity analysis. Also, it became clear that this investigation's simple matching or Dice similarity coefficients had no bearing on the outcomes. Furthermore, the population structure analysis data will be valuable for conducting association mapping in mango for a cultivar of characteristics.

On the other hand, it became evident that the 'Zebda' cultivar differed genetically from the other mango cultivars evaluated and was further distinct from them. It will probably participate in breeding and enhancement efforts for mangoes in Egypt as a good parent. In particular, considering that it produces fruit of high quality and is resistant to diseases and unfavorable circumstances in the environment. It is important to note that every result from this study will be helpful information for identifying markers for future studies, describing germplasm, breeding, and managing mango germplasm.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

The authors received support from The Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, Grant No. [IFPRC-187-130-2020]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors: The Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah: IFPRC-187-130-2020.

Competing Interests

Diaa Abd El-Moneim is an Academic Editor for PeerJ.

Author Contributions

- Mohammed A.A. Hussein conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Manal Eid conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.

- Mehdi Rahimi conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Faten Zubair Filimban conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Diaa Abd El-Moneim conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.

Data Availability

The following information was supplied regarding data availability: The raw data is available in the Supplemental Files.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.15722#supplemental-information.

REFERENCES

- Abd El-Moneim D, Elsarag EIS, Aloufi S, El-Azraq AM, Alshamrani SM, Safhi FAA, Ibrahim AA. 2021. Quinoa (*Chenopodium quinoa* Willd.): genetic diversity according to ISSR and SCoT markers, relative gene expression, and morpho-physiological variation under salinity stress. *Plants* 10:2802 DOI 10.3390/plants10122802.
- Abdelsalam NR, Ali HM, Salem MZ, Ibrahem EG, Elshikh MS. 2018. Genetic and morphological characterization of *Mangifera indica* L. growing in Egypt. *HortScience* 53:1266–1270 DOI 10.21273/HORTSCI13084-18.
- Adhikari, Sinchan. 2015. Development of molecular signature in three cultivated species of Cymbopogon and identification of sex-linked markers in Trichosanthes dioica roxb and *Cucumis sativus* L based on RAPD_ISSR and AFLP analysis. *Available at http://hdl.handle.net/10603/195853*.
- Avramidou E, Ganopoulos I, Mylona P, Abraham EM, Nianiou-Obeidat I, Osathanunkul M, Madesis P. 2023. Comparative analysis of the genetic diversity of Faba Bean (*Vicia faba* L). *Sustainability* 15:1016 DOI 10.3390/su15021016.
- Bally IS, Bombarely A, Chambers A, Cohen Y, Dillon N, Innes D. 2021. The 'Tommy Atkins' mango genome reveals candidate genes for fruit quality. *BMC Plant Biology* 108:1–18.
- **Bassam BJ, Caetano-Anollés G. 1993.** Silver staining of DNA in polyacrylamide gels. *Applied Biochemistry and Biotechnology* **42**:181–188 DOI 10.1007/BF02788051.
- **Botstein D, White RL, Skolnick M, Davis RW. 1980.** Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* **32**:314–331.

- Bukhari B, Mehvish A, Babar M, Raza S, Ullah SI, Qureshi MA, Zafar MS, Naz A, Rauf S, Shaukat FR, Rauf N. 2022. Genetic diversity estimating among thirteen late maturing mango varieties using SSR markers. *Plant Cell Biotechnology and Molecular Biology* 23(21–22):32–41.
- Denčić S, Depauw R, Momčilović V, Aćin V. 2016. Comparison of similarity coefficients used for cluster analysis based on SSR markers in sister line wheat cultivars. *Genetika* 48:219–232 DOI 10.2298/GENSR1601219D.
- Deshpande AB, Anamika K, Jha V, Chidley H, Oak P, Kadoo N, Pujari K, Giri A, Gupta V. 2017. Transcriptional transitions in Alphonso mango (*Mangifera indica* L.) during fruit development and ripening explain its distinct aroma and shelf life characteristics. *Scientific Reports* 7(1):8711 DOI 10.1038/s41598-017-08499-5.
- **Dokupilová I, Migliaro D, Mihálik D, Crespan M, Kraic J. 2014.** Genotyping of Vitis vinifera L. within the Slovak national collection of genetic resources. *Central European Journal of Biology* **9**:761–767.
- **Duval M-F, Bunel J, Sitbon C, Risterucci A-M. 2005.** Development of microsatellite markers for mango (Mangifera indica L.). *Molecular Ecology Notes* **5**:824–826 DOI 10.1111/j.1471-8286.2005.01076.x.
- **Earl DA, von Holdt BM. 2012.** STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**:359–361 DOI 10.1007/s12686-011-9548-7.
- Eiadthong W, Yonemori K, Kanzaki S, Sugiura A, Utsunomiya N, Subhadrabandhu S. 2000. Amplified fragment length polymorphism analysis for studying genetic relationships among Mangifera species in Thailand. *Journal of the American Society for Horticultural Science* 125:160–164 DOI 10.21273/JASHS.125.2.160.
- Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14:2611–2620 DOI 10.1111/j.1365-294X.2005.02553.x.
- **Fitmawati Hartana A, Purwoko BS. 2010.** Diversity of Indonesian mango (*Mangifera indica*) cultivars based on morphological and RAPD markers. *SABRAO Journal of Breeding and Genetics* **42**:84–95.
- Gadaleta A, Mangini G, Mule G, Blanco A. 2007. Caracterization of dinucleotide and trinucleotide EST-derived microsatellites in the wheat genome. *Euphytica* 153(1–2):73–85 DOI 10.1007/s10681-006-9243-z.
- Garcia AA, Benchimol LL, Barbosa AM, Geraldi IO, Souza Jr CL, Souza APd.
 2004. Comparison of RAPD, RFLP, AFLP and SSR markers for diversity studies in tropical maize inbred lines. *Genetics and Molecular Biology* 27:579–588
 DOI 10.1590/S1415-47572004000400019.
- Honsho C, Nishiyama K, Eiadthong W, Yonemori K. 2005. Isolation and characterization of new microsatellite markers in mango (*Mangifera indica*). *Molecular Ecology Notes* 5:152–154 DOI 10.1111/j.1471-8286.2004.00866.x.

- Islam SR, Karuna K, Kumari A, Prasad BD, Mir H, Kumar P. 2018. Evaluation of genetic diversity among mango (*Mangifera indica* L.) hybrids and their parentage using RAPD markers. *Chemical Science Review and Letters* 7:513–519.
- Karihaloo J, Dwivedi Y, Archak S, Gaikwad AB. 2003. Analysis of genetic diversity of Indian mango cultivars using RAPD markers. *The Journal of Horticultural Science* and Biotechnology 78:285–289 DOI 10.1080/14620316.2003.11511619.
- Kumari S, Chakraboraty SK, Bhowmick PK, Singh VJ, Prased ASM. 2020. validation of hybride rice (*Oryza sativa* L.) seed vigour traits using SSR markers. *Indian Journal of Genetics* 80(2):204–208 DOI 10.31742/IJGPB.80.2.1.
- Liu BH. 1998. *Statistical genomics: linkage, mapping and QTL analysis*. Boca Raton: CRC Press.
- Luo C, He X, Chen H, Ou S, Gao M, Brown J, Tondo C, Shenell R. 2011. Genetic diversity of mango cultivars estimated using SCoT and ISSR markers. *Biochemical Systematics and Ecology* 39:676–684 DOI 10.1016/j.bse.2011.05.023.
- Meyer AdS, Garcia AAF, Souza APd, Souza Jr CLd. 2004. Comparison of similarity coefficients used for cluster analysis with dominant markers in maize (*Zea mays* L). *Genetics and Molecular Biology* 27:83–91 DOI 10.1590/S1415-47572004000100014.
- Mirimin L, Brett M, Sven K, Stephen L, Aletta B, Paul C, Paulette B, Rouvay R. 2015. Genetic analyses reveal declining trends and low effective population size in an overfished South African sciaenid species, the dusky kob (Argyrosomus japonicus). *Marine and Freshwater Research* **67**:266–276.
- Mollet P, Kery M, Gardner B, Pasinelli G, Royle JA. 2015. Estimating population size for capercaillie (*Tetrao urogallus* L.) with spatial capture-recapture models based on genotypes from one field sample. *PLOS ONE* 10:e0129020 DOI 10.1371/journal.pone.0129020.
- Nei M. 1987. Molecular evolutionary genetics. New York: Columbia University Press.
- Pachauri V, Taneja N, Vikram P, Singh NK, Singh S. 2013. Molecular and morphological characterization of Indian farmers rice varieties (*Oryza sativa* L.). *Australian Journal of Crop Science* 7:923–932.
- Padmakar B, Dinesh M, Ravishankar K. 2017. Marker-trait association for fruit characters in mango (*Mangifera Indica* L.) cultivars. *Journal of Horticultural Sciences* 11:170–178.
- Peakall R, Smouse PE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28:2537–2539 DOI 10.1093/bioinformatics/bts460.
- **Porebski S, Bailey LG, Baum BR. 1997.** Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Molecular Biology Reporter* **15**:8–15 DOI 10.1007/BF02772108.
- **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**:225–238 DOI 10.1007/BF00564200.

- Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959 DOI 10.1093/genetics/155.2.945.
- Ramadasappa P, Mallaiah S, Ramakrishnaiah TN. 2021. Fluorescein Amidite (FAM) based RAPD barcodes for genetic analyses of mango (*Mangifera indica* L.) genotypes. *Advances in Bioresearch* 12(5 September):70–77.
- Ravishankar KV, Bommisetty P, Bajpai A, Srivastava N, Mani BH, Vasugi C, Rajan S, Dinesh MR. 2015. Genetic diversity and population structure analysis of mango (*Mangifera indica*) cultivars assessed by microsatellite markers. *Trees* 29:775–783 DOI 10.1007/s00468-015-1155-x.
- **Reif JC, Melchinger AE, Frisch M. 2005.** Genetical and mathematical properties of similarity and dissimilarity coefficients applied in plant breeding and seed bank management. *Crop Science* **45**:1–7.
- **Rohlf F. 2000.** NTSYS-pc: numeric taxonomy and multivariate analysis system. Version 2.2. New York: Exeter Software, Setauket.
- Schnell R, Olano C, Quintanilla W, Meerow A. 2005. Isolation and characterization of 15 microsatellite loci from mango (*Mangifera indica* L.) and crossspecies amplification in closely related taxa. *Molecular Ecology Notes* 5:625–627 DOI 10.1111/j.1471-8286.2005.01018.x.
- Sherman A, Rubinstein M, Eshed R. 2015. Mango (*Mangifera indica* L.) germplasm diversity based on single nucleotide polymorphisms derived from the transcriptome. *BMC Plant Biology* 15:277 DOI 10.1186/s12870-014-0410-4.
- **Spangelo L, Hsu C, Fejer S, Watkins R. 1971.** Inbred line × tester analysis and the potential of inbreeding in strawberry breeding. *Canadian Journal of Genetics and Cytology* **13**:460–469 DOI 10.1139/g71-070.
- Srivastava N, Bajpai A, Chandra R, Rajan S, Muthukumar M, Srivastava MK. 2012. Comparison of PCR based marker systems for genetic analysis in different cultivars of mango. *Journal of Environmental Biology* **33**:159–166.
- **Tessier C, David J, This P, Boursiquot JM, Charrier A. 1999.** Optimization of the choice of molecular markers for varietal identification in *Vitis vinifera* L. *Theoretical and Applied Genetics* **98**:171–177 DOI 10.1007/s001220051054.
- Viruel M, Escribano P, Barbieri M, Ferri M, Hormaza J. 2005. Fingerprinting, embryo type and geographic differentiation in mango (*Mangifera indica* L. Anacardiaceae) with microsatellites. *Molecular Breeding* 15:383–393 DOI 10.1007/s11032-004-7982-x.
- Wang P, Luo Y, Huang J. 2020. The genome evolution and domestication of tropical fruit mango. *Genome Biology* 21(1):-60 DOI 10.1186/s13059-019-1906-x.
- Williams JG, Kubelik AR, Livak KJ, Rafalski JA, Tingey SV. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* 18:6531–6535 DOI 10.1093/nar/18.22.6531.
- Yeh FC. 1999. POPGENE version 1.31: Microsoft Windows–based freeware for population genetic analysis, quick user guide. Edmonton: Center for International Forestry Research, University of Alberta.