Magna Scientia Advanced Research and Reviews

eISSN: 2582-9394

Cross Ref DOI: 10.30574/msarr

Journal homepage: https://magnascientiapub.com/journals/msarr/

(RESEARCH ARTICLE)

Check for updates

Genetic diversity of selected *Moringa oleifera* lam. Provenances from the coastal region of Kenya

Sarah Kwamboka Ondieki [1, *], Carol W. Hunja [2], Geoffrey M. Muluvi [2], Jackson C. Korir [1], Festus Mutiso [3], Johnson U. Kitheka [3], Danson Kioko [3], Josphert Kimatu [2], James K. Ndufa [4] and Kennedy Mutati [5]

[1] *School of Natural and Applied Sciences, Department of Biological Sciences, Masinde Muliro University of Science and Technology, P.O Box 190-50100 Kakamega, Kenya.*
[2] *School of Science and Computing, South Eastern Kenya University, P. O Box 170-90200 Kitui, Kenya.*
[3] *School of Agriculture, Environment, Water and Natural Resources, South Eastern Kenya University, P.O Box 170-90200 Kitui, Kenya.*
[4] *Kenya Forestry Research Institute, P.O. Box 20412-00200 Nairobi, Kenya.*
[5] *Department of Water & Agriculture, County Government of Kitui, P.O. Box 33-90200 Kitui, Kenya.*

Magna Scientia Advanced Research and Reviews, 2023, 08(02), 118–128

Publication history: Received on 13 June 2023; revised on 22 July 2023; accepted on 25 July 2023

Article DOI: https://doi.org/10.30574/msarr.2023.8.2.0113

Abstract

Moringa oleifera Lam (syn. *M. Pterygosperma Gaertn*) is one of the most economically important crops widely cultivated globally. Traditionally, it has been used as a remedy for conditions such as diabetes, inflammation, cancer, bacterial, viral, and fungal infections, joint pain and heart disease. *Moringa oleifera* has also been utilized for human consumption, as forage and for water purification. Considering it is a cross pollinated plant, *M. oleifera* has huge variability. Due to the fast-growing nature and adaptability to variable climatic conditions, *M. oleifera* has been popularly cultivated in Coastal and Eastern Kenya. Studying the genetic diversity of the drumstick tree is essential in selecting valuable genotypes and for improvement of its cultivars for improved breeding programs. The use of marker-assisted selection of cultivar for the desired traits will benefit faster breeding programs. Single nucleotide polymorphisms (SNPs), being the most recent molecular markers, were used in this study. Seventeen *M. oleifera* provenances from Coastal Kenya were analyzed by genotyping by sequencing (GBS). Briefly, *M. oleifera* leaves were homogenized, crushed and DNA extracted, genome complexity reduction and library preparation was carried out followed by barcoding, electrophoresis, pooling, preparation of sequencing libraries and sequencing. Next generation sequencing (NGS) technologies, Illumina Hiseq 2500, was used for sequencing and SNP mining. Binning was carried out to remove noise. *In silico* assembly of sequence contigs was undertaken. Data analysis was undertaken using DArT R and KDCompute Platforms. Genetic characterization was achieved by carrying out cluster analysis, principle coordinate analysis (PCoA), 3D plot and phylogenetic tree. One hundred and sixty four (164) genotypes were identified. It was evident from the analysis undertaken that the genotypes consistently clustered into four clades/groups. However, the clusters obtained by each analysis tool comprised different genotypes due to the assumptions employed by each tool. The similarity coefficient from Hierarchical clustering ranged from 63% to 100% indicating that the genotypes had low variability. An interactive and highly robust 3D plot was produced using DArT R (Adegenet package). Interestingly, Pwani University genotypes clustered separately. It is noteworthy that these genotypes consistently clustered uniquely and independently irrespective of the tool used. This is possibly due to the seeds being of unique origin. Nevertheless, the high similarity between the genotypes could be attributed to the *M. oleifera* plants in the various provenances having the same ancestry. Given the high frequency of SNPs and their involvement as a source of allele variations, this research could contribute to the discovery of associations between gene allelic forms and phenotypes, enabling the alleles to be linked to desirable traits (fast growth and high seed production).

* Corresponding author: Ondieki Kwamboka Sarah

Keywords: Moringa; Genetic variation; GBS; Kenyan provenances

1. Introduction

Moringa oleifera Lam (syn. *M. Pterygosperma Gaertn*) is a member of the Moringaceae, a monogenetic family of shrubs and trees (Ramachandran et al., 1980); (Stephenson & Fahey, 2004). *M. oleifera* is the most commonly cultivated *Moringa* plant species, with a diploid chromosome (2n = 28). It's used as a medicinal plant, human food, animal fodder, antimicrobial agent, and biofuel resource (Gopalakrishnan et al., 2016). Its resistance to abiotic stresses makes it an ideal tree for arid-zone cultivation (Sciences, 2015). *M. oleifera* is a nutritionally and medicinally important *Moringa* species that is grown in over 60 countries around the world (Leone et al., 2015). From India, *M. oleifera* was introduced to a number of tropical and subtropical countries (Nouman et al., 2014).

Moringa oleifera is a highly cross-pollinated crop that has amassed disparate genes from various species. Any genetic improvement and production of superior cultivars requires an understanding of genetic diversity among advanced breeding lines. Classification, conservation, and cultivar growth all require a thorough understanding of genetic variation.

(Saini et al., 2013) used molecular markers such as RAPD, ISSR, and the Cytochrome P450 marker system to evaluate genetic variation among common breeding lines of *M. oleifera*. The use of morphological and cytological methods to determine genetic diversity in breeding lines is not only time consuming, but it can also result in errors on several levels (Monica et al. 2010). Polymorphism detection at the genome level indicates the existence of variation, allowing for the effective identification of genetic diversity within a population (Wu et al., 2010); (S.-T. He et al., 2021). As a result, in recent years, the use of genetic markers for crop improvement has become increasingly prevalent (GarciaI et al. 2004). The use of molecular markers for indirect selection of traits of interest in crops is known as marker-assisted selection (MAS). MAS has been widely used in plant breeding to improve crop yield, quality, and resistance to biotic and abiotic stresses as an essential and efficient process. GBS (genotyping-by-sequencing) is a new MAS technique that can help speed up plant breeding and crop improvement (J. He et al., 2014).

Numerous studies on *Moringa oleifera* plant genetic diversity have been carried out including (J. P. da Silva et al., 2010), Brazil, (Shahzad et al., 2013), Pakistan, (Ojuederie et al., 2013), Nigeria, (Rajalakshmi et al., 2019) and (Saini et al., 2013), India and (Kleden et al., 2017) using *Moringa oleifera* from Indonesia. However, few or no robust studies have been conducted in Kenya.

Comprehending the genetic diversity within advanced breeding lines could subsequently be used to enhance varietal selection and cross breeding (Kumar et al., 2017). This study therefore determined the genetic diversity of *M. oleifera* provenances obtained from the Kenyan Coastal region to enhance the genetic information of this valuable plant in order to make informed decisions on utilization, conservation and management of this important species.

2. Material and methods

2.1. Study Area

The *M. oleifera* provenance trial was established in South Eastern Kenya University (SEKU) located 15 km off Kwa Vonza Market, along the Kitui-Machakos main road, Yatta/Kwa Vonza location, Lower Yatta Sub-County, Kitui County. The site lies at latitudes and longitudes of 037.755460 E 01.313580S respectively. The site elevation is 1173m a.s.l. The experimental design was randomized complete block design (RCBD) (Festus et al., 2022).

2.2. Experimental Design

The species was originally introduced in the Coast and Eastern region of Kenya where the provenances were selected (Fig. 1). One hundred and sixty four (164) most productive plants were selected (9 to 10 plants per provenance shown in Table 1) (Festus et al., 2022).

Table 1 No of Trees sampled per provenance

	Provenances	**No. of trees sampled**
1	Mwakiki	10
2	Pwani University	10
3	Shika Adabu	10
4	Miasenyi	10
5	Diani-Ukunda	10
6	Maungu	10
7	India	10
8	Kilifi	10
9	Likoni	10
10	Kibwezi Town	10
11	Mbololo	10
12	Gede	9
13	Arabuko Sokoke	9
14	Mackinon	9
15	Waa Kwale	9
16	Vipingo	9
17	Samburu	9
	Total	164

2.3. Sampling Method

Approximately 10 *M. oleifera* leaves per sampled tree were selected and placed in a microtube. The samples were freeze dried (lypholization) overnight using a lypholizer. The dried samples were then grinded using a Geno Grinder at 1300 rpm for 3 minutes. The grounded samples were used for DNA isolation.

2.4. Genotyping by Sequencing (GBS)

A high-throughput genotyping technique utilizing Diversity Arrays Technology Pty Ltd's (Canberra, Australia) DArT-SeqTM technology was used for sample analysis in this study. DArT-SeqTM technology and HiSeq2500 for next-generation sequencing, utilizes a complexity reduction technique to enhance genomic representations with single copy sequences. SNPs and presence-absence sequence variations are both detected using DArTSeq, which is also known as DArT-Seq markers (Raman et al., 2014).

2.5. DNA isolation/ extraction

A plant genomic DNA isolation summary by DArT was utilized based on plant User manual NucleoMag® 96 Plant May 2014. The reversible adsorption of nucleic acids to paramagnetic beads under suitable buffer conditions was the basis of the NucleoMag® 96 Plant method.

2.5.1. Agarose Gel Electrophoresis

Agarose gel electrophoresis of the DNA was utilized to show the integrity of the DNA. Quantification of the isolated DNA was carried out using 0.8% Agarose gel to visualize the bands.

2.6. Library Preparation

Through combining two restriction enzymes (Pst1-CTGCAG-, Mse I) and adapters linked by barcode to identify each sample to operate within a single lane on the Illumina HiSeq2500 instrument, a genomic representation of the samples was created. The DNA barcoded adapters were added along with DNA ligase. PST 1 enzyme was utilized as a barcode adapter while MSE 1 utilized as a common adapter. The site for MSE I was effectively PCR amplified.

Quality control for the GBS library was undertaken using 1.2 % Agarose gel electrophoresis. Equimolar amounts of the amplified fragments were pooled using Tecan, an equipment used to pool large samples. The pooled samples were then applied in the c-Bot bridging PCR (Illumina), and then sequenced using the Illumina HiSeq2500 system.

The preparation of sequencing libraries was carried out in accordance to optimized SEQART Africa protocols. Illumina Hiseq 2500 was utilized in the sequencing of the sequence DNA libraries. Seventy seven (77) cycles of sequencing (single read) were completed. The analytical program developed and patented by DArT Pvt. Ltd., Australia, was used to generate two types of data, scores for ''presence/absence'' markers (dominant), called SilicoDArTs and SNP markers (Spinoso-Castillo et al., 2020). High quality markers were identified for this particular investigation by filtering the FASTQ data based on quality parameters.

2.7. Data analysis

Polymorphism information content (PIC) was calculated using *silico* DArT. Principal component analysis (PCA) was computed based on the genetic distance matrix.

3. Results and discussion

3.1. Characterization of *M. oleifera* provenances

Through a combination of cluster analysis, PCoA and 3D Plot, the genotypes were characterized. Clustering generated varied phylogenetic trees arising from both simple and hierarchical clustering.

3.2. Simple Cluster analysis

The *M. oleifera* SNP profile was clustered into 4 clades, summarized in Table 2 and as depicted in Figure 1. Clade III had the highest number of genotypes, which amounted to 58. Clade II consisted of 47 genotypes. Clade IV had 38 genotypes whereas clade I had the least genotypes (20) as summarized in Table 2.

Table 2 A summary of the genotypes obtained in Fig. 1 by simple cluster analysis

Clade	**I**	**II**	**III**	**IV**
Provenance	Mackinon – 9 India – 6 Samburu – 2 Gede – 1 Arabuko Sokoke -1	Shika Adabu – 10 Arabuko Sokoke – 6 Gede – 6 Kibwezi – 6 Waa Kwale – 5 Samburu – 4 India – 3 Kilifi – 2 Mbololo - 2 Diani-Ukunda – 2 Vipingo - 1	Mwakiki – 10 Miasenyi – 10 Likoni – 10 Mbololo – 7 Vipingo – 5 Waa Kwale – 4 Diani-Ukunda- 3 Pwani University – 2 Kilifi – 2 Maungu – 2 Gede – 1 Samburu – 1 India - 1	Pwani University – 8 Kilifi – 6 Diani-Ukunda -5 Kibwezi Town -4 Vipingo – 3 Samburu – 2 Arabuko Sokoke -1 Mbololo - 1
Total	20	47	58	38

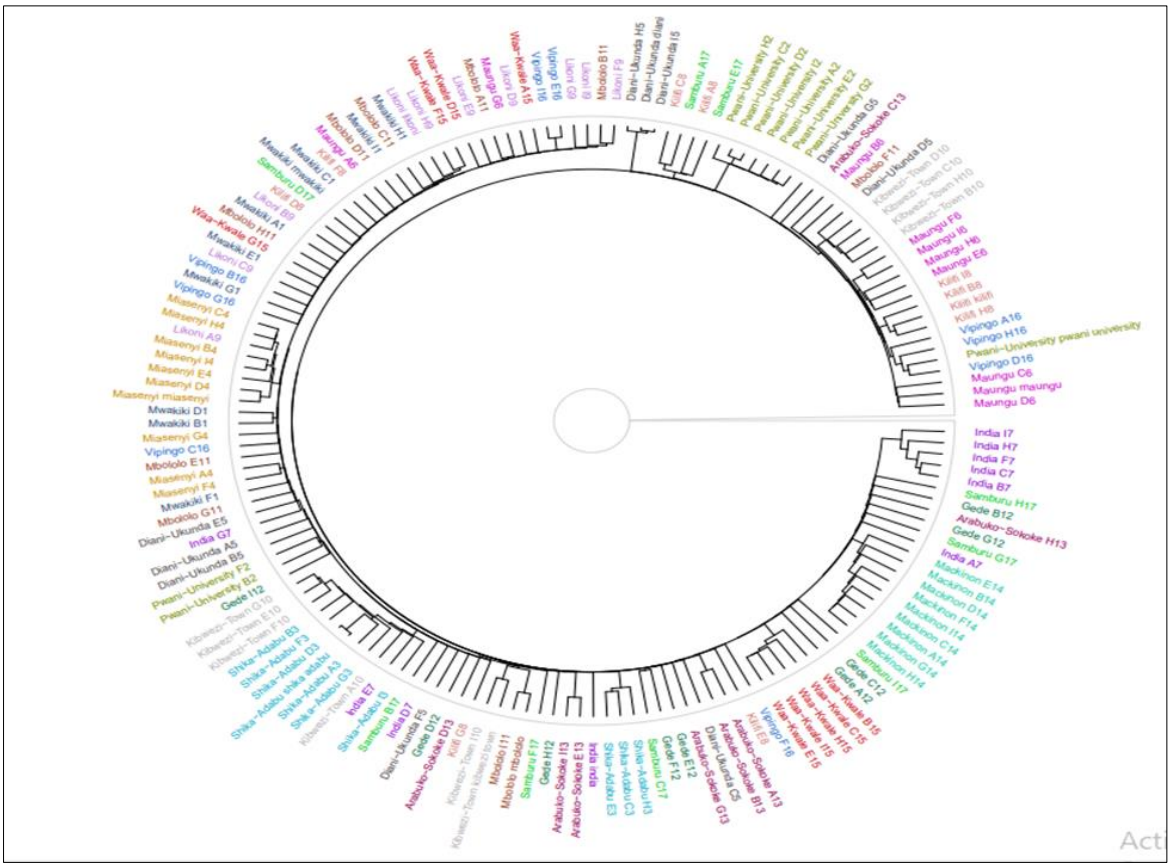

Figure 1 Phylogenetic tree showing diversity and variance between M. oleifera genotypes obtained from simple cluster analysis

3.3. Hierarchical Cluster analysis

Table 3 A summary of genotypes obtained by Hierarchical clustering dendrogram (Figure 2) with similarity coefficients indicated

Clade	Genotypes per Provenance	Sub-clades	Similarity Coefficient (%)		Clade	Genotypes	Sub-clades	Similarity Coefficient (%)	
I	Miasenyi - 10	Miasenyi - 2	100	94	III	Mackinon - 9	Samburu - 2	97	97
	Likoni - 1	Miasenyi - 8	94			Gede - 9	Kilifi - 1		
		Likoni - 1				India - 10	Maungu - 6	79	
II	Mwakiki - 10	Shika Adabu - 5	94	96		Pwani University - 10	Diani-Ukunda - 1		
	Likoni - 9	Kibwezi Town - 4				Samburu - 8	Kibwezi Town - 4		
	Mbololo - 7	India - 1				Arabuko Sokoke - 7	Vipingo - 1		
	Shika Adabu - 5	Diani- Ukunda - 1	79			Waa Kwale - 7	Pwani University - 7	79	
	Vipingo - 5	Mwakiki - 1				Diani-Ukunda - 7	Maungu - 3		
	Kibwezi Town - 4	Samburu - 1				Kilifi - 6	Kilifi - 4		
	Diani-Ukunda - 3	Likoni - 2	79			Maungu - 6	Arabuko Sokoke - 2		
	Maungu - 3	Mbololo - 2				Shika Adabu - 5	Shika Adabu - 2		
	Waa Kwale - 3	Kilifi - 1				Kibwezi Town - 4	Diani-Ukunda - 1		
	Kilifi - 2	Mbololo - 2	90			Vipingo - 4	Mackinon - 9		
	India - 1	Waa Kwale - 2				Mbololo - 2	Waa Kwale - 4		
	Samburu - 1	Likoni - 5					Vipingo - 4		
	Arabuko Sokoke - 1	Maungu - 1					Mbololo - 1		
		Kilifi - 1					Pwani University - 2	92	
		Arabuko Sokoke - 1					Samburu - 6		
		Vipingo - 2					Gede - 7		
		Mbololo - 3	97				Diani-Ukunda - 5		
		Samburu - 1					Shika Adabu - 1		
		Maungu - 2					Arabuko Sokoke - 3		
		Mwakiki - 9					Kilifi - 1		
		Vipingo - 2					Gede - 2		
		Waa Kwale - 1					India - 2		
		Likoni - 1					India - 7	96	
							Arabuko Sokoke - 4		
							Mbololo - 2		
							Kilifi - 1		
IV	Kibwezi Town - 2	Kibwezi Town - 2	100	97					
	Mbololo - 1	Mbololo - 1	91						
	Shika Adabu - 1	Shika Adabu - 1							
	Pwani University - 1	Pwani University - 1							

Further, hierarchical clustering was performed to obtain the similarity coefficient, essential in determining the level of variability. A hierarchical cluster dendrogram was constructed with height against distance using two p values, the number on the left (red) (fig. 2) indicates an approximately unbiased p-value (Au) and is computed by multiscale bootstrap resampling while the number on the right (green) indicates a bootstrap probability p-value (Bp) and is computed by normal bootstrap resampling. The hierarchical cluster dendrogram exhibited four clades with the highest similarity coefficient being 1 and the lowest being 0.63.

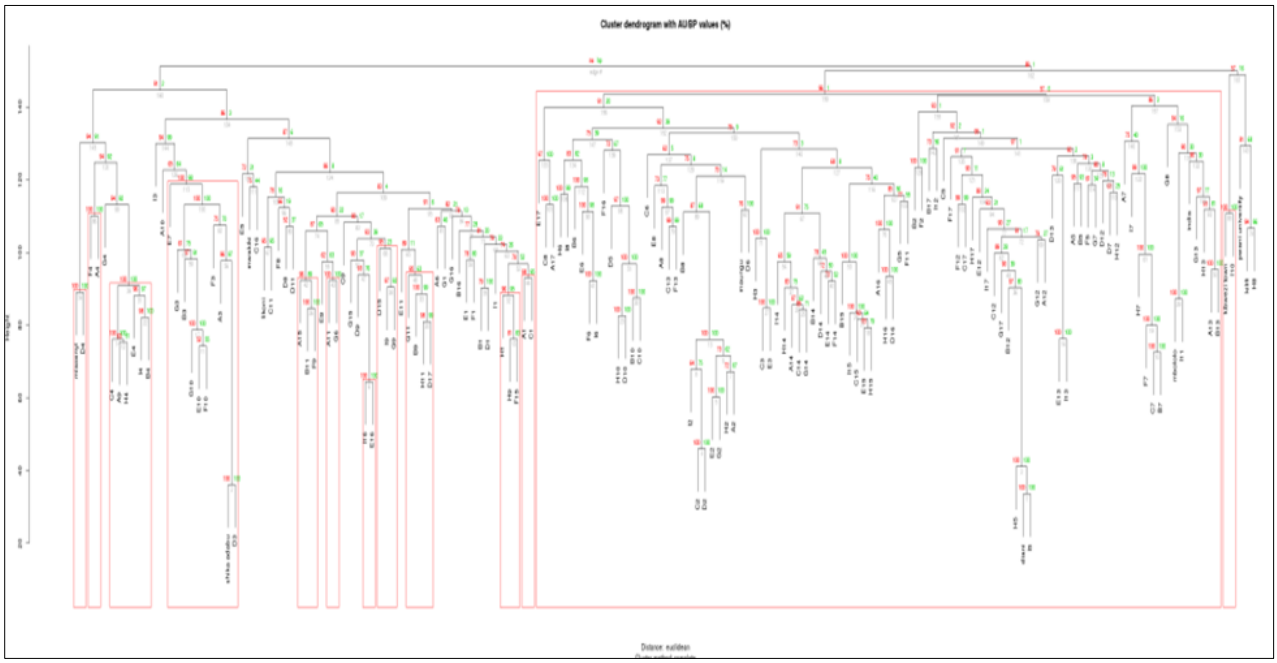

Figure 2 Hierarchical clustering dendrogram

3.4. Principal Component Analysis

Table 4 Principal Coordinate Analysis (PCoA) showing the number of genotypes per coordinates summarized from Figure 3

Coordinates	I: PC1 -ve, PC2 +ve	II: PC1 -ve, PC2 -ve	III: PC1 +ve, PC2 -ve	IV: PC1 +ve, PC2 +ve
Genotypes	Mackinon - 7	Maungu - 8	Shika Adabu - 7	Likoni - 10
	Pwani University - 6	Arabuko Sokoke - 6	Diani-Ukunda - 5	Mwakiki - 8
	India - 4	Kilifin- 6	Kibwezi Town - 5	Mbololo - 7
	Waa Kwale - 4	Kibwezi Town - 5	Miasenyi - 4	Miasenyi - 6
	Diani-Ukunda - 3	Samburu - 5	Gede - 4	Vipingo - 5
	Gede - 2	Shika Adabu - 3	India - 3	Waa Kwale - 4
	Kilifi - 1	India - 3	Arabuko Sokoke - 2	Kilifi - 2
	Mbololo - 1	Gede - 3	Mwakiki - 2	Maungu - 2
	Arabuko Sokoke - 1	Vipingo - 3	Samburu - 2	Pwani University - 1
	Vipingo - 1	Pwani University - 2	Pwani University - 1	Samburu -1
	Samburu - 1	Diani-Ukunda - 2	Kilifi - 1	
		Mackinon - 2	Mbololo - 1	
		Mbololo - 1		
		Waa Kwale - 1		
Total	31	50	37	46

The Moringa oleifera SNP profile classified the 164 genotypes once again into four major groups. This grouping pattern corresponds with that of clustering analysis as shown by the clustering dendrograms (Figures 1 & 2). The genetic similarity of individuals and populations were visualized by a principal coordinate analysis (PCoA) plot, which represented the samples based on the eigenvalues in the two-dimensional graph with respect to the first two principal coordinates as shown in figure 3.

Principal component (PC) 1 amounted to a 71% variance and the PC2 had 15% variance as evident in the PCoA plot (Figure 3).

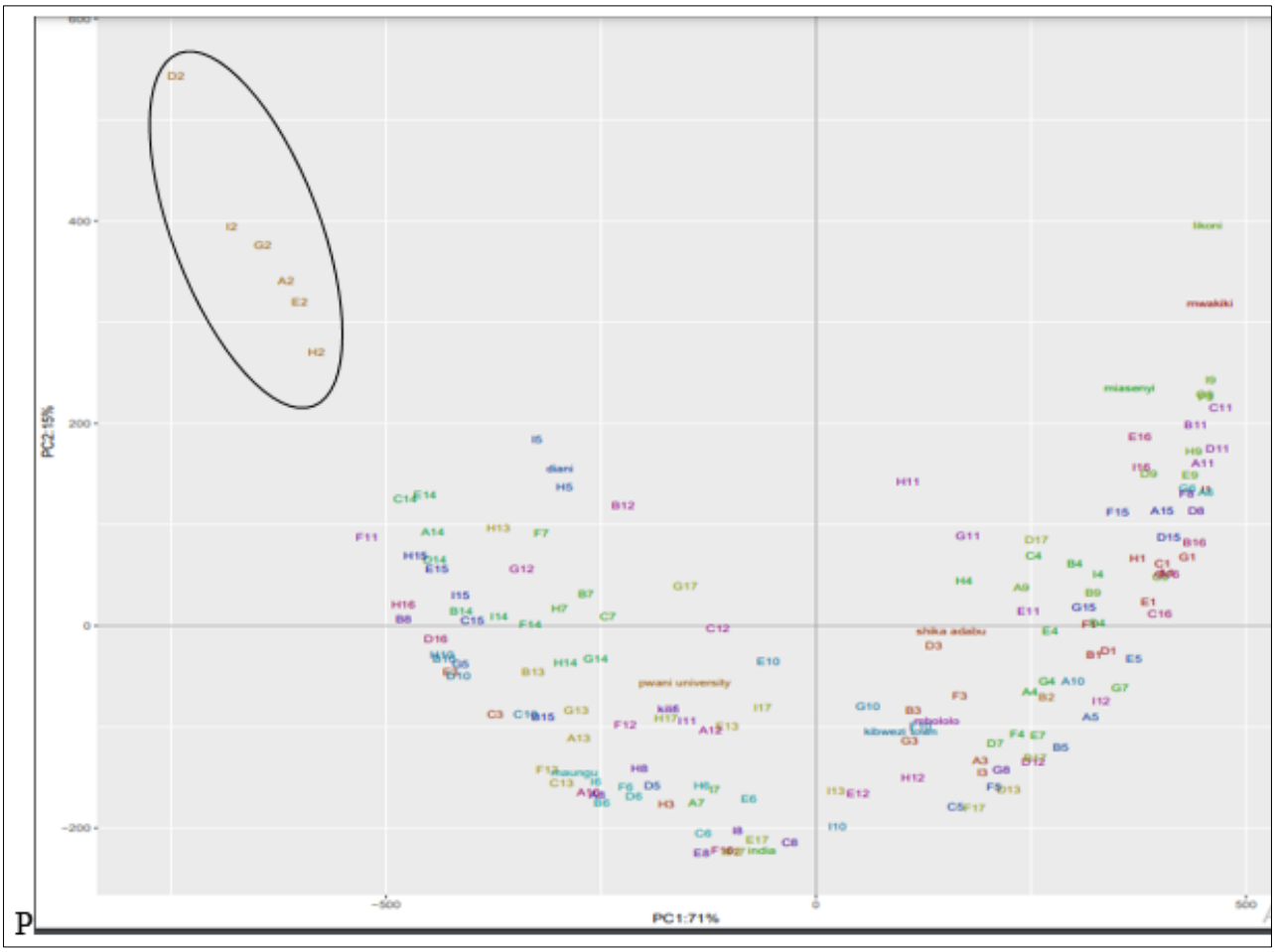

Figure 3 Principal coordinate analysis plot with six uniquely clustered Pwani University genotypes as shown in the ellipse

3.5. 3D Plot Results

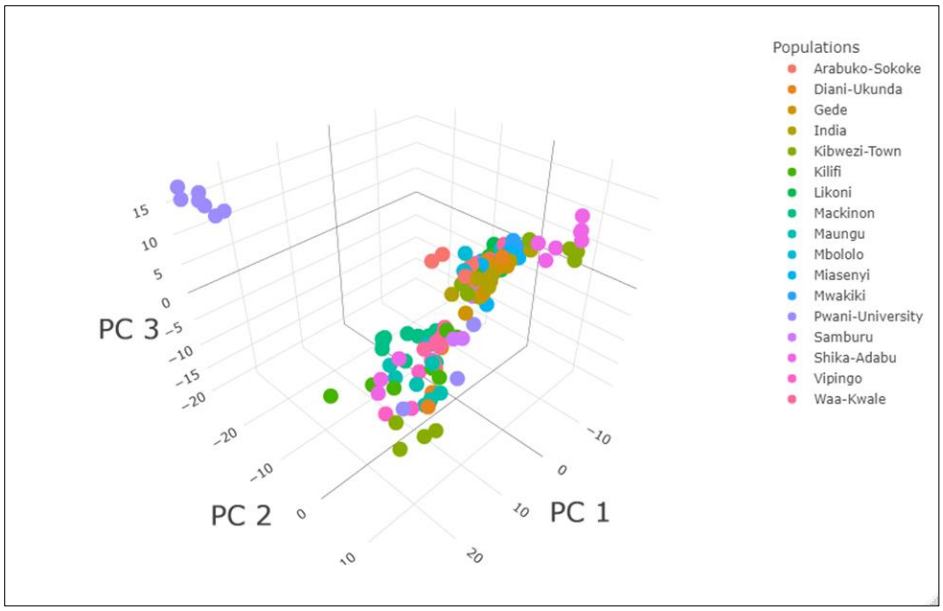

Figure 4 Three-dimensional plot showing a unique clustering of the Pwani University genotypes (PC3)

A more robust and highly sensitive analysis was conducted where an interactive 3D plot was produced using DArT R (Adegenet package) due to the fact that 3 dimensions were indicated by the scree plot (Gruber et al., 2019).

4. Discussion

A variation to DArT is the DArTseq technology which entails using the Illumina system's next-generation sequencing to replace the microarray hybridization stage (Sohail et al., 2012). It is an effective diagnostic technique for examining genetic diversity. The approach yields higher polymorphic markers, including codominant SNPs and dominant silicoDArT (Tomkowiak et al., 2019).

4.1. Simple and hierarchical clustering

Simple cluster analysis makes the assumption that the data are discontinuous. It shows the pattern of genetic similarity between genotypes based on phenotypic characteristics or evolutionary links. It is used to distinguish between distinct groups and clusters related lines or germplasm in one group. It is based on Unweighted paired group method using arithmetic mean (UPGMA) (Bhandari et al., 2017). In this analysis, four clades were obtained with no clear clustering by geographic region (figure 2).

The hierarchical cluster analysis (Figure. 3) also identified four clades similarly to the simple cluster analysis. Further, similarity coefficients were obtained by hierarchical clustering essential in determining the level of diversity between genotypes. The analysis displayed varied levels of genetic variation among the genotypes with similarity coefficients ranging from 63% to 100%, which indicates a high level of similarity between the studied genotypes. Other studies identified comparable levels of similarity coefficients with *M. oleifera* genotypes obtained from within a country. Fifty four per cent (54%) to 96% from Costal regions of Tanzania (Mgendi et al., 2010); 62% from the United States (A. V. C. da Silva et al., 2012); 48.68% from India (Saini et al., 2013); 81.50% from Nigeria (Ojuederie et al., 2013); 82.86% from India (kumar Ganesan et al., 2014); 55.9% from Timor Island, Indonesia (Kleden et al., 2017) and 88.25% from India (Ravi et al., 2020).

Multiple regional analysis identified similarity coefficients of 66.5% from India, Malawi and Kenya (Muluvi et al., 2004). Rufai et al., (2013) obtained similarity coefficients of 38% to 89% for seeds obtained from six different countries. Ridwan et al., (2021) studied genotypes from six Indonesian Islands with a polymorphism percentage of 81.40%.

Interestingly, the within and between counties analysis of *M. oleifera* rendered high similarity coefficients that cannot be attributed to their geographic distance. This observation also suggests that it is not possible to attribute the demographic divergence seen in these populations to isolation by distance. The absence of geographic clustering suggests that the genetic differences between individuals from various geographic locations are not very significant. This might be brought on by the movement of seeds, cuttings, and/or high rates of gene flow between the nearby populations (kumar Ganesan et al., 2014). The lack of clustering according to geographic locations can also be due to lack of divergence brought about by evolution considering not many years have passed since the introduction of *M. oleifera* in the Kenyan Coast.

Comparing and contrasting the two clustering techniques revealed some discrepancies. From the study it is evident from UPGMA clustering analyses there were 4 clades which differed in the number of genotypes in each clade per cluster analysis. For instance, in simple clustering analysis (table 2) all 10 Miasenyi genotypes clustered in clade III whereas in hierarchical clustering (table 3) they clustered in clade I. All 10 Mwakiki genotypes were clustered in clade II as shown in table 3 while they were clustered in clade III in table 2. There were several noticeable sub clades too from the hierarchical clustering analysis. This was also reported by (A. V. C. da Silva et al., 2012). Some authors (Costa et al., 2011; Dos Santos et al., 2011) have employed regrouping as an alternative to make the results more insightful. The employment of multiple grouping techniques prevents erroneous conclusions due to variations in group classification, optimization, and sorting. As a result, this choice is made when allocating materials among a specific collection of genotypes (Arriel et al., 2006). According to (Shahzad et al., 2013), *M. oleifera* accessions from 9 distinct countries (India, Senegal, Tanzania, Mozambique, Zimbabwe, Florida, Belize, Mexico, and Haiti) shared a high degree of genetic similarity with one accession from Pakistan (Punjab accession) and exhibited a low genetic diversity with each other. In light of this, they made the assumption that the accessions from the nine various countries came from Punjab and were transferred by humans to the other nine countries. This is evident in our study as the 164 genotypes exhibited a low genetic diversity and can be attributed to the *M. oleifera* plant in Coastal Kenya having come from a common ancestry, that is, India. It could also be attributed to the plant's domestication. Usually, when compared to the diversity of the species in the wild, domestication results in a decrease in genetic diversity (Osawaru et al., 2015).

It is notable from simple clustering dendrogram (from DArT R) that there were 163 samples as opposed to the 164 samples that were analyzed. This could be due to absolute similarity from sequence duplicity or close homology of two of the genotypes which could have been algorithmically collapsed into one sample.

4.2. Principal coordinate analysis

One approach to describing principal coordinate analysis (PCoA) is as a data-reduction method that works with quantitative types of data. Multi-correlated data are converted by PCoA into another set of uncorrelated variables for further research. Linear combinations of the original variables make up this new collection of variables. Its foundation is the growth of eigenvalues and mutually independent eigenvectors (principal components) arranged in descending order of variance size. These components produce scatter plots of data that have the optimal characteristics for analyzing underlying variability and correlation (Bhandari et al., 2017).

The PCoA frequently determines the eigenvalues and eigen-vectors of a matrix that represents all of the distances between the data points, measured using the Euclidean or Gower distance. It creates a 2 or 3 dimensional scatter map of the samples with the least amount of distortion possible, with the distances between the samples in the plot reflecting their genetic distances (Bhandari et al., 2017). Four groups were identified which corresponds with UPGMA clustering. This is similar to the results by (A. V. C. da Silva et al., 2012). The 7 genotypes of Pwani University were still grouped uniquely than the rest of the genotypes.

In PCoA, all accessions were designated with various colors based on their different population numbers to indicate their region specificity (Fig. 4). The intermixing of color across the coordinates further supports the UPGMA trees (Fig. 2 & 3), which show that there was no location-specific grouping of the genotypes. (kumar Ganesan et al., 2014) also found out that there was no geographic specific grouping in their study and concluded that lack of clustering in accordance to geographical areas was indicative of individuals from different geographical areas lacking a significant difference genetically. This reiterates the high genetic similarity between the genotypes studied.

Evidently, seven genotypes from Pwani University consistently clustered together and separately from the rest of the genotypes across all the analyses. These genotypes seem to be unique requiring further analysis. It is possible that they represent independently sourced germplasm in comparison to the other provenances that indicate a single seed source.

5. Conclusion

The genotypes consistently clustered into four clades or groups regardless of the analysis used. It is also evident that the genotypes were not consistent in the clades, that is, different analyses exhibited different genotypes in the four clades. This is resultant from the type of tool for UPGMA tree used.

It was evident from this study that there was less genetic diversity. Domestication of plants can also cause reduction in genetic diversity. *M. oleifera* in Kenyan Coast originated from India and has since been domesticated. It was also evident that the genotypes did not cluster according to their geographical location.

Recommendation

This research may help identify relationships between gene allelic forms and phenotypes, enabling the alleles to be linked to desired features like rapid growth and high seed output. This is because SNPs are highly frequent and involved in allele variations.

Compliance with ethical standards

Disclosure of conflict of interest

There are no conflict of interest in relation to this manuscript.

Funding

The authors acknowledge SEQART, Kenya for providing the laboratory and equipment for conducting this study.

References

[1] Arriel, N. H. C., Di Mauro, A. O., Di Mauro, S. M. Z., Bakke, O. A., Unêda-Trevisoli, S. H., Costa, M. M., Capeloto, A., & Corrado, A. R. (2006). Multivariate techniques for the determination of genetic diversity in sesame using RAPD markers. Pesquisa Agropecuária Brasileira, 41, 801–809.

[2] Bhandari, H. R., Bhanu, A. N., Srivastava, K., Singh, M. N., & Shreya, H. A. (2017). Assessment of genetic diversity in crop plants-an overview. Adv. Plants Agric. Res, 7(3), 00255.

[3] Costa, T. S., Silva, A. V. C. da, Lédo, A. da S., Santos, A. R. F. dos, & Silva Júnior, J. F. da. (2011). Genetic diversity of access to the mangaba germplasm bank in Sergipe. Brazilian Agricultural Research, 46, 499–507.

[4] da Silva, A. V. C., dos Santos, A. R. F., da Silva Lédo, A., Feitosa, R. B., Almeida, C. S., da Silva, G. M., & Rangel, M. S. A. (2012). Moringa genetic diversity from germplasm bank using RAPD markers. Tropical and Subtropical Agroecosystems, 15(1), 31–39.

[5] da Silva, J. P., Serra, T. M., Gossmann, M., Wolf, C. R., Meneghetti, M. R., & Meneghetti, S. M. (2010). Moringa oleifera oil: Studies of characterization and biodiesel production. Biomass and Bioenergy, 34(10), 1527–1530.

[6] Dos Santos, A. R. F., Silva-Mann, R., Ferreira, R. A., & de Souza Brito, A. (2011). Water pre-hydration as priming for Moringa oleifera Lam. Seeds under salt stress. Tropical and Subtropical Agroecosystems, 14(1), 201–207.

[7] Festus, M., Danson, K., Josphert, K., & Kennedy, M. (2022). Growth and fruiting of selected provenances of Moringa oleifera Lam. In South Eastern region of Kenya. Magna Scientia Advanced Research and Reviews, 5(2), 008–018.

[8] Gopalakrishnan, L., Doriya, K., & Kumar, D. S. (2016). Moringa oleifera: A review on nutritive importance and its medicinal application. Food Science and Human Wellness, 5(2), 49–56. https://doi.org/10.1016/j.fshw.2016.04.001

[9] He, J., Zhao, X., Laroche, A., Lu, Z.-X., Liu, H., & Li, Z. (2014). Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. Frontiers in Plant Science, 5, 484.

[10] He, S.-T., Yang, J., Wei, J., Wu, J.-C., Zheng, Y.-X., Zhang, Y.-P., & Peng, X.-M. (2021). Genetic diversity of eleven Moringa oleifera Lam. Germplasm introduced to Yunnan, southwest China and their backgrounds during worldwide cultivation. Genetic Resources and Crop Evolution, 1–12.

[11] Kleden, M. M., Soetanto, H., Kusmartono, K., & Kuswanto, K. (2017). Genetic diversity evaluation of Moringa oleifera, lam from east flores regency using marker random amplified polymorphic DNA (RAPD) and its relationship to chemical composition and in vitro gas production. AGRIVITA, Journal of Agricultural Science, 39(2), 219–231.

[12] kumar Ganesan, S., Singh, R., Choudhury, D. R., Bharadwaj, J., Gupta, V., & Singode, A. (2014). Genetic diversity and population structure study of drumstick (Moringa oleifera Lam.) using morphological and SSR markers. Industrial Crops and Products, 60, 316–325.

[13] Kumar, P., Dolkar, R., Manjunatha, G., & Pallavi, H. M. (2017). Molecular fingerprinting and assessment of genetic variations among advanced breeding lines of Moringa oleifera L. by using seed protein, RAPD and Cytochrome P450 based markers. South African Journal of Botany, 111, 60–67.

[14] Leone, A., Spada, A., Battezzati, A., Schiraldi, A., Aristil, J., & Bertoli, S. (2015). Cultivation, Genetic, Ethnopharmacology, Phytochemistry and Pharmacology of Moringa oleifera Leaves: An Overview. International Journal of Molecular Sciences, 16(6), Article 6. https://doi.org/10.3390/ijms160612791

[15] Mgendi, M. G., Manoko, M. L., & Nyomora, A. (2010). Genetic diversity between cultivated and non-cultivated Moringa oleifera Lam. Provenances assessed by RAPD markers.

[16] Muluvi, G. M., Sprent, J. I., Odee, D., & Powell, W. (2004). Estimates of outcrossing rates in Moringa oleifera using Amplified fragment length polymorphism (AFLP). African Journal of Biotechnology, 3(2), Article 2. https://doi.org/10.4314/ajb.v3i2.14932

[17] Nouman, W., Basra, S. M. A., Siddiqui, M. T., Yasmeen, A., Gull, T., & Alcayde, M. A. C. (2014). Potential of Moringa oleifera L. as livestock fodder crop: A review. Turkish Journal of Agriculture and Forestry, 38(1), 1–14.

[18] Ojuederie, O. B., Igwe, D. O., Okuofu, S. I., & Faloye, B. (2013). Assessment of genetic diversity in some Moringa oleifera Lam. Landraces from Western Nigeria using RAPD markers. The African Journal of Plant Science Biotechnology, 7(1), 15–20.

[19] Osawaru, M. E., Ogwu, M. C., & Aiwansoba, R. O. (2015). Hierarchical approaches to the analysis of genetic diversity in plants: A systematic overview. University of Mauritius Research Journal, 21.

[20] Rajalakshmi, R., Rajalakshmi, S., & Parida, A. (2019). Genetic diversity, population structure and correlation study in Moringa oleifera Lam. Using ISSR and SRAP markers. Proceedings of the National Academy of Sciences, India Section B: Biological Sciences, 89(4), 1361–1371.

[21] Ramachandran, C., Peter, K. V., & Gopalakrishnan, P. K. (1980). Drumstick (Moringa oleifera): A multipurpose Indian vegetable. Economic Botany, 276–283.

[22] Raman, H., Raman, R., Kilian, A., Detering, F., Carling, J., Coombes, N., Diffey, S., Kadkol, G., Edwards, D., & McCully, M. (2014). Genome-wide delineation of natural variation for pod shatter resistance in Brassica napus. PloS One, 9(7), e101673.

[23] Ravi, R. D., Siril, E. A., & Nair, B. R. (2020). The efficiency of Cytochrome P450 gene-based markers in accessing genetic variability of drumstick (Moringa oleifera Lam.) accessions. Molecular Biology Reports, 47(4), 2929–2939.

[24] Saini, R. K., Saad, K. R., Ravishankar, G. A., Giridhar, P., & Shetty, N. P. (2013). Genetic diversity of commercially grown Moringa oleifera Lam. Cultivars from India by RAPD, ISSR and cytochrome P 450-based markers. Plant Systematics and Evolution, 299(7), 1205–1213.

[25] Shahzad, U., Khan, M. A., Jaskani, M. J., Khan, I. A., & Korban, S. S. (2013). Genetic diversity and population structure of Moringa oleifera. Conservation Genetics, 14(6), 1161–1172. https://doi.org/10.1007/s10592-013-0503-x

[26] Sohail, Q., Shehzad, T., Kilian, A., Eltayeb, A. E., Tanaka, H., & Tsujimoto, H. (2012). Development of diversity array technology (DArT) markers for assessment of population structure and diversity in Aegilops tauschii. Breeding Science, 62(1), 38–45.

[27] Spinoso-Castillo, J. L., Escamilla-Prado, E., Aguilar-Rincón, V. H., Ramos, V. M., de los Santos, G. G., Pérez-Rodríguez, P., & Corona-Torres, T. (2020). Genetic diversity of coffee (Coffea spp.) in Mexico evaluated by using DArTseq and SNP markers. Genetic Resources and Crop Evolution, 67(7), 1795–1806.

[28] Stephenson, K. K., & Fahey, J. W. (2004). Development of tissue culture methods for the rescue and propagation of endangered Moringa spp. Germplasm. Economic Botany, 58(1), S116–S124.

[29] Tomkowiak, A., Bocianowski, J., Wolko, \Lukasz, Adamczyk, J., Miko\lajczyk, S., & Kowalczewski, P. \Lukasz. (2019). Identification of Markers Associated with Yield Traits and Morphological Features in Maize (Zea mays L.). Plants, 8(9), 330.

[30] Wu, J.-C., Yang, J., Gu, Z.-J., & Zhang, Y.-P. (2010). Isolation and Characterization of Twenty Polymorphic Microsatellite Loci for Moringa oleifera (Moringaceae). HortScience, 45(4), 690–692. https://doi.org/10.21273/HORTSCI.45.4.690