

Full Length Research Paper

## Assessment of the genetic diversity of Kenyan coconut germplasm using simple sequence repeat (SSR) markers

Maurice E. Oyoo<sup>1\*</sup>, Najya Muhammed<sup>2</sup>, Kimani N. Cyrus<sup>3</sup> and Simon M. Githiri<sup>3</sup>

<sup>1</sup>Department of Crops, Horticulture and Soil Sciences, Egerton University, P. O. Box 536-20115, Egerton, Kenya.

<sup>2</sup>Department of Chemistry and Biochemistry, Pwani University, P. O. Box 195-80108, Kilifi, Kenya.

<sup>3</sup>Jomo Kenyatta University of Agriculture and Technology, P. O. Box 62000-00100, Nairobi, Kenya.

Received 13 July 2016, Accepted 7 September, 2016.

Genetic diversity and relationship among 48 coconut individuals (*Cocos nucifera* L.) collections from the Coastal lowland of Kenya were analyzed using 15 simple sequence repeat (SSR) primer pairs. Diversity parameters were calculated using Popgene Software version 1.31. The gene diversity values ranged from 0.0408 (CAC68) to 0.4861 (CAC23) with a mean value of 0.2839. The polymorphic information content (PIC) values ranged from 0.0400 to 0.3680 with a mean value of 0.2348. Marker CAC23 had the highest PIC and revealed highest gene diversity values in this study. Analysis of the molecular variation indicated that within individual variation was 98% while among materials, variation was low at 2% suggesting that molecular variation was not defined by region of production. Cluster analysis was constructed using DARwin program version 6.0. Forty eight (48) coconut individuals were clustered into three groups.

**Key words:** Coconut palm (*Cocos nucifera* L., Arecaceae), germplasm, simple sequence repeat (SSR) markers, cluster analysis, diversity.

### INTRODUCTION

The coconut palm (*Cocos nucifera* L., Arecaceae) is a crop mostly cultivated in the coastal lowlands of Kenya and plays an important role in the economy. However, Kenya has been lagging behind in technology development for product diversification and by product utilization. Beside this, the coconut grows well in agro-ecological coastal lowlands (CL), where frequent droughts in these zones have been affecting the coconut

yields with most trees drying up or tipping off. The unavailability of drought resistant varieties and poor access to quality planting materials are another major hindrance to growth and productivity in the coconut industry in Kenya (Muhammed et al., 2013). Furthermore, the slow growth and long pre-breeding period of palm does not promote the genetic enhancement of coconut palm for productivity and tolerance to biotic and abiotic

\*Corresponding author. Email: mauriceoyoo464@gmail.com.

**Table 1.** Coastal AEZ characteristics.

AEZ	Temperature (°C)	Altitude (m)	Rainfall (mm)	Natural vegetation	Number of individuals sampled
CL1	24-30	0-900	1100-2700	Moist and forest	12
CL2	24-30	0-900	1000-1600	Moist and dry forest	26
CL3	24-30	0-900	800-1400	Dry forest, moist woodland	06
CL4	24-30	0-900	600-1100	Dry woodland and highland	04

stresses (Rajesh et al., 2008). Current estimate put total number of plants in Kenyan coastal lowlands at about 4.4 million with an average nut yield of 1.5 t ha<sup>-1</sup>, while that of copra is as low as 0.45 t ha<sup>-1</sup>. Germplasm collections, containing significant amount of genetic diversity within and among coconut populations, are essential for an effective crop improvement programme. Therefore, the assessment of genetic diversity within coconut populations becomes increasingly significant in germplasm conservation and utilization. Since morphological characterization in the coast of Kenya has shown diversity (Oyoo et al., 2015), the coconut genetic diversity in Kenya shall provide sufficient scientific data for germplasm management and utilization. Diversity analysis in coconut palm has been done by morphological traits, biochemical and molecular markers (Rajesh et al., 2005). Morphological and biochemical markers are not desirable due to the long juvenile phase of palms, high cost of conducting investigations, long-term of field evaluation, influence of environment factors on the phenotype and limited number of available phenotypic markers (Manimekalai et al., 2006). Molecular markers, detect variation at the DNA level and are detectable at all stages of development. Since they can also cover the entire genome, DNA markers overcome most limitations of morphological and biochemical markers (Ashburner et al., 1997; Teulat et al., 2000; Upadhyay et al., 2004; Manimekalai et al., 2006). Among available molecular marker techniques, simple sequence repeat (SSR) markers provide good signal in evaluating genetic diversity and genetic relationships in plants. The increased number of coconut SSR markers greatly improves the previously established genetic relationships among coconut populations.

Three types of coconut palms have been described in Kenya, the East African tall (EAT) Dwarf types and the hybrids. The EAT yield good quality copra and toddy but with inferior coconut water as compared to the Dwarf coconut (Gethi and Malinga, 1997). Just as in Asia, the Talls are naturally cross-pollinating types, vigorous growing, comparatively late to flowering and the fruits are with intermediate colors of brown, green, yellow, orange among individual palms. Dwarfs, in contrast, are naturally self-pollinating types with reduced growth habitat, early flowering and produce large number of medium to small, distinctly colored (green or yellow or orange or brown) fruits (Dasanayaka et al., 2009). To date, little information

is available on the genetic diversity among Kenyan coconut palms. For sustainable breeding, adoption and conservation *in situ*, it is important to develop a strategy to use coconut diversity for socio-economic benefits (Batugal and Oliver, 2003). In the current study, 48 coconut individuals from four administrative units of Kenyan coast were assessed using 15 SSR markers.

In Kenya, coconut is cultivated in six counties; Kwale, Kilifi, Mombasa, Tana River, Lamu and Taita Taveta. With the exception of Taita Taveta, the other five counties have a high concentration of coconut tree population and a total coastline of 640 km of the Indian Ocean. Coastal Kenya covers areas from the sea with ample amount of rainfall to the far west with 600 mm of rainfall per year, usually with poor distribution. Rainfall is bimodal with major rainy season beginning in April and lasts till July. Short rainy season is seen in October to November (Jaetzold and Schmidt, 1983). The coastal lowlands zone is divided into five sub zones called coastal lowlands (CL): CL2-CL6. These zones are sub-divided based majorly on topography, soil and water which influence agricultural development (Jaetzold and Schmidt, 1983).

The assessment of genetic diversity and structure of germplasm is essential for the efficient organization of breeding material. The objective of this research was to estimate genetic variation of coconut germplasm in Kenyan coast and suggest how that knowledge might lay the groundwork for the genetic improvement and breeding of coconut in Kenya.

## MATERIALS AND METHODS

### Germplasm sampling

Coconut leaf samples were collected *in situ* and coded as previously described by Oyoo et al. (2015). Coding the palms sampled was done to reflect the county, district, division, sub division, village and the collection number. If collection was from Kilifi county, Magarini district, Magarini division, Gogoni location, Ngarite village and it was the first palm sampled, the code name given was: KLF/MAG/MAG/GOG/NGA/01. Sampling was done at different agro ecological zones (Table 1) and the GPS data was recorded. The focus was in areas where palms grown were morphologically different with a marked change in altitude or cropping systems, a formidable barrier such as a mountain or a river exists, or local people ethnically different (in terms of dialect) from previous collection site. The priority of taking sample was farmers' field while avoiding collecting duplicates. A major challenge was that farmers had no clear distinct local names for coconut at the coast and insisted '*nazi ni nazi tu*' (coconuts are just

**Table 2.** List of coconut-specific microsatellite primer pairs with their forward and reverse sequences.

Primer	Repeat type	F	R	Size range	Tm
CAC02	(CA) <sub>15</sub> (AG) <sub>7</sub>	AGCTTTTTTCATTGCTGGAAT	CCCCTCCAATACATTTTTTCC	225–240	49
CAC04	(CA) <sub>19</sub> (AG) <sub>17</sub>	CCCCTATGCATCAAAACAAG	CTCAGTGTCCGTCTTTGTCC	185–207	53
CAC06	(AG) <sub>14</sub> (CA) <sub>9</sub>	TGTACATGTTTTTGGCCAA	CGATGTAGCTACCTTCCCC	146–164	49
CAC10	(AG) <sub>13</sub> (CA) <sub>9</sub>	GATGGAAGGTGGTAATGCTG	GGAACCTCTTTTGGGTCATT	156–163	53
CAC21	(CA) <sub>11</sub>	AATTGTGTGACACGTAGCC	GCATAACTCTTTCATAAGGGA	149–151	53.7
CAC23	(CA) <sub>8</sub>	TGAAAACAAAAGATAGATGTCAG	GAAGATGCTTTGATATGGAAC	170–179	53.9
CAC56	(CA) <sub>14</sub>	ATTCTTTTGGCTTAAAACATG	TGATTTTACAGTTACAAGTTTGG	138–162	53.9
CAC68	(CA) <sub>13</sub>	AATTATTTTCTTGTACATGCATC	AACAGCCTCTAGCAATCATAG3	130–146	54
CAC71	(CA) <sub>17</sub>	ATAGCTCAAGTTGTTGCTAGG3	ATATTGTCATGATTGAGCCTC3	172–283	54
CAC72	(CA) <sub>18</sub>	TCACATTATCAAATAAGTCTCACA	GCTCTCTTCTCATGCACA	124–132	54
CAC84	(CA) <sub>13</sub>	TTGGTTTTTGTATGGAACCTCT	AAATGCTAACATCTCAACAGC	150–163	54
CN11E6	(CT) <sub>21</sub>	TACTTAGGCAACGTTCCATTC	TAACCAGAAAGCAAAAAGATT	85–128	50
CN1C6	(CT) <sub>1</sub> TT(GT) <sub>5</sub>	AGTATGTGAGTAGGATTATGG	TTCCTTGGACCCTTATCTCTT	175–184	52
CN1H2	(GA) <sub>18</sub>	TTGATAGGAGAGCTTCATAAC	ATCTTCTTTAATGCTCGGAGT	230–321	52
CNZ01	(CT) <sub>15</sub> (CA) <sub>9</sub>	ATGATGATCTCTGGTTAGGCT	AAATGAGGGTTTGAAGGATT	109–131	52

The SSR primers were developed by Perera et al. (1999, 2000) and Teulat et al. (2000)

coconuts). In addition, the general structure of parents and progeny was a mixture in a population. Estimating the age of the population was not easy. The fields were uncultivated with low input or none at all. The collected leaf samples were labeled and stored in silicon gels.

#### DNA extraction

DNA was extracted from dry leaf samples using a CTAB protocol modified from the Doyle and Doyle (1990) method. The modification involved omission of the ammonium acetate step and a longer DNA precipitation time of 12 h. The quality and quantity of the extracted DNA samples were ascertained by running them on a 1% agarose gel and by using a nanodrop spectrophotometer. The DNA samples were then diluted to a working concentration of 30 ng/μl. The primer sequences and associated information are given in Table 2.

#### SSR markers amplification and electrophoresis

The PCR amplification was performed in a 10 μl volume mix consisting of 5 U Dreamtaq polymerase enzyme (Thermo scientific corp, Lithuania), x 6 Dreamtaq buffer (Thermo scientific corp, Lithuania), 2.5 mM of each dNTPS (Bioneer corp, Republic of Korea), MgCl<sub>2</sub>, 5 μM of each primer (Inqaba biotec, S.A) and 30 ng DNA template in an Applied Biosystems 2720 thermocycler (Life Technologies Holdings Pte Ltd, Singapore). The PCR cycles consisted of initial denaturation at 94°C for 5 min followed by 35 cycles of denaturation at 94°C for 30 s, annealing at 54–59.7°C (depending on the primer), extension at 72°C for 1 min followed by one cycle of final extension at 72°C for 10 min. The amplicons were mixed with 6x Orange DNA loading dye (Thermo scientific corp, Lithuania) and separated on a 2% agarose gels (Duchefa, Netherlands) stained with Invitrogen life technologies ethidium bromide (Invitrogen corp, U.S.A) in a 0.5x TBE buffer. The separated amplicons were visualized on an Ebox-VX5 gel visualization system (Vilber Lourmat inc, France). The alleles were

scored as absent or present based on the size of the amplified product using a 100 bp O'geneRuler ready to use DNA Ladder (Thermo Scientific Corp, Lithuania).

#### Data analysis

##### Cluster and principal component analyses

Genetic dissimilarities between all the germplasm was calculated using DARwin version 6.0 (Perrier et al., 2003; Perrier and Jacquemoud-Collet, 2006) using simple matching coefficient. The dissimilarity coefficients were then used to generate an unweighted neighbour-joining tree (Saitou and Nei, 1987) with Jaccard's similarity coefficient. Bootstrapping value of 1,000 was used. Principal component analyses were also estimated by using Darwin version 6.0.

##### Analysis of diversity parameters

The amplified SSR markers were scored as present (1) or absent (0), and then recorded into a binary matrix as discrete variables. Markers which could be amplified and interpreted consistently were omitted from the analysis. The diversity studies of the 48 coconut ecotypes (from different AEZ) were conducted using Popgene version 1.31. The following parameters were estimated: number of alleles per locus (*N*), polymorphic information content (*pic*), *m*\* = major allele frequency, *ne* = effective number of alleles (Kimura and Crow, 1964), \* *h* = Nei's (1973) gene diversity, *I* = Shannon's information index (Lewontin, 1972) and the percentage of polymorphic loci. Other diversity parameters measured were *Ht* = gene diversity in the total population, *Hs* = gene diversity within and between sub populations *Gst* = gene differentiation relative to the total population, *Nm* = estimate of gene flow from *Gst*, which in this case is equivalent to Wright *Fst* statistics (*Fst*) (Nei, 1978). Analysis of molecular variance (AMOVA) was computed using Powermarker software version 3 (Liu and Muse, 2005).

**Table 3.** Diversity parameters for 15 SSR loci used to analyze genetic diversity of coconut germplasm in Kenya.

Locus	Sample size	Allele no.	m*	ne*	h*	i*	PIC	Amplified bands (no.)
CAC02	48	2.0	0.8333	1.3846	0.2778	0.4506	0.2392	40
CAC04	48	2.0	0.6667	1.8000	0.4444	0.6365	0.3457	32
CAC06	48	2.0	0.8958	1.2295	0.1866	0.3341	0.1692	43
CAC10	48	2.0	0.6875	1.7534	0.4297	0.6211	0.3374	43
CAC21	48	2.0	0.9583	1.0868	0.0799	0.1732	0.0767	46
CAC23	48	2.0	0.5833	1.9459	0.4861	0.6792	0.368	28
CAC56	48	2.0	0.8125	1.4382	0.3047	0.4826	0.2583	39
CAC68	48	2.0	0.9792	1.0425	0.0408	0.1013	0.04	47
CAC71	48	2.0	0.875	1.2800	0.2188	0.3768	0.1948	42
CAC72	48	2.0	0.9167	1.1803	0.1528	0.2868	0.1411	44
CAC84	48	2.0	0.8542	1.3318	0.2491	0.4154	0.2181	41
CNZ01	48	2.0	0.6042	1.3846	0.2778	0.4506	0.3639	40
CN1C6	48	2.0	0.7708	1.5463	0.3533	0.5383	0.2909	37
CN11E6	48	2.0	0.8333	1.3846	0.2778	0.4506	0.2392	40
CN1H2	48	2.0	0.8333	1.9168	0.4783	0.6713	0.2392	29
Mean		2.0	0.8069	1.4470	0.2839	0.4445	0.2348	39.4
SD				0.288	0.137	0.173		

m\* = Major allele frequency; \*ne = Effective number of alleles [Kimura and Crow (1964)]. \*h = Nei's (1973) gene diversity; \*i = Shannon's Information index [Lewontin (1972)]; Smith et al., (1997) polymorphic information content (PIC). The number of polymorphic loci is: 15. The percentage of polymorphic loci is 100.00%.

## RESULTS AND DISCUSSION

### Diversity parameters

The major allele frequency value ranged from 0.5833 for CAC23 to 0.9792 for CAC68 with a mean of 0.8069 (Table 3). The number of effective alleles values ranged from 1.0425 to 1.9459 with a mean value of 1.4470. It was observed that marker CAC68 had the lowest values while marker CAC23 had the highest value. The gene diversity values ranged from 0.0408 to 0.4861 with a mean value of 0.2839. Marker CAC68 had the lowest value while marker CAC23 had the highest value, suggesting that CAC23 loci could be useful in revealing genetic diversity of coconut accessions in Kenya. This observation was also confirmed by Shannon's information index at locus CAC23 ( $p = 0.6792$ ), which had the highest value as compared to the lowest value of  $p = 0.1013$  at locus CAC68. Marker CAC23 also showed greatest PIC (0.3680%) as opposed to marker CAC68 which revealed the lowest PIC values (0.0400%). An average of 34.9 bands were amplified with the highest number of bands amplified being observed for marker CAC68. Least number of bands (29) was amplified by marker CN1H2.

The weighted average of estimated haplotype diversities in the total population ( $Ht = 0.4884$ ) and weighted average of estimated haplotype diversities in the subpopulation ( $Hs = 0.4686$ ) were best revealed at the CAC23 locus. The lowest ( $Ht = 0.408$  and

$Hs = 0.0382$ ) were recorded at the CAC68 locus (Table 4). The highest gene differentiation relative to the total population studied,  $Gst = 0.2016$  recorded for marker CAC04, and the lowest  $Gst = 0.0378$  was recorded at CN11E6 locus. The highest estimate of gene flow from  $Gst$ ,  $Nm = 28.0937$ , was recorded for marker CAC56, while the least  $Nm = 1.9801$  was recorded at the CAC04 locus

Results presented in Table 5 also reveal that genetic similarity were very high in the subpopulations. It ranged from 0.9502 (genetic identity between Kilifi and Lamu counties) to 0.9891 (between Lamu and Kilifi counties). Genetic distances were very low among the subpopulations and they ranged from 0.011 (Kwale and Lamu) to 0.511 (Lamu and Kilifi).

Further, when the genetic variation of the coconut germplasm was partitioned by analysis of molecular variance (AMOVA), 2% of the variation was revealed among the germplasm that were produced within the counties (Table 6), while the within population variation was 98%, further suggesting genetic redundancy of coconut at the coastal Kenya on the basis of where grown.

### Cluster analysis

The dendrogram was constructed from the genetic dissimilarities among all the individuals using DARwin 6. The 48 germplasm were clustered into three distinct groups (Figure 1). Groups II and III were further

**Table 4.** Nei's (1978) analysis of gene diversity in subdivided and total populations.

Locus	Sample size	Ht	Hs	Gst	Nm*
CAC02	48	0.2477	0.2208	0.1083	4.1153
CAC04	48	0.394	0.3146	0.2016	1.9801
CAC06	48	0.1665	0.1494	0.1026	4.374
CAC10	48	0.4217	0.3807	0.0972	4.6424
CAC21	48	0.0703	0.0675	0.0394	12.1961
CAC23	48	0.4884	0.4689	0.0399	12.0332
CAC56	48	0.3073	0.3019	0.0175	28.0937
CAC68	48	0.0408	0.0382	0.0638	7.3333
CAC71	48	0.2491	0.2315	0.0707	6.5759
CAC72	48	0.1298	0.1253	0.0352	13.6981
CAC84	48	0.2156	0.2041	0.0533	8.8874
CNZ01	48	0.4639	0.4334	0.0657	7.1065
CN1C6	48	0.3567	0.3315	0.0705	6.592
CN11E6	48	0.2874	0.2765	0.0378	12.714
CN1H2	48	0.2491	0.2049	0.1777	2.3137
Mean	48	0.2726	0.25	0.0829	5.5296
St. Dev		0.0186	0.0159		

Ht = Gene diversity in the total population; Hs = gene diversity within and between sub populations; Gst= gene differentiation relative to the total population; Nm = estimate of gene flow from Gst.

**Table 5.** Genetic identity (above diagonal) and genetic distance (below diagonal).

Population ID	Kilifi	Kwale	Lamu	Tana River
Kilifi	****	0.9563	0.9502	0.9638
Kwale	0.0447	****	0.9891	0.9697
Lamu	0.0511	0.011	****	0.9565
Tana River	0.0369	0.0308	0.0445	****

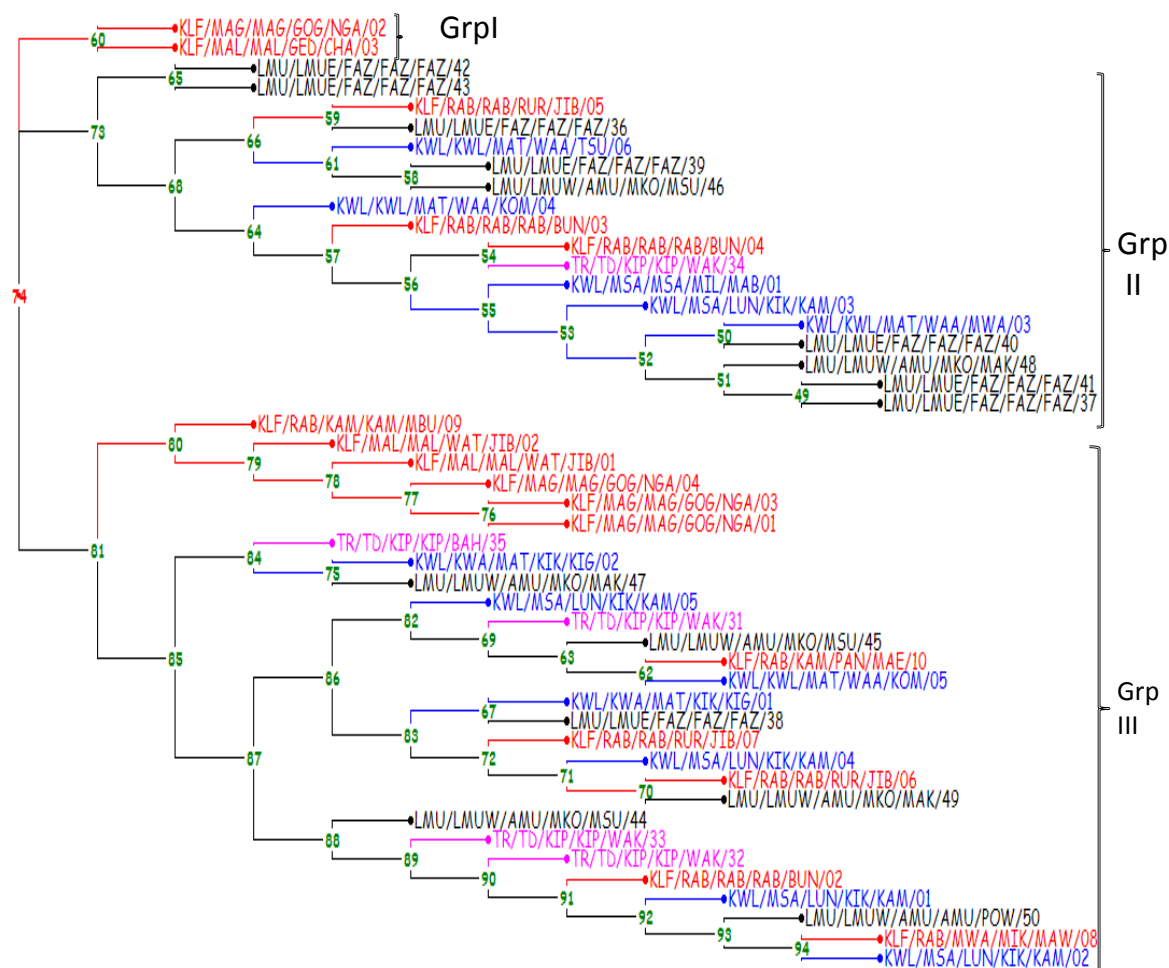
**Table 6.** Analysis of molecular variance (AMOVA) for 48 coconut individuals sampled from four counties of the coastal lowlands of Kenya.

Source	df	SS	MS	Est. Var.	%
Among populations	3	7.871	2.624	0.042	2
Within populations	44	94.317	2.144	2.144	98
Total	47	102.188		2.185	100

subdivided into two groups with one sub group in group III consisting of solely collections from Kilifi County. The dendrogram revealed that materials sampled from same counties were not assigned to the same group. Group I consisted of only two germplasm lines from Kilifi; KLF/MAG/MAG/NGA/02 and KLF/MAL/MAL/GED/CHA/03. Group II was composed of germplasm sampled from Lamu county and a few materials from Kwale and three from Kilifi. Only one material, TR/TD/KIP/KIP/WAK/34 from Tana River clustered in this group, while the rest of materials from

this county were clustered in group III. Group III consisted of germplasm from all the four counties; however samples from Kilifi formed the majority with up to six samples all clustering from node 76 to 80. Clustering failed to group the germplasm solely on the basis of the origin.

Principal coordinate analysis (PCoA) (Figure 2) was carried out to determine genetic relationship among the coconut germplasm studied. The first three axes accounted for the majority of total cumulative variation (61.64%) (Table 7) with the three axes accounting for

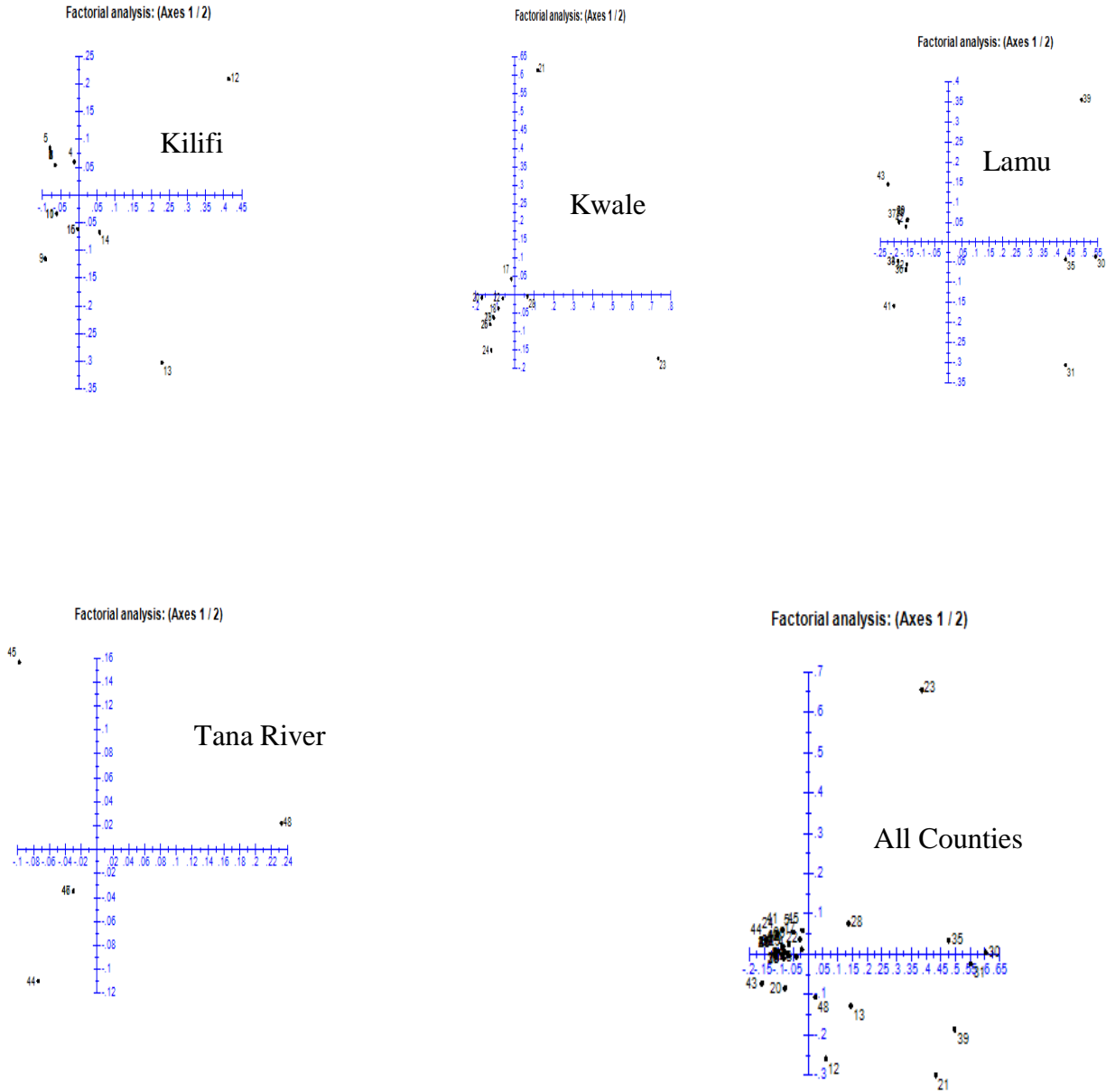


**Figure 1.** Dendrogram based on genetic dissimilarities between 48 individuals sampled at the Kenyan coast. Numbers represent bootstrap confidence limits for 1000 replicates.

37.61, 14.02 and 10.02%, respectively. Individual counties total cumulative percent variations were higher than the overall trial percent cumulative variation. Total cumulative variation was highest in Tana River (97.88%); and it was explained by only two axes. The county was followed by Lamu (91.15%), Kwale (90.51%) and Kilifi (90.19%) in that order. These results could be corroborated by results (Figure 2) that shows that germplasm did not cluster around each other in Tana River. Coconut is an important tropical economic crop, but not many studies have been conducted in relation to diversity based on molecular markers. Due to a lack of available molecular markers, Akpan (1994) and Sugimura et al. (1997) indicated that previous assessment and characterization of coconut germplasms relied mostly on morphological and agronomic traits. However, these traits do not generally provide an accurate measure of genetic diversity, because many show complex inheritance and are easily influenced by the external environment. To mitigate, this problem (Rivera et al., 1999) described the

development and characterization of microsatellite markers for *Cocos nucifera* using genomic-SSR hybrid screening. Recently, Xiao et al. (2013) used computational methods to utilize sequence information from next-generation sequencing as described by Fan et al. (2013) to develop SSR markers.

UPGMA dendrogram was drawn to visualize relationships among 48 coconut studied. The dendrogram of cluster analysis of individuals clearly illustrated population structure of the tested varieties. It showed the presence of subpopulations in all the cluster and different levels of similarity between the localities. It however failed to cluster germplasm per counties grown as expected. This could be as a result of farmers exchanging planting materials within the coastal Kenya. As described by Aremu et al. (2007), principal component analysis (PCA) reveals the pattern of character variation among genotypes. In this study, PCA failed to cluster coconut materials based on their counties where they are grown. This could be explained by the low coefficient of



**Figure 2.** PCoA of axes 1 and 2 based on dissimilarity of 15 loci across 48 coconut individuals from different counties.

variation realized in coconut at the Coastal Kenya, probably suggesting high level of genetic similarity and very low genetic distances as shown in Table 6. This conclusion could also be deduced from the analysis of molecular variance (Table 7), which suggested that more variation existed within individual coconut as opposed to among groups based on the areas where they are produced.

PIC provides an estimate of discriminatory power of a marker by taking into account not only the number of alleles at a locus, but also their relative frequencies. Table 4 shows that loci CAC23 (0.368), CAC10 (0.3374),

CCZ01 (0.3639) and CAC04 (0.3457) revealed a higher PIC levels than other loci, indicating that these primers are suitable for detecting the genetic diversity of coconut accessions in Kenyan Coast. These markers also revealed a higher weighted average of estimated haplotype diversities in the total population (*Ht*). The *Ht* values for these markers as compared to others are relevant in view of the fact that there are only two alleles per locus. Furthermore, SSR analysis also showed that the mean gene flow, *Gst* of coconut in Kenya was 0.0829, suggesting that the gene exchange between individuals was minimal. Based on the fact there are

**Table 7.** Eigen value and percentage of total variation accounted for by the first three principal component axes.

Axis	Eigen value	Proportion (%)	Cumulative (%)
<b>Kilifi County</b>			
1	0.01721	48.65	48.65
2	0.01191	33.68	82.33
3	0.00278	7.86	90.19
<b>Kwale County</b>			
1	0.05584	46.56	46.56
2	0.03757	31.33	77.89
3	0.01513	12.62	90.51
<b>Lamu County</b>			
1	0.08289	68.27	62.27
2	0.1956	16.11	84.38
3	0.0822	6.77	91.15
<b>Tana River County</b>			
1	0.1422	62.93	62.93
2	0.0079	34.95	97.88
<b>All Counties</b>			
1	0.03838	37.61	37.61
2	0.01429	14.01	51.62
3	0.01023	10.02	61.64

only two alleles per locus in this study, *Gst* is identical to *Fst*, Wright F statistics. *Gst* is independent of gene diversity within subpopulation and can be used to compare differentiation in different organisms. For this reason, *Hs* could be very small but the absolute *Gst* is small. Genetic analysis (Table 6) also revealed that these genetic similarities were very high while genetic distances were very low. This is an indication that coconuts within a region were not genetically diverse from those of the other regions. Coconut germplasm were sampled from random mating groups or individuals, thus gene diversity may be replaced with heterozygosity while gene identity may be referred to as homozygosity.

## Conclusion

Coconut sampled from different ecological zones failed to cluster per county or AEZ from where they were sampled. Molecular variability within the individuals was high while within individuals was low. This is an indication of the narrow genetic diversity in Kenya. Coconut breeding scheme should therefore be initiated to enhance variability. Such a scheme should include introductions from other areas while conserving the Kenyan coconut germplasm for future crop improvement.

## Conflict of interests

The authors have not declared any conflict of interest.

## ACKNOWLEDGEMENT

The authors are greatly indebted to the National Council for Science, Technology and Innovation (NACOSTI) Kenya for the financial support AND Kenya coconut Development Authority (KCDA) for the collection of the coconut samples.

## REFERENCES

- Akpan EEJ (1994). Evaluation of tall coconut (*Cocos nucifera* L.) genotypes within the Nigerian coconut germplasm bank. *Oleagineux*. 49:13-20.
- Aremu CO, Adebayo MA, Ariyo OJ, Adewale BB (2007). Classification of genetic diversity and choice of parents for hybridization in cowpea *Vigna unguiculata* (L.) Walp for humid savanna ecology. *Afr. J. Biotechnol.* 6(20):2333-2339.
- Ashburner GR, Thompson WK, Halloran GM (1997). RAPD analysis of South Pacific coconut palm populations. *Crop Sci.* 37(3):992-997.
- Batugal P, Oliver JT (2003). Poverty reduction in coconut growing communities, vol.I: The Framework and Project Plan, International Plant Genetic Resources Institute Regional Office for Asia, the Pacific and Oceania (IPGRI-APO), Serdang, Selangor, Malaysia. pp. 1-350.



- Dasanayaka PN, Everard JMDT, Karunanayaka EH, Nandadasa HG (2009). Analysis of coconut (*Cocos nucifera* L.) diversity using microsatellite markers with emphasis on management and utilization of genetic resources. *J. Natl. Sci. Found.* 37:99-109.
- Doyle JJ, Doyle JL (1990). Isolation of plant DNA from fresh tissue. *Focus.* 12:13-15.
- Fan H, Xiao Y, Yang Y, Xia W, Mason AS, Xia Z, Qiao F, Zhao S, Tang H (2013). RNA-Seq Analysis of *Cocos nucifera*: Transcriptome sequencing and *De Novo* assembly for Subsequent Functional Genomics Approaches. *PLOS ONE* 8:e59997.
- Gethi JG, Malinga WK (1997). Kenya Coconuts: Present Past and Future: Pp155-158. In: CP Topper. PDS Caligari, AK Kallaya, SH Shomari, PAL Masawe and A. Mpunani (Eds). Proceedings of the International Cashew and Coconut Conference: 'Trees for Life-Key to development'. Bio-Hybrids International, Reading, UK.
- IPGRI (1995). Descriptors for Coconut (*Cocos nucifera* L.). International Plant Genetic Resource Institute, Rome, Italy.
- Jaetzold R, Helmut S (1983). Farm Management Handbook of Kenya. 4 vols. Nairobi: Kenya Ministry of Agriculture.
- Kimura M, Crow JF (1964). The number of alleles that can be maintained in a finite population. *Genetics* 49:725-38.
- Lewontin RC (1972). Testing the theory of natural selection. *Nature* 236:181-182.
- Lewontin RC (1972). The apportionment of human diversity. *Evol. Biol.* 6:381-398.
- Liu K, Muse SV (2005). Power Marker: an integrated analysis environment for genetic marker analysis. *Bioinform.* 21:2128-2129.
- Manimekalai R, Nagarajan P (2006). Assessing genetic relationships among coconut (*Cocos nucifera* L.) accessions using inter simple sequence repeat markers. *Sci. Hort.* 108:49-54.
- Manimekalai R, Nagarajan P (2007). Use of Simple Sequence Repeat markers for estimation of genetic diversity in coconut (*Cocos nucifera* L.) germplasm accessions. *J. Plant Biochem. Biotechnol.* 16(1):29-33.
- Muhammed N, Nyamota R, Hashim S, Malinga JN (2013). Zygotic embryo *in vitro* culture of *Cocos nucifera* L. (sv. East African Tall variety) in the coastal lowlands of Kenya. *Afr. J. Biotechnol.* 12(22):3435-3440.
- Nei M (1973). Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA* 70:3321-3323.
- Nei M (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics.* 89(3):583-90.
- Oyoo ME, Muhammed N, Githiri SM, Ojwang PO, Muniu FK, Masha E, Owuochi J (2015). *In-situ* morphological characterization of coconut in the coastal lowlands of Kenya. *Afr. J. Crop Sci.* 9(2):65-74.
- Perera L, Russell JR, Provan J, Powell W (2000). Use of microsatellite DNA markers to investigate the level of genetic diversity and population genetic structure of coconut (*Cocos nucifera* L.). *Genome* 43(1):15-21.
- Perera L, Russell JR, Provan J, Powell W (1999). Identification and characterization of microsatellites in coconut (*Cocos nucifera* L.) and the analysis of coconut population in Sri Lanka. *Molecular. Ecology* 8(2):344-346.
- Perrier X, Jacquemoud-Collet JP (2006). DARwin software.
- Rajesh MK, Arunachalam V, Nagarajan P, Lebrun P, Samsudeen K, Thamban C (2008). Genetic survey of 10 Indian coconut landraces by simple sequence repeats (SSRs). *Scientia Horticulturae* 118(4):282-287.
- Rajesh MK, Sabana AA, Rachana KE, Shafeeq Rahman, Jerard BA, Anitha K (2005). Genetic relationship and diversity among coconut (*Cocos nucifera* L.) accessions revealed through SCoT analysis. *3 Biotech.* 5(6):999-1006.
- Rivera R, Edwards KJ, Barker JHA, Arnold GM, Ayad G, Hodgkin T, Karp A (1999). Isolation and characterization of polymorphic microsatellites in *Cocos nucifera* L. *Genome.* 42:668-675.
- Smith J, Chin E, Shu H, Smith O, Wall S, Senior M, Mitchell S, Kresovich S and Ziegler J (1997). An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays* L.): Comparisons with data from RFLPs and pedigree. *Theor. Appl. Genet.* 95:163-173.
- Sugimura Y, Itano M, Salud CD, Otsuji K, Yamaguchi H (1997). Biometric analysis on diversity of coconut palm: cultivar classification by botanical and agronomical traits. *Euphytica* 98:29-35.
- Teulat B, Aldam C, Trehin R, Lebrun P, Barker JHA, Arnold GM, Karp A, Baudouin L, Rognon F (2000). An analysis of genetic diversity in coconut (*Cocos nucifera* L.) populations from across the geographic range using sequence-tagged microsatellites (SSRs) and AFLPs. *Theor. Appl. Genet.* 100(5):764-771.
- Upadhyay A, Jayadev K, Manimekalai R, Parthasarathy VA (2004). Genetic relationship and diversity in Indian coconut accessions based on RAPD markers. *Sci. Hort.* 99(3-4):353-362.
- Xiao Yong, YiLuo, YaodongYang, HaikuoFan,Wei Xia, Annaliese S. Mason, SonglinZhao, Ross Sager FeiQiao (2013). Development of microsatellite markers in *Cocos nucifera* and their application in evaluating the level of genetic diversity of *Cocosnucifera*. *Plants Omics J.* 6(3):193-200.