

Full Length Research Paper

Genetic diversity of *Dacryodes edulis* provenances used in controlled breeding trials**Josephine Therese Makueti^{1*}, Gordon Otieno², Zac Tchoundjeu¹, Alice Muchugi³, Alain Tsobeng¹, Ebenezer Asaah⁴ and Robert Kariba³**¹Tree Diversity, Domestication and Delivery, World Agroforestry Centre (ICRAF), P. O. Box 16317, Yaounde, Cameroon.²Department of Biochemistry and Biotechnology, Kenyatta University, P. O. BOX 43844 00100, Nairobi Kenya.³Tree Diversity, Domestication and Delivery, World Agroforestry Centre (ICRAF), P. O. Box 30677, 00100, Nairobi, Kenya.⁴Tree Diversity, Domestication and Delivery, World Agroforestry Centre (ICRAF), P. O. Box 210, Freetown, Sierra Leone.

Received 11 March, 2015; Accepted 29 June, 2015

Genetic diversity characterization among *Dacryodes edulis* accessions is important in developing genetically diverse and superior cultivars. This study assessed the genetic diversity of three *D. edulis* provenances used in controlled breeding trials at ICRAF-Cameroon based on SSR marker technique. Genomic DNA isolated from individual *D. edulis* samples were primed with six SSR primers (CB09, LD06, CG11, LB12 CE09 and CC01), amplified through PCR and the products resolved using the ABI 3730 DNA analyzer. Alleles using the Gene Mapper software were called. The polymorphism information content (PIC), gene diversity and heterozygosity estimates were computed by using PowerMarker software while GenAlEx software was employed to estimate gene flow levels, to reveal partitioning of variation across the populations and display the genetic relationships among the populations. DARwin software was used to construct a dendrogram that also displayed the genetic relationships among the populations. Analysis of Molecular Variance (AMOVA) was used to assess the structure of genetic diversity among provenances. All the 6 primer pairs used in this study were polymorphic. Gene diversity values were calculated for all analyzed loci ranged from 0.2763 to 0.8846 for the loci CG11 and CE09, respectively, with a mean of 0.5851. Heterozygosity values ranged from 0.0220 and 0.5385 for the loci CG11 and LD06, respectively, with an average of 0.3132. A significant low level of genetic differentiation ($F_{ST} = 0.018$, $p=0.01$) was detected among the populations. AMOVA analysis indicated that only 2% of genetic variation existed in populations. The neighbor (joining tree) and PCoA analysis could not reveal a clear separation between populations. Consequently, this study established that the high common level of gene flow obtained could help in maintaining low genetic differentiation and closing genetic relationships of *D. edulis* populations. Thus introgression of new alleles may be required to broaden the genetic base of the *D. edulis* provenances. This study also highlighted the need of isolation and characterization of more DNA markers in *D. edulis* and their uses in advanced studies such as gene discovery, trait linked marker assisted selection and gene mapping.

Key words: Controlled breeding, DNA markers, gene flow, heterozygosity, provenance.

INTRODUCTION

The progressively deteriorating situation of the environment under tropics has often been associated to

declining rainfall and over-exploitation of phytogenetic resources by local communities in the developing

countries (Megevand et al., 2013). Moreover, the degradation of this natural ecosystems due to subsistence agriculture, demographic pressure, urbanization, infrastructure and mining developments, unsustainability harvesting of both timber and non-timber forest products from trees such as *Dacryodes edulis* (G.Don) H.J.Lam contribute to accelerated deforestation (Alemagi and Kozak, 2010; Phelps et al., 2013). Immediate consequences are lack of availability of wild fruits, medicinal plants and other plant and animal products. This exposes the most vulnerable segment of the communities, the aged, the poor, women and children, to malnutrition and reduced income, as traditionally their livelihoods partly depend on forest products (Cribb, 2010).

D. edulis is an edible fruit tree indigenous to the Gulf of Guinea and central Africa regions. However, due to human activities, its current geographical distribution extends beyond its area of origin to as far as tropical Asia (Kengue, 2002). It belongs to the Burseraceae family, a tropical tree family known for fragrant resins, which include frankincense, myrrh, and copal (Onana, 2008). The species contributes to rural incomes, supplements the local diet and is used in traditional and modern therapies (Anegbeh et al., 2005; Schreckenberget al., 2006).

In the last few years, agroforestry has been more fully recognized as a land-use practice that can contribute to both biodiversity conservation and livelihood development (Pye-Smith, 2010; Leakey, 2012), and this has led to greater interest in understanding the genetic variation of trees within these systems (Muchugi et al., 2006). This is especially so as the roles of intraspecific diversity in underpinning wider ecosystem functions have become evident. This interest needs to be underpinned by increased genetic research on farm populations of agroforestry tree species such as *D. edulis*. For most species, how genetic diversity is distributed within geographical space is unclear. Furthermore, how the distribution of this variation depends on the breeding system of a species (Assogbadjo et al., 2011), and on processes such as climate change, floral evolution, forest management and cultivation, is largely unknown (Dawson et al., 2009). Lack of understanding is a disincentive to crop improvement programs and the conservation of genetic resources, because it's difficult or impossible to choose in the different options that are available for management of a particular species, only some options will prove it to be sustainable (Atangana, 2010). Therefore, understanding the molecular basis of this essential biological phenomenon in plants is crucial for the efficient conservation, management and utilization of

plant genetic resources (Mondini et al., 2009). It has been demonstrated that DNA markers are the most promising techniques used to differentiate among genotypes at species and subspecies level (Kumar et al., 2009). Furthermore, DNA based markers have become methods of choice in genetic diversity studies, as they analyse variation at DNA level. This excludes all environmental influences and time specificity, since analysis can be performed at any growth using any plant part and requires only small amounts of material (Rao, 2004, Khanam et al., 2012). In this study therefore genetic diversity analysis was done using simple sequence repeats (SSR) DNA based markers. This approach was chosen because it had already been used with success to determine genetic diversity of many tropical species including *Allanblackia* spp. (Russell et al., 2009; Atangana et al., 2010); *Citrus* spp (El-Mouei et al., 2011) and *Carica papaya* (Madarbokus and Sanmukhiya, 2012). Furthermore, comparative studies in plants have shown that SSR markers are more valuable than other DNA based markers and provide an effective means for discriminating between genotypes (Powell et al., 1996; Li et al., 2001). Hence, the limited knowledge regarding the genetic diversity of the *D. edulis* provenances used in controlled breeding trials at ICRAF has been a disincentive to the breeding program. Thus, the objective of this study was to assess the genetic diversity of three *D. edulis* provenances used in controlled breeding trials at ICRAF-Cameroon based on SSR marker technique, in order to provide crucial background genetic diversity information for the efficient management and utilization of the collections in the breeding programme.

MATERIALS AND METHODS

Plant material

Ninety one genotypes, 40 collected from Boumyebel, 13 collected from Limbe and 38 collected from Kekem were used in the study. Accessions were divided into three populations according to their geographical origin: Boumyebel, Limbe and Kekem. The different sites of Cameroon where the provenances were collected can be distinct in agro-ecological zones namely: (i) humid forest with monomodal rainfall; (ii) humid forest with bi-modal rainfall and (iii) western high lands savannah (Table 1; Figure 1).

DNA isolation

DNA was isolated from *D. edulis* leaf samples preserved in silica gel. About 200 mg of *D. edulis* leaf tissue was ground to fine paste. A mixture of approximately 550 µl of CTAB extraction buffer (2% CTAB), 100 mM Tris-HCL (pH 7.5), 1.4 M NaCl, 20 mM EDTA (pH

*Corresponding author. E-mail: jmakueti@cgiar.org. Tel: +237 696 68 21 12.

Table 1. Characteristics of the agro-ecological zones (Adapted from Ambassa-Kiki, 2000).

Population	Agro-ecological zone	Altitude (m)	Latitude	Longitude	Temperature	Rainfall (mm)	Soil type
Boumnyebel Limbe	HFZ with two raining seasons	500-1000	2°6' - 5°48'	10°30' - 16°12'	< 25°C	1500 - 2000	Acidic pH 4-5.5
Kekem	HFZ with one raining season	2000-4095	2°6' - 6°12'	8°48' - 10°30'	22 - 29°C	2500 - 4000	AcidicH 3.5-4.8

Legend: HFZ=Humid Forest Zone

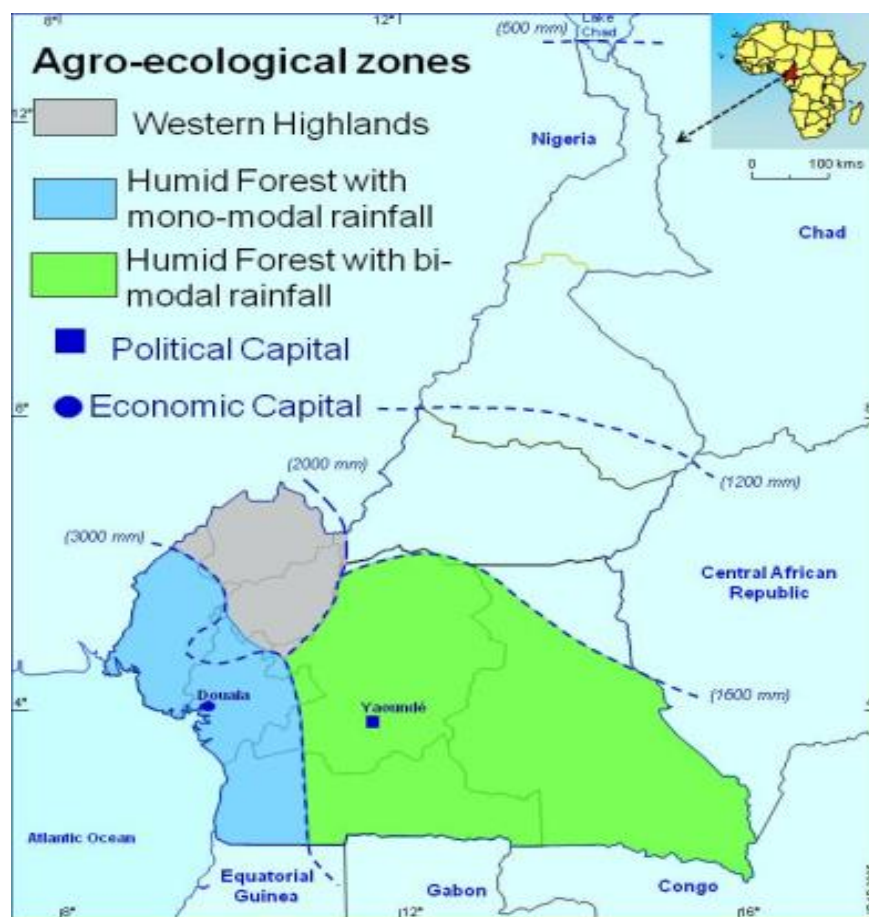


Figure 1. Corresponding agro-ecological zones where *D. edulis* naturally grows in Cameroon (adapted from Amabassa-Kiki, 2000).

Table 2. Characteristics of the primer pairs used for analysis. H_O: observed heterozygosity; H_E: expected heterozygosity.

Marker	Allele size range (bp)	Repeat motif	Dye*	Primer sequences (5' – 3')	H _O	H _E
CB09	163 – 200	(GA) 10	6-FAM	F: CAGTGGTAAAGGTAGGATGG R: GGTTTGTGTTGGTTGTTTTTC	0.64	0.71
LD06	152 – 175	(GA) 10	PET	F: CATGTGCATCTTTTGTATCTC R: TAGAGGGGAAAGGTAGAAGG	0.56	0.64
CG11	170 – 185	(CT) 8	VIC	F: CACACGATCAAGAAAGACC R: CCCGTAATGACTCTCTAATTG	0.07	0.07
LB12	205 – 230	(GA) 14	PET	F: AGACCACAAGAACCTTTGC R: GGCCACTTAATTTGATTTTG	0.27	0.28
CE09	125 – 168	(GA) 16	VIC	F: GTAGCTGGGAGGAGATCAGG R: ACAGGCCCCAAAAGCAC	0.84	0.85
CC01	325 – 375	(CT) 9	PET	F: ATGTTCCGAGGTAGTTCTTGC R: AGAAGATGTGGAAAATGGTG	0.42	0.42

8.0), 0.2% Mercaptoethanol and a little amount of polyvinyl pyrrolidone (PVP) preheated at 65 °C was added to the ground tissue and mixed gently in a 2 ml Eppendorf tube. The CTAB and plant extract mixture were then incubated for about 30 min at 65°C in a re-circulating water bath. After incubation, 550 µl of Chloroform:Iso Amyl Alcohol (24:1) was added to each tube and the solution mixed by inversion. After mixing, the tubes were centrifuged at 14000 rpm for 15 min at room temperature to spin down cell debris. The upper aqueous phase (contains DNA) was then transferred to a clean 1.5 ml Eppendorf tube. To each tube, 0.7 volume of ice cold isopropanol was added to precipitate the DNA. The solution was mixed by inversion followed by overnight incubation at -20°C. The precipitated DNA was then isolated by spinning the tubes at 14000 rpm for 10 min at 4°C to form a pellet.

The precipitated DNA stuck to the bottom of the Eppendorf tube and was visible as a clear solid precipitate. The upper phase was carefully discarded leaving the DNA precipitate at the bottom of the Eppendorf tube. To wash the DNA, 500 µl of ice cold 70% ethanol was added to each tube and mixed gently by inversion. After mixing, the tubes were centrifuged at 14000 rpm for 10 min to pellet

the DNA at the bottom of the tube. The supernatant was removed and the DNA washed again with ice cold 70% ethanol. The supernatant was discarded and the DNA pellet allowed to air dry for approximately 45 min. The DNA was re-suspended in 100 µl ddH₂O). To remove RNA, 4µl of RNase Cocktail (RNase A (500 U/ml) and RNase T1 (20,000 U/ml) was added and the solution incubated at 37°C for 30 min. The isolated DNA was stored at 4°C until analysis.

DNA quality and concentration analysis

Quality and quantity of DNA samples were determined by comparison to a DNA standard marker (Lambda DNA, Promega USA). 2 µl of each DNA sample was mixed with 1 ul of SYBR® green fluorescent dye and 2 ul loading buffer (xylene cyanol, sucrose and bromophenol blue). Each mixture was then loaded into a well prepared on 0.8% agarose gel. Electrophoresis was then carried for 1 h at 80V and DNA sample band intensities were compared with that of the DNA standard marker. Moreover, the concentration and purity of the DNA was determined using

the spectrophotometer at an absorbance ratio of 260 nm (A₂₆₀) and 280 nm (A₂₈₀). Quantification of double-stranded (ds) DNA was achieved by using 1 µl of each extract and scanning its absorbance over a 200 to 300 nm range with Nucleic Acid settings applied for the instrument. Based on this reading all extracts were standardised to a concentration of 20 ng/µl using necessary dilution.

SSR analysis

A total of 6 polymorphic SSR markers published for *D. edulis* by Benoit et al. (2011) were used to screen 91 *D. edulis* DNA samples (Table 2). PCR amplification reactions were performed with MyTaq DNA polymerase (Bioline). The PCR reactions were set up as follows: 4 µl of 5x MyTaq Reaction Buffer Colorless comprising 5 mM dNTPs solution and 15 mM MgCl₂, 2 µl of 10 µM forward primer labelled with either 6-Fam or Vic or Ned or Pet (Applied Biosystems), 2 µl of 10µM reverse primer (unlabeled) solution, 0.4 µl MyTaq DNA polymerase, 3 µl of diluted 20 ng/µl DNA, and 8.6 µl of ddH₂O to bring the total volume

to 20 μ l. The amplification was carried out using GeneAmp® PCR System 9700 thermal cycler (Applied Biosystems, USA). For each SSR amplification process, an initial denaturation of DNA at 94°C for 3 min was followed by 35 cycles of denaturation at 94°C for 20 s, annealing at 57°C for 1 min and extension at 72°C for 1 min with the final extension at 72°C for 10 min. The PCR products were then stored at 4°C awaiting analysis.

Electrophoresis and PCR product visualization

The PCR products were separated on 2% agarose gels using 1X TBE buffer (44.5 mM boric acid, 1 mM EDTA and 44.5 mM Tris base). The DNA fragments were stained with 1 μ l of SYBR® green fluorescent dye and mixed with 2 μ l loading buffer (xylene cyanol, sucrose and bromophenol blue). Each well was loaded with 5 μ l of the mixture (PCR products, SYBR® green fluorescent dye and loading buffer). 1 μ l of 100 bp DNA Ladder (Invitrogen) was loaded in the first well of each gel. Electrophoresis was carried out for 11/2 h at 80 V. Subsequently, the DNA bands were visualized using a UV transilluminator and photographed. The sizes of the amplified DNA fragments were determined by comparing the migration distance of amplified fragments relative to the molecular weight of the 100 base pairs (bp) DNA ladder.

Scoring of DNA fragments

DNA fragments were pooled together along with internal size standard (GeneScan™ 500 LIZ® from Applied Biosystems, USA) based on dye label. The PCR products were co-loaded based on dye label, to reduce the cost of genotyping. ABI 3730xl Genetic Analyzer (Applied Biosystems, USA) was then used to carry out Capillary electrophoresis. Genetic Analyzer (Applied Biosystems, USA) employs a fluorescent based capillary detection system that uses polymer as the separation matrix (Kitavi et al., 2014). This facilitates the accurate sizing of the alleles. Based on the relative migration of the internal size standard, data generated by the ABI 3730xl Genetic Analyzer was then analyzed using GeneMapper 3.5 software (Applied Biosystems, Foster City, California, USA) and the allele sizes scored in base pairs (bp) based on predefined analysis protocols. Only alleles with a relative fluorescent unit of > 500 were scored.

Gene diversity, heterozygosity, and polymorphism

Gene diversity, heterozygosity, and polymorphism information content (PIC) for all accessions were calculated using PowerMarker Version 3.25 (Liu and Muse, 2005). Gene diversity was defined as the probability that two randomly selected alleles from the accessions are different. PIC was used to deduce the polymorphic informativeness of each SSR primer as described by Weir (1996). PowerMarker Version 3.25 was also used to generate the major allele frequency. The major allele frequency is the frequency of the most common allele at a locus (Lu et al., 2011).

Analysis of molecular variance

The analysis of molecular variance (AMOVA) was done by employing GenAlEx software version 6.5 (Peakall and Smouse, 2012) in the aim of assessing the structure of genetic diversity among *D. edulis* provenances. AMOVA revealed the partitioning of genetic variation among populations, among individuals, within populations and within individuals. The variance components were used to calculate F-statistics which is useful for analysis of gene flow levels, where F_{ST} reflects genetic differentiation; F_{IT} is the

inbreeding coefficient within individuals relative to the total and F_{IS} is the inbreeding coefficient within individuals relative to the subpopulation (Wright, 1978; Kiambi et al., 2005; Kahiu et al., 2013). F-statistics values were tested by 99 permutations for significance in all analyses (Peakall and Smouse, 2012).

Cluster analysis

Cluster analysis is the partitioning set of objects into groups so that objects within a group are similar and objects in different groups are dissimilar. It is efficient in grouping objects with similar characters (Hodgkin et al., 1995). Genetic relationships among the provenances were displayed by classification using dissimilarity matrix (dissimilarity = 1 – similarity). The dendrogram was constructed with the weighted neighbor joining method on the basis of allelic data from 6 SSR markers across the *D. edulis* accessions using DARwin software Version 5.0 (Perrier et al., 2003; Perrier and Jacquemoud-Collet, 2006). Average linkage cluster between provenances was analyzed. Clusters were defined based on their unique characters.

Principle coordinate analysis

The genetic relationships among the genotypes were also analyzed using PCoA. PCoA was based on standardized covariance of genetic distances calculated for codominant markers. Genetic distances were computed using GenAlEx software version 6.5 (Peakall and Smouse, 2012). Genetic distance matrices generated from the SSR data sets were then subjected to Principal Coordinate Analysis by employing GenAlEx software version 6.5 (Peakall and Smouse, 2012).

RESULTS AND DISCUSSION

DNA quality and quantity

The quality and quantity of extracted genomic DNA is vital for successful SSR analysis. Agarose gel 0.8% was used for quality control of genomic DNA. Bands were visualized using a UV transilluminator and photographed. The quality and quantity of DNA extracted from samples were determined by comparison to a 50 ng/ μ l DNA standard marker (Lambda DNA, Promega USA). The extracted DNA produced strong bands of high molecular weight with little or no smear and thus the samples were of the required quality and quantity suitable for PCR amplifications (Figure 2).

DNA concentration and purity

To ensure that a sufficient amount of DNA template is used in analysis it is important to determine the concentration and purity of DNA of samples. However, the presence of residual RNA leads to overestimation of the concentration and purity of DNA hence correct RNA elimination is important. The 260/280 ratio is a good estimate of extract purity and therefore the concentrations and purity of the DNA samples were determined using the NanoDrop 3300 spectrophotometer

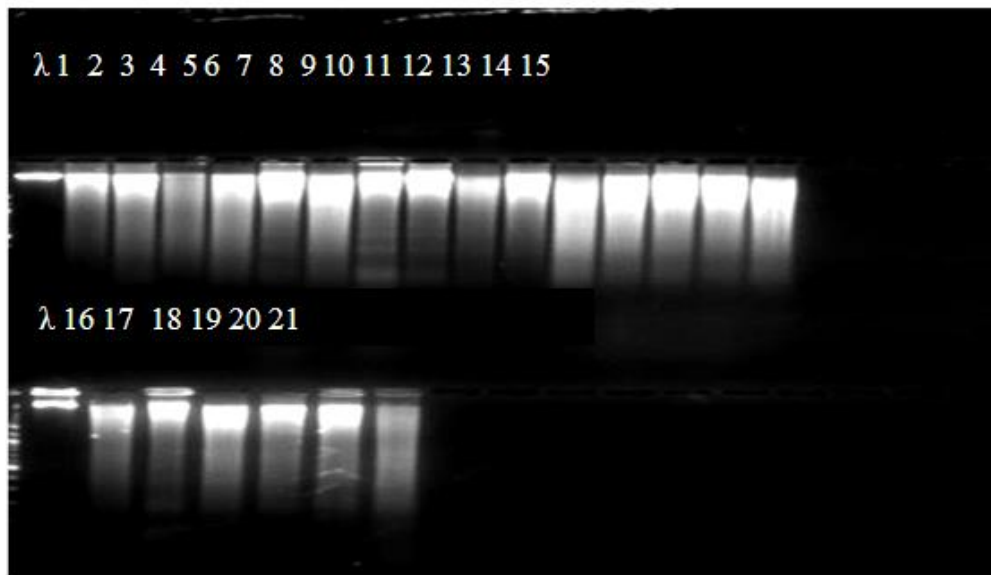


Figure 2. Genomic DNA extracted from *D. edulis* genotypes using CTAB extraction method. The first well of each row represents a molecular size marker (Lambda DNA, Promega USA), whereas the rest represent *D. edulis* DNA samples.

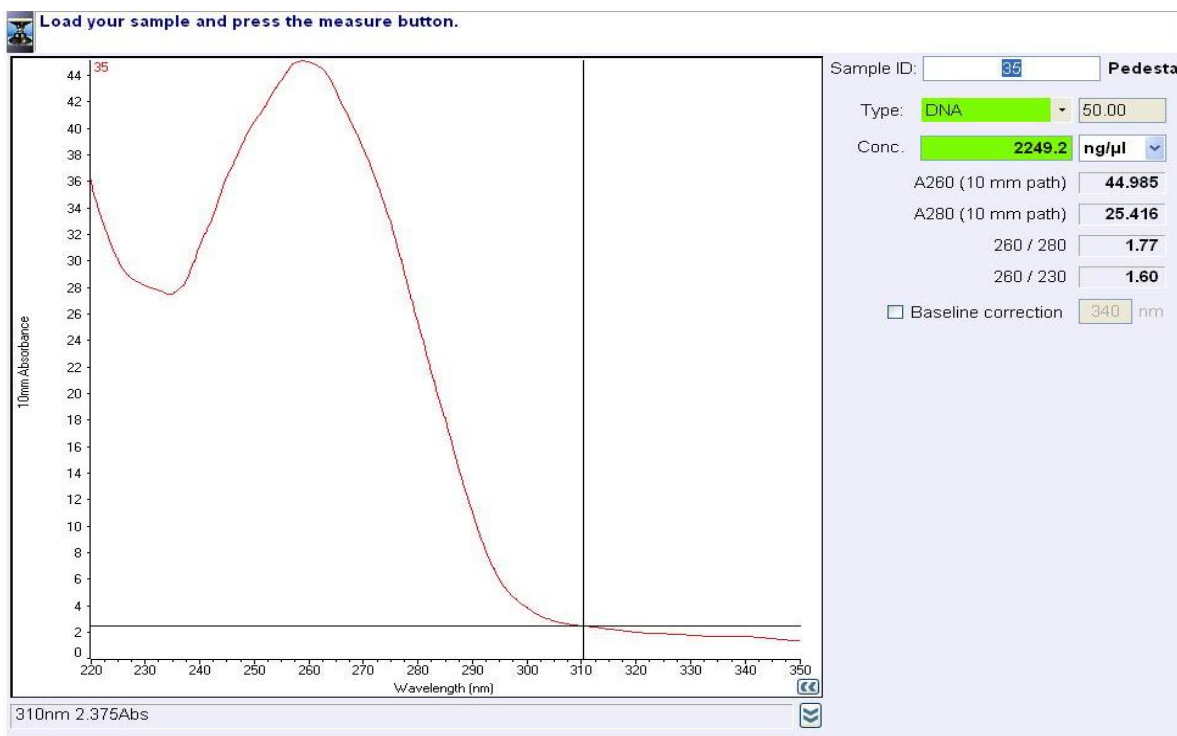


Figure 3. Screenshot of the spectrophotometer DNA concentration output for sample 35.

(Applied Biosystems) at an absorbance ratio of 260 nm (A260) and 280 nm (A280) (Figure 3). No extracts displayed a 260/280 ratio > 2.0 which suggested there

was minimum levels of residual RNA present (Table 3, Appendix 1). The DNA concentrations ranged from 941.3 to 6738.8 ng/μL, which was sufficient for dilution and

Table 3. Spectrophotometer output of DNA concentrations of selected samples.

Sample ID	Nucleic acid concentration	Unit	260/280	Factor
1	3265.4	ng/μl	1.73	50
2	3555.2	ng/μl	1.76	50
3	3618.9	ng/μl	1.79	50
4	3058.5	ng/μl	1.79	50
5	3119.3	ng/μl	1.66	50

subsequent amplification of SSR regions. A few examples of the spectrophotometer outputs of DNA concentrations of selected samples are presented in Table 3. Upon quantification, all samples were normalized to a final concentration of 20 ng/μl by adding variable volumes of double distilled water to a final volume of 100 μl.

PCR product visualization

The success of PCR amplification was checked through electrophoresis. Agarose gel 2% was used during electrophoresis. Bands were visualized using a UV transilluminator and photographed. Successful amplification of the amplified DNA fragments was determined by comparing the migration distance of amplified fragments relative to the molecular weight of the 100 base pairs (bp) DNA ladder. Successful amplification was noted in all the six primer pairs used in the analysis. Strong bands were seen on the gel images and these bands were within the required base pair size ranges relative to the 100 base pairs (bp) DNA ladder bands and thus suitable product yields were obtained. Moreover, primer dimers were not observed thus amplification was efficient and the amplification products were suitable for genotyping. Below is an agarose gel electrophoresis image showing PCR Amplification Products with Primer LB12 (Figure 4).

Scoring of amplified DNA fragments

GeneMapper 3.5 software was used to display SSR profiles generated on capillary sequencer and to score allele sizes of all the 91 *D. edulis* PCR amplification products for the 6 SSR primers used in this study. SSR profiles had good peaks, high heterozygote peak height ratios and were free of artefacts (Figure 5). The peaks were sized and the alleles called based on predefined analysis protocols.

Polymorphism, heterozygosity, and gene diversity

The number of alleles at a determined SSR locus (allelic

richness) is the simplest measure of genetic diversity (Saavedra et al., 2013). The six SSR markers used in this study produced a total of 61 alleles, with an average of 11.1667 alleles per marker. Marker CE09 detected the highest number of alleles with 17 alleles while CBE09, LD06, LB12 and CC01 had 14, 12, 11 and 6 alleles, respectively. Marker CG11 detected the least number of alleles with 6 alleles (Table 4). The average number of alleles per marker detected in this study is comparable to that reported by Benoit et al. (2011) for *D. edulis*. The number of alleles per locus observed by Benoit et al. (2011) ranged from 3 to 15, with an average of 7.3.

The polymorphism information content (PIC) was defined by Botstein et al. (1980) as the measure to calculate the discrimination power and informativeness of the SSR markers. PIC value is therefore a measure of polymorphism among genotypes for a marker locus used in genetic diversity analysis since it reflects allelic diversity and frequency among the genotypes. Microsatellites, which provide these polymorphic regions of the genome, allow estimation of such parameters as average observed and expected heterozygosities. All the 6 primers were polymorphic across the 91 *D. edulis* samples. PIC values varied greatly for all tested SSR loci and these values ranged from 0.2652 to 0.8737 for the loci CG11 and CE09, respectively, with a mean of 0.5632 (Table 4). The highest PIC value 0.8737 was obtained for CE09 followed by (0.7525) for CB09. The results of Todou et al. (2013) are in agreement with this finding, since they reported that the most polymorphic loci were CB09 and CE09. Furthermore, according to Ganapathy et al. (2012), markers with PIC values exceeding 0.5 are very efficient in discriminating genotypes and extremely useful in detecting the polymorphism rate at a particular locus. In this study four markers (CC01, LD06, CB09 and CE09) had PIC values that exceeded 0.5 and therefore were the most useful in detecting the rate of polymorphism.

The observed heterozygosity (H_o , also known as heterozygosity) values ranged from 0.0220 to 0.5385 for the loci CG11 and LD06, respectively, with an average of 0.3132. Meanwhile, gene diversity (H_e , also known as expected heterozygosity) was 0.5851 on average and ranged from 0.2763 (CG11) to 0.8846 (CE09). The estimates of gene diversity are comparable with those estimated by Benoit et al. (2011). Benoit et al. (2011)

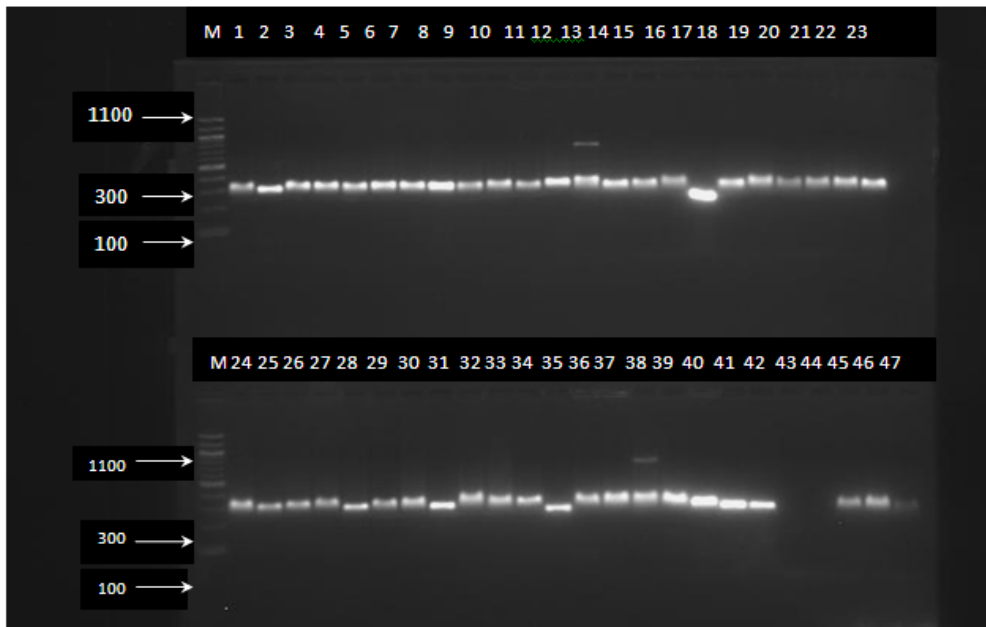


Figure 4. *D. edulis* PCR amplification products with Primer LB12. M designates a 100 bp ladder.

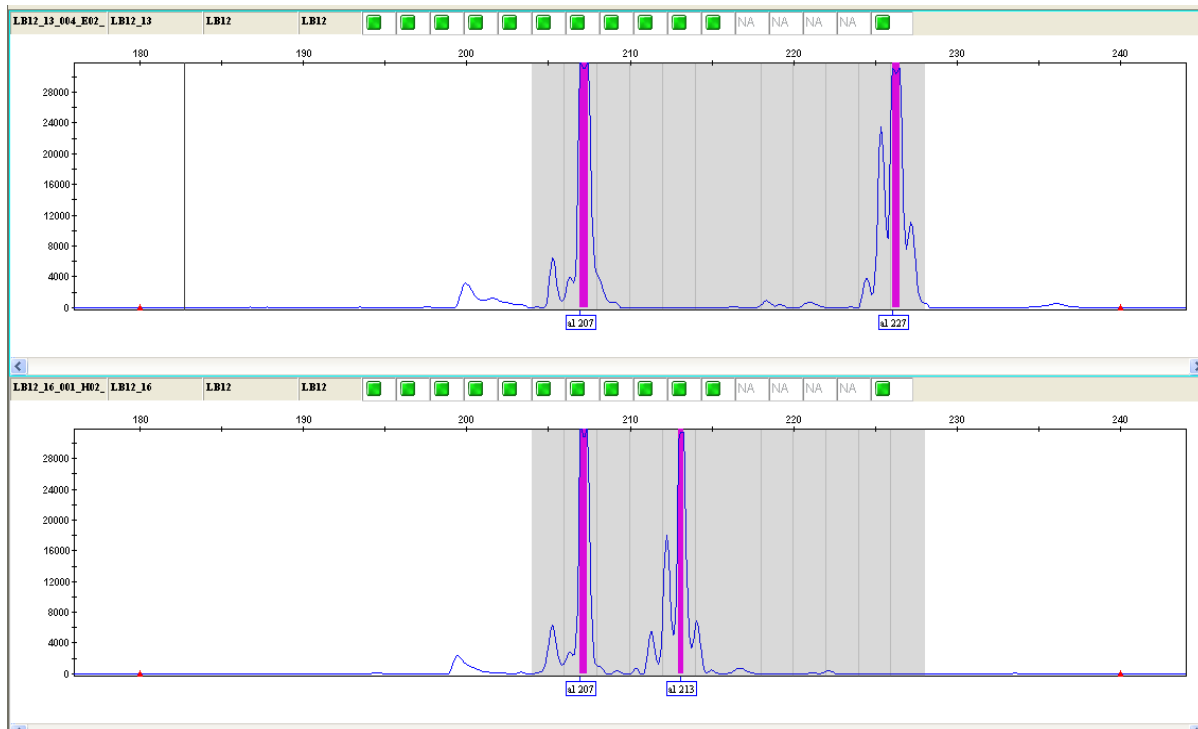


Figure 5. Electropherograms illustrating SSR profiles generated on capillary sequencer and allele sizes of two *D. edulis* genotypes for primer LB12 scored using GeneMapper software. The numbers at the bottom of each peak indicate allele size in base pairs (bp).

reported that polymorphism was widely variable among loci and that expected heterozygosities ranged from 0.06

to 0.84 with a mean value of 0.49. Consequently, an average level of genetic diversity was observed at

Table 4. Gene diversity, heterozygosity, and polymorphism at 6 SSR loci in three *D. edulis* populations.

Marker	Major Allele Frequency	Genotype N°	Allele N°	Gene diversity	Heterozygosity	PIC
LD06	0.4780	13.0000	12.0000	0.6477	0.5385	0.5894
CE09	0.1703	39.0000	17.0000	0.8846	0.5055	0.8737
LB12	0.7747	13.0000	11.0000	0.3926	0.2198	0.3840
CB09	0.4066	24.0000	14.0000	0.7743	0.4505	0.7525
CG11	0.8462	6.0000	6.0000	0.2763	0.0220	0.2652
CC01	0.6648	13.0000	7.0000	0.5351	0.1429	0.5145
Mean	0.5568	18.0000	11.1667	0.5851	0.3132	0.5632

Table 5. Summary Analysis of Molecular Variance for the *D. edulis* samples.

Source	df	Sum of squares	Mean square	Estimated variance	%
Among populations	2	8.684	4.342	0.032	2
Among individuals within populations	88	225.294	2.560	0.810	45
Within individuals	91	85.500	0.940	0.940	53
Total	181	319.478		1.782	100

Table 6. F-statistics.

F-Statistics	Value	P(rand >= data)
F _{ST}	0.018	0.010
F _{IS}	0.463	0.010
F _{IT}	0.473	0.010

population levels. Generally these findings illustrates that although the 6 *D. edulis* SSR markers used in this study present a high discriminating power, however, they are under selective pressure.

This is typically and/or ideally, only the best individuals are being brought into domestication programmes by farmers, so domestication is generally considered to reduce the genetic diversity of the species being domesticated (Miller and Gross, 2011; Graefe et al., 2013;). Moreover, the expected heterozygosity across the 6 loci was higher than the observed heterozygosity. The major differences between the observed heterozygosity and gene diversity for all the microsatellite markers was therefore an indication of the presence of inbreeding among the accessions.

Analysis of molecular variance

Analysis of molecular variance (AMOVA) was chosen for partitioning the total molecular variation present at the inter- and intra-population levels. A low inter-population variation was found (2%). Furthermore, the AMOVA results (Table 5) showed that most of the molecular

variance (98%) was distributed within populations with 45% of this variation partitioned among individuals and 53% within individuals. The results of Todou et al. (2013) are also in agreement with this finding, since they reported that only 3% of the total detected genetic diversity was inter-population.

The above variance components were used to calculate F-statistics. F-statistics values are useful for the analysis of the magnitude of genetic differentiation, where F_{ST} provides a measure of the genetic differentiation between subpopulations; F_{IT} is the inbreeding coefficient within individuals relative to the total and F_{IS} is the inbreeding coefficient within individuals relative to the subpopulation (Kiambi et al., 2005 in Wright, 1978). The observed overall inbreeding coefficient F_{IT} (0.473) revealed that the *D. edulis* populations studied were highly inbred overall. This inbreeding can partially be attributed to inbreeding within individual accessions F_{IS} (0.463). Consequently a low level of genetic differentiation was detected as exhibited by the low fixation index (F_{ST} = 0.18) (Table 6). This F_{ST} value was slightly lower than that observed in a genetic diversity study of different populations of *D. edulis* from Cameroon and Gabon (Todou et al., 2013). Todou et al. (2013)

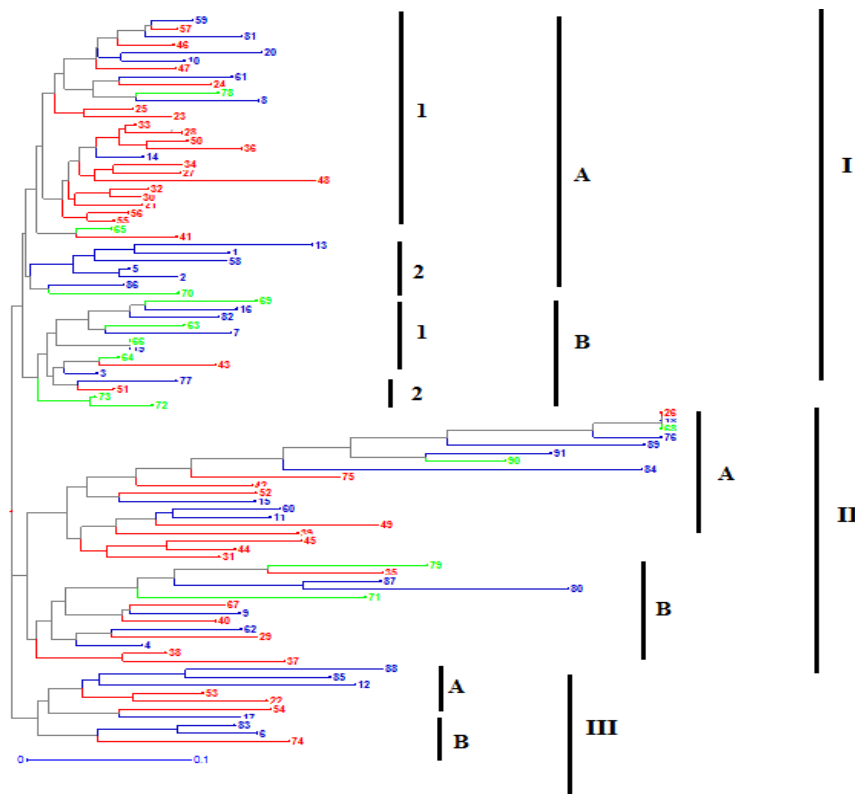


Figure 6. Dendrogram of *D. edulis* genotypes based on SSR markers. Populations have been represented with different colors (Kekem in blue, Boumnyebel in red and Limbe in green).

reported $F_{ST} = 0.03$ for the two populations of *D. edulis* from Cameroon and Gabon on the basis of five different SSR loci.

Wright (1978) cited by Kiambi et al. (2005) suggested that an F_{ST} ranging between 0.15 and 0.25 indicates large differentiations, 0.05 to 0.15 moderate, and 0 to 0.05 little differentiations. F_{ST} exceeding 0.25 suggest extremely large differentiations. Hence this study revealed low levels of genetic differentiation among populations of *D. edulis* accessions studied. Additionally, these findings indicate that most of genetic diversity detected in this study is intra-population. The low level of genetic differentiation among the populations studied and the fact that most of the genetic diversity detected in this study is intra-population can be attributed to high levels of gene flow. This is because populations which have a high common level of gene flow display lower F_{ST} values than populations of a lower gene flow. This result is fairly in line with that pointed out by Flynn (2009).

Cluster analysis

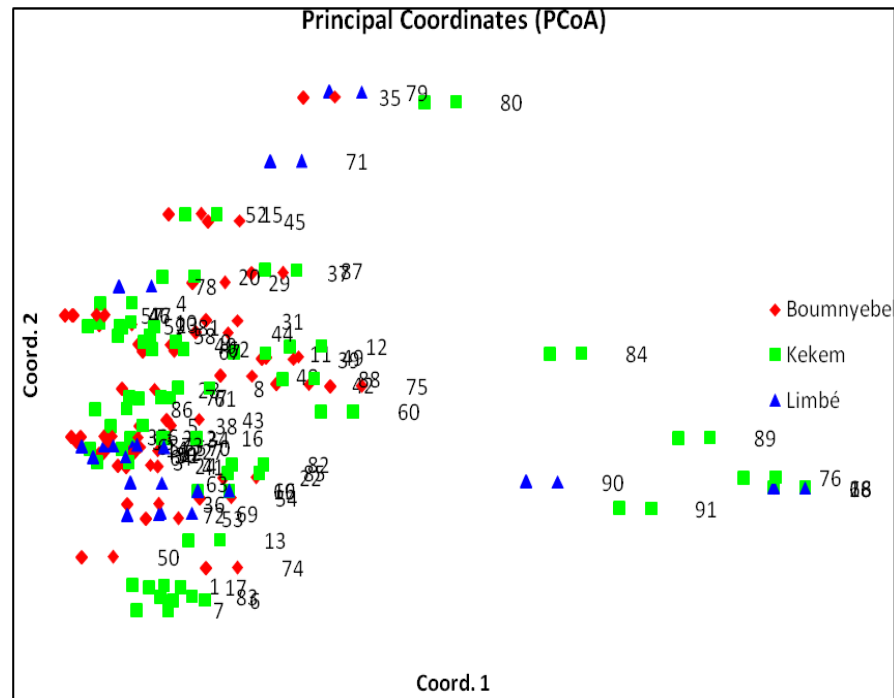
Traditionally, plant breeders have always evaluated plant origins and pedigrees as an extra indication of

the characteristics that were acquired by a new genotype. With the advent of molecular marker techniques, new tools that have the capacity to reveal similarities in the genome of related plant species or cultivars directly became available. Knowledge of genetic similarity between genotypes is useful in any breeding program because it facilitates efficient sampling and utilization of germplasm resources. The breeder can use genetic similarity information to make informed decisions regarding the choice of genotypes to cross for the development of populations or to facilitate the identification of diverse parents to cross in hybrid combinations in order to maximize the expression of heterosis (Hemeida et al., 1998).

Cluster analysis separated the accessions based on genetic variability. Accessions with similar genetic traits were grouped together irrespective of the collection region. Genetic relationships among the populations were displayed by classification using dissimilarity matrix (dissimilarity = 1 – similarity) and a dendrogram constructed (Figure 6). The 91 *D. edulis* accessions were divided into three major Clusters I, II and III. The first major Cluster I had two sub clusters, A and B: (i) Sub-cluster A had two groups 1 and 2 and comprised a total of thirty five accessions. Group 1 had the highest number

Table 7. Percentage of variation explained by the first 2 axes.

Axis	1	2
%	20.78	8.80
Cum %	20.78	29.58

**Figure 7.** A two-dimensional plot of the Principal Coordinate Analysis (PCoA) of SSR data showing the clustering of populations of *D. edulis*. The first and second principal coordinates account for 20.78 and 8.80% of total variation respectively.

of accession among the two groups of sub-cluster A of major Cluster I; (ii) Sub-cluster B comprised two groups 1 and 2 consisting only fourteen accessions. Major Cluster II had two sub-clusters, A and B comprising a total of thirty-two accessions. Major Cluster III also had two sub-clusters, A and B. However it had the least number of accessions among the three major clusters. *D. edulis* genotypes from Boumnyebel and Kekem were placed in all the three major clusters. On the other hand, none of the accessions from Limbe was placed in major cluster III. From the neighbor joining tree, it was clear that there was a close genetic relationship between the *D. edulis* populations analyzed in this study. The results of cluster analysis also showed that there was low relationship between genetic divergence and geographical origins.

Principle coordinate analysis

Principal coordinate analysis (PCoA) was also used to

visualize the genetic association of genotypes, which is primarily explained by the first two principal coordinates. This multivariate approach was chosen to complement the cluster analysis information since cluster analysis is more sensitive with closely related individuals, while PCoA is more informative regarding distances among major groups (Ghosh et al., 2014). The first and second coordinates extracted 20.78% and 8.80% of the total molecular variation, respectively (Table 7). PCoA also showed no clear separation of the populations and thus was consistent with cluster analysis. The genotypes were scattered in no distinct and meaningful groups, either based on population and/or geographical origins (Figure 7). The two-dimensional plot of the Principal Coordinate Analysis (PCoA) of *D. edulis* genotypes constructed from the 6 SSR marker data is similar in many respects to that of an AFC (factorial analysis of correspondence) plot constructed by Todou et al. (2013) from 5 SSR marker data.

In general, very little differentiation was observed

among the populations with respect to the F-statistics values. In addition, AMOVA revealed that most of genetic diversity detected in this study was intra-population while cluster analysis and Principal Coordinate Analysis revealed close genetic relationships among the populations. Consequently these results clearly illustrated the contribution of gene flow to genetic differences among populations. In low input farming systems, gene flow can be considered a function of pollen flow and seed exchange, which, in different ways, are influenced by natural and human selection pressures. *D. edulis* is dioecious species (Kengue, 2002) but there are individuals who carry both female flowers and male flowers, and individuals who bear hermaphrodite flowers and this therefore favors self-pollination. *D. edulis* fruits and seedlings are also usually exchanged between relatives, neighbors, communities and/or villages facilitating seed dispersal (Degrande et al., 2013; Takoutsing et al., 2013).

In addition, transactions allow the transport of commercial fruits (seeds) for long distances in search of markets and thus farmers select fruits of good quality in order to create markets for their crops (Degrande et al., 2012). This contributes to a significant deficit in genetic diversity and genetic differentiation among populations. Nonetheless, it is important to note that the markers used in this study were not linked to any traits of interest and this could be the next step in *D. edulis* genomics. To allow visualization of genotypes that harbor important traits such as shorter juvenile phase to first fruit production, less competitive rooting system if intercropped and quality fruits of desired traits, there is need carry out further diversity studies on these genotypes using markers that are linked to specific traits. This will allow scientists to determine if there are multiple or different mechanisms that control these traits and which genotypes have the traits in order to transfer these traits to the best yielding and most popular varieties. Hence, markers linked to these traits will allow the pyramiding of these traits into a select few *D. edulis* genotypes. The planned sequencing of *D. edulis* genome by the African Orphan Crops Consortium (AOCC) will be a major step towards achieving this objective.

Conclusion

This study provided a detailed analysis and quantification of genetic diversity in three *D. edulis* provenances from Cameroon. Moreover, the present study established that the high common level of gene flow obtained could help in maintaining low genetic differentiation and closing genetic relationships of *D. edulis* populations. This can be attributed to inbreeding and the exchange of seeds between farmers either through relatives, and/or markets. Hence, broadening the genetic base of the studied *D. edulis* genotypes may be achieved through introgression

of new alleles. Furthermore, the study also highlighted the need of isolation and characterization of more DNA markers in *D. edulis* and their uses in advanced studies such as gene discovery, trait linked marker assisted selection and gene mapping.

Conflict of Interests

The authors have not declared any conflict of interests.

REFERENCES

- Alemagi D, Kozak RA (2010). Illegal logging in Cameroon: Causes and the path forward. *Forest Policy and Econ.* 12(8):554- 561. <http://www.africad.ubc.ca>
- Ambassa-Kiki R (2000). Caractérisation biophysique succincte des différentes zones agro-écologiques du Cameroun. Institut de Recherches Agronomiques pour le Développement, Yaoundé, P 4.
- Anegbah P, Ukafor V, Usoro C, Tchoundjeu Z, Leakey RRB, Schreckenberg K (2005). Domestication of *Dacryodes edulis*: 1. Phenotypic variation of fruit traits from 100 trees in southeast Nigeria. *New Forests* 29(2):149-160.
- Assogbadjo AE, Glèlè Kakai R, Edon S, Kyndt T, Sinsin B (2011). Provenance variation in fruit characteristics seed germination and seedling growth of *Adansonia digitata* L. in Benin. *New Forests* 41:113-125. <http://leabenin-fsauac.net/publications/>
- Atangana AR (2010). Phenotypic diversity in fruit and seed traits, and neutral genetic diversity in *Allanblackia floribunda*. PhD Thesis. University of Laval. P 102.
- Atangana AR, Beaulieu J, Khasa DP (2010). Wild genetic diversity preservation in a small-sized first generation breeding population of *Allanblackia floribunda* (Clusiaceae). *Tree Genetic and Genome* 6:127-136. <http://www.genomiqueforestiere.chaire.ulaval.ca>
- Benoit L, Born C, Vignes H, Chevallier MH, Todou G, Debain C, Oly HI (2011). Polymorphic microsatellite loci from *Dacryodes edulis* (Burseraceae), a Central African rainforest and fruit-tree species. *Am. J. Bot.* 98(4):74-75. <http://dx.doi.org/10.3732/ajb.1000463>
- Botstein D, White RL, Skolnick M, Davis RW (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Human Genet.* 32:314-331. <http://dx.doi.org/0002-9297/80/3203-0013>
- Cribb J (2010). The Coming Famine: the Global Food Crisis and What We Can Do To Avoid it. University of California Press, Los Angeles, USA, 248 p. ISBN: 9780643100404
- Dawson IK, Lengkeek A, Weber JC (2009). Managing genetic variation in tropical trees: linking knowledge with action in agroforestry ecosystems for improved conservation and enhanced livelihoods. *Biodivers. Conserv.* 18:969-986. <http://dx.doi.org/10.1007/s10531-008-9516-z>
- Degrande A, Franzel S, Yeptiep SY, Asaah E, Tsobeng A, Tchoundjeu Z (2012). Effectiveness of grassroots organizations in the dissemination of agroforestry innovations. *Agroforestry for Biodiversity and Ecosystems Services-Science and Practice.* pp. 141-164.
- Degrande A, Tadjio P, Takoutsing B, Asaah E, Tsobeng, Tchoundjeu Z (2013). Getting trees into farmers' fields: Success of rural nurseries in distributing high quality planting materials in Cameroon. *Small-Scale Forestry* 12(3):403-422. <http://dx.doi.org/10.1007/s11842-012-9220-4>
- El-Mouei R, Choumane W, Dway F (2011). Molecular characterization and genetic diversity in genus *Citrus* from Syria. *Int. J. Agri. Biol.* 13(2):351-356. 10-516/DYX <http://www.fspublishers.org>
- Flynn P (2009). Characterization of rare Irish cattle breeds by comparative molecular studies using nuclear and mitochondrial DNA markers. Doctoral dissertation, University College Dublin.
- Graefe S, Dufour D, van Zonneveld M, Rodriguez F, Gonzalez A (2013). Peach palm (*Bactris gasipaes*) in tropical Latin America: implications for biodiversity conservation, natural resource

- management and human nutrition. *Biodivers. Conserv.* 22:269-300. <http://dx.doi.org/10.1007/s10531-012-0402-3>
- Ganapathy KN, Gomashe SS, Rakshit S, Prabhakar B, Ambekar SS, Ghorade RB, Patil JV (2012). Genetic diversity revealed utility of SSR markers in classifying parental lines and elite genotypes of sorghum (*Sorghum bicolor* L. Moench). *Austr. J. Crop Sc.* 6(11):1486-1495. <http://search.informit.com.au/documentSummary;dn=908197893835275;res=IELHSS>; ISSN: 1835-2693.
- Ghosh RK, Wongkaew A, Sreewongchai T, Nakasathien S, Phumichai C (2014). Assessment of genetic diversity and population structure in Jute (*Corchorus* spp.) Using Simple Sequence Repeat (SSR) and Amplified Fragment Length Polymorphism (AFLP) markers. *Kasetsart J. Nat. Sci.* 48:83-94. http://kasetsartjournal.ku.ac.th/kuj_files/2014/A1404181048051562
- Hodgkin TA, Brown HD, van Hintum ThJL, Morales EAV (1995). Core collections of plant genetic resources. John Wiley & Sons, New York, USA.
- Hemeida AA, Hassan MT, James AS (1998). The identification of some Acacia species by Amplified Fragment Length Polymorphisms (AFLP) molecular markers. In Proceedings of the International Conference of Genetic Engineering and its Applications, Sharm El-Sheikh, South Sinai, Egypt (8–11 April 2004), pp.197-211.
- Kahiu N, Maswili R, Muchira C (2013). Assessing relationships in Kenyan Sorghum landraces by use of Simple Sequence Repeat Molecular Markers. *Int. J. Sci. Res. Knowledge* 1(8):276-287. <http://dx.doi.org/10.12983/ijsrk>
- Khanam S, Sham A, Bennetzen JL, Aly MAM (2012). Analysis of molecular marker-based characterization and genetic variation in date palm (*Phoenix dactylifera* L.). *Austr. J. Crop Science* 6(8):1236-1244.
- Kengue J (2002) The African bush butter tree: *Dacryodes edulis* G. Don. International Centre for Underutilised Crops, Southampton, UK.
- Kiambi D, Newbury HJ, Ford-Llord BV, Dawson I (2005). Contrasting genetic diversity among *Oryza longistaminata* (A. Chev et Roehr) proportions from different geographic origins using AFLP. *Afr. J. Biotechnol.* 4(4):308-317. <http://www.academicjournals.org/AJB>
- Kitavi M N, Muluvi GM, Kiambi DK, Hausman B, Kairichi M, Machuka J (2014). Assessment of the genetic diversity and pattern of relationship of West African sorghum accessions using microsatellite markers. *Afr. J. Biotechnol.* 13(14):1503-1514. <http://dx.doi.org/10.5897/ajb08.1030>
- Kumar P, Gupta VK, Misra AK, Modi DR, Pandey BK (2009). Potential of molecular markers in plant biotechnology. *Plant Omics* 2(4):141. <http://www.pomics.com>
- Leakey RRB (2012). Living with the trees of life: towards the transformation of tropical agriculture. CAB International, Cambridge, USA, 200pp.
- Li CD, Fatokun CA, Ubi B, Singh BB, Scoles G (2001). Determining genetic similarities among cowpea breeding lines and cultivars by microsatellites markers. *Crop Sci.* 41:189-197.
- Liu K, Muse SV (2005). PowerMarker: Integrated analysis environment for genetic marker data. *Bioinformatics* 21:2128-2129. <http://www.ncbi.nlm.nih.gov/pubmed/15705655>
- Lu L, Pomper KW, Lowe JD, Crabtree SB (2011). Genetic variation in pawpaw cultivars using microsatellite analysis. *J. Am. Soc. Hort. Sci.* 136:415-421. <http://www.bioone.org/doi/abs/10.3101/1098-7096-73.2.113>
- Madarbokus S, Ranghoo-sanmukhiya VM (2012). Identification of genetic diversity among papaya varieties in Mauritius using morphological and molecular markers. *Int. J. Life Sci. Biotechnol. Pharma Res.* 1(4):125-163.
- Megevand C, Mosnier A, Hourticq J, Sanders K, Doetin N, Streck C (2013). Deforestation trends in the Congo Basin: Reconciling economic growth and forest protection. World Bank, Washington, DC. P 38.
- Miller AJ, Gross BL (2011). From forest to field: Perennial fruit crop domestication. *Am. J. Bot.* 98:1389-1414. <http://www.landinstitute.org/wp-content/uploads>
- Mondini L, Noorani A, Pagnotta MA (2009). Assessing plant genetic diversity by molecular tools. *Diversity* 1(1):19-35. <http://dx.doi.org/10.3390/d1010019>
- Muchugi A, Lengkeek AG, Kadu CAC, Muluvi GM, Njagi ENM, Dawson KI (2006). Genetic variation in the threatened medicinal tree *Prunus africana* in Cameroon and Kenya: implications for current management and evolutionary history. *South Afr. J. Bot.* 72:498-506. <http://dx.doi.org/10.1016/j.sajb.2006.01.003>
- Onana JM (2008). A synoptic version of *Dacryodes edulis* (Burseraceae) in Africa, with a new species from central Africa. *Kew Bulletin.* 63:385-400.
- Peakall R, Smouse PE (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. *Bioinformatics* 28(19):2537-2539.
- Perrier X, Flori A, Bonnot F (2003). Methods for data analysis. In: Hamon P, Seguin M, Perrier X, Glazmann JC (eds) Genetic diversity of cultivated tropical plants. Science Publishers, Inc and Cirad, Montpellier, pp. 31-63.
- Perrier X, Jacquemoud-Collet JP (2006). DARwin software. Available from <http://www.darwin.cirad.fr/darwin>
- Phelps J, Carrasco LR, Webb EL, Koh LP, Pascual U (2013). Agricultural intensification escalates future conservation costs. *Proceeding of the National Academy of Sciences.*
- Powell W, Morgante M, Andre C, Hanafey M, Vogel MJ, Tingey SV, Rafalski A (1996). The comparison of RFLP, RAPD, AFLP and SSR (microsatellites) markers for germplasm analysis. *Molecular Breeding* 2:225-235.
- Pye-Smith C (2010). A window on a better world. An innovative agroforestry development programme is transforming lives and landscapes in rural Cameroon. ICRAF Trees for Change N°5; Nairobi, World Agroforestry Centre, P 34.
- Rao KN (2004). Plant genetic resources: Advancing conservation and use through biotechnology. *Afr. J. Biotechnol.* 3:136-145. <http://www.academicjournals.org/AJB>
- Russell JR, Kadu CAC, Jamnadass R, Booth A, Cordeiro NJ (2009). Genetic diversity in the African fruit tree *Allanblackia*: implications for conservation and use of a genus newly subject to domestication. *Tree Genetics and Genomics* 5:517-527. <http://dx.doi.org/10.1007/s11295-009-0205-1>
- Saavedra J, Silva TA, Mora, F, Scapim CA (2013). Bayesian analysis of the genetic structure of a Brazilian popcorn germplasm using data from simple sequence repeats (SSR). *Chilean J. Agric. Res.* 73(2):99-107.
- Schreckenber K, Awono A, Degrande A, Mbosso C, Ndoye O, Tchoundjeu Z (2006). Domesticating indigenous fruit trees as a contribution to poverty reduction. *Forests, Trees and Livelihoods* 16:35-51.
- Takoutsing B, Tchoundjeu Z, Degrande A, Asaah E, Gyau A, Nkeumoe F, Tsobeng A (2013). Assessing the quality of seedlings in small-scale nurseries in the highlands of Cameroon: the use of growth characteristics and quality thresholds as indicators. *Small-Scale Forestry* 12(1):1-15. <http://dx.doi.org/10.1007/s11842-013-9241-7>
- Todou G, Benoit L, D'Eeckenbrugge GC, Joly HI, Onana JM, Akoa A (2013). Comparaison des diversités génétiques de *Dacryodes edulis* (G. Don) H.J. Lam. et de *Dacryodes buettneri* (Engel.) H.J. Lam. (Burséracées), deux espèces forestières utiles en Afrique centrale. *Int. J. Biol. Chem. Sci.* 7(3):1243-4254. <http://dx.doi.org/10.4314/ijbcs.v7i3.30>
- Weir BS (1996). Genetic Data Analysis II: Methods for Discrete Population Genetic Data. Sinauer Associates Inc., Sunderland, MA, USA. P 445.
- Wright MH (1978). Evolution and Genetics of population. University of Chicago Press, Chicago.