

# Determination of the Genetic Structure of the Oleaginous *Lagenaria siceraria* of the Nangui Abrogoua University Germplasm Collection

Ahou Anique Gbotto<sup>1\*</sup>, Kouamé Kevin Koffi<sup>1</sup>, Jean-Pierre Baudoin<sup>2</sup>, Irié Arsène Zoro Bi<sup>1</sup>

<sup>1</sup>Laboratoire de Génomique Fonctionnelle et Amélioration Génétique (LaGeFAGe), UFR-SN, Université Nangui Abrogoua, Abidjan, Côte d'Ivoire

<sup>2</sup>Unité de Phytotechnie Tropicale et d'Horticulture, Gembloux Agro Bio Tech, Université de Liège, Gembloux, Belgium

Email: \*aniquegbotto@yahoo.fr

Received 14 September 2015; accepted 14 December 2015; published 17 December 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

Thirty accessions of *Lagenaria siceraria* from the Nangui Abrogoua University germplasm collection were analyzed using three microsatellite markers. The average Polymorphism Information Content (PIC) value was 0.61. The average observed heterozygosity value ( $H_0 = 0.631$ ) compared to the average expected heterozygosity value ( $H_e = 0.645$ ) did not show significant differences in the selected accessions, which suggested random mating in the set of accessions. Within accession inbreeding estimate ( $F_{IS}$ ), was 40% and was not significantly different from zero. The reduction of heterozygotes was likely the result of presence of null alleles. Analysis of Molecular Variance (AMOVA) within and among 30 accessions of *L. siceraria* revealed that 39% of the total variation resides among accessions and 61% within accessions. Accession structuring pattern derived from Bayesian clustering analysis revealed two clusters. Based on the genetic structure of the accessions analyzed, a sampling strategy to collect and conserve genetic resources of *L. siceraria* was suggested.

## Keywords

Genetic Structure, *Lagenaria siceraria*, Null Alleles, Oleaginous, SSR

---

\*Corresponding author.

## 1. Introduction

A germplasm collection is a means of preserving the genetic diversity of a cultivated species before that diversity is lost as a result of implementing high input crop monoculture systems. Such collections serve as a genetic bank from which valuable genes can be selected. An important source of germplasm is the gene pool of landraces that farmers' fields constitute in terms of specific ecological adaptations, usefulness in breeding programs and/or crop improvement. However, to understand the dynamics of diversity in agroecosystems, genetic variability must be investigated [1]. Therefore, for maintenance of the diversity and identification of valuable genes, evaluation of collections is essential. Such studies should not be neglected, particularly for minor or orphan crops such as indigenous edible-seeded cucurbits.

Cucurbits are present in both the New and Old World and are among the most important plant families that supply humans with edible products and useful fibers. Cucurbits are divided into five sub-families: Fevilleae, Melothrieae, Cucurbitaceae, Sicyoideae, and Cyclanthereae.

*Lagenaria siceraria* is a member of the cucurbit family (Cucurbitaceae) which includes several other economically important species such as cucumber and melon that belong to the *Cucumis*, as well as squash and pumpkin that belong to the genus *Cucurbita*, and watermelon that belongs to the genus *Citrullus* [2].

Cultivated *L. siceraria* (Malign) Stanley is commonly known as the white-flowered bottle gourd, but called "Bebu" in Côte d'Ivoire and "Egusi" in Benin, Nigeria, and Ghana. This species is a diploid ( $2n = 22$ ) belonging to the genus *Lagenaria*. Worldwide, *L. siceraria* is grown for its fruit either being harvested mature young and used as a vegetable or harvested mature and used as bottle, utensil, or pipe. In West Africa, oleaginous cucurbits are cultivated for their seeds where they are reported to make an important social and cultural contribution [3]. Another recent utilization of *L. siceraria* is as rootstocks for watermelon against soil-borne diseases and low soil temperature [4] [5]. This plant was one of the first crops to be domesticated. Based on archeological evidence, *L. siceraria* is presumed to have been domesticated in Africa [6] and might have dispersed to the New World by ocean currents or by human migration in pre-historic times [7]. Africa is believed to be the centre of genetic diversity for bottle gourd, although wild progenitors of bottle gourd have not been identified there [8].

The oleaginous and nutritious seeds of *Lagenaria siceraria*, are important in the social and cultural lives of several people [3]. For example dried, slightly toasted and ground seeds of the indigenous *L. siceraria* are used as soup thickener. Achu *et al.* [9] showed that egusi (*L. siceraria*) had a high nutritional value: protein ( $34.19\% \pm 0.85\%$ ); fat ( $50.08\% \pm 1.23\%$ ), provided good quality oil and good groundcover. In addition, commonly found in many traditional cropping systems, the plant is well adapted to extremely divergent agro-ecosystems and various cropping systems characterized by minimal inputs [10] [11]. *L. siceraria* thus represented an excellent plant model for which improved cropping systems implementation could insure the economic prosperity of rural women from tropical Africa. To our knowledge, no detailed study has been devoted to the genetic diversity and reproduction biology. However, investigations reported for others species, suggested that cucurbit family was predominantly outcrossing [12]. Such expectations are based on the fact that indigenous edible-seeded cucurbits are generally monoecious and entomophilous [13]. Neither the occurrence of auto-incompatibility, nor the reproductive mechanisms of this plant have been clearly demonstrated.

The regeneration of the cucurbits in gene banks is mainly done through botanical seeds. Unfortunately, the seeds quickly lose their capacity to germinate and cannot be preserved for more than one year [14]. Thus, regular plant regeneration is required to avoid genetic resource depletion. Also, maintaining the true type plant is a major problem, because it is both hard and time consuming due to the numerous practical precautions required by the mating system. For *Lagenaria siceraria*, the task is also complicated by the creeping behavior of the target species, making appropriate harvesting tedious. Therefore, there is a need to identify a reduced number of *L. siceraria* accessions that can be managed efficiently.

Based on fruits shape, two distinct cultivars have been described [11]. The first one, with round-fruited, is characterized by the presence of a cap on the distal side of seeds. There is no cap on seeds, from elongated fruits of the second cultivar. In spite of the nutritional and agronomic potentials of *L. siceraria*, in depth basic investigations on the crop are scant [15]. For example, to our knowledge, only genetic characterizations of *L. siceraria* accessions from Nangui Abrogoua University collection were done using isozyme [15]. This study using few accessions did not detect the genetic structure of the material analyzed. The weakness of isozyme markers is that each of the proteins that are being scored may not be expressed in the same tissue and at the same time in development. Therefore, several samplings of the genetic population need to be made. To refine these studies and understand the mechanism responsible for genetic variance at both inter- and intra-accession levels, we used

more accessions and microsatellite markers, which are widely employed in the analysis of genetic diversity in *L. siceraria* [16] and have proven to be polymorphic.

The objective of this study was 1) to estimate the amount of genetic diversity within and among watermelon accessions, 2) to determine the genetic structure among accessions and cultivars (or morphotypes), of the oleaginous *L. siceraria*.

## 2. Material and Methods

### 2.1. Plant Material

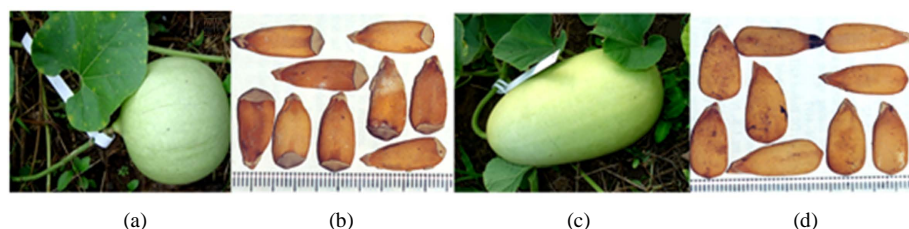
Plant materials were selected from a collection of *Lagenaria siceraria* maintained at Nangui Abrogoua University (Abidjan, Côte d'Ivoire). The seed samples of thirty accessions were collected mainly in five geographical zones (South, East, Northeast, North, and Centre) of Côte d'Ivoire. The selected accessions were representative of two cultivars. The first one, with round fruit is characterized by the presence of a cap on the distal side of seeds (C) and the second cultivar (SC), with elongated fruits and characterized by seeds without a cap (**Figure 1**).

The cultivar "C" contained 22 accessions and the cultivar "SC", 8 accessions, according to seeds availability in each cultivar. A complete description of these accessions is given in **Table 1**. Ten seeds per accession, in total 300 seeds were analyzed.

**Table 1.** Characteristics of 30 *Lagenaria siceraria* accessions used for genetic diversity analysis.

Identification number	Collection zone	Collection site	Cultivar	Fruit shape
Ls005	South	Danguira 2	C	Round fruit
Ls008	South	Danguira 1	C	Round fruit
Ls015	South	GarndAlépé	C	Round fruit
Ls020	East	Assalekro	C	Round fruit
Ls026	East	Ahounan	C	Round fruit
Ls031	East	M'Batto	C	Round fruit
Ls035	East	Akakro	C	Round fruit
Ls041	East	AssièKoumassi	C	Round fruit
Ls047	East	Tagadi	C	Round fruit
Ls053	East	Kamala	C	Round fruit
Ls063	East	Marawi	C	Round fruit
Ls068	East	Flakièdougou	C	Round fruit
Ls078	East	Tefrôh	C	Round fruit
Ls088	East	Kongossoh	C	Round fruit
Ls096	East	KouassiN'dawa	C	Round fruit
Ls104	East	Sianoudi	C	Round fruit
Ls112	East	Nafambeni	C	Round fruit
Ls119	East	Gouméré	C	Round fruit
Ls124	Northeast	Sépidouo	C	Round fruit
Ls130	Northeast	Assoum 1	C	Round fruit
Ls142	Northeast	Djémitédouo	C	Round fruit
Ls147	Northeast	Ondéfidouo	C	Round fruit
Ls152	North	Dikodougou	SC	Elongated fruit
Ls154	North	Tapré	SC	Elongated fruit
Ls160	North	Karakpo	SC	Elongated fruit
Ls165	North	Nerkéné	SC	Elongated fruit
Ls166	North	Kpon	SC	Elongated fruit
Ls179	North	Nogota	SC	Elongated fruit
Ls202	North	Ploro	SC	Elongated fruit
Ls207	Centre	Saria	SC	Elongated fruit

C presence of a cap; LS, *Lagenaria siceraria*; SC without a cap.



**Figure 1.** Seeds from the two cultivar of *Lagenaria siceraria* oleaginous type. Round fruit (a); Seeds with a cap (b); Elongated fruit (c) ; Seeds without a cap (d).

## 2.2. DNA Extraction

The young leaves of each seedling were collected and stored at  $-80^{\circ}\text{C}$  until use. These samples were used for DNA isolation and PCR analysis. DNA isolation was carried out according to procedure described by Levi and Thomas [17] with a few modifications. Fresh leaf (0.1 g) tissue was finely ground in 1.5 ml microtubes (ependorfs) in liquid nitrogen and resuspended in 700  $\mu\text{l}$  CTAB extraction buffer [0.1 M Tris-Base, 1.4 M NaCl, 2.5% cetyltrimethylammonium bromide (CTAB), 20 mM EDTA-dissodium, 0.2% sodium dodecyl sulfate (SDS), 0.5% Sarkosyl, 250 mg polyvinylpyrrolidone (MW 40 (PVP-40) and 250 mg polyvinylpyrrolidone (PVP)]. Each tube was mixed by gentle agitation and then incubated for 30 min at  $65^{\circ}\text{C}$ . The supernatant was taken and 350  $\mu\text{l}$  of isopropanol were added to precipitate the DNA. The DNA pellet was washed in absolute ethanol and dried. Then the pellet was resuspended in TE to a final concentration of ca  $100\text{ ng}\cdot\mu\text{l}^{-1}$  containing  $10\text{ g}\cdot\text{ml}^{-1}$  RNase. The DNA solution was stored at  $-20^{\circ}\text{C}$  until use. DNA concentration was measured by a Nanodrop ND-1000 spectrophotometer (NanoDrop Technologies, Inc).

## 2.3. PCR Conditions

The microsatellite markers had been set up for *Lagenaria siceraria* previously [16]. The PCR reaction condition used was as follows: genomic DNA samples (15 ng) were amplified in a 15  $\mu\text{l}$  reaction volume containing  $1\times$  ThermoPol Reaction buffer (20 mM Tris HCl, 10 mM  $(\text{NH}_4)_2\text{SO}_4$ , 10 mM KCl, 0.1% Triton X-100 pH 8.8 @  $25^{\circ}\text{C}$ ), 0.2 mM each of the four dNTPs, 2 mM  $\text{MgCl}_2$ , 0.5 mM of each forward and reverse primer, and 0.5 U of Taq polymerase (BioLabsInc, NEW ENGLAND). The amplifications were performed in a thermocycler (Biometra) programmed as follows: an initial cycle at  $94^{\circ}\text{C}$  for 3 min, followed by 40 cycles at  $94^{\circ}\text{C}$  for 30 s,  $52^{\circ}\text{C}$  -  $55^{\circ}\text{C}$  for 30 s and  $72^{\circ}\text{C}$  for 1 min. Cycling was followed by a final extension at  $72^{\circ}\text{C}$  for 8 min, and a soak at  $4^{\circ}\text{C}$ .

## 2.4. Electrophoresis

PCR products were separated in denaturing 6% polyacrylamide gels prepared using an acrylamide/bisacrylamide ratio of 19:1, 0.53 TBE (Tris boric acid ethylenediaminetetraacetic acid) buffer, 0.1% ammonium persulfate (APS), and 8.33% tetramethylethylenediamine (TEMED). Polyacrylamide gels were cast in a vertical gel casting plate. Immediately after addition of APS, 70 ml of the gel solution was poured directly into the gel casting plate. The plate with gel solution was then kept at room temperature for approximately 1.5 hours to allow polymerization. The amplified DNAs were mixed with 20  $\mu\text{l}$  of formamide dye (98% formamide, 10 mM EDTA pH 8.0, 1% xylene cyanol and 1% bromophenol blue) before denaturation by heating for 3 min at  $90^{\circ}\text{C}$ . Three microliters of each denaturated DNA mixture were loaded onto a pre-warmed polyacrylamide gel. Electrophoresis was performed at 55 W for 2 hours. The separated DNA bands were revealed using a silver staining method as described by Creste *et al.* [18] which was slightly modified.

## 2.5. Genetic Data Analysis

Fourteen primer pairs used to evaluate 44 entries of Chinese bottle gourd were tested to estimate the genetic diversity among *Lagenaria siceraria* accessions in the present study.

The genetic diversity was evaluated based on genotype and allele frequencies, using the level of polymorphism 0.95 criterion. There should be at least two alleles each with a frequency of at least 0.05. Only one allele has a frequency of 0.95 and the rest of the alleles have less than 0.05. Hence, the locus cannot be considered po-

lymorphic. To evaluate the informativeness of each marker, polymorphic information content (PIC) of an SSR locus was calculated, based on the allele frequencies [19]. The number of alleles per locus, estimates of observed and expected heterozygosity, and Shannon's Information Index were calculated for each population at each locus using GenAlEx v. 6.1 [20]. Comparison between observed and expected heterozygosities were examined according to Mann-Whitney U test using software STATISTICA version 7.1 [21]. The fixation indices were estimated at each polymorphic locus and tested for significant deviation using an exact test performed by the software Genepop [22]. Within each accession, null allele frequencies were estimated using the maximum likelihood estimator based on the EM algorithm of Dempster *et al.* [23] and implemented in Genepop 4.0 [22].

Analysis of molecular variance (AMOVA) was calculated for the sampled accessions to estimate the partitioning of genetic variation at different levels and then to investigate the hierarchical level upon which genetic variation can be attributed. Significance of AMOVA was tested using a nonparametric permutation approach with 999 permutations [24].

A model-based on clustering algorithm in order to search for the most likely number of accessions sampled was used. This algorithm assigns individuals to accessions and also assesses accessions heterogeneity as implemented in the STRUCTURE program [25]. The STRUCTURE analysis was conducted at five replications of  $K$  (assumed number of subpopulations), ranging from 1 to 10, with 100,000 repetitions of Markov Chain Monte Carlo (MCMC) and a burn-in period of 50,000, using the admixture model. Each assessment of  $K$  was repeated five times to check the repeatability of the results.

An UPGMA tree based on the Nei's genetic distances matrix was constructed in PHYLIP package version 3.6 [26]. Cluster analysis was used to describe the relationships among and within different *L. siceraria* accessions. First, 1000 times bootstrapping was performed on SEQBOOT program to generate confidence in the dataset. Then, biased genetic distance from gene frequencies on GENDIST program was computed [27]. The cluster analysis tree was produced with the NEIGHBOUR program which use a matrix of pairwise distances (based on gene frequency genetic distances) between all pairs of accessions and CONSENSUS program.

Confidence in tree topology was assessed by bootstrapping over loci (1000 iterations) and the phylogenetic tree was visualized in TREEVIEW 1.6.6 [28].

### 3. Results

#### 3.1. Estimation of the Informativeness of SSR Markers

Among fourteen primer pairs tested, only three showed polymorphism in this studied. The allelic composition of each marker in each genotype was determined to calculate a PIC value. The average PIC value was 0.61 with a maximum of 0.65 observed with LSR030 and a minimum of 0.55 observed with LSR020 (Table 2).

#### 3.2. Genetic Diversity and Accession-Level Heterozygosity

The population statistics generated by the three microsatellites has been summarized in Table 3. A total of 116 alleles were detected across the three loci. The mean effective number of alleles per locus ( $A$ ), varied respectively from 1 (Ls020) to 5.56 (Ls166) with a mean of 2.88 (Table 3). The average observed heterozygosity ( $H_o$ ) was 0.631, ranging from 0 (Ls020) to 1 (LS207) and the average expected heterozygosity ( $H_e$ ) was 0.645, ranging from 0 (Ls020) to 0.863 (LS166) (Table 3). Mann-Whitney U test indicated that there is no significant difference ( $p > 0.05$ ) between observed heterozygosity and expected heterozygosity. Indeed, of the 89 inbreeding coefficients calculated, only 36 (40.5%) were significantly different from zero ( $p < 0.05$ ). Eighteen inbreeding coefficients showed negative indices. The average inbreeding coefficients ( $F_{IS} = 0.040$ ) was significantly different from zero ( $p < 0.05$ ) for the analyzed accessions. Null alleles frequencies estimates ranged from 0 (Ls005) to 34.15% (Ls147), and were consistent with the  $F_{IS}$  estimates. Overall, the three markers investigated appeared to be affected by at least one null allele. The mean accession diversity using the Shannon Information Index ( $I$ ) was 1.113. Accession Ls166 was the most diverse ( $I = 1.805$ ) and the least diverse accession was Ls020 ( $I = 0$ ).

Analysis of molecular variance (AMOVA) within and among 30 accessions of *Lagenaria siceraria* revealed that 39% of the total variation resides among accessions and 61% within accessions (Table 4). Calculations carried out separately for differentiation among cultivars, exhibited similar trends of AMOVA taking into account no prior grouping of accessions. We found more genetic diversity within cultivars (90%) than among cultivars (10%).

**Table 2.** Simple sequence repeat markers selected, their motif, primer sequence and calculated PIC value.

SSR name	Repeat type	Primer sequence (5'-3')	PIC
LSR011	ATT	TTCGCCTCAGTCCATCTAGTTT ATGTCGTACCTTTTCCCCTTTT	0.55
LSR020	G	AACTGAAACCATTAACGAAGGC AATAAGCAGCAACCATGTCAAC	0.64
LSR030	AT	GGAGACAAAACCAACAACGAA GAAAATGCAGACAAAGAAAGCC	0.65
Mean			0.61

PIC, polymorphism information content.

**Table 3.** Effective number of alleles, observed ( $H_o$ ) and expected ( $H_e$ ) heterozygosities, fixation index ( $F_{IS}$ , following Weir and Cockerham [33], and estimated frequency of null alleles ( $F_{null}$ ) per locus and accession, and Shannon's Information Index of 30 accessions of *Lagenaria siceraria* oleaginous type from SSR markers analysis.

Accession	Locus	A	$H_o$	$H_e$	$F_{IS}^a$	$F_{null}$ [CI] <sup>b</sup>	I
LS005	LSR011	2.062	0.600	0.542	-0.1134	0 [no CI]	0.886
	LSR020	2.381	1.000	0.611	-0.6981*	0 [no CI]	0.943
	LSR030	2.410	0.300	0.616	0.0031*	0.1868 [0.0613 - 0.3129]	0.975
LS008	LSR011	1.681	0.300	0.426	0.2198	0.0980 [no CI]	0.731
	LSR020	3.333	1.000	0.737	0.0334*	0 [no CI]	1.280
	LSR030	2.899	0.600	0.689	0.4423	0.0178 [0 - 0.2062]	1.081
LS015	LSR011	2.899	0.600	0.689	0.1360	0.0433 [0 - 0.2650]	1.081
	LSR020	4.444	1.000	0.816	-0.2414	0 [no CI]	1.622
	LSR030	3.077	0.500	0.711	0.3077	0.1310 [0.0057 - 0.3252]	1.235
LS020	LSR011	1.000	0.000	0.000	-	-	0.000
	LSR020	2.817	1.000	0.679	-0.5158	0 [no CI]	1.165
	LSR030	2.857	0.400	0.684	0.5200*	0.1368 [0 - 0.2823]	1.161
LS026	LSR011	1.220	0.000	0.189	1.0000	0.2037 [no CI]	0.325
	LSR020	2.985	1.000	0.700	-0.4634*	0 [no CI]	1.192
	LSR030	2.985	0.800	0.700	-0.1520*	0 [no CI]	1.096
LS031	LSR011	2.941	0.600	0.695	0.1429*	0.0611 [no CI]	1.089
	LSR020	3.175	0.900	0.721	-0.2656	0 [no CI]	1.224
	LSR030	3.846	1.000	0.779	-0.3043***	0 [no CI]	1.366
LS035	LSR011	2.174	0.400	0.568	0.3077	0.0910 [0 - 0.2699]	0.898
	LSR020	4.444	0.900	0.816	-0.1096*	0 [no CI]	1.544
	LSR030	2.667	0.900	0.658	-0.3966*	0 [no CI]	1.106



## Continued

LS041	LSR011	2.667	1.000	0.658	−0.5652 <sup>*</sup>	0 [no CI]	1.040
	LSR020	3.030	1.000	0.705	−0.4516 <sup>*</sup>	0 [no CI]	1.202
	LSR030	4.348	1.000	0.811	−0.2500	0 [no CI]	1.609
LS047	LSR011	2.632	0.600	0.653	0.0847	0 [no CI]	1.030
	LSR020	3.704	1.000	0.768	−0.3235 <sup>*</sup>	0 [no CI]	1.495
	LSR030	3.846	1.000	0.779	−0.3043 <sup>*</sup>	0 [no CI]	1.441
LS053	LSR011	2.632	0.600	0.653	0.0847	0.0163 [0 - 0.1899]	1.030
	LSR020	5.000	0.900	0.842	−0.0728 <sup>*</sup>	0 [no CI]	1.680
	LSR030	2.985	1.000	0.700	−0.4634 <sup>*</sup>	0 [no CI]	1.192
LS063	LSR011	1.780	0.111	0.464	0.7714 <sup>*</sup>	0.2509 [no CI]	0.778
	LSR020	1.724	0.200	0.442	0.5610	0.1649 [no CI]	0.611
	LSR030	1.835	0.300	0.479	0.3864	0.1110 [0 - 0.2967]	0.647
LS068	LSR011	2.299	0.600	0.595	−0.0093	0 [no CI]	0.927
	LSR020	1.835	0.300	0.479	0.3864	0.1110 [0 - 0.2967]	0.647
	LSR030	1.923	0.400	0.505	0.2174	0.0553 [0 - 0.2623]	0.673
LS078	LSR011	1.969	0.125	0.525	0.7742	0.2471 [no CI]	0.685
	LSR020	1.724	0.200	0.442	0.5610	0.1649 [no CI]	0.611
	LSR030	4.255	0.300	0.805	0.6400 <sup>***</sup>	0.2656 [no CI]	1.515
LS088	LSR011	2.247	0.400	0.584	0.3271	0.1113 [no CI]	0.938
	LSR020	2.381	0.200	0.611	0.6842 <sup>*</sup>	0.2511 [no CI]	0.943
	LSR030	3.571	0.400	0.758	0.4857 <sup>**</sup>	0.2096 [0.0899 - 0.3583]	1.332
LS096	LSR011	2.597	0.400	0.647	0.3950	0.1175 [no CI]	1.010
	LSR020	1.471	0.200	0.337	0.4194	0.1106 [no CI]	0.500
	LSR030	1.905	0.500	0.500	−0.0000	0 [no CI]	0.914
LS104	LSR011	1.835	0.500	0.479	−0.0465	0 [no CI]	0.647
	LSR020	2.219	0.667	0.582	−0.1566	0 [no CI]	0.868
	LSR030	2.941	0.200	0.695	0.7231 <sup>**</sup>	0.2811 [no CI]	1.089
LS112	LSR011	2.532	0.700	0.637	−0.1053	0 [no CI]	0.999
	LSR020	2.198	0.300	0.574	0.4906 <sup>*</sup>	0.1445 [no CI]	0.967
	LSR030	3.704	0.600	0.768	0.2286 <sup>*</sup>	0.0716 [no CI]	1.345
LS119	LSR011	2.632	1.000	0.653	−0.5789 <sup>*</sup>	0 [no CI]	1.030
	LSR020	1.852	0.400	0.484	0.1818	0 [no CI]	0.802
	LSR030	2.410	0.600	0.616	0.0270	0 [no CI]	1.013
LS124	LSR011	3.846	1.000	0.779	−0.3043 <sup>*</sup>	0 [no CI]	1.448
	LSR020	2.174	0.700	0.568	−0.2475	0 [no CI]	0.898
	LSR030	2.778	0.700	0.674	−0.0413	0 [no CI]	1.330

## Continued

LS130	LSR011	1.724	0.400	0.442	0.1000	0 [no CI]	0.895
	LSR020	3.175	1.000	0.721	-0.4173	0 [no CI]	1.258
	LSR030	3.774	0.600	0.774	0.2340*	0.0668 [0 - 0.2766]	1.431
LS142	LSR011	2.597	0.700	0.647	-0.0862	0 [no CI]	1.136
	LSR020	3.175	1.000	0.721	-0.4173*	0 [no CI]	1.331
	LSR030	3.175	0.300	0.721	0.5970*	0.2407 [0.0709 - 0.4014]	1.258
LS147	LSR011	3.176	0.111	0.725	0.8545***	0.3415 [no CI]	1.259
	LSR020	2.817	1.000	0.679	-0.5126	0 [no CI]	1.165
	LSR030	4.082	0.500	0.795	0.3836	0.1596 [0.0319 - 0.3523]	1.583
LS152	LSR011	2.857	0.400	0.684	0.4286*	0.1273 [0 - 0.2814]	1.161
	LSR020	3.077	1.000	0.711	-0.4400	0 [no CI]	1.235
	LSR030	3.175	0.500	0.721	0.3182	0.1414 [0.0291 - 0.3337]	1.376
LS154	LSR011	1.923	0.400	0.505	0.2174	0.0553 [0 - 0.2623]	0.673
	LSR020	3.922	1.000	0.784	-0.2950	0 [no CI]	1.474
	LSR030	3.636	0.300	0.763	0.6197*	0.2496 [no CI]	1.415
LS160	LSR011	2.899	0.500	0.689	0.2857	0.0813 [0 - 0.2494]	1.081
	LSR020	3.175	1.000	0.721	-0.4173	0 [no CI]	1.331
	LSR030	2.899	0.400	0.689	0.4331	0.1810 [0.0325-0.3741]	1.208
LS165	LSR011	3.279	1.000	0.732	-0.3953	0 [no CI]	1.240
	LSR020	4.082	1.000	0.795	-0.2766	0 [no CI]	1.479
	LSR030	2.532	0.600	0.637	0.0609	0.0810 [no CI]	1.340
LS166	LSR011	2.985	1.000	0.700	-0.4634*	0 [no CI]	1.192
	LSR020	5.556	0.900	0.863	-0.0452**	0 [no CI]	1.805
	LSR030	5.000	0.600	0.842	0.2987***	0.0817 [no CI]	1.752
LS179	LSR011	2.941	1.000	0.695	-0.4754	0.1983 [no CI]	1.194
	LSR020	2.985	1.000	0.700	-0.4634	0.3041 [no CI]	1.192
	LSR030	5.000	0.600	0.842	0.2987*	0.1005 [0 - 0.2872]	1.680
LS202	LSR011	2.597	0.700	0.647	-0.0862	0 [no CI]	1.136
	LSR020	4.651	1.000	0.826	-0.2245	0.1017 [no CI]	1.567
	LSR030	1.504	0.400	0.353	-0.1429	0 [no CI]	0.613
LS207	LSR011	2.062	0.600	0.542	-0.1134	0 [no CI]	0.886
	LSR020	3.175	1.000	0.721	-0.4173*	0 [no CI]	1.331
	LSR030	1.504	0.400	0.353	-0.1429	0 [no CI]	0.613
Average		2.876	0.631	0.645	0.040	0.069	1.113
SE		0.098	0.032	0.016	0.035	0.009	0.035

<sup>a</sup>*p*-value of the score test for heterozygote deficiency, with: \*1% *p*-value \5%; \*\*0.1% *p*-value \1%; \*\*\**p*-value \0.1%; <sup>b</sup>95% confidence intervals (CI) for null allele frequencies provided by Genepop. A effective number of alleles per polymorphic locus; *H<sub>o</sub>* observed heterozygosity; *H<sub>e</sub>* expected heterozygosity under Hardy-Weinberg equilibrium; *F<sub>IS</sub>* Fixation index; *F<sub>null</sub>* null alleles frequencies per locus and accession; *I* Shannon's Information Index and *SE* standard error of sample mean.



**Table 4.** Partitioning of genetic variation using AMOVA on SSR data taking into account (a) no prior grouping of accessions; (b) among cultivars.

Level	Source	df	SS	MS	Est. Var.	%D	P value
(a) PA	Among accessions	29	426.007	14.690	1.268	39%	
	Within accessions	270	542.300	2.009	6.243	61%	0.001
	Among cultivars	1	45.140	45.140	0.358	10%	
(b) PCu	Within accessions per cultivar	298	923.167	3.098	3.098	90%	0.001

PA, partitioning all accessions; PCu, partitioning per cultivar; df, degrees of freedom; SS, sum of square; MS, mean square; Est. Var. estimated variance and %D distribution of total variance. The probability was estimated computing 999 permutation.

### 3.3. Cluster and Assignment Analysis

The Bayesian analysis using the software STRUCTURE indicated the presence of two main clusters in the entire set of accessions. The highest value for  $\Delta K$ , the rate of change in the log probability of the data between successive potential numbers of clusters, was obtained for  $K = 2$ . Estimated log probability of the data was higher under  $K = 2$  (−3576.96) than under  $K = 1$  (−4170.1). High proportions of admixed individuals were observed in the region of six accessions with assigned membership different from one to another individual. The results were plotted to evaluate the geographical relationships of the accessions and the cultivars in different genetic clusters (Figure 2). The first cluster was composed of ten accessions. All of them are characterized by the presence of a cap on the distal side of seeds. The second was composed of 20 accessions. Twelve of these accessions are characterized by the presence of a cap on the distal side of seeds, eight showed the absence of a cap on the distal side of seeds. Accessions from southern part of Côte d'Ivoire are exclusively member of the first cluster and accessions from northeastern, northern and center part of Côte d'Ivoire are exclusively member of the second cluster. Only accessions from the eastern part of Côte d'Ivoire are member of the two clusters.

The UPGMA phylogenetic tree based on the Nei's genetic distances matrix from SSR data is shown in Figure 3. The dendrogram consisted of two major clusters. The branch separating these two clusters was well supported (bootstrap = 1000). The first cluster contained a group of twenty-seven accessions (I), the second one (II), consisted of three accessions from eastern part of Côte d'Ivoire (one from Gontougo and the two others from Moronou).

## 4. Discussion

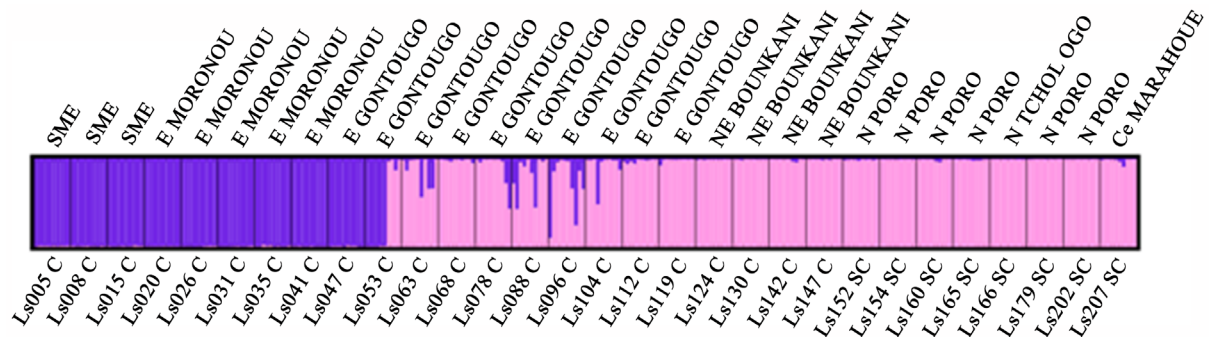
### 4.1. Estimation of the Informativeness of SSR Markers

The PIC value is a measure of the polymorphism level detected by a particular marker and is dependent on the number of alleles detected and their distribution in the population. The number of alleles (1 to 7) and PIC mean value (0.61) observed in this study suggested that the SSR markers selected were highly informative with sufficient discriminatory power.

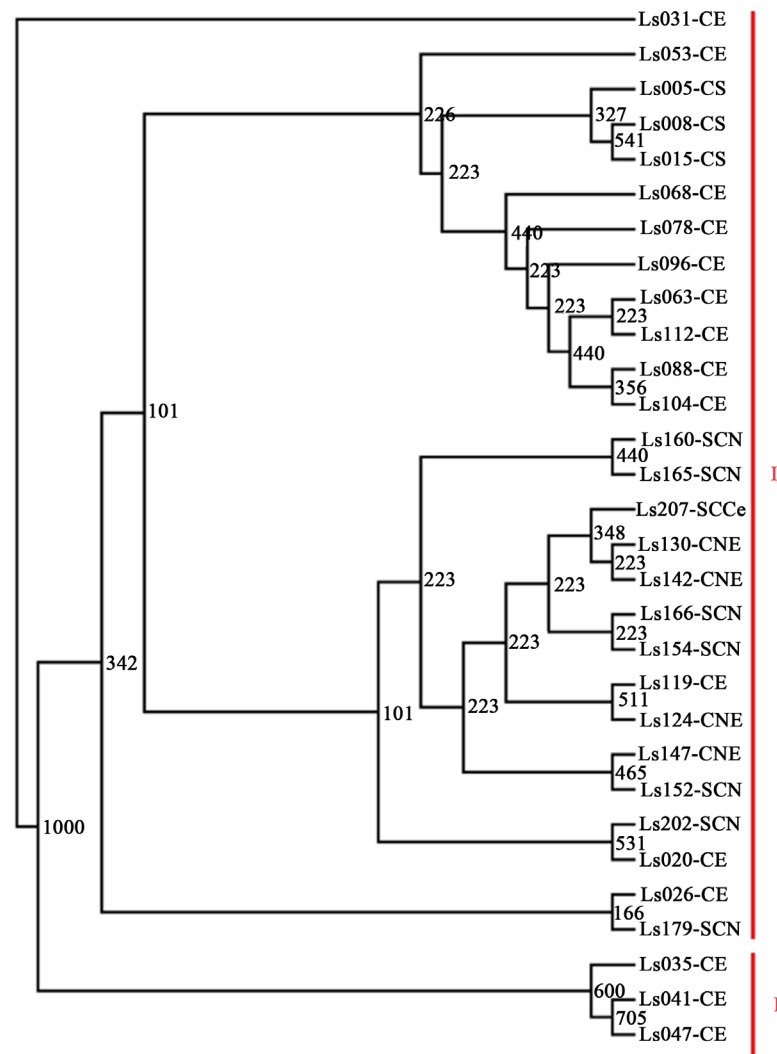
Indeed, based on the PIC values, all the markers were highly informative ( $PIC > 0.50$ ). Such a result indicated the high utility of used set of markers for genetic diversity analysis. A similar PIC value was reported in eight *C. lanatus* (another cucurbit species) accessions collected in Zimbabwe [29] using nine SSR primers. A similar trend was observed in investigations carried out by Ram *et al.* [30] in five germplasm lines of bottle gourd (*Lagenaria siceraria*) using six RAPD primers. These results also indicated profound polymorphism in bottle gourd landraces.

The used microsatellites with wide range of heterozygosity reduced the risk of overestimating genetic variability, which might occur with microsatellites exhibiting only high heterozygosity. Although varying across the loci, the mean values of observed heterozygosity were lower than the expected mean heterozygosity values. However, failure of significant differences between observed and expected heterozygosities according to Mann-Whitney U test ( $p > 0.05$ ) suggested random mating in *Lagenaria siceraria*.

The number of alleles at different marker loci serves as measure of genetic variability having a direct impact on the differentiation of accessions. The allelic variation in this study was lower than those obtained by Gürçan *et al.* [31] in 60 Turkish bottle gourd accessions. This discrepancy could be attributed to the methods used in



**Figure 2.** Genetic structure of microsatellites across 300 individuals from 30 accessions. Bar plot showing clustering of individuals by Structure with  $K = 2$  [25]. Each color represents one accession, each accession is represented by a vertical bar, each individual is represented by a single vertical line broken into  $K$  colored segments, with lengths proportional to each of the  $K$  inferred clusters. E east; C presence of a cap on the distal side of seeds; Ce, center; Ls, *Lagenaria siceraria*; N, North; NE, Northeast, S, south; SC, absence of a cap.



**Figure 3.** Dendrogram UPGMA showing relationships among 30 accessions of *Lagenaria siceraria* oleaginosa type based on three microsatellites loci. CE presence of a cap, Eastern part; CN presence of a cap, Northern part; CNE presence of a cap, Northeastern part; SCCe absence of a cap, Central part ; SC absence of a cap, Northern part.

sample genotyping. Indeed, Gürcan *et al.* [31] used the capillary electrophoresis method for allele determination, whereas we used a direct reading on polyacrylamide gels. The capillary electrophoresis method is known to be the most efficient approach in sample genotyping [32]. The gene diversity indices obtained from the present study were higher than those ( $H_e = 0.073$ ;  $H_o = 0.053$ ) reported for 30 *L. siceraria* accessions from Nangui Abrogoua University germplasm using allozyme markers [15]. This discrepancy could be due to the fact that molecular markers are more polymorphic than isozymes. Significant deviations from Hardy-Weinberg expectations due to heterozygotes deficits were observed in 11 accessions, confirming random mating in the plant material studied. The same results had been reported by Koffi *et al.* [15]. A relative high prevalence of null alleles has been observed in the concerned accessions. Null alleles affect population parameter estimates. The observed heterozygosity would be largely underestimated [33].

We found a significantly low level (39%) of genetic differentiation among accessions and within them the estimated genetic variation was 61% (AMOVA,  $P = 0.001$ ). The results of this study were different from the findings of Minsart *et al.* [34] in *Citrullus lanatus* with, 88% variation among accessions and 12% within them. These results show that the structuring of the genetic diversity could be depend on the mode of sampling. For the present study, we selected ten seeds and a relatively higher number of accessions (30) while in the study conducted by Minsart *et al.* [34], 20 seeds were randomly chosen per accession and only three accessions were selected. Such contrasted sampling schemes should have resulted in inverted trends in accession's genetic structure as revealed by AMOVAs. Accession genetic structure observed in this study was similar to that reported in the previous study [29]. Performing AMOVA within and among seven accessions of watermelons divided into two major groups (cow-melons and sweet watermelons). Authors demonstrated that only 0.8% of the total variation resided between the two groups, 10% between accessions within groups and 89.2% within accessions.

Overall, results from accession genetic structure showed that *L. siceraria* maintained a high level of variability within cultivars in accordance with its mating system, coupled with farmer's seed management approaches. Indeed, at the collecting sites a few seeds are usually saved from the previous season's harvest, or obtained from neighboring farmers or local markets, resulting in the gradual depletion of genetic variability. Similar results have been reported in *Cucumeropsis mannii*, another oleaginous cucurbits cultivated in Côte d'Ivoire [35].

## 4.2. Cluster and Assignment Analysis

Accession structuring pattern derived from Bayesian clustering analysis revealed two clusters.

Data collected show a clustering according to geographical location. This result indicated that besides the forces such as exchange of genetic stock, genetic drift, spontaneous variation, natural and artificial selection, geographical origin also is responsible for genetic diversity.

A dendrogram established based on SSR genotyping of 30 accessions, also detect clustering by geographical location, which is not in agreement with Yetişir *et al.* [36] in which clustering of bottle gourd accessions from Turkey was based around fruit morphology much more than on geographical origin. No significant grouping pattern based on cultivar was observed in group I, on other hand, in group II, all the accessions are characterized by the presence of a cap on the distal side of seeds. However, according to the findings of Koffi *et al.* [15], the UPGMA cluster analysis of morphological differentiation among cultivars of *Lagenaria siceraria* showed that the used two cultivars were well separated. Consistent with the result from Uluturk *et al.* [37], morphological and molecular genetic diversity are distinct factors and must be considered separately in germplasm characterization. This is especially important for crops like cucurbits which have limited molecular genetic diversity.

## 5. Conclusion and Orientations for Future Research

Microsatellite markers have proven to be useful tools in this study in estimating the genetic variation within and among *Lagenaria siceraria* accessions. The relatively high level of the genetic diversity within accessions and cultivars was also in accordance with the mating system of *L. siceraria* and it suggested these accessions can be regarded as potential sources of genetic tank for *in situ* conservation. Furthermore, a clear understanding of the genetic diversity with explicit analyses of genetic structure of *L. siceraria* accessions is important; it can help in understanding the remarkable morphologic diversity existing among edible-seeded *L. siceraria* accessions genetic resources especially in terms of fruit and seed characters. An assessment of genetic diversity based only on morpho-agronomic traits might be biased, because distinct morphotypes can result from only a few mutations while they share a common genetic base. Therefore, molecular markers have the potential to complement al-

ready existing estimations of diversity, and to be used to construct core collections for effective genebank management. The relatively high genetic diversity within accessions supported the sampling scheme proposed.

## Acknowledgements

This research was financially supported by the International Center for Research in Agroforestry (ICRAF), and the Académie de Recherche et d'Enseignement Supérieur (ARES, Brussels, Belgium) and supervised by the Comité Universitaire pour le Développement (CUD, Brussels, Belgium).

## References

- [1] Mujaju, C. and Nybom, H. (2011) Local-Level Assessment of Watermelon Genetic Diversity in a Village in Masvingo Province, Zimbabwe: Structure and Dynamics of Landraces on Farm. *African Journal of Agricultural Research*, **6**, 5822-5834.
- [2] Beevy, S.S. and Kuriachan, P. (1996) Chromosome Numbers of South Indian Cucurbitaceae and a Note on the Cytological Evolution in the Family. *The Journal of Cytology and Genetics*, **31**, 65-71.
- [3] Burkill, H.M. (1985) The Useful Plants of West Tropical Africa. Royal Botanic Gardens, Kew, Richmond, Surrey, UK.
- [4] Lee, J.M. (1994) Cultivation of Grafted Vegetables I. Current Status, Grafting Methods and Benefits. *Hort Science*, **29**, 235-239.
- [5] Yetişir, H. and Sari, N. (2003) Effect of Different Rootstock on Plant Growth, Yield and Quality of Watermelon. *Australian Journal of Experimental Agriculture*, **43**, 1269-1274. <http://dx.doi.org/10.1071/EA02095>
- [6] Whitaker, T.W. (1971) Endemism and Pre-Columbian Migration of Bottle Gourd, *Lagenaria siceraria* (Mol.) Standl. In: Riley, C.L.K.J.C., Pennington, C.W. and Runds, R.L., Eds., *Man across the Sea*, University of Press, Austin, 64-69.
- [7] Morimoto, Y., Maundu, P., Fujimaki, H. and Morishima, H. (2005) Diversity of Landraces of the White-Flowered Gourd (*Lagenaria siceraria*) and Its Wild Relatives in Kenya: Fruit and Seed Morphology. *Genetic Resources and Crop Evolution*, **52**, 737-747. <http://dx.doi.org/10.1007/s10722-004-6119-8>
- [8] Whitaker, T.W. (1971) Endemism and Pre-Columbian Migration of Bottle Gourd, *Lagenaria siceraria* (Mol.) Standl. In: Riley, C.L., Kelley, J.C., Pennington, C.W. and Runds, R.L., Eds., *Man across the Sea*, University of Texas Press, Austin, 78-218.
- [9] Achu, M.B., Fokou, E., Tchiégang, C., Fotso, M. and Tchouanguep, M.F. (2005) Nutritive Value of Some Cucurbitaceae Oilseeds from Different Regions in Cameroon. *African Journal of Biotechnology*, **4**, 1329-1334.
- [10] Achigan-Dako, E.G., Fanou, N., Kouke, A., Avohou, H., Vodouhe, S.R. and Ahanchede, A. (2006) Evaluation agronomique de trois espèces de Egusi (Cucurbitaceae) utilisées dans l'alimentation au Bénin et élaboration d'un modèle de prédiction du rendement. *Biotechnology, Agronomy, Society and Environment*, **10**, 121-129.
- [11] Zoro Bi, I.A., Koffi, K.K., Djè, Y., Malice, M. and Baudoin, J.P. (2006) Indigenous Cucurbits of Côte d'Ivoire: A Review of Their Genetic Resources. *Sciences & Nature*, **3**, 1-9.
- [12] Montes-Hernandez, S. and Eguiarte, L.E. (2002) Genetic Structure and Indirect Estimates of Gene Flow in Three Taxa of *Cucurbita* (Cucurbitaceae) in Western Mexico. *American Journal of Botany*, **89**, 1156-1163. <http://dx.doi.org/10.3732/ajb.89.7.1156>
- [13] Gusmini, G. and Wehner, T.C. (2003) Polygenic Inheritance of Some Vine Traits in Two Segregating Watermelon Families. *Cucurbit Genetics Cooperative Report*, **26**, 32-35.
- [14] Al-Maskri, A.Y., Khan, M.M., Iqbal, M.J. and Abbas, M. (2004) Germinability, Vigour and Electrical Conductivity Changes in Accelerated Aged Watermelon (*Citrullus lanatus* T.) Seeds. *Journal of Food, Agriculture and Environment*, **2**, 100-103.
- [15] Koffi, K.K., Anzara, G.K., Malice, M., Djè, Y., Bertin, P., Baudoin, J.P. and Zoro Bi, I.A. (2009) Morphological and Allozyme Variation in a Collection of *Lagenaria siceraria* (Molina) Standl. from Côte d'Ivoire. *Biotechnology, Agronomy, Society and Environment*, **13**, 257-270.
- [16] Xu, P., Wu, X., Luo, J., Wang, B., Liu, Y., Ehlers, J.D., Wang, S., Lu, Z. and Li, G. (2011) Partial Sequencing of the Bottle Gourd Genome Reveals Markers Useful for Phylogenetic Analysis and Breeding. *BMC Genomics*, **12**, 467. <http://dx.doi.org/10.1186/1471-2164-12-467>
- [17] Blair, M.W., Panaud, O. and Mc Couch, S.R. (1999) Inter-Simple Sequence Repeat (ISSR) Amplification for Analysis of Microsatellite Motif Frequency and Fingerprinting in Rice (*Oryza sativa* L.). *Theoretical and Applied Genetics*, **98**, 780-792. <http://dx.doi.org/10.1007/s001220051135>
- [18] Creste, S., Neto, A. and Figuera, A. (2001) Detection of Single Sequence Repeat Polymorphisms in Denaturing Polyacrylamide Sequencing Gels by Silver Staining. *Plant Molecular Biology Reporter*, **19**, 299-306.

- <http://dx.doi.org/10.1007/BF02772828>
- [19] Guo, X. and Elston, R.C. (1999) Linkage Information Content of Polymorphic Genetic Markers. *Human Heredity*, **49**, 112-118. <http://dx.doi.org/10.1159/000022855>
  - [20] Peakall, R. and Smouse, P.E. (2006) GENALEX 6: Genetic Analysis in Excel. Population Genetic Software for Teaching and Research. *Molecular Ecology Notes*, **6**, 288-295. <http://dx.doi.org/10.1111/j.1471-8286.2005.01155.x>
  - [21] StatSoft (2005) STATISTICA, logiciel d'analyse de données. 7.1. Edition. [www.statsoft.fr](http://www.statsoft.fr)
  - [22] Rousset, F. (2008) Genepop'007: A Complete Re-Implementation of the Genepop Software for Windows and Linux. *Molecular Ecological Resources*, **8**, 103-106. <http://dx.doi.org/10.1111/j.1471-8286.2007.01931.x>
  - [23] Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977) Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **39**, 1-38.
  - [24] Excoffier, L., Smouse, P.E. and Quattro, J.M. (1992) Analysis of Molecular Variance Inferred from Metric Distances among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data. *Genetics*, **131**, 479-491.
  - [25] Pritchard, J.K., Stephens, M. and Donnelly, P. (2000) Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, **155**, 945-959.
  - [26] Felsenstein, J. (1995) Phylogeny Inference Package (PHYLIP). 3.57c Edition, University of Washington, Seattle.
  - [27] Nei, M. (1972) Genetic Distance between Populations. *American Naturalist*, **106**, 283-292. <http://dx.doi.org/10.1086/282771>
  - [28] Page, R.D.M. (1996) Tree View: An Application to Display Phylogenetic Trees on Personal Computers. *Computer Applications in the Biosciences*, **12**, 357-358.
  - [29] Mujaju, C., Sehic, J., Werlemark, G., Garkava-Gustavsson, L., Fatih, M. and Nybom, H. (2010) Genetic Diversity in Watermelon (*Citrullus lanatus*) Landraces from Zimbabwe Revealed by RAPD and SSR Markers. *Hereditas*, **147**, 142-153. <http://dx.doi.org/10.1111/j.1601-5223.2010.02165.x>
  - [30] Ram, H.H., Sharma, K. and Jaiswal, H.R. (2006) Molecular Characterization of Promising Genotypes in Bottle Gourd including a Novel Segmented Leaf Type through RAPD. *Vegetable Science*, **33**, 1-4.
  - [31] Gürçan, K., Say, A., Yetişir, H. and Denli, N. (2015) A Study of Genetic Diversity in Bottle Gourd [*Lagenaria siceraria* (Molina) Standl.] Population, and Implication for the Historical Origins on Bottle Gourds in Turkey. *Genetic Resources and Crop Evolution*, **62**, 321-333. <http://dx.doi.org/10.1007/s10722-015-0224-8>
  - [32] Sverdlow, H. and Gestland, R. (1990) Capillary Gel Electrophoresis for Rapid, High Resolution DNA Sequencing. *Nucleic Acids Research*, **18**, 1415-1419. <http://dx.doi.org/10.1093/nar/18.6.1415>
  - [33] Sefc, K.M., Regner, F., Tureschek, E., Glössi, J. and Steinkellner, H. (1999) Identification of Microsatellite Sequences in *Vitis riparia* and Their Applicability for Genotyping of Different *Vitis* Species. *Genome*, **42**, 367-373. <http://dx.doi.org/10.1139/gen-42-3-367>
  - [34] Minsart, L.-A., Zoro Bi, I.A., Dje, Y., Baudoin, J.-P., Jacquemart, A.-L. and Bertin, P. (2011) Set up of Simple Sequence Repeat Markers and First Investigation of the Genetic Diversity of West-African Watermelon (*Citrullus lanatus* ssp. *vulgaris* Oleaginous Type). *Genetic Resources and Crop Evolution*, **58**, 805-814. <http://dx.doi.org/10.1007/s10722-010-9617-x>
  - [35] Koffi, K.K., Gbotto, A.A., Malice, M., Djè, Y., Bertin, P., Baudoin, J.P. and Zoro Bi, I.A. (2008) Morphological and Allozyme Variation in a Collection of *Cucumeropsis mannii* Naudin (Cucurbitaceae) from Côte d'Ivoire. *Biochemical Systematics and Ecology*, **36**, 777-789. <http://dx.doi.org/10.1016/j.bse.2008.07.008>
  - [36] Yetişir, H., Şakar, M. and Serçe, S. (2008) Collection and Morphological Characterization of *Lagenaria siceraria* Germplasm from the Mediterranean Region of Turkey. *Genetic Resources and Crop Evolution*, **55**, 1257-1266. <http://dx.doi.org/10.1007/s10722-008-9325-y>
  - [37] Uluturk, Z.I., Frary, A. and Doganlar, S. (2011) Determination of Genetic Diversity in Watermelon [*Citrullus lanatus* (Thunb.) Matsum & Nakai] Germplasm. *Australian Journal of Crop Science*, **5**, 1832-1836.