Journal of Agricultural Science; Vol. 7, No. 10; 2015
ISSN 1916-9752 E-ISSN 1916-9760
Published by Canadian Center of Science and Education

Molecular Categorization of Some Water Yam (*Dioscorea alata* L.) Germplasm in Ghana Using Microsatellite (SSR) Markers

Emmanuel Otoo[1], Michael Anokye[2], P. A. Asare[2] & J. P. Tetteh[2]

[1] Crops Reseach Institute, The Council for Scientific and Industrial Research, Kumasi, Ghana

[2] University of Cape Coast, Cape Coast, Ghana

Correspondence: Emmanuel Otoo, Crops Reseach Institute, The Council for Scientific and Industrial Research, P.O. Box 3785, Kumasi, Ghana. E-mail: otooemmanuel@gmail.com

Received: July 29, 2015 Accepted: August 25, 2015 Online Published: September 15, 2015

doi:10.5539/jas.v7n10p225 URL: http://dx.doi.org/10.5539/jas.v7n10p225

Abstract

The yam species, *Dioscorea alata* has an advantage for sustainable cultivation due to its comparatively good agronomic characteristics. Breeders are therefore keen to improve the food quality of the species. Nevertheless, published data on molecular classification and genetic diversity of this crop are scanty. This research therefore investigated genetic variability and relationships among some collected Ghanaian *D. alata* accessions (35) together with 14 introductions from IITA in Nigeria. The true genetic identity and population structure of the accessions were determined using 14 Simple Sequence Repeats primer pairs available for yams. Dimensional scatter diagram of the principal coordinates showed a wide dispersion among the accessions. Cluster analysis using unweighted neighbour-joining method clearly separated the 49 accessions into five main groupings. From the allele frequency analysis, the size of amplified alleles ranged from 100 to 510 bp. The mean polymorphic information content (PIC) values for all markers used was 0.91 and ranged between 0.86 and 0.94 in loci YM13 and YM30 respectively. Gene diversity was high and the average observed heterozygosity was 0.77. Gene diversity was high ranging from 0.87 in YM13 to 0.94 in YM30 with a mean of 0.92. Generally, the allele frequency of all the primers was below 0.95 indicating that they were all polymorphic in character. The findings of this study confirm that SSR molecular markers are able to identify closely related materials within species.

Keywords: *Dioscorea alata*, germplasm, SSRs, Ghana, genetic diversity

1. Introduction

Yams (*Dioscorea* spp.) are the world's fourth most important tuber crop in economic terms (Dansi et al., 2013) after potatoes (*Solanum tuberosum* L.), cassava (*Manihot esculenta* Crantz) and sweet potatoes (*Ipomoea batatas* (L.) Poir.). Although they are cultivated in most tropical countries, West Africa alone produces over 95% of the world's output (FAO, 2011). They serve as the staple carbohydrate source for millions of people (Mignouna, Dansi, & Zok, 2002; Adejumo, Okundare, Afolayan, & Balogun, 2013). The genus Dioscorea comprises over 600 species (Sesay et al., 2013) but only 10 of them are cultivated. These are: *Dioscorea alata* L., *Dioscorea esculenta* L, *Dioscorea batatas* Decne or *D. opposite* Thumb. originating from Asia, *Dioscorea bulbifera* L., *Dioscorea cayenensis-rotundata* complex, *Dioscorea dumetorum* Kenth originating from Africa, *Dioscorea trifida* L. originating from America, *Dioscorea nummularia* Lam. and *Dioscorea pentaphylla* L. originating from both Asia and Oceania (Girma, Korie, Dumet, & Franco, 2012). Of all these species, *D. alata* and the *D. cayenensis – D. rotundata* complex are the most widely cultivated and have real economic significance in Africa (Norman, Tongoona, Danson, & Shanahan, 2012).

The Asiatic *D. alata* Linn, introduced to West Africa some hundred years ago, is fairly widely grown (Mignouna & Dansi, 2003). Very few varieties of the species are used for major food products in West Africa, or further processed. This is as a result of its perceived unimpressive food quality trait such as its less suitability for the preferred cohesive and elastic dough in *fufu* or pounded yam. Several cultivars are also susceptible to pests and diseases and lack the aesthetic values of smooth skin and elegant tuber shape that appeal to consumers in the market (Obidiegwu, Asiedu, Ene-Obong, Muoneke, & Kolesnikova-Allen, 2009). Even though *Dioscorea alata* is also eaten as boiled, it is less preferred to *Dioscorea rotundata* varieties. Breeders are keen to improve the food quality of the species as it has good agronomic flexibility and productive potential (Obiediegwu et al., 2009). The problem facing researchers currently is that, traditional cultivars have not been adequately

characterized. In view of this, breeding and selection of *Dioscorea alata* cultivars with novel or improved characteristics do not exist. This also makes reference to varieties ambiguous, unreliable and impossible to determine the true genetic variation *D. alata* (Otoo, Akromah, & Kolesnikova-Allen, 2009).

Additionally, although guidelines exist for the safe movement of yams, Lebot (2009), there has been little official sharing of germplasm between countries. The informal nature of the yam trade and exchange of planting materials among farmers, have also led to duplication of planting materials whose ethnic or local names have changed in different localities. Thus, the same material may be called differently in another region (Otoo, Akromah, & Kolesnikova-Allen, 2009). In order to exploit *Dioscorea alata* for diverse uses, it has to be well characterized for further investigation so as to influence usage of the crop. The identification of promising *Dioscorea alata* materials is therefore essential in order to improve production and productivity of the crop.

Before the advent of molecular markers, plant breeders relied on phenotypic traits as markers for cultivar identification (Elias, McKey, Panaud, Anstett, & Robert, 2001; Zacarias, Botha, Labuschagne, & Benesi, 2004). These markers are still used in Africa because they are readily available for use; particularly where the capacity to use molecular markers is not yet fully developed (Fregene, 2000). Indeed many released cultivars in Ghana were developed based on morphological descriptors (Asare, Galyuon, Sarfo, & Tetteh, 2011). With the development and application of molecular (DNA) markers, the estimation of plant genetic diversity has become much more simple and reliable. This is because, in contrast to morphological or biochemical marker techniques, DNA-based methods are independent of environmental factors and highly polymorphic for each loci (Karp, Kresovich, Bhat, Ayad, & Hodgkin, 1997). DNA markers, though expensive have proven to be effective tools for distinguishing between closely related genotypes.

A number of DNA marker techniques are available and are important tools for genetic identification in plant breeding and germplasm management (Mba et al., 2001). These DNA markers used in diversity analysis include random amplified polymorphic DNA (RAPDs), restriction fragment length polymorphisms (RFLPs), amplified fragment length polymorphisms (AFLPs) and simple sequence repeats (SSRs). Molecular characterization of agronomically important traits within the *D. alata* germplasm, using different markers will be useful in the molecular breeding programmes. Among the developed molecular markers, SSR markers are being considered as the markers of choice as they are able to detect variation in allele frequency at many unlinked loci (Moyib, Odunloa, & Dixon, 2007). SSR markers are particularly attractive to study because they are abundant in plants; they have high level of polymorphism, and are adaptable to automation (Mba et al., 2001). In Ghana, SSR markers have been used to search for duplicates, core collection and to analyze variation in natural populations of putative progenitors of yam (Otoo, Akromah, & Kolesnikova-Allen, 2009). This study was therefore designed to assess the genetic diversity and relationships among some Ghanaian *Dioscorea alata* germplasm based on SSR markers in order to provide information that could be used to improve upon the current accessions through breeding and selection for desirable characteristics.

2. Materials and Methods

2.1 DNA Extraction and Quantification

Total genomic DNA was extracted from young freshly harvested leaves of 49 *D. alata* yam accessions (Table 1) from the experimental fields of CSIR-CRI, Fumesua, Ghana. DNA extraction was carried out at the Crops Research Institute (CRI) Molecular Biology Laboratory, Fumesua, Ashanti Region, Ghana. DNA was extracted using a modified protocol method (Dellaporta, Woods, & Hicks, 1983).

Table 1. 'Water yam' accessions collected and used for the study

Accessions	**Sources**
AGA 97 023	CRI
TA 97 148	CRI
FA 89 026	CRI
AGA 97 136	CRI
SO 89 149	CRI
TA 97 106	CRI
TA 97 025	CRI
TA 97 116	CRI
FA 89 036	CRI
TA 97 121	CRI
82/318	CRI
82/526	CRI
SO 89 066	CRI
SO 89 100	CRI
SO 89 103	CRI
SO 89 028	CRI
TA 97 113	CRI
AGA 97 115	CRI
TA 97 065	CRI
S P R	CRI
AGA 97 224	CRI
TA 97 130	CRI
SO 89 039	CRI
FA 89 039	CRI
AGA 97 204	CRI
SO 89 120	CRI
GHA 89 107(A)	CRI
TA 97 144	CRI
FA 89 019	CRI
AGA 97 066	CRI
TA 97 131	CRI
TA 97 143	CRI
TDa 00/0003	IITA
TDa 01/00046	IITA
TDa 98/01168	IITA
TDa 00/00045	IITA
TDa 98/01174	IITA
TDa 291	IITA
TDa 98/01176	IITA
TDa 02/00151	IITA
TDa 98/01166	IITA
TDa 00/00046	IITA
TDa 297	IITA
TDa 02/00012	IITA
TDa 01/0004(A)	IITA
TDa 01/0004(B)	IITA
Matches(A)	CRI
Matches(B)	CRI
Matches(C)	CRI

Note. NB: The last five italicized accessions are checks.

The DNA quantity and quality were determined using a spectrophotometer (Biochrom Libra S12) and taking the absorbance reading at 260 nm and 280 nm (A_{260} and A_{280} respectively) levels. The quality of DNA was assessed using the absorbance ratio at 260 to that at 280 wavelengths (A260/A280).

DNA quantity was calculated according to Weising, Nybom, Wolff, and Kahl (2005) as:

$$\text{DNA } (\mu g/\mu l) = A_{260} \times 50 \quad (1)$$

Where, A_{260} is the absorbance at 260 nm.

Thus the concentration of DNA in µg/mL was calculated as:

$$\text{DNA } (\mu g/mL) = [A_{260} \times 50] \times DF \quad (2)$$

Where, DF is the dilution factor.

From the quantities of DNA calculated, the appropriate volumes were pipette into samples tubes and topped up with sterile distilled water (SDW) to make concentration of 10 ng/µl used for polymerase chain reaction amplifications.

Running the DNA samples on 0.8% agarose gels stained with ethidium bromide also assessed the integrity of the DNA. Each well contained a mixture of 2 µl of 2X loading dye and 1 µl of genomic DNA sample. The gels were run with 1X TAE buffer from the cathode to the anode with a constant voltage of 100 V for 45 minutes. They were visualized after electrophoresis with a UV transilluminator (UVP Inc., USA) and photographed with a canon digital (Canon, Power Shoot A4000 IS 16 MP) Camera. High quality DNA samples usually appear as thick bands. For samples with very weak concentration, no further dilution was done.

2.2 Molecular Markers and Polymerase Chain Reactions

Fourteen highly polymorphic SSR DNA markers (Table 2), procured from Industricord (South Africa), which are widely distributed in the yam genome, were used in genotyping the accessions. Amplifications were carried out in Multigene Gradient thermal cycler (Labnet International Inc., California, USA) and GeneAmp PCR System 9700 (Applied Biosystems, USA) of 96-well plates with heated lid to reduce evaporation. The DNA from 49 accessions were fingerprinted using SSR markers in a 10 µl reaction volume of mastermix containing, 1.0 µl of buffer (10X), 0.8 µl of $MgCl_2$ (25 mM), 0.16 µl of dNTPs (Deoxynucleotide Triphosphates) (10 mM), 1 µl of both forward and reverse primer (10 µM), 0.06 µl Taq polymerase, 2 µl of genomic DNA template all together with 3.98µl of nuclease free PCR water.

Reactions were conducted at an initial denaturation step at 95 °C for 2 mins, a touchdown procedure of 94 °C for 30secs denaturation, annealing step of 65 °C for 20 secs reducing at -1 °C per cycle for 10 times (annealing temperature depending on marker), followed by 94 °C for 30secs, 55 °C for 30 secs, 72 °C for 1 min for 35 cycles and a final extension/elongation step at 72 °C for 5 min and then stored (holding) at 4 °C. The amplified products were stored at -20 °C until required to run gels.

Table 2. Set of yam microsatellite markers used in DNA fingerprinting

S/NO	Marker name	Primer sequence (5'-3')a	No. of Bases
1	Da1A01	AACTATAATCGGCCAGAGG (F)	19
		TGTTGGAAGCATAGAGAATT (R)	20
2	Da1C12	GCCTTTGTGCGTATCTGA (F)	18
		AATCGGCTACACTCATCTC (R)	19
3	Da1D08	GATGCTATGAACACAACTAA (F)	20
		ATTTGACAGTGAGAATGGA (R)	19
4	Da1F08	CAGAATGCTTCGTAATCCAAC (F)	21
		AACTATAAGGAATTGGTGCC (R)	20
5	YM13	TTCCCTAATTGTTCCTCTTGTTG (F)	23
		GTCCTCGTTTTCCCTCTGTGT (R)	21
6	YM15	TACGGCCTCACTCCAAACACTA (F)	22
		AAAATGGCCACGTCTAATCCTA (R)	22
7	YM30	GGTCCTCTTCTATCCCAACAA (F)	21
		CACGTATTAACTCCATCTATCCAA (R)	24
8	YM26	ACTCGACAACTCAATGAAACAAAA (F)	24
		CGCTGGGGGTGGCTTAT (R)	17
9	D9	CTCACTAGCGTGGGGAAGAC (F)	20
		TGAAAGAAAGAAGCGCAAGG (R)	20
10	D16	TGGGGTAAGACAAGGGACAC (F)	20
		TGCACCACCAACTGTCTAGC (R)	20
11	D14	TGTCGTTAGGCGTGAAGTTGC (F)	20
		AAATCCAATGCTTCGACCAG (R)	20
12	D17	TCGTGAAAGGATGCAATGAG (F)	20
		TGCGGGTGGTAAAAGAAGAC (R)	20
13	D19	GCCGTAGCACGAACTAGACC (F)	20
		GAATTCACGGGCTTTCAGAC (R)	20
14	D22	CCTTGAGCTGGACCTCATTC (F)	20
		AGCAGCGTGTTTTTCAAGGT (R)	20

Note. NB: F: Forward primer, R: reverse primer.

2.3 PAGE and Silver Nitrate Staining

The PCR products were separated using horizontal polyacrylamide gel electrophoresis (PAGE). The amplified DNA fragments were separated on 6% Polyacrylamide gel at 100v for 30–45 minutes in TBE (Tris-borate-ethylenediaminetetraacetic acid) (1X) using a cell electrophoretic apparatus (MS Major Science, UK) and BIO RAD (Criterion^{TM} Cassettes). 6X DNA loading dye was added to the PCR products for visual tracking of DNA migration during electrophoresis. A 100bp DNA marker (gene rule) was used as a reference to estimate the size of specific DNA bands in the PCR amplified products.

After the PAGE, the DNA fragments were visualized in the gel by silver nitrate staining procedure (Creste, Neto, & Figueira, 2001) with modification and photo documented with a digital camera. The gel was stored in distilled water. PCR products were scored for presence (1) or absence (0) of bands.

2.4 Gel Scoring of DNA Fragments

After the silver staining of PCR products, size matching/calling was done using a reference standard of AXYGEN (Biosciences) 100 bp ladder DNA marker which ranges from 100 bp–3000 bp. The bands on the gel were scored for presence (1) or absence (0) of bands together with their respective sizes. For each marker, alleles for the data set were scored according to size of base pairs of the 100 bp ladder DNA marker. This procedure was conducted for each marker until all alleles were scored with the smallest and largest sized alleles representing the start of the first scoring and the end of the last scoring, respectively.

2.5 Data Analysis

The 14 SSR markers (Table 2) and 49 *D. alata* accessions were subjected to gene diversity and genetic differentiation analysis. To determine the relationship among accessions based on hierarchical cluster analysis, the presence (1) or absence (0) of individual allele was scored for each genotype across all SSR markers used for the study. The data in this form were used to calculate genetic distances between pairs of *D. alata* accessions from the comparisons of the band scores. Pairwise distance matrices were computed using the GenStat Discovery Edition 3 software (Jaccard's similarity coefficient) to generate structure tree.

Power Marker computer programme, version 3.25, Liu and Muse (2005) was used in conducting allelic frequency analysis. To estimate genetic diversity among the accessions, SSR loci were recorded as diploids with single bands taken to indicate the presence of two identical alleles. Genetic diversity was estimated using five statistics averaged over loci; polymorphic information content (PIC); mean number of alleles per locus or allelic richness (A); the average observed heterozygosity (H_o); and the average gene diversity (H_e) was computed, according to Nei (1983). The structure of the genetic diversity within population was further analyzed by factor analysis (PCoA). Analysis was performed using DARwin5 Version: 5.0.158 software (developed by CIRAD) (Perrier & Jacquemoud-Collet, 2006).

3. Results

3.1 Allele Frequency Analysis

Table 3 shows the results of allele numbers detected together with their range of sizes for the 14 SSRs which amplified the microsatellites used in screening the 49 genotypes of *D. alata*. The size of amplified alleles ranged from 100 to 510 bp. SSR Marker YM15, YM30, and D17 recorded the highest number of alleles detected (24 alleles each). This was followed by Da1D08 (23 alleles) and YM26, and D22 (22 alleles each) in that order, the least being YM13 (11 alleles). A total of 273 alleles were amplified with 14 SSR loci analyzed with an average of 19.5 alleles per locus varying from 11 to 24 alleles observed per locus for YM13 and (YM15, YM30, D17 each), respectively (Table 3).

The genetic diversity revealed by expected heterozygosity (H_E) ranged from 0.34 for Da1F08 to 1.00 for Da1D08 with a mean of 0.77 (Table 4). Further, the discriminative power of each SSR primer was assessed by calculating polymorphic information content (PIC) values. Mean PIC values for all markers used were 0.91 and ranged between 0.86 and 0.94 in loci YM13 and YM30, respectively. Markers with more alleles but lower allele frequency had larger PIC as found in YM30 (24 alleles and the highest PIC of 0.94 followed by Da1D08 (23 alleles and the PIC of 0.93 respectively. Again, it can be establish from the results that markers D9, YM15, D17, Da1A01, YM26 and YM13 had the highest frequencies of 0.23, 0.19, 0.18, 0.18, 0.17 and 0.17 respectively, while YM30, Da1D08 and D22 had the lowest frequencies of the predominant allele (0.10, 0.12 and 0.13 respectively). The average frequency of the predominant allele was 0.15 (Table 4).

Gene diversity was high ranging from 0.87 in YM13 to 0.94 in YM30 with a mean of 0.92. Generally, the allele frequency of all the primers was below 0.95 indicating that they were all polymorphic in character (Table 4).

Figure 1 shows the results of SSR markers profile generated for some of the *D. alata* accessions used in the study after silver staining of PCR products. Generally, the pattern of movement of the DNA bands on the gel distinguishes one accession from the other.

Table 3. Range of sizes and allele numbers detected for the 14 SSRs which amplified the microsatellites used in screening *D. alata* accessions

SSR marker name	Allele size range (bp)	Published allele size range (bp)*	Number of alleles detected
Da1A01	190-270	212-225	21
Da1C12	160-215	140-160	16
Da1D08	190-270	223-337	23
Da1F08	180-240	166-179	14
YM13	200-275	175-250	11
YM15	160-248	170-293	24
YM26	100-173	102-174	22
YM30	200-390	Unknown	24
D9	160-320	Unknown	17
D16	210-510	Unknown	19
D14	250-360	Unknown	15
D17	201-310	Unknown	24
D19	135-400	Unknown	21
D22	165-288	Unknown	22

Note. *Publish allele size ranges for the various SSR products were obtained from Otoo, Akromah, and Kolesnikova-Allen (2009).

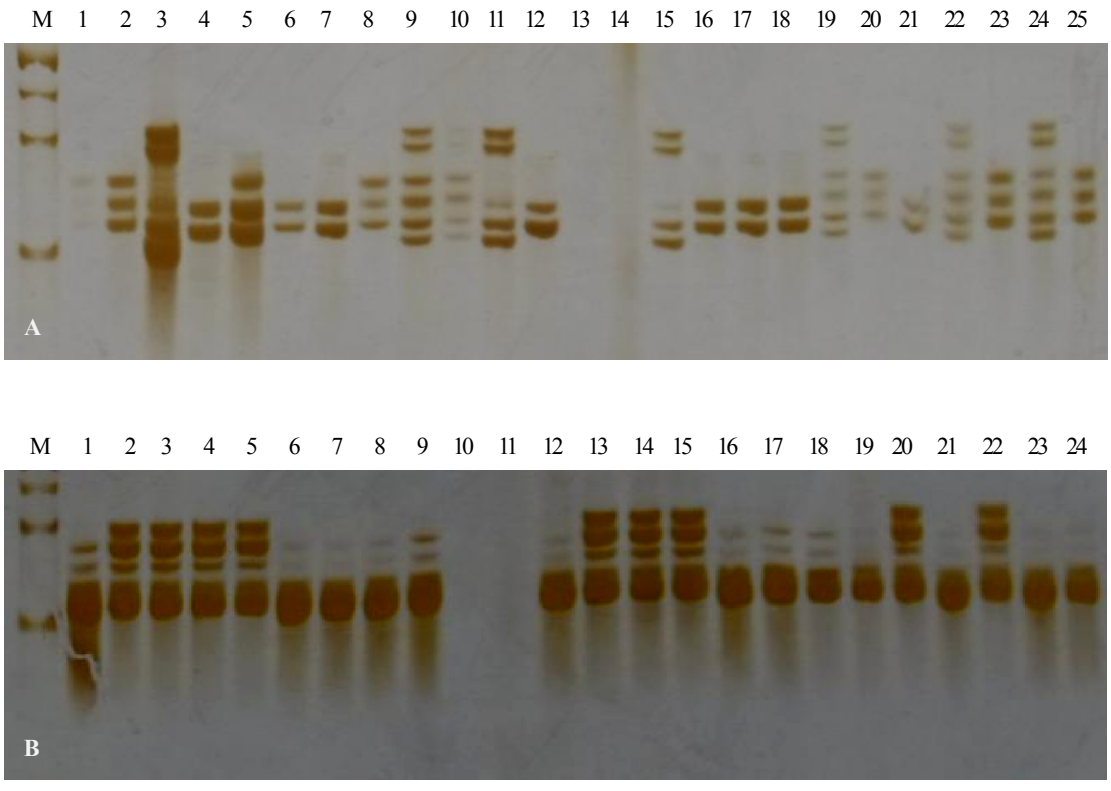

Figure 1. SSR markers profile of *D. alata* accessions generated after silver nitrate staining of PCR products. A: Da1D08 and B: YM13

Table 4. Summary of genetic differentiation parameters generated by 14 SSR primers on 49 genotypes of *D. alata*

Marker	Allele Frequency	Allele No	Gene Diversity	Heterozygosity	PIC
YM 13	0.17	11	0.87	0.94	0.86
YM 26	0.17	22	0.93	0.82	0.92
YM 15	0.19	24	0.92	0.95	0.92
Da1D08	0.12	23	0.94	1.00	0.93
Da1F08	0.14	14	0.91	0.34	0.90
YM30	0.10	24	0.94	0.69	0.94
Da1C12	0.13	16	0.92	0.89	0.91
D22	0.13	22	0.93	0.62	0.93
D17	0.18	24	0.92	0.68	0.91
D9	0.23	17	0.89	0.55	0.88
D14	0.14	15	0.91	0.79	0.91
D19	0.13	21	0.93	0.84	0.92
Da1A01	0.18	21	0.91	0.84	0.90
D16	0.13	19	0.93	0.80	0.93
Mean	**0.15**	**19.50**	**0.92**	**0.77**	**0.91**

3.2 Ordination Analysis

3.2.1 Factorial Analysis

Generally, factorial analysis aim to give an overall representation of diversity and not in the individual effects. On the other hand, tree methods tend to represent individual relations faithfully and may be less accurate for the global structure. Thus, the two ways of viewing data must be considered complementary rather than concurrent (Perrier, Flori, & Bonnot, 2003). Hence in this study, ordination method was combined with clustering procedure to ascertain the diversity in 49 *D. alata* accessions.

Results from principal coordinates analysis (PCoA) of the molecular data showed that the first five coordinates were important (Table 5). PCoA axis 1, 2, 3, 4 and 5 accounted for 27.11% of observed variation. The genetic distances were generated using DARwin5 Version 5.0.158 Software in generating the PCoA plots. For the observed variation recorded for individual axis, axis 1 recorded the highest (8.42%) percentage variation. This was followed by axis 2 (5.84%) and axis 3 (4.83%) in that order (Table 5).

Table 5. Principal coordinates analysis (PCoA) of molecular data showing percentage of variation explained by each axis

Principal coordinates	Individual (%)	Cumulative (%)
Axis 1	8.42	8.42
Axis 2	5.84	14.26
Axis 3	4.83	19.09
Axis 4	4.19	23.28
Axis 5	3.83	27.11

PCoA1 versus PCoA2 of SSR allelic data for *D. alata* accessions used in the study:

The results of factorial analysis of molecular data confirmed generally that, the accessions were grouped into four clusters corresponding to the four quadrants (Figure 2). Quadrant I, II, III, IV had 11, 13, 15 and 10 accessions respectively. Generally, with the exception of quadrants III and IV which is made up of a combination

of collections from CSIR-CRI and IITA in their respective quadrants, majority of the accessions in quadrant I and II are local collections from CSIR-CRI which clustered together. The checks (Matches and TDa 01/0004) were scattered within quadrants I, II and III respectively.

3.2.2 Cluster Analysis

Genetic diversity tree of 49 *D. alata* accessions of molecular data based on Jaccard's index genetic similarity:

Further, tree analysis of the molecular data using GenStat Discovery Edition 3 statistical software and tree construction procedure based on Jaccard's index genetic similarity matrix approach showed large number of intra-specific polymorphisms that facilitated reliable discrimination between accessions (Figure 3). The band scores of microsatellite alleles and calculated genetic distances were used to generate dendrogram showing relationship between the accessions used for the study. From the results, the similarity coefficients ranged from 0.0 to 1.0. Dendrogram for the 49 accessions evaluated with the SSR markers revealed three distinct major clusters (I, II, III) and two minor (IV and V) ones at a similarity coefficient of 0.07 as shown by the results from Figure 3. The sub clusters size varied from 1 to 18 accessions. Clusters IV and V were made up of one (1) and two (2) accessions respectively. The largest group, cluster III, had 18 genotypes comprising accessions from both CSIR-CRI and IITA collections followed by cluster II with 17 accessions, mainly from CSIR-CRI collections, the least being cluster I, consisting of 11 accessions. Cluster I had 5 and 6 of its germplasm from CSIR-CRI and IITA collections respectively.

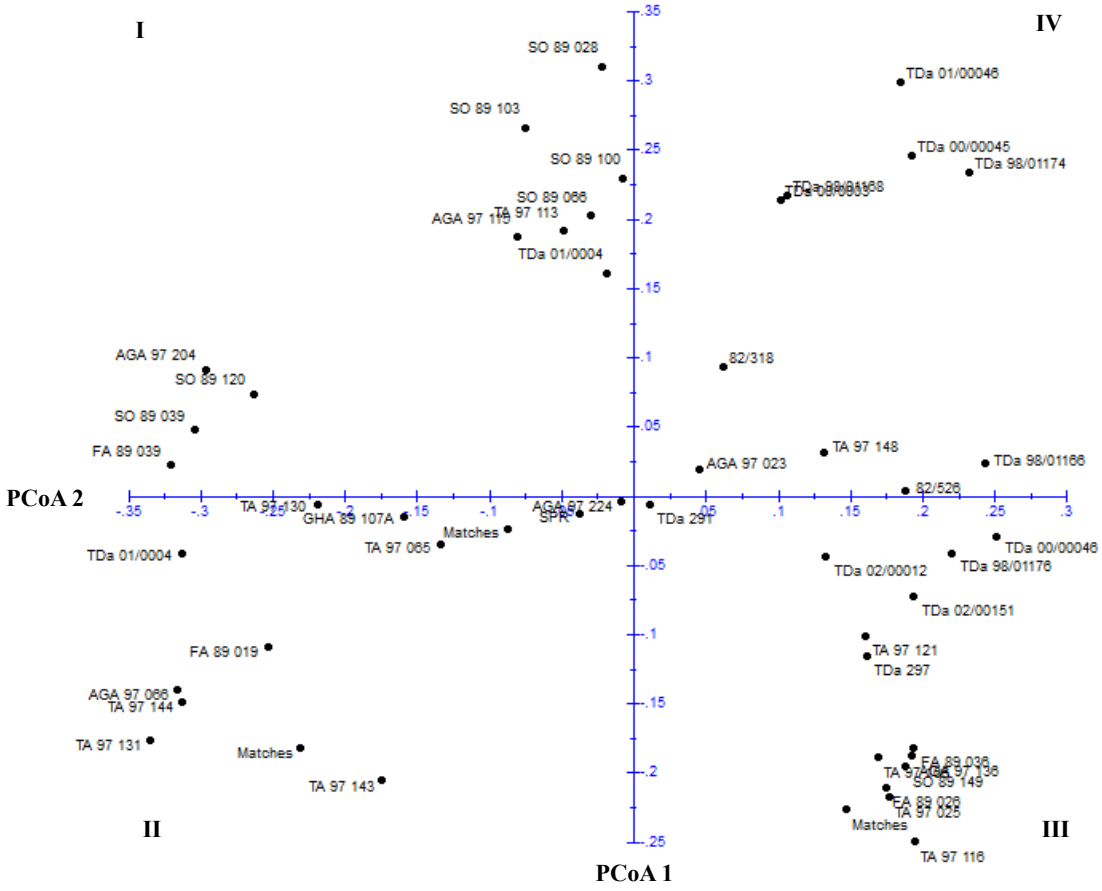

Figure 2. PCoA1 versus PCoA2 of SSR allelic data for *D. alata* accessions used in the study

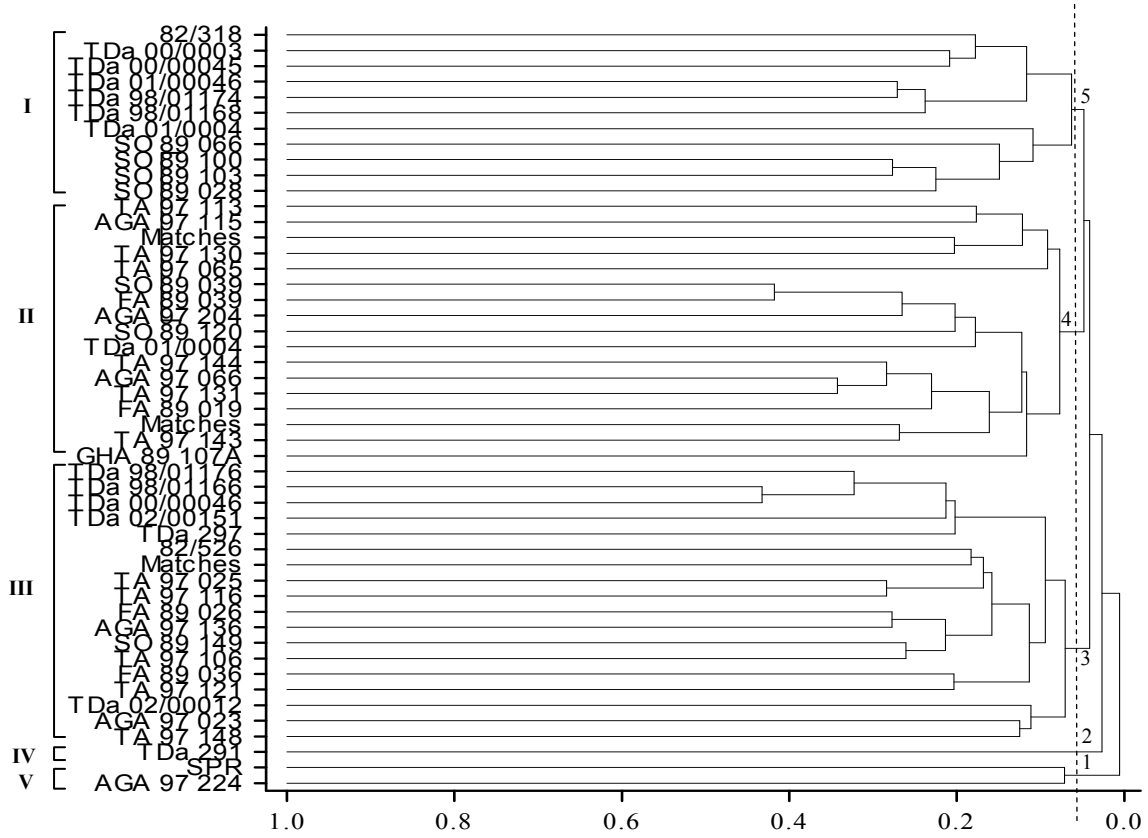

Figure 3. Genetic diversity tree of 49 *D. alata* accessions of molecular data based on Jaccard's index genetic similarity using UPGMA

4. Discussion

4.1 Allele Frequency Analysis

Molecular diversity of the accessions studied was evaluated by using 14 polymorphic SSR primers. In essence, the allele frequency analysis calculates two common measure of variation for each locus namely: expected heterozygosity and polymorphic information content (PIC). The expected heterozygosity measure is helpful in establishing the informativeness of a locus. Loci with expected heterozygosity of 0.5 or less are not very useful for large-scale parentage analysis (Otoo, Akromah, & Kolesnikova-Allen, 2009). The results of observed heterozygosity of the loci was greater than 0.5, in all the markers used for the study (Table 4) except in Da1F08 (0.34) signifying that a good parentage analysis can be obtained from the molecular analysis. This observation also illustrates that, at a single locus, any two alleles, chosen at random from the population are different from each other (IPGRI and Cornell University, 2003). Additionally, the high observed heterozygosity values in this study confirm the heterozygote nature of most of the accessions studied (Obidiegwu et al., 2009). Nevertheless, the average heterozygosity over all loci therefore estimates the extent of genetic variability in the population. The mean heterozygosity value of 0.77 reveals that there was some degree of genetic variation among the population (IPGRI and Cornel University, 2003).

Further, the number of alleles estimated the efficiency of each primer and discriminating power was calculated by the assessment of the polymorphic information content. PIC is regarded as one of the important features of molecular markers and can be used to evaluate the differentiation ability of the markers within the population (Junjian, Colowit, & Mackill, 2002). PIC is a measure of informativeness related to expected heterozygosity and is calculated from allele frequencies (Norman et al., 2012). Such measurement is useful in linkage mapping studies. The results from the allelic frequency analysis generally implied that the loci revealed high polymorphism verified by elevated PIC values (0.86-0.94) (Table 4). Thus, the SSR markers used were efficient in discriminating the species. The amount of PIC is a function of detected alleles and the distribution of their frequency (Moghaddam et al., 2009). Thus, markers with more alleles and low allele frequency had larger PIC as

found in YM30 (24 alleles and the highest PIC of 0.94 followed by Da1D08 (23 alleles and the PIC of 0.93) respectively indicating a better distinction of the accessions (Table 4). These results confirmed the utility of the PIC as a measure of the capacity of a marker to discriminate among closely related individuals as pointed out by Prevost and Wilkinson (1999) and Escandón, Zelener, Dela Torre, and Soto (2007). The average PIC of the SSR markers used for the study was 0.91 and ranged between 0.86 and 0.94 in loci YM13 and YM30, respectively (Table 4) indicating the ability of the utilized markers to differentiate the *D. alata* accessions. PIC values demonstrated that the SSRs used in the study presented, on average, high level of information. The mean PIC value recorded in this study however differs from results obtained from previous studies (Obidiegwu et al., 2009), where a study of 89 accessions of *D. alata* collected from Benin, Congo, Côte d'Ivoire, Equatorial Guinea, Gabon, Ghana, Nigeria, Sierra Leone and Togo with 13 SSR markers gave (a mean PIC value of 0.65, ranging from 0.30 to 0.82) and (Obidiegwu et al., 2009) when assessing *D. cayenensis/D. rotundata* accessions with 15 SSR loci had (PIC = 0.65, on average, varying from 0.37 to 0.80). Otoo, Akromah, and Kolesnikova-Allen (2009) also had different PIC values (PIC = 0.53, on average, varying from 0.00 to 0.89) when they conducted a study on 'Delineation of Pona complex in Ghana' using 13 SSR markers. The mean PIC value observed in this study is higher than what has been reported in previous studies using SSR markers. The result of this study therefore showed that all the primers were highly informative and can be used for genetic diversity studies and the study of phylogenetic relationship.

The markers D9, YM13, Da1A01, and YM15 (Tables 4) had the highest frequencies of 0.23, 0.17, 0.18, and 0.19 respectively, while YM30, Da1D08 and D22 (Tables 4) had the lowest frequencies of the predominant allele (0.10, 0.12 and 0.13 respectively). Low frequency of the predominant allele reveals the suitable allelic distribution among the accessions. SSR markers with the higher number of alleles per locus showed the lowest frequency of the predominant allele, thus, markers with lower frequency of the predominant allele have more differentiation ability than other markers.

Large number of alleles per locus observed (19.5, on average, varying from 11 to 24 alleles) (Table 3) in this study is an indication of considerable allelic variants per locus (genetic diversity present) among the *D. alata* accessions under investigation (Moghaddam et al., 2009).

A gene is said to be polymorphic if the frequency of one of its alleles is less than or equal to 0.95 or 0.99 (IPGRI and Cornell University, 2003). Results observed from allelic frequency analysis proved that all the 14 primers were polymorphic (Table 4). No rare alleles (alleles with allelic frequencies of less than 0.005) were obtained. This was possibly due to the genetic closeness of the genotypes studied.

Gene diversity values of 0.92 on average were also observed. This demonstrates genetic polymorphism in *D. alata* germplasm studied. This high diversity might also be due to the fact that this is a vegetatively propagated crop, which usually maintains high heterozygosity levels (Siqueira et al., 2012). According to Obidiegwu et al. (2009), yams are dioecious plants and spontaneous hybridization must have contributed to the ancestry of some of the accessions, although the selection of somatic mutants might have been the main source of variability used by farmers in their plant improvement practices. This genetic variation offers high potential for genetic improvement because it implies high amount of genetic variance upon which selection could be made for breeding (Obidiegwu et al., 2009).

Principal coordinates analysis (PCoA) of the molecular data showed that the first five coordinates were important (Table 5). PCoA axis 1, 2, 3, 4 and 5 accounted for 27.11% of observed variation. Axis 1 recorded the highest (8.42%) percentage variation. This was followed by axis 2 (5.84%) and axis 3 (4.83%) in that order. However, the plots of PCoA1 versus PCoA2 demonstrated a wide dispersion of accessions along the four quadrants (Figure 2). Quadrant I had 11 accessions which are mainly CSIR-CRI collections with one IITA check (TDa 01/0004). However, two major distinct clusters were observed within that group. One of the cluster is made up of AGA 97 204, SO 89 120, SO 89 039, and FA 89 039 whereas the other cluster is made up of SO 89 028, SO 89 103, SO 89 100, SO 89 066, TA 97 113, AGA 97 115 and TDa 01/0004 (Figure 2). Similar observations were made in quadrant II. Quadrant II had 13 accessions including two checks (Matches) from CSIR-CRI and one IITA check (TDa 01/0004) with TDa 01/0004 as the most distinct member of this group. Accessions AGA 97 224 and TA 97 130 were very close to the horizontal line separating Quadrants I and II. Quadrant III had the most (15) accessions comprising collections from both CSIR-CRI and IITA collections together with one check (Matches) from local collection (Figure 2). TDa 291 was the most distinct accession in this quadrant and was closer to the midpoint of the horizontal line separating quadrant III and IV. The rest of the accessions were grouped in quadrant IV and had the least accessions (10). Here, it can also be observed that, the collections were both from CSIR-CRI and IITA with 82/318 as the most distinct member of this quadrant. The observed grouping in quadrant III and IV where the SSR primer failed to distinguish the accessions from both CSIR-CRI and IITA

collections indicates that these are duplicates in the collection with different entries. The groupings within the accessions (Figure 2) which was observed in almost every quadrant (I, II, III and IV) indicates a possible sub groups among members in the respective quadrants.

Further, it was observed from the PCoA analysis that, the SSR primer was able to differentiate the checks (Matches and TDa 01/0004) indicating the efficiency of the microsatellite markers to separate the accessions. Generally, the high polymorphism revealed by each of the primers taken separately is not surprising since molecular markers can differentiate between closely related individuals.

Further analysis of the molecular data using tree analysis concept, the SSR marker profiles resulted in two main clusters at a similarity level of 0.02 (Figure 3). However, 5 sub clusters were identified from the allelic data from molecular analysis at 0.07 similarity coefficient (Figure 3). The similarity matrix coefficient of the molecular data ranged from 0.0 to 1.00. Again, some of the accessions from CSIR-CRI clustered with IITA collections and vice versa reflecting their genetic similarity. This observation could also imply that, geographical location has not played a major role in the differentiation of the species. The result in this study is in agreement with Obidiegwu et al. (2009), when they evaluated 89 accessions from nine African countries with SSR and found non-distinction between country cultivars of *D. alata.* Lebot, Trilles, Noyer, and Modesto (1998), also examining the genetic relationship among 269 cultivars of *D. alata* from the South Pacific, Asia, Africa and the Caribbean with isozymes, concluded that the most widespread *D. alata* cultivars exhibited a narrow genetic base.

According to these authors, these accessions must have been distributed over great distances as clones during centuries of human migration and it is possible that some of them share common origins. The authors also stated that the majority of accessions within clusters are most likely clones of a common source. This observation agrees with this study, since most of the accessions within clusters (Figure 3) from CSIR-CRI and IITA materials are probably clones originated from a common source. The cluster analysis conducted by Malapa, Arnau, Noyer, and Lebot (2005) using AFLPs revealed the existence of three major groups of genotypes within *D. alata*, each assembling accessions from distant geographical origins and distinct ploidy levels.

The failure of the markers to distinguish between accessions TA 97 113 and AGA 97 115, FA 89 026 and AGA 97 136 as well as SO 89 039 and FA 89 039 collections from CSIR-CRI (Figure 3) signifies that these accessions may be duplicates in the local collection. Detection of duplicates in a collection is critical for effective management of germplasm (Asare et al., 2011).

4.2 Selection of Core Collection

According to Dansi et al. (2000), assessment of genetic diversity within *Dioscorea* spp. will serve as an instrument to identify cultivar misclassification, help to understand the relationships between cultivars, and assist in identifying putative duplicates towards the establishment of an accurate core collection. Based on the molecular data collected, the following core *D. alata* accessions comprising: SO 89 039, SO 89 066, AGA 97 224, FA 89 026, TDa 98/01168, TDa 291, and TDa 98/01166 were selected.

5. Conclusions

1) The 49 *D. alata* accessions used in the study were genetically variable and therefore clustered in groups based on their close relationships or associations.

2) The SSR markers were efficient in discriminating the 49 *D. alata* accessions into 5 clusters.

3) Identical materials may have different names in different collections and areas due to numerous vernacular names hence duplicates were identified and removed.

4) The SSR markers namely; YM13, YM15, YM26, YM30, Da1C12, Da1A01, Da1D08, D9, D14, D16, D17, D19, D22 and Da1F08 facilitated the grouping of the accessions.

5) Core collections comprising SO 89 039, SO 89 066, AGA 97 224, FA 89 026, TDa 98/01168, TDa 291 and TDa 98/01166 were selected to represent the 49 *D. alata* accessions used in this study.

References

Adejumo, B. A., Okundare, R. O., Afolayan, O. I., & Balogun, S. A. (2013). Quality attributes of yam flour (Elubo) as affected by blanching water temperature and soaking time. *International Journal of Engineering Science (IJES), 2*(1), 216-221. http://dx.doi.org/10.1023/b:qual.0000040333.21667.0c

Asare, P. A., Galyuon, I. K. A., Sarfo, J. K., & Tetteh, J. P. (2011). Morphological and molecular based diversity studies of some cassava (*Manihot esculenta* Crantz) germplasm in Ghana. *African Journal of Biotechnology, 10*(63), 13900-13908. http://dx.doi.org/10.5897/AJB11.929

Creste, S., Neto, A. T., & Figueira, A. (2001). Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. *Plant Molecular Biology Reporter, 19*(4), 299-306. http://dx.doi.org/10.1007/bf02772828

Dansi, A., Dantsey-Barry, H., Dossou-Aminon, I., N'Kpenu, E. K., Agré, A. P., Sunu, Y. D., & Vodouhè, R. (2013). Varietal diversity and genetic erosion of cultivated yams (*Dioscorea cayenensis* Poir - *D. rotundata* Lam complex and *D. alata* L.) in Togo. *International Journal of Biodiversity and Conservation, 5*(4), 223-239. http://dx.doi.org/10.5897/IJBC12.131

Dansi, A., Mignouna, H. D., Zoundjihékpon, J., Sangaré, A., Ahoussou, N., & Asiedu, R. (2000). Identification of some Benin republic's Guinea yam (*Dioscorea cayenensis/Dioscorea rotundata* complex) cultivars using random amplified polymorphic DNA. *Genetic Resource and Crop Evolution, 47*, 619-625. http://dx.doi.org/10.1023/A:1026589702426

Dellaporta, S. L., Woods, J., & Hicks, J. B. (1983). A plant DNA mini preparation: Version II. *Plant Molecular Biology Reporter, 1*, 19-21. http://dx.doi.org/10.1007/bf02712670

Elias, M., McKey, D., Panaud, O., Anstett, M. C., & Robert, T. (2001). Traditional management of cassava morphological and genetic diversity by the Makushi Amerindians (Guyana, South America): Perspectives for on farm conservation of crop genetic resources. *Euphytica, 120*, 143-157. http://dx.doi.org/10.1023/a:1017501017031

Escandón, A. S., Zelener, N., De La Torre, M. P., & Soto, S. (2007). Molecular identification of new varieties of Nierembergia linariaefolia (Graham), a native Argentinean ornamental plant. *Journal of Applied Genetics, 48*(2), 115-123. http://dx.doi.org/10.1007/bf03194668

Food and Agriculture Organization (FAO). (2011). *Production Statistics*. Retrieved from http://www.faostat.fao.org/site/342/default.aspx

Fregene, M., Bernal, A., Duque, M., Dixon, A., & Tohme, J. (2000). AFLP analysis of African cassava (*Manihot esculenta* Crantz) germplasm resistant to the cassava mosaic disease (CMD). *Theoretical and Applied Genetics, 100*, 678-685. http://dx.doi.org/10.1007/s001220051339

Girma, G., Korie, S., Dumet, D., & Franco, J. (2012). Improvement of accession distinctiveness as an added value to the global worth of the yam (*Dioscorea* spp) genebank. *International Journal of Conservation Science, 3*(3), 199-206. http://dx.doi.org/10.1111/j.1365-2621.2011.02896.x

IPGRI and Cornell University. (2003). Genetic diversity analysis with molecular marker data. *Learning module* (pp. 2-71).

Junjian, N., Colowit, P. M., & Mackill, D. (2002). Evaluation of genetic diversity in rice subspecies by microsatellite markers. *Crop Science, 42*, 601-607. http://dx.doi.org/10.2135/cropsci2002.0601

Karp, A., Kresovich, S., Bhat, K. V., Ayad, W. G., & Hodgkin, T. (1997). *Molecular tools in plant genetic resources conservation: A guide to the technologies*. IPGRI technical bulletins, Rome, Italy. Retrieved from http://198.93.227.125/publicat/techbull/TB2

Lebot, V. (2009). Tropical root and tuber crops: Cassava, sweet potato, yams and aroids. *Crop production science in horticulture series: 17* (pp. XIX-413). Wallingford: CABI Press. http://dx.doi.org/10.1017/s0014479709007832

Lebot, V., Trilles, B., Noyer, J. L., & Modesto, J. (1998). Genetic relationships between *Dioscorea alata* L. cultivars. *Genetic Resources and Crop Evolution, 45*, 499-509. http://dx.doi.org/10.1017/s0014479709007832

Liu, K., & Muse, S. V. (2005). PowerMaker: Integrated analysis environment for genetic marker data. *Bioinformatics, 21*(9), 2128-2129. http://dx.doi.org/10.1093/bioinformatics/bti282

Malapa, R., Arnau, G., Noyer, J. L., & Lebot, V. (2005). Genetic diversity of the greater yam (*Dioscorea alata*) and relatedness to *D. nummularia* Lam. and *D. transversa* Br. as revealed with ALFP markers. *Genetic Resources and Crop Evolution, 52*, 919-929. http://dx.doi.org/10.1007/s10722-003-6122-5

Mba, R. E. C., Stephenson, P., Edwards, K., Melzer, S., Mkumbira, J., Gullberg, U., & Fregene, M. (2001). Simple Sequence Repeats (SSR) Markers Survey of the Cassava (*Manihot esculenta* Crantz) genome: Towards an SSR-based molecular genetic map of cassava. *Theoretical and Applied Genetics, 102*, 21-31. http://dx.doi.org/10.1007/s001220051614

Mignouna, H. D., & Dansi, A. (2003a). Yam (*Dioscorea* ssp.) domestication by the Nago and Fon ethnic groups

in Benin. *Genetic Resources and Crop Evolution, 50*, 512-528. http://dx.doi.org/10.1023/a:1023990618128

Mignouna, H. D., Dansi, A., & Zok, S. (2002). Morphological and isozymic diversity of the cultivated yams (*Dioscorea cayenensis*/*Dioscorea rotundata* complex) of Cameroon. *Genetic Resource and Crop Evolution, 49*, 21-29. http://dx.doi.org/10.1023/a:1013805813522

Mignouna, H. D., Ellis, N. T. H., Knox, M. R., Asiedu, R., & Quat, N. N. (1998). Analysis of genetic diversity in Guinea yams (*Dioscorea* spp) using AFLP fingerprinting. *Tropical Agriculture (Trinidad), 75*, 2.

Moghaddam, M., Mohammmadi, S. A., Mohebalipour, N., Toorchi, M., Aharizad, S., & Javidfar, F. (2009). Assessment of genetic diversity in rapeseed cultivars as revealed by RAPD and microsatellite markers. *African Journal of Biotechnology, 8*, 3160-3167.

Moyib, O. K., Odunloa, O. A., & Dixon, A. G. O. (2007). SSR markers reveal genetic variation between improved cassava cultivars & landraces within a collection of Nigerian cassava germplasm. *African Journal of Biotechnology, 6*(23), 266-2674.

Nei, M. (1983). Estimation of genetic distance and phylogenetic trees from DNA analysis. *Proceedings of 5th World Congress on Genetics Applied to Livestock Production, 21*, 405-412.

Norman, P. E., Tongoona, P., Danson, J., & Shanahan, P. E. (2012). Molecular characterization of some cultivated yam (*Dioscorea* spp.) genotypes in Sierra Leone using simple sequence repeats. *International Journal of Agronomy and Plant Production, 3*(8), 265-273.

Obidiegwu, J. E., Asiedu, R., Ene-Obong, E. E., Muoneke, C. O., & Kolesnikova-Allen, M. (2009). Genetic characterization of some water yam (*Dioscorea alata* L.) accessions in West Africa with some simple sequence repeats. *Journal of Food, Agriculture and Environment, 7*, 634-638.

Otoo, E., Akromah, R., Kolesnikova-Allen, M., & Asiedu, R. (2009). Delineation of pona complex of yam in Ghana using SSR markers. *International Journal of Genetics and Molecular Biology, 1,* 006-016.

Perrier, X., & Jacquemoud-Collet, J. P. (2006). *DARwin software*. Retrieved from http://darwin.cirad.fr/darwin

Perrier, X., Flori, A., & Bonnot, F. (2003). Data analysis methods. In P. Hamon., M. Seguin., X. Perrier & J. C. Glaszmann (Eds.), *Genetic diversity of cultivated tropical plants* (pp. 43-76). Montpellier: Enfield, Science Publishers. http://dx.doi.org/10.1017/s0014479704252152

Prevost, A., & Wilkinson, M. J. (1999). A new system of comparing PCR primers applied to ISSR fingerprinting of potato cultivars. *Theoretical and Applied Genetics, 98*(1), 107-112. http://dx.doi.org/10.1007/s001220051046

Sesay, L., Norman, P. E., Massaquoi, A., Kobba, F., Allieu, A. P., Gboku, M. L., & Fomba, S. N. (2013). Assessment of farmers' indigenous knowledge and selection criteria of yam in Sierra Leone. *Sky Journal of Agricultural Research, 2*(1), 1-6.

Siqueira, M. V. B. M., Dequigiovanni, G., Corazon-Guivin, M. A., Feltran, J. C., & Veasey, E. A. (2012). DNA fingerprinting of water yam (*Dioscorea alata*) cultivars in Brazil based on microsatellite markers. *Horticultura Brasileira, 30*, 653-659. http://dx.doi.org/10.1590/s0102-05362012000400015

Weising, K., Nybom, H., Wolff, K., & Kahl, G. (2005). *DNA Fingerprinting in plant Principles, Methods, and Applications* (2nd ed.). New York: CRC Press. http://dx.doi.org/10.1093/aob/mcj057

Zacarias, A. M., Botha, A. M., Labuschagne, M. T., & Benesi, I. R. M. (2004). Characterisation and genetic distance analysis of cassava (*Manihot esculenta* Crantz) germplasm from Mozambique using RAPD fingerprinting. *Euphytica, 138*, 49-53. http://dx.doi.org/10.1023/b:euph.0000047057.05642.74