Annals of Applied Biology ISSN 0003-4746

RESEARCH ARTICLE

Genetic differentiation analysis of African cassava (*Manihot esculenta*) landraces and elite germplasm using amplified fragment length polymorphism and simple sequence repeat markers

A.A.J. Raji[1], I. Fawole[2], M. Gedil[1] & A.G.O. Dixon[1]

1 Central Biotechnology Laboratory, International Institute of Tropical Agriculture (IITA), Ibadan, c/o IITA Ltd, Carolyn House, 26 Dingwall Road, Croydon, UK
2 Department of Crop Protection and Environmental Biology, University of Ibadan, Nigeria

Keywords
Elite cultivars; genetic diversity; germplasm; landraces; molecular markers.

Correspondence
A.A.J. Raji, Central Biotechnology Laboratory, International Institute of Tropical Agriculture, Ibadan, c/o IITA Ltd., Carolyn House, 26 Dingwall Road, Croydon CR9 3EE, UK.
Email: a.raji@cgiar.org

Received: 22 August 2008; revised version accepted: 27 January 2009.

doi:10.1111/j.1744-7348.2009.00329.x

Abstract

Molecular-marker-aided evaluation of germplasm plays an important role in defining the genetic diversity of plant genotypes for genetic and population improvement studies. A collection of African cassava landraces and elite cultivars was analysed for genetic diversity using 20 amplified fragment length polymorphic (AFLP) DNA primer combinations and 50 simple sequence repeat (SSR) markers. Within-population diversity estimates obtained with both markers were correlated, showing little variation in their fixation index. The amount of within-population variation was higher for landraces as illustrated by both markers, allowing discrimination among accessions along their geographical origins, with some overlap indicating the pattern of germplasm movement between countries. Elite cultivars were grouped in most cases in agreement with their pedigree and showed a narrow genetic variation. Both SSR and AFLP markers showed some similarity in results for the landraces, although SSR provided better genetic differentiation estimates. Genetic differentiation (*F*st) in the landrace population was 0.746 for SSR and 0.656 for AFLP. The molecular variance among cultivars in both populations accounted for up to 83% of the overall variation, while 17% was found within populations. Gene diversity (H_e) estimated within each population varied with an average value of 0.607 for the landraces and 0.594 for the elite lines. Analyses of SSR data using ordination techniques identified additional cluster groups not detected by AFLP and also captured maximum variation within and between both populations. Our results indicate the importance of SSR and AFLP as efficient markers for the analysis of genetic diversity and population structure in cassava. Genetic differentiation analysis of the evaluated populations provides high prospects for identifying diverse parental combinations for the development of segregating populations for genetic studies and the introgression of desirable genes from diverse sources into the existing genetic base.

Introduction

Molecular markers are informative tools for the measurement of genetic variation and the identification of valuable genetic resources. They provide useful and powerful methods for the effective conservation of biodiversity and can add substantial value to the exploitation of genetic resources in plant breeding systems. The application of molecular marker technologies in cassava has provided a means for the reliable assessment of genetic diversity

 © 2009 The Authors
Journal compilation © 2009 Association of Applied Biologists

(Chavarriaga–Aguirre *et al.*, 1999; Cabral *et al.*, 2000; Raji, 2003; Fregene *et al.*, 2003; Elias *et al.*, 2004; Lokko *et al.*, 2006; Balyejusa Kizito *et al.*, 2007) and mapping of important traits of interest through marker-assisted selection (Akano *et al.*, 2002; Lokko *et al.*, 2005; Okogbenin *et al.*, 2006). The application of molecular markers in genetic diversity studies in combination with pedigree and performance information offers a great potential for enhancing germplasm utilisation and facilitates the selection of parents for developing mapping populations for various traits of interest in cassava breeding. Efforts have been made in the recent past to develop several simple sequence repeat (SSR) markers for the genetic analysis of the cassava genome with at least 136 markers placed on the cassava genetic map to date (Chavarriaga–Aguirre, 2000; Mba *et al.*, 2001; Okogbenin *et al.*, 2006). In addition to the genomic SSR markers, expressed sequence-tag-based microsatellites have also been recently developed (Tangphatsornruang *et al.*, 2008; Raji et al., unpublished data) to further expand the SSR markers available for a more efficient analysis of the cassava genome. Other PCR-based markers, such as amplified fragment length polymorphism (AFLP), have been utilised in various studies in cassava. For example, several reports have described the importance and role of AFLP in the determination of genetic variability in cassava germplasm collections, conservation, varietal identification and the use of genetic resources (Roa *et al.*, 1997; Wong *et al.*, 1999; Elias *et al.*, 2000; Fregene *et al.*, 2000).

Africa is the secondary centre of diversity for cassava. It was introduced in the 16th century and has been grown there ever since, almost solely for human consumption. In recent times, however, cassava has been rapidly transformed into an industrial and cash crop with increasing use in starch, feeds and energy production. Africa is the highest producer of cassava in the world with an annual production of 115 million tonnes in 2005 (FAO, 2006). The International Institute of Tropical Agriculture (IITA), Nigeria, over the years has collected and conserved a number of African cassava landraces with diverse attributes including agronomic and quality traits, resistance to pests and diseases, desirable plant architecture and other adaptation characteristics (Raji *et al.*, 2007). Through population improvement schemes, improved cassava germplasm has been developed at IITA for high yield and other key traits, such as low cyanide content, drought resistance and early maturity. Considering the importance of cassava as a major staple feeding several millions in sub-Saharan Africa, there is a crucial need to fully understand the genetic diversity of this crop and explore its inherent genetic potential for breeding purposes. Molecular marker analysis of genetic diversity in African cassava landraces has been minimally exploited for genetic improvement studies. Molecular markers have unique attributes, advantages and disadvantages specific for each marker system because of their different properties in terms of genome coverage. A number of studies in cassava have used SSR and AFLP markers, but information on the comparative assessment of these markers for population genetic structure in the African cassava population is rather limited. This study aims to address the relative usefulness of AFLP and SSR markers for detecting genetic variation and deducing the genetic structure among collections of cassava landrace germplasm and elite lines. The availability of this information will facilitate the designing of efficient strategies for cassava improvement and future collection expeditions while enhancing the utilisation of broad-based genetic resources.

Materials and methods

Plant material

Materials used in this study were 36 African cassava landrace cultivars (Table 1) selected from a prior classification of 500 landraces based on agro–morphological characterisation and 27 elite lines (Table 2) selected from 400 elite cassava lines maintained at IITA, Ibadan.

DNA isolation

Genomic DNA was isolated from fresh, young apical leaves of cassava using the Dellaporta protocol (Dellaporta *et al.*, 1983) with some modifications as follows: 0.5 g of ground frozen leaf tissue was suspended in DNA extraction buffer containing 5 μL β-mercaptoethanol. The suspension was mixed and incubated at 65°C for 30 min, followed by chloroform–isoamyl alcohol (24:1) extraction and subsequent precipitation with 0.67 volume isopropanol at −20°C. The pellet was formed after centrifugation at 10 000 g. for 10 min and washed with 70% ethanol. The DNA pellet was resuspended in 10 mM Tris–HCl (pH 8.0) and 0.1 mM ethylenediaminetetraacetic acid (EDTA) (TE) buffer and quantified. The quality of the DNA was measured in a 1% (w/v) agarose gel stained in ethidium bromide by comparing the intensity of the band with a known standard DNA fragment. The purity and concentration were determined by measuring the absorbance at 260 nm (A_{260}) and 280 nm (A_{280}) in a spectrophotometer.

Simple sequence repeat marker analysis

A total of 140 SSR markers (Mba *et al.*, 2001) were screened for polymorphism and reproducibility,

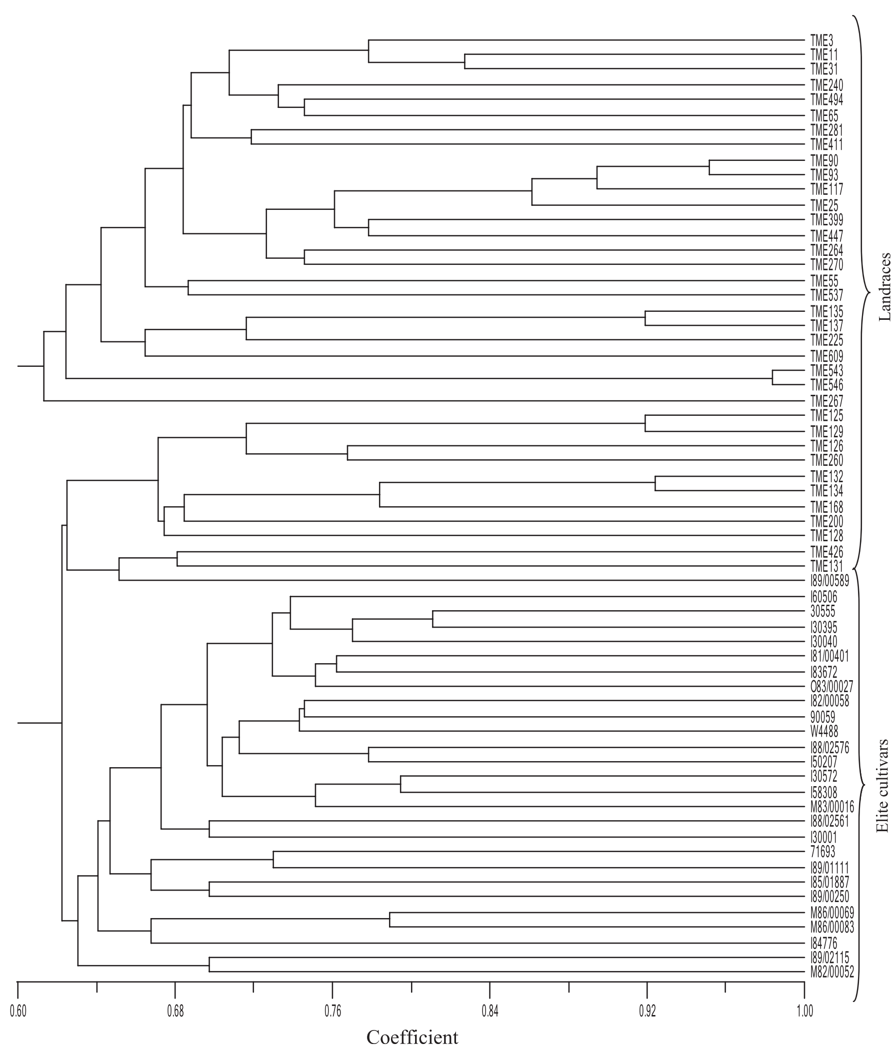

Figure 1A Dendrogram showing variation between African cassava landraces and elite cultivars based on SSR markers using UPGMA.

resulting in the selection of the 50 most informative markers used for genotyping the populations. PCR amplification was carried out in a 25 μL reaction mixture containing 25 ng genomic template DNA, 10× PCR reaction buffer [10 mM Tris–HCl (pH 9.0) at 25°C; 25 mM KCl and 0.1% Triton], 1 unit Taq DNA polymerase (Promega Corp, Madison, WI, USA), 2.5 mM $MgCl_2$, 0.1 mM each of dNTPs (dATP, dGTP, dCTP and dTTP) and 0.2 μM of each primer. Amplification of the reaction was performed in a programmable Peltier Thermal Cycler (PTC-200; MJ Research, Inc., Waltham, MA, USA) using the following temperature cycling parameters: an initial denaturation at 94°C for 4 min, followed by 29 cycles of denaturation at 94°C for 45 s, annealing at 55°C for 1 min and extension at 72°C for 1 min, and then a final extension at 72°C for 7 min after the cycles. Formamide dye (98% formamide, 10 mM EDTA and 0.25% each of bromophenol blue and xylene cyanol FF) was added to products from PCR amplification, denatured at 94°C for 4 min and separated on 6% (w/v) denaturing polyacrylamide gels at 60 W for about 2 h. After electrophoresis, the gel was stained using the silver sequence DNA staining kit (Promega Corp) following the instructions provided.

Amplified fragment length polymorphism marker analysis

Amplified fragment length polymorphism reaction was performed as described by Vos *et al.* (1995) with some modifications. Approximately 150–200 ng of DNA was used for the analysis of each sample. Digestion of genomic DNA was conducted with 5 units each of *Eco*RI

Table 1 African cassava landraces, country of origin and cluster assignment based on SSR marker fingerprinting

Cultivar	Country	Cluster[a]
TME 3	Nigeria	1
TME 11	Nigeria	1
TME 31	Nigeria	1
TME 65	Nigeria	1
TME 494	Nigeria	1
TME 281	Bénin	1
TME 240	Togo	1
TME 411	Togo	1
TME 90	Nigeria	2
TME 93	Nigeria	2
TME 117	Nigeria	2
TME 25	Nigeria	2
TME 399	Togo	2
TME 447	Bénin	2
TME 264	Ghana	2
TME 270	Ghana	2
TME 55	Nigeria	3
TME 537	Malawi	3
TME 135	Cape Verde	4
TME 137	Cape Verde	4
TME 225	Bénin	4
TME 609	Niger Republic	5
TME 543	Uganda	6
TME 546	Uganda	6
TME 267	Ghana	7
TME 125	Cape Verde	8
TME 129	Cape Verde	8
TME 126	Cape Verde	8
TME 260	Ghana	8
TME 132	Cape Verde	9
TME 134	Cape Verde	9
TME 168	Cameroon	9
TME 200	Bénin	9
TME 128	Cape Verde	9
TME 426	Togo	9
TME 131	Cape Verde	9

SSR, simple sequence repeat.
[a]Cluster analysis grouping resulting from SSR marker analysis.

(recognition sequence 5′ GAATTC 3′) and *Mse*I (5′ TTAA 3′). The mixture was incubated at 37°C for 3 h. The reaction buffer contained 5× restriction ligation buffer, 50 mM Tris–HCl (pH 7.5), 50 mM MgAc, 250 mM KAc, 25 mM dithiothreitol (DTT) and 250 ng μL^{-1} bovine serum albumin (BSA). To the double-digested DNA sample, a total volume of 10 μL of ligation reaction mixture was added including 5 pmoles of *Eco*RI adapter and 50 pmoles of *Mse*I adaptor, 10 mM ATP, 1 unit of T4-DNA ligase, with the same reaction buffer described above. The ligation reaction mixture was incubated overnight at 20°C. Digested and adapter-ligated DNA samples were diluted 20-fold with TE buffer. Preamplification of adapter-ligated template DNA was performed in a 20 μL reaction containing 5 μL of diluted template mixture, 75 ng of both *Mse*I +1 primer and *Eco*RI +1 primer, 200 μM dNTPs, 10 mM Tris–HCl (pH 8.3), 1.5 mM $MgCl_2$, 50 mM KCl and 1 unit Taq polymerase. The PCR was performed on a PTC-200 thermal cycler at initial denaturation temperature of 94°C for 2 min, followed by 30 cycles of 94°C for 30 s, 60°C for 30 s, 72°C for 1 min, followed by extension at 72°C for 5 min. The preamplified product was diluted 20-fold with TE buffer and used as a template for selective amplification to generate AFLPs. Selective amplification was performed in a final volume of 20 μL containing 2 μL of diluted preamplification products and 30 ng of each selective *Mse*I +3 and *Eco*RI +3 primers, 200 μM dNTPs, 10 mM Tris–HCl (pH 8.3), 1.5 mM $MgCl_2$, 50 mM KCl and 1 unit Taq polymerase using the cycling conditions described by Vos *et al.* (1995). Formamide dye was added to products from selective amplification and separated on 6% denaturing polyacrylamide gels, as previously described.

Data analyses

Allele frequency and polymorphism were evaluated in each SSR locus, and AFLP fragments were recorded in a binary matrix format as presence and absence of fragments. Genetic distance was determined according to Nei's (1972) estimates using the NTSYS-pc software (Rohlf, 1990). Dendrograms were constructed using the unweighted pair group method with arithmetic mean (UPGMA) (Sneath & Sokal, 1973) and the TREE subprogram of the NTSYS-pc software. The Popgene software (Yeh *et al.*, 1997) was used to measure the extent of between and within-population genetic differentiation of the cassava samples using the Wright's F statistic estimates (Wright, 1951) for all SSR loci (Weir & Cockerham, 1984). Furthermore, principal component analysis (PCA), using the SAS software (SAS, 2004) and the multidimensional scaling (MDS) subprogram of the NTSYS-pc software were used to obtain a graphical representation of the genetic structure of the populations (Schiffman *et al.*, 1981).

Using the codominant data analysis option of the GenAlEx software (version 6.031) for the SSR data, the following were computed: number of alleles (N_a), effective number of alleles (N_e) (Kimura & Crow, 1964), expected (H_e) and observed heterozygosity (H_o); fixation index (measured as $1 - H_o/H_e$) to test for departure from Hardy–Weinberg equilibrium values (Wright, 1978) and Shannon's information index (I) for the measurement of allelic and genetic diversity for each locus (Peakall & Smouse, 2006).

Analysis of molecular variance using the methods of Excoffier *et al.* (1992) was performed to partition the

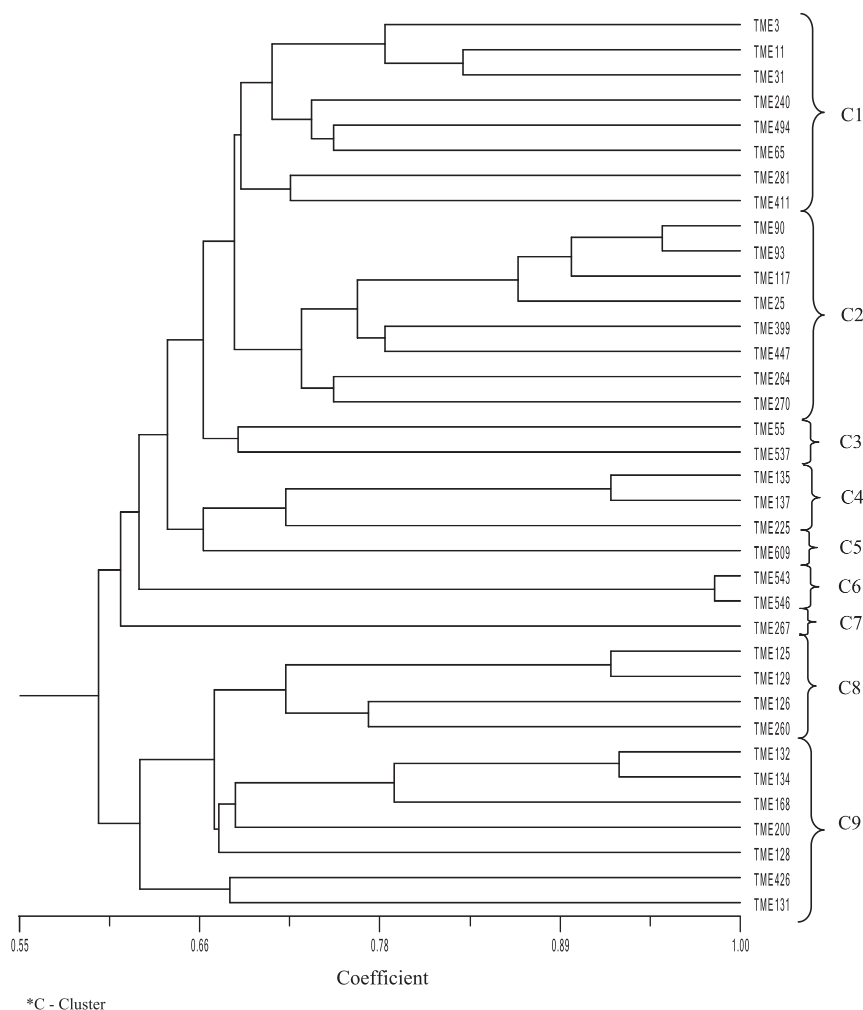

Figure 1B Dendrogram showing within-population variation between African cassava landraces based on SSR markers using UPGMA.

population level genetic differentiation variance observed among the landraces and elite cassava cultivars. Finally, the Mantel test (Mantel, 1967) using the NTSYS-pc software was conducted with 1000 permutations for the hypothesis testing of correlations between the distance matrices resulting from AFLP and SSR markers for both populations.

Results

Genetic differentiation based on amplified fragment length polymorphism and simple sequence repeat allele diversity

Based on the identification matrix generated from 50 SSR primer pairs, the elite cultivars were differentiated into seven groups and landraces into nine groups (Fig. 1A and Fig. 1B). The number of alleles produced by the SSR markers ranged from one to six per locus with a total of 185 polymorphic, reproducible and clear fragments. Twenty AFLP primer combinations produced 39–94 fragments for each primer combination, of which 5–21 polymorphic fragments were scored. The dendrogram showing the clustering pattern of elite and landrace populations based on AFLP markers is shown in Fig. 2. Sample profiles of PCR-amplified products obtained from both AFLP and SSR marker analysis are shown (Figs 3 and 4). The genetic similarity coefficient of landraces resulting from SSR markers ranged from 0.52 to 0.97 and from 0.51 to 0.81 for elite cultivars (data not shown). The AFLP coefficients varied from 0.46 to 0.91 for landraces and from 0.45 to 0.73 for elite cultivars (data not

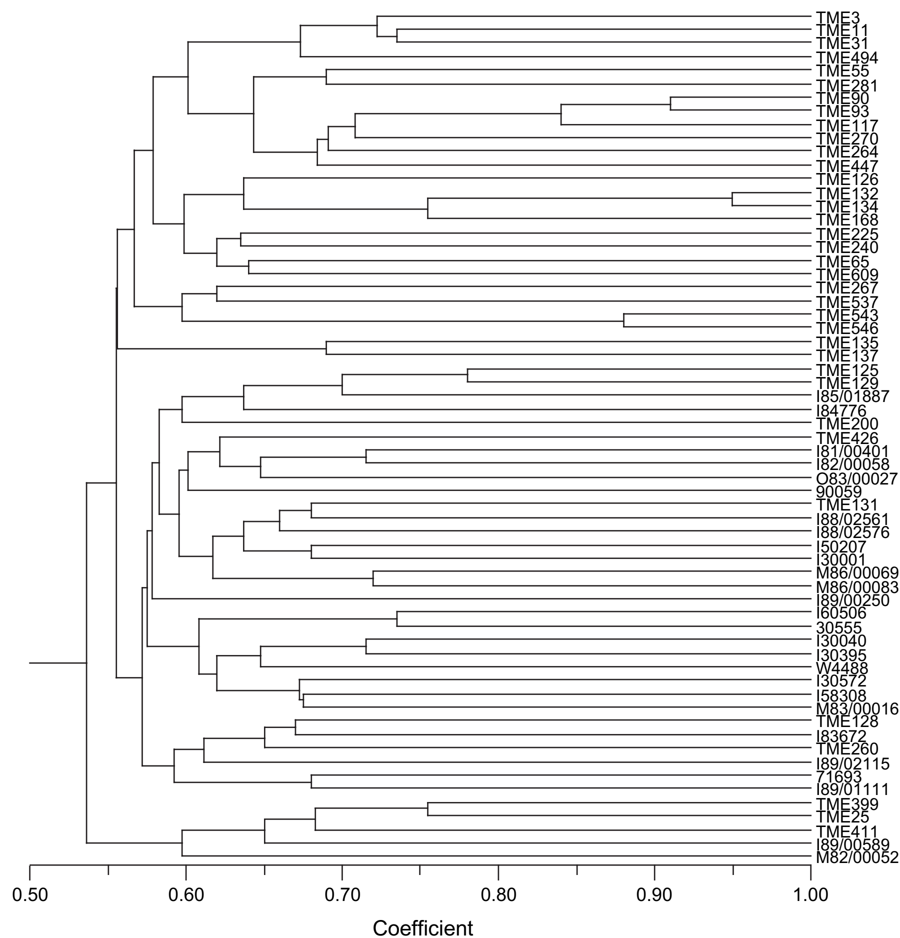

Figure 2 Dendrogram showing variation between African cassava landraces and elite cultivars based on AFLP markers using UPGMA.

shown). In addition, comparison between the genetic similarity matrices of both markers, using the matrix comparison test for a significant relationship by random permutation, showed a significant correlation between markers for both the landraces ($r = 0.58$, $P \leq 0.001$) and the elite cultivars ($r = 0.40$, $P \leq 0.001$).

Population genetics parameters, namely, the observed number of alleles, number of effective alleles, Shannon's information index, observed heterozygosity, expected heterozygosity and fixation index, are shown in Table 3 for both populations. The observed heterozygosity averaged across both the elite and the landrace populations ranged from 0.174 to 1.000 with a mean value of 0.730, while the expected heterozygosity varied from 0.298 to 0.762 with an average value of 0.630 (Table 3). The fixation index for each SSR locus approximated across the elite and landrace populations is presented (Table 3). The mean fixation index value across all loci was −0.150 (Table 3). Genetic differentiation estimates between populations for all loci are also presented (Table 4). Negative fixation index values were estimated for some loci, and positive values were observed in others, indicating some deviations from Hardy–Weinberg equilibrium (Table 3). Negative fixation index values are probably indicative of an excess of heterozygotes, while considerably high positive values could be a result of inbreeding or null alleles (Hartl & Clark, 1997).

Genetic relatedness

Cluster analysis provided an adequate grouping of the population based on Nei's (1972) estimates of genetic coefficients, and PCA displayed a graphical overview of the genetic diversity distribution within and between landraces and elite cultivars. As expected, most of the elite cultivars were grouped in accordance to their pedigree (Fig. 1, Table 2). The landraces were grouped in agreement with their geographical origin, with a few overlaps

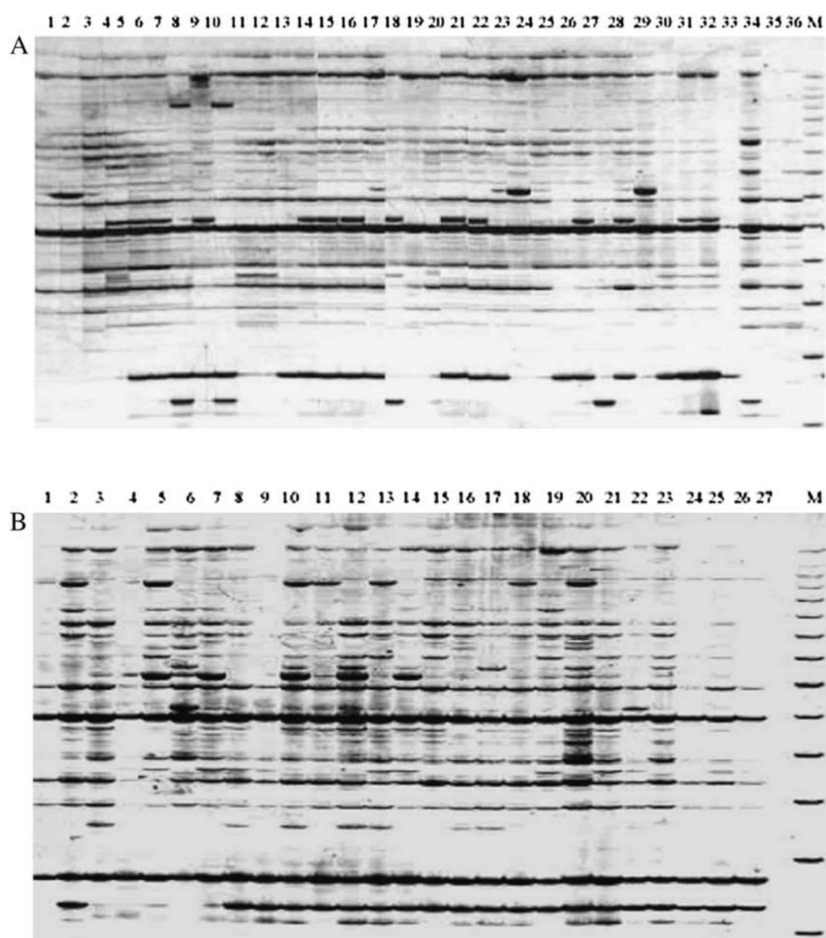

Figure 3 Amplified fragment length polymorphic profiles of (A) 36 cassava landraces and (B) 27 elite cassava cultivars obtained from primer pair E-AAC/M-CTC.

between the cultivars according to proximity between countries (Fig. 2, Table 1). Clustering patterns obtained from both AFLP and SSR showed an acceptable level of consistency when compared with similar studies where both marker systems have been used (Basu *et al.*, 2004; Prakash *et al.*, 2005; Geleta *et al.*, 2006). Apart from the main groups identified within the population, subgroups were also evident, as shown on the dendrograms, particularly in the landrace population (Fig. 1B). The elite cultivars, on the other hand, did not show as much within-population differentiation as the landraces. Analysis of molecular variance showed that the overall variance was partitioned into 83% within and 17% between elite and landrace populations (Table 5).

The SSR markers appeared more efficient at revealing both the between-population and the within-population differentiation in the present study. The elite cassava cultivars were clearly separated from the African landraces by the SSR markers, while the dendrogram resulting from AFLP markers showed some overlap between the two populations (Figs 1 and 5). The level of within-population differentiation, shown by the landraces based on SSR markers in the present study, is comparable to the results of Fregene *et al.* (2003) where a similar level of extensive subgrouping was observed within some of the African cassava landraces in their study.

The PCA provided a visual representation of genetic differentiation and distribution within and between each population studied. Distribution of cultivars was observed on the basis of within-population and between-population estimates. The first three principal axes of the PCA plot accounted for up to 72% for SSR marker data within the landrace population (Fig. 5B). Three-dimensional PCA graphs showing the genetic distribution between the landraces and the elite cultivars are presented (Figs 5 and 6). Populations were grouped based on the eigenvalues of the first three axes. The PCA plots resulting from the first three principal components accounted for

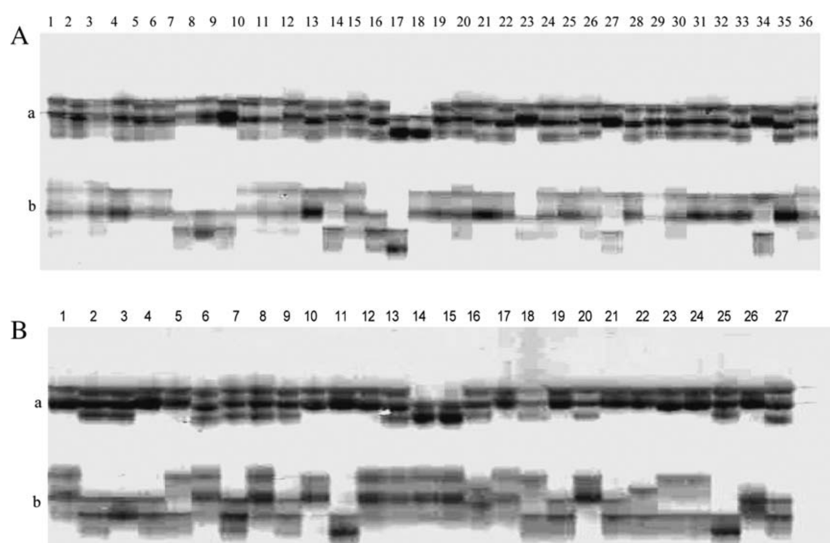

Figure 4 Simple sequence repeat loci showing polymorphism of (A) 36 cassava landraces and (B) 27 elite cassava cultivars generated using SSRY 105(a) and SSRY 61(b).

Table 2 Elite cassava populations and their respective pedigree information assayed using 50 SSR and 20 AFLP markers

Elite Cultivar	Pedigree and Selection Information
60506	Nigerian improved clone in the 1950s ex Moor Plantation, Ibadan, Nigeria
30555	58308 × OYARUGBA DUDU
30395	58308 OP
30040	30555 HS (58308 × OYARUGBA DUDU)
81/00401	50193 × (59 × IACT–127)
83672	COMPOSITE B
O83/00027	58308 × ISUNIKANKIYAN
82/00058	30572 HS
90059	CIAT 294/2 OP
W4488	58308 × 58308
88/02576	KR 685 × very low cyanide population (VLB)
50207	59 × IAC7–127 OP
30572	58308 × BRANCA DE SANTA CATARINA
58308	Genetic stock for CMD, CBB resistance and low cyanide, ex Moor Plantation
M83/00016	58308 × BRANCA DE SANTA CATARINA
88/02561	KR 685 × very low cyanide population (VLB)
30001	Lost pedigree
71693	COMPOSITE B
89/01111	30001 × (58308 × OYARUGBA FUNFUN)
85/01887	58308 × BRANCA DE SANTA CATARINA
89/00250	(59 × IAC7–1227) HS
89/00589	30001 × (58308 × OYARUGBA FUNFUN)
M86/00069	60076 (KR 651 OP) × high CMD/CBB res. population
M86/00083	60076 (KR 651 OP) × high CMD/CBB res. population
84776	58308 × ISUNIKANKIYAN
89/02115	53101 × high yield population
M82/00052	58308 × BRANCA DE SANTA CATARINA

AFLP, amplified fragment length polymorphism; CBB, cassava bacterial blight; CMD, cassava mosaic disease.

70% of the total eigenvalue for the SSR markers and 58% for AFLP markers to account for the variation between the populations. Many of the subgroups identified from the cluster analysis of the elite and landraces were also apparent on the PCA plots (Figs 5 and 6). MDS analysis conducted to further explore the variation structure in both populations also confirmed some of the genetic structure revealed by PCA and cluster analyses (data not shown). However, grouping of the population by MDS, while exhibiting a pattern similar to PCA and cluster analyses on a general scale, showed some discrepancies in the general distribution and classification along the first and second coordinates.

Discussion

The genetic diversity and population genetic structure observed in the present study indicate a high level of genetic differentiation within the landraces. The genetic differentiation values of 0.746 (SSR) and 0.656 (AFLP) suggest that geographical or regional variation could be responsible for most of the genetic differentiation observed. Balyejusa Kizito *et al.* (2007) reported a similar observation in Ugandan cassava landraces. Negative fixation index values observed may indicate an excess of heterozygosity, probably because of assortative mating for landraces and possibly attributable to heterosis (Hartl & Clark, 1997) in the elite lines. The genetic similarities observed among cultivars in countries with close proximity (notably, Bénin, Ghana, Nigeria and Togo) indicate possibilities of exchange/movement of planting

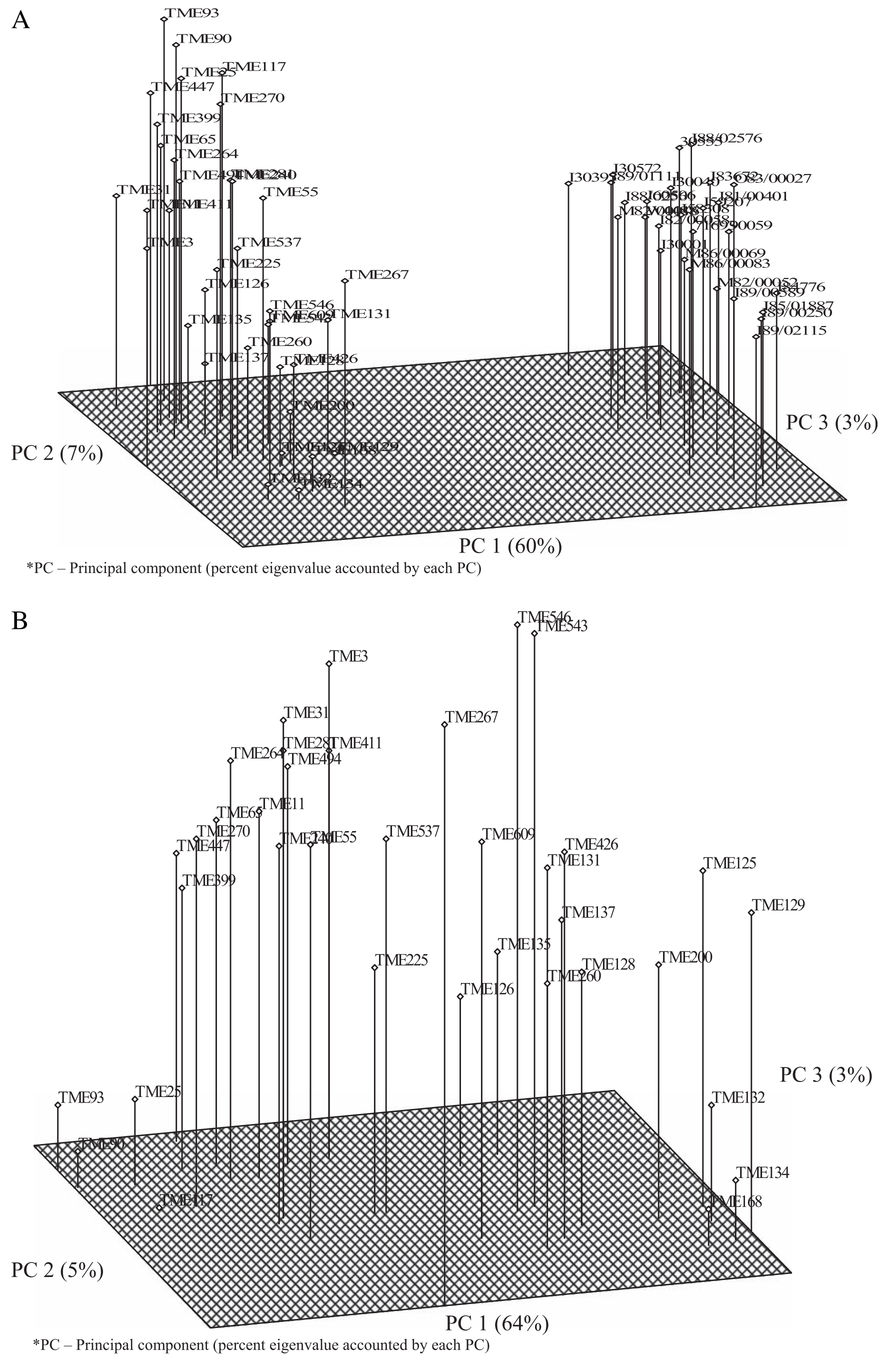

Figure 5 Principal component plot showing (A) genetic differentiation between African cassava landraces and elite cultivars based on SSR markers (B) within-population genetic differentiation in the African cassava landraces.

materials between farmers and research centres in these countries. The level of genetic differentiation within the elite cultivars was moderate to low, which is a direct reflection of the relatively narrow genetic base and the limited number of germplasm resources available to cassava breeders during the initial development of elite cultivar breeding. In addition, breeding pressures over the past two decades for selected traits of economic importance and the vegetative propagation nature of cassava are other critical factors affecting the genetic diversity of elite cassava germplasm. Results from the present study indicate that the clustering pattern of

*PC – Principal component (percent eigenvalue accounted by each PC)

Figure 6 Principal component plot showing genetic differentiation between African cassava landraces and elite cultivars based on AFLP markers.

most elite cultivars is in accordance with their pedigree information. A widely adopted IITA elite cultivar 30572 derived from 58308 and a Latin American cultivar (Branca de Santa Catarina) were closely clustered together with one of its parents. Although most of the elite genotypes are related by descent, a few did not cluster as closely together with their parents, progeny or half-sibs as expected; this was most evident from the clustering resulting from the AFLP markers and may possibly be a result of the dominant nature of the marker with less information content compared with the SSR markers. Dominant markers are unable to discriminate homozygous from heterozygous individuals simply by the presence or absence of fragments, unlike codominant markers that permit the explicit differentiation of these individuals. However, the application of AFLP and other dominant markers remains relevant in many genetic diversity studies because they are cheap (Piepho & Koch, 2000) and are able to analyse many loci per experiment without requiring previous sequence information.

In the past two decades, most of the elite cassava germplasm cultivars were developed using 58308 and its derivatives. The elite cultivar 58308 until recently has been the singular source of genetic stock for low cyanide content, resistance to cassava mosaic disease and cassava bacterial blight (Dixon *et al.*, 2001; Nweke, 2004; Ortiz, 2006; Raji *et al.*, 2008). Other limited genetic resources from Latin America, wild cassava relatives and a few African landraces have also been used in elite cultivar development (Nassar & Ortiz, 2007). Recently, breeding efforts at IITA have been geared towards the diversification of germplasm sources to develop new cultivars with novel traits, such as high starch content, combined disease resistance and high contents of beta-carotene and protein. Efforts through reverse genetics and crosses between cassava and castor bean (*Ricinus communis*), another genus in the Euphorbiaceae family, are currently being explored for novel genetic combinations.

In general, the genetic diversity of landraces in a particular geographical region is often affected by cultural traditions, undocumented movement and the exchange of materials by farmers, resulting in identical cultivars being given different names and vice versa (Mignouna & Dixon, 1997; Raji *et al.*, 2007). The genetic distribution pattern observed in this study is helpful for ongoing and future cassava breeding efforts for the selection of genetically distinct cultivars among the landraces for genetic mapping purposes and suggests the need for broadening the genetic base of the elite cassava breeding pool. The range of genetic differentiation observed among the African landraces in the present study and from several reports based on morphological characterisation (Raji *et al.*, 2007) and molecular markers (Fregene *et al.*, 2003; Lokko *et al.*, 2006) for special traits suggests that the landraces may prove very useful for the identification of new and unique alleles. Analysis of a more extensive collection of cassava germplasm is needed to define a more genetically diverse core set and to further refine the genetic structure of the cassava genepool.

Table 3 Genetic differentiation parameters averaged across the landraces and elite cassava population for each SSR locus

Locus	N_a	N_e	I	H_o	H_e	Fix. Index
SSRY1	5	3.650	1.407	0.730	0.726	−0.006
SSRY3	3	2.477	0.974	0.857	0.596	−0.437
SSRY4	4	3.301	1.278	0.698	0.697	−0.001
SSRY8	3	2.762	1.054	0.253	0.637	0.601
SSRY9	3	1.939	0.776	0.571	0.484	−0.179
SSRY16	5	4.218	1.512	0.968	0.762	−0.269
SSRY19	3	2.503	1.000	0.634	0.600	−0.057
SSRY21	4	2.627	1.093	0.380	0.619	0.384
SSRY23	4	3.153	1.250	0.190	0.682	0.715
SSRY27	4	3.625	1.335	0.682	0.724	0.057
SSRY28	5	3.623	1.383	0.952	0.723	−0.315
SSRY30	4	3.660	1.339	0.968	0.726	−0.332
SSRY32	3	2.688	1.040	0.952	0.627	−0.516
SSRY33	3	2.735	1.049	0.920	0.634	−0.451
SSRY34	3	1.971	0.806	0.666	0.492	−0.353
SSRY35	3	2.443	0.968	0.904	0.590	−0.531
SSRY42	4	3.295	1.285	0.873	0.696	−0.253
SSRY45	4	3.707	1.349	0.761	0.730	−0.043
SSRY49	4	3.010	1.149	0.936	0.667	−0.402
SSRY51	3	2.021	0.859	0.507	0.505	−0.005
SSRY59	4	3.082	1.188	0.746	0.675	−0.104
SSRY61	4	2.425	0.475	0.206	0.298	0.308
SSRY63	4	2.821	1.140	1.000	0.645	−0.548
SSRY70	2	1.950	0.680	0.492	0.487	−0.009
SSRY72	4	2.571	1.098	0.333	0.611	0.454
SSRY84	3	2.314	0.937	0.793	0.567	−0.397
SSRY91	2	1.746	0.619	0.365	0.427	0.145
SSRY95	4	3.359	1.273	1.000	0.702	−0.423
SSRY99	3	2.031	0.878	0.492	0.507	0.031
SSRY100	4	3.542	1.317	0.952	0.717	−0.327
SSRY101	5	4.087	1.495	1.000	0.755	−0.323
SSRY105	4	2.967	1.216	0.730	0.663	−0.101
SSRY106	4	2.673	1.120	0.492	0.625	0.213
SSRY109	4	3.455	1.304	0.857	0.710	−0.206
SSRY110	4	2.836	1.153	0.936	0.647	−0.446
SSRY113	5	3.199	1.313	0.952	0.687	−0.385
SSRY122	3	2.956	1.091	0.920	0.661	−0.391
SSRY134	3	2.856	1.073	0.825	0.649	−0.270
SSRY135	3	2.481	1.000	0.634	0.597	−0.063
SSRY158	4	2.846	1.120	0.968	0.648	−0.492
SSRY161	3	2.832	1.068	0.539	0.646	0.165
SSRY164	2	1.727	0.612	0.539	0.421	−0.281
SSRY170	4	3.093	1.230	1.000	0.676	−0.477
SSRY172	5	3.699	1.426	0.984	0.729	−0.348
SSRY173	5	3.379	1.344	1.000	0.704	−0.420
SSRY174	6	3.532	1.436	0.984	0.716	−0.372
SSRY177	3	2.670	1.036	0.650	0.625	−0.040
SSRY179	4	2.882	1.138	0.968	0.653	−0.482
SSRY182	3	2.624	1.018	0.174	0.618	0.718
SSRY185	3	2.123	0.870	0.698	0.529	−0.320
Mean	3.66	2.863	1.111	0.733	0.630	−0.150
SD	0.89	0.63	0.24	0.26	0.09	0.34

Fix. index, fixation index (deficiencies of heterozygotes relative to Hardy–Weinberg expectations); H_e, expected heterozygosity; H_o, observed heterozygosity; I, Shannon's information index; N_a, observed number of alleles; N_e, effective number of alleles; SSR, simple sequence repeat.

Table 4 Genetic differentiation indicators estimated for each population and F-statistics summary calculated across populations for all SSR loci

Mean	Landraces	Elite
No. of alleles	3.620 ± 0.124	3.500 ± 0.125
Na Freq. ≥5%[a]	3.300 ± 0.100	3.260 ± 0.117
No. of effective alleles/locus[b]	2.705 ± 0.090	2.623 ± 0.090
Information index	1.067 ± 0.033	1.034 ± 0.035
No. private alleles	0.160 ± 0.060	0.040 ± 0.028
H_e	0.607 ± 0.015	0.594 ± 0.015
F-statistics across population		
Mean *F*is	−0.200	
Mean *F*it	−0.153	
Mean *F*st	0.746 (SSR) 0.656 (AFLP)	

H_e, Nei's heterozygosity corrected for small sample sizes (Nei, 1978); *F*is, measures the correlation of genes in individuals within populations; *F*it, estimates the correlation of genes within individuals overall populations; *F*st, measures the correlation of genes of different individuals in the same population; SSR, simple sequence repeat.
[a]No. of alleles (Na), with frequency >5%.
[b]No. of effective alleles (Kimura & Crow, 1964).

Genetic differentiation was found to be lower in the elite lines than among the landraces, probably as a result of high selection pressure during breeding in the development of the elite cultivars, coupled with genetic drift. Populations derived from crosses with large differences in polymorphic markers may be utilised to map some of the useful traits of interest that have already been identified within the germplasm population. The estimated mean genetic similarity among the African cassava landraces resulting from both AFLP and SSR marker data suggested a considerable level of diversity from which diverse parental combinations could be generated. Moreover, the minimal contribution of the African landraces to the elite germplasm population development was also evident from the genetic differentiation displayed in a joint analysis of both populations by cluster and PCA plots on the basis of both AFLP and SSR marker systems.

The significant correlation observed between the clusters obtained from the two molecular marker systems could be an indication that the markers assessed some identical

Table 5 Analysis of molecular variance showing within-population and among-population variance for genetic differentiation for the landraces and elite cassava population

Source	d.f.	SS	MS	Est. Var.	%Var.	Value	*P* value
Among pops	1	183.09	183.09	5.14	17		
Within pops	61	1493.73	24.49	24.49	83	0.173	0.001
Total	62	1676.83	207.58	29.63			

Among pops, % variance between landraces and elite cassava cultivars; Est var., estimated variance; MS, mean square; ss, sum of squares; Within pops, % variance among both populations.

regions in the cassava genome. Similar results have been reported by Roa *et al.* (2000) where statistically significant correlations and agreements were observed between the results obtained from both AFLP and SSR markers.

Our studies showed that, in general, these two marker systems are powerful and able to provide genetic data that are reflective of pedigree background and geographical origin. Results obtained from AFLP and SSR did not conform on a 100% scale, but a reasonable degree of agreement was found between distance matrices, based on the Mantel test. The effectiveness of SSR and AFLP markers in the definition of germplasm diversity and cultivar associations was demonstrated.

Acknowledgements

The authors are grateful to R. Umelo for editing the manuscript and anonymous reviewers for their useful comments and suggestions. This research was supported by IITA, Ibadan, Nigeria.

References

Akano A.O., Dixon A.G.O., Mba C., Barrera E., Fregene M. (2002) Genetic mapping of a dominant gene conferring resistance to the cassava mosaic disease (CMD). *Theoretical and Applied Genetics*, **105**, 521–525.

Balyejusa Kizito E., Chiwona–Karltun L., Egwang T., Fregene M., Westerbergh A. (2007) Genetic diversity and variety composition of cassava on small scale farms in Uganda: an interdisciplinary study using genetic markers and farmer interviews. *Genetica*, **130**, 301–318.

Basu A., Ghosh M., Meyer R., Powell W., Basak S.L., Sen S.K. (2004) Analysis of genetic diversity in cultivated jute determined by means of SSR markers and AFLP profiling. *Crop Science*, **44**, 678–685.

Cabral G.B., Carvalho L.J.C.B., Schaal B.A. (2000) Relationship analysis of closely related species to cassava (*Manihot esculenta* Crantz) based on microsatellite–primed PCR. In *Proceedings of the Fourth International Scientific Meeting of the Cassava Biotechnology Network*, pp. 36–50. Bahia, Brazil: (CBN) Salvador.

Chavarriaga–Aguirre P. (2000) Review on the use of molecular markers to characterise cassava and its wild relatives' germplasm. In *Proceedings of the Fourth International Scientific Meeting of the Cassava Biotechnology Network*, pp. 180–191. Bahia, Brazil: (CBN) Salvador.

Chavarriaga–Aguirre P., Maya M.M., Tohme J., Duque M.C., Iglesias C., Bonierbale M.W., Kresovich S., Kochert G. (1999) Using microsatellites, isozymes and AFLPs to evaluate genetic diversity and redundancy in the cassava core collection and to assess the usefulness of DNA-based markers to maintain germplasm collections. *Molecular Breeding*, **5**, 263–273.

Dellaporta S.L., Wood J., Hicks J.R. (1983) A plant DNA mini preparation, version II. *Plant Molecular Biology Reporter*, **1**, 19–21.

Dixon A.G.O., Whyte J.B.A., Mahungu N.M., Ng S.Y.C. (2001) Tackling the cassava mosaic disease (CMD) challenge in sub-Saharan Africa: the role of host-plant resistance and germplasm deployment. In C. M. Faquet & N. J. Taylor (eds), *Cassava: An ancient crop for modern times* (compact disc). *Proceedings of the Fifth International Scientific Meeting of the Cassava Biotechnology Network, 4–9 November 2001*, pp 189. St. Louis, Missouri, USA.

Elias M., Panaud O., Robert T. (2000) Assessment of genetic variability in a traditional cassava (*Manihot esculenta* Crantz) farming system, using AFLP markers. *Heredity*, **85**, 219–230.

Elias M., Mühlen G.S., McKey D., Roa A.C., Tohme J. (2004) Genetic diversity of traditional South American landraces of cassava (*Manihot esculenta* Crantz): an analysis using microsatellites. *Economic Botany*, **58**, 242–256.

Excoffier L., Smouse P.E., Quattro J.M. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction sites. *Genetics*, **131**, 479–491.

FAO. (2006) *Food Outlook – No. 1 June 2006 available via* [WWW document]. URL http://www.fao.org/docrep/009/j7927e/j7927e00.HTM. Accessed 10 July 2008.

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.

Fregene M.A., Suarez M., Mkumbira J., Kulembeka H., Ndedya E., Kulaya A., Mitchel S., Gullberg U., Rosling H., Dixon A.G.O., Dean R., Kresovich S. (2003) Simple sequence repeat marker diversity in cassava landraces: genetic diversity and differentiation in an asexually propagated crop. *Theoretical and Applied Genetics*, **107**, 1083–1093.

Geleta N., Labuschagne M.T., Viljoen C.D. (2006) Genetic diversity analysis in sorghum germplasm as estimated by AFLP, SSR and morpho–agronomical markers. *Journal of Biodiversity and Conservation*, **15**, 3251–3265.

Hartl D.L., Clark A.G. (1997) *Principles of Population Genetics*. 3rd edn. Sunderland, MA, USA: Sinauer Associates, Inc. 519 pp.

Kimura M., Crow J.F. (1964) The number of alleles that can be maintained in a finite population. *Genetics*, **49**, 725–738.

Lokko Y., Danquah E.Y., Offei S.K., Dixon A.G.O., Gedil M.A. (2005) Molecular markers associated with a new source of resistance to the cassava mosaic disease. *African Journal of Biotechnology*, **4**, 873–881.

Lokko Y., Dixon A., Offei S., Danquah E., Fregene M. (2006) Assessment of genetic diversity among African cassava *Manihot esculenta* Crantz accessions resistant to the cassava mosaic virus disease using SSR markers. *Genetic Resources and Crop Evolution*, **53**, 1441–1453.

Mantel N. (1967) The detection of disease clustering and regression approach. *Cancer Research*, **27**, 209–220.

Mba R.E.C., Stephenson P., Edwards K., Melzer S., Mkumbira J., Gullberg U., Apel K., Gale M., Tohme J., Fregene M. (2001) Simple sequence repeat (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.

Mignouna H.D., Dixon A.G.O. (1997) Genetic relationships among cassava clones with varying levels of resistance to African mosaic disease using RAPD markers. *African Journal of Root and Tuber Crops*, **2**, 28–32.

Nassar N.M.A., Ortiz R. (2007) Cassava improvement: challenges and impacts. *Journal of Agricultural Science*, **145**, 163–171.

Nei M. (1972) Genetic distance between populations. *American Naturalist*, **106**, 283–293.

Nei M. (1978) Estimation of average heterozygosity and genetic distances from a small number of individuals. *Genetics*, **89**, 583–590.

Nweke F. (2004) New challenges in the cassava transformation in Nigeria and Ghana. EPTD discussion papers 118, International Food Policy Research Institute (IFPRI) [WWW document]. URL http://www.ifpri.org/divs/eptd/dp/papers/eptdp118.pdf. Accessed 15 october 2008.

Okogbenin E., Marin J., Fregene M. (2006) An SSR–based molecular genetic map of cassava. *Euphytica*, **147**, 433–440.

Ortiz R. (2006) Improving cassava for enhancing yield, minimizing pest losses and creating wealth in sub-Saharan Africa, cassava improvement to enhance livelihoods in sub-Saharan Africa and Northeastern Brazil. In *First International Meeting on Cassava Breeding, Biotechnology and Ecology, Brasilia, Brazil, 11-15 November 2006*, pp. 129–141. Eds Ortiz, R. and N.M.A. Nassar. Brazil: Universidade de Brasilia.

Peakall R., Smouse P.E. (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288–295.

Piepho H.P., Koch G. (2000) Codominant analysis of banding data from a dominant marker system by normal mixtures. *Genetics*, **155**, 1459–1468.

Prakash N.S., Combes M.C., Dussert S., Naveen S., Lashermes P. (2005) Analysis of genetic diversity in Indian robusta coffee genepool (*Coffea canephora*) in comparison with a representative core collection using SSRs and AFLPs. *Genetic Resources and Crop Evolution*, **52**, 333–343.

Raji A.A. (2003) *Assessment of genetic diversity and heterotic relationships in African improved and local cassava.* (Manihot esculenta, *Crantz) germplasm*. PhD Thesis. University of Ibadan, Nigeria: 218 pp.

Raji A.A., Dixon A.G.O., Ladeinde T.A.O. (2007) Agronomic traits and tuber quality attributes of farmer grown cassava (*Manihot esculenta*) landraces in Nigeria. *Journal of Tropical Agriculture*, **45**, 9–13.

Raji A.A., Dixon A.G.O., Ladeinde T.A.O. (2008) Screening landraces for additional sources of field resistance to cassava mosaic disease and green mite for integration into the cassava improvement program. *Journal of Integrative Plant Biology*, **50**, 311–318.

Roa A.C., Maya M.M., Duque M.C., Tohme J., Allem A.C., Bonierbale M.W. (1997) AFLP analysis of relationships among cassava and other *Manihot* species. *Theoretical and Applied Genetics*, **95**, 741–750.

Roa A.C., Chavarriaga-Aguirre P., Duque M.C., Maya M.M., Bonierbale M.W., Iglesias C., Tohme J. (2000) Cross-species amplification of cassava (*Manihot esculenta*) (Euphorbiaceae) microsatellites: allelic polymorphism and degree of relationship. *American Journal of Botany*, **87**, 1647–1655.

Rohlf F.J. (1990) *NTSYS–pc Numerical Taxonomy and Multivariate Analysis System, version 2.02j*. New York, USA: Exeter Publications.

SAS. (2004) *SAS Institute User's Guide: Statistics, version 9.0*. Cary, NC, USA: SAS Institute. 1028 pp.

Schiffman S.S., Reynold M.L., Young F.W. (1981) *Introduction to Multidimensional Scaling: Theory, Methods and Applications*. New York, USA: Academic Press.

Sneath P.H.A., Sokal R.R. (1973) *Numerical Taxonomy–The Principles and Practice of Numerical Classification*. San Francisco, CA, USA: W.H. Freeman & Co. 573 pp.

Tangphatsornruang S., Sraphet S., Singh R., Okogbenin E., Fregene M., Triwitayakorn K. (2008) Development of polymorphic markers from expressed sequence tags of *Manihot esculenta* Crantz. *Molecular Ecology Resources*, **8**, 682–685.

Vos P., Hogers R., Bleeker M., Reijans M., Van de Lee T., Hornes M., Frijters A., Pot J., Peleman J., Kuiper M., Zabeau M. (1995) AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, **21**, 4407–4414.

Weir B.S., Cockerham C.C. (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **3**, 1358–1370.

Wong H.L., Yeoh H.H., Lim S.H. (1999) Customisation of AFLP analysis for cassava varietal identification. *Phytochemistry*, **50**, 919–924.

Wright S. (1951) The genetical structure of populations. *Annals of Eugenics*, **15**, 323–354.

Wright S. (1978) *Variability Within and Among Natural Populations*. Vol. 4. Chicago, USA: The University of Chicago Press. 580 pp.

Yeh F.C., Yang R.C., Boyle T.B.J., Ye Z.H., Mao J.X. (1997) *POPGENE, the User–Friendly Shareware for Population Genetic Analysis. Molecular Biology and Biotechnology Centre*. Canada: University of Alberta.