

# Molecular diversity of Cuban cassava (*Manihot esculenta* Crantz) cultivars assessed by simple sequences repeats (SSR)

Yoel Beovides <sup>(1)</sup>, Martin Fregene <sup>(2)</sup>, Janneth P. Gutiérrez <sup>(3)</sup>, Marilys D. Milián <sup>(1)</sup>, Orlando Coto <sup>(4)</sup>, Charles Buitrago <sup>(3)</sup>, José A. Cruz <sup>(1)</sup>, Elianet Ruiz <sup>(1)</sup>, Milagros Basail <sup>(1)</sup>, Aymé Rayas <sup>(1)</sup>, Daniel Rodríguez <sup>(1)</sup>, Arletys Santos <sup>(1)</sup>, Jorge López <sup>(1)</sup>, Víctor Medero <sup>(1)</sup>

<sup>(1)</sup> Instituto de Investigaciones de Viandas Tropicales (INIVIT). Apdo. 6. Santo Domingo. 53000 Villa Clara (Cuba). E-mail: biomol.biotech@inivit.cu

<sup>(2)</sup> Donald Danforth Plant Science Center. 975 N Warson Rd. Creve Coeur. St. Louis, MO 63132 (United States).

<sup>(3)</sup> Internacional Center of Tropical Agricultura (CIAT). Km 17. Recta Cali-Palmira. Apartado Aéreo 6713. Cali (Colombia).

<sup>(4)</sup> Instituto de Investigaciones en Fruticultura Tropical. 7ma No. 3005 e/ 30 y 32. Miramar. Playa, Habana (Cuba).

Received on September 9, 2013; accepted on June 29, 2015.

A total of 36 microsatellites (SSR) markers were used to analyze the genetic diversity of 163 accessions of cultivated cassava (*Manihot esculenta* Crantz), 94 accessions of them from the Cuban Cassava Germplasm Collection and 69 genotypes from different countries and conserved at the International Center for Tropical Agriculture (Colombia). This study was carried out to determine genetic diversity within and between all accessions to promote their better use and conservation strategies. Thirty-four of those markers were used for the genetic diversity study based on their higher polymorphism. The Cuban cultivars showed the highest average allele number per loci with 5.8 and 100% of the loci were polymorphic, as well as those from Guatemala. The average proportion of individual heterozygosity observed ( $H_o$ ) was high ( $0.5918 \pm 0.0351$ ), while the highest  $H_o$  rates were observed in groups of genotypes from Cuba (0.6016) and Tanzania (0.6459). The total heterozygosity ( $H_T$ ) was high ( $0.6538 \pm 0.1770$ ), but only 7.4% ( $G_{ST} = 0.0740 \pm 0.0377$ ) was due to differences between the five countries studied. Genetic differentiation coefficients (estimated by F-statistics) were low to moderate ( $F_{ST} > 0.04$ ) and 17 unique alleles with low frequency were found in Cuban cultivars. The results provide the first molecular characterization of Cuban cassava genotypes and showed a wide diversity among landraces from Cuba. Application of this valuable information can be used for genetic diversity conservation and genotype identification studies for the genetic breeding program of cassava.

**Keywords.** *Manihot esculenta*, genetic variation, genetic markers, microsatellites, germplasm.

## Diversité moléculaire de cultivars cubains de manioc (*Manihot esculenta* Crantz) analysés par séquences simples répétées (SSR)

Un total de 36 marqueurs microsatellites (SSR) ont été utilisés pour analyser la diversité génétique de 163 accessions cultivées de manioc (*Manihot esculenta* Crantz), 94 d'entre elles à partir de la collection de matériel génétique de manioc cubain et 69 autres génotypes de différents pays d'Afrique, d'Asie et d'Amérique latine, conservés au Centre International d'Agriculture Tropicale (Colombie). L'étude a été menée afin de déterminer la diversité génétique au sein et entre toutes les accessions pour promouvoir une meilleure conservation et utilisation. Trente-quatre SSR étaient polymorphes et utiles dans les études sur la diversité génétique. Les cultivars cubains ont montré le plus grand nombre de loci avec en moyenne 5,8 allèles et 100 % étaient des loci polymorphes, comme ceux du Guatemala. L'hétérozygotie moyenne observée ( $H_o$ ) était élevée ( $0,5918 \pm 0,0351$ ), avec les valeurs les plus élevées de  $H_o$  pour les génotypes de Cuba (0,6016) et de Tanzanie (0,6459). L'hétérozygotie globale ( $H_T$ ) était également élevée ( $0,6538 \pm 0,1770$ ), mais seulement 7,4 % ( $G_{ST} = 0,0740 \pm 0,0377$ ) était dû à des différences entre les cinq pays étudiés. Les coefficients de la différenciation génétique (estimé par F-Statistics) étaient bas à modérés et 17 allèles uniques ont été trouvés avec une faible fréquence d'occurrence entre les génotypes cubains. Les résultats fournissent la première caractérisation moléculaire des génotypes de manioc conservés à Cuba et confirment leur large variabilité génétique. L'exploitation de cette information importante contribue à améliorer les stratégies nationales de conservation et d'utilisation du patrimoine génétique, en particulier dans les programmes d'amélioration génétique du manioc.

**Mots-clés.** *Manihot esculenta*, différence génétique, marqueur biologique, microsatellite, matériel génétique.

## 1. INTRODUCTION

Cassava (*Manihot esculenta* Crantz) is an important crop that is commonly grown in tropical and subtropical areas. It grows very well in a less fertile soil in contrast to many other crops, even in marginal areas and it is able to express a good yield of roots, under favorable production conditions or not; millions of people around the world depend on it to subsist (Vásquez et al., 2012). However, its potentialities are far from being adequately exploited and knowledge of its genetic diversity is limited.

The application of advanced molecular breeding technologies could be useful for cassava research studies. This has been demonstrated through research on worldwide important crops (Raji et al., 2009). Molecular marker technologies can assist conventional breeding efforts as valuable tools for the analysis of genetic relatedness, the identification and selection of desirable genotypes for crosses as well as for germplasm conservation.

Molecular genetic markers can be used to map genomes, identify regions of the genome controlling a trait, and follow a segment of interest of the genome in a plant breeding scheme (Kunkeaw et al., 2010; Bang et al., 2011; Ferguson et al., 2012). Various research studies on cassava genetic variation were based on molecular markers such as Restriction Fragment Length Polymorphisms (RFLP) (Fregene et al., 1997) or Random Amplified Polymorphic DNA (RAPD) (Vieira et al., 2011). Nevertheless, Simple Sequence Repeats (SSRs) or microsatellites are particularly attractive for diversity studies (Siqueira et al., 2009; Turyagyenda et al., 2012) because they are abundant in plants and they have a high level of polymorphism (Mba et al., 2001).

Microsatellites are powerful markers for both genetic diversity studies and marker assisted selection (MAS) and they have been used in many plant species (Roubus et al., 2010; Lekha et al., 2011). In addition, SSRs markers were used to evaluate the molecular diversity of some cassava germplasm (Raghu et al., 2007; Kunkeaw et al., 2011; Montero-Rojas et al., 2011) or to compare genetic diversity between improved cassava cultivars and landraces from Africa, Latin America, Asia and four *Manihot* wild species (Whankaew et al., 2012).

In Cuba, cassava serves as one of the main sources of energy-rich food for humans and animals. The “Instituto de Investigaciones de Viandas Tropicales” (INIVIT) preserves an important germplasm collection with 540 cassava genotypes, from different collects made in Cuba (nationwide), some of them are improved cultivars and others were introduced in Cuba from other countries by germplasm’s exchange. Previous morphological and agronomic characterization

revealed a wide genetic variability (Milián et al., 2000). Current advance in the development of methods using DNA polymorphisms as molecular markers provides alternative methods for landraces characterization and especially, microsatellites could help to reveal or to evaluate the genetic variability in the Cuban germplasm.

Genetic diversity of cassava cultivars was assessed in 94 Cuban accessions from the germplasm held at INIVIT, as a mean of designing better conservation approaches and identifying progenitors with a wide genetic base for breeding. Therefore, the objectives of this study are:

- to assess the genetic diversity of 94 Cuban cassava accessions with an appropriate number of SSR markers;
- to determine genetic relationships between Cuban accessions and other 69 genotypes from different geographic origin.

## 2. MATERIALS AND METHODS

### 2.1. Plant material

A total of 163 cassava accessions from different geographical origin were used (**Table 1**), 94 of them are Cuban (collected or obtained by breeding programs in Cuba) and were selected from the Cuban cassava germplasm collection held at “Instituto de Investigaciones de Viandas Tropicales” (INIVIT); these were selected according to their genetic or productive potential according to previous morphoagronomic studies (Milián et al., 2000), and breeders and producers criteria. The remaining 69 accessions were chosen from the International Centre of Tropical Agriculture (CIAT) germplasm collection in Colombia; these accessions represent the main cassava producing countries. These genotypes were used before in similar studies of diversity with the same SSR set we used in this investigation (Fregene et al., 2003). The genotypes ‘TMS30572’ and ‘CM2177-2’ were used as controls because of their very well-known genetic pattern for the evaluated microsatellites.

Woody stakes of each cultivar, 20 to 30 cm length, were harvested from all Cuban genotypes and planted in a greenhouse, the young leaf tissue from the buds generated from cuttings were used to obtain the genomic DNA for each Cuban accession. The DNA from the others 69 accessions was kindly given by the Cassava Breeding Program of CIAT.

### 2.2. DNA extraction

Two hundred mg of young and healthy leaves from each Cuban accession, collected from small buds

**Table 1.** List of 163 cassava (*Manihot esculenta*) accessions from different geographic origins used for genetic diversity in this study with 36 SSR markers — *Liste de 163 accessions de manioc (Manihot esculenta) d'origines géographiques différentes utilisées dans cette étude de la diversité génétique au moyen de 36 marqueurs SSR.*

No.	Accession	Origin	No.	Accession	Origin	No.	Accession	Origin
1	TMS30572 *	Nigeria	56	MCOL146	Colombia	111	TME 1068	Nigeria
2	CM2177-2 *	Colombia	57	CUB830 (C)	Cuba	112	CUB663 (C)	Cuba
3	CUB622	Cuba (W)	58	CUB831 (E)	Cuba	113	TME 1119	Nigeria
4	CUB624	Cuba (C)	59	TME(299)	Nigeria	114	TME 1134	Nigeria
5	34 (2002)	Guatemala	60	CUB869 (E)	Cuba	115	CUB671 (U)	Cuba
6	CUB629	Cuba (C)	61	CUB898 (U)	Cuba	116	TME 1169	Nigeria
7	CUB630	Cuba (C)	62	CUB897 (C)	Cuba	117	CUB678 (C)	Cuba
8	GUA149	Guatemala	63	CUB619 (W)	Cuba	118	TZ126	Tanzania
9	CUB634	Cuba (C)	64	CUB900 (W)	Cuba	119	CUB685 (U)	Cuba
10	CUB636	Cuba (C)	65	CUB901 (W)	Cuba	120	CUB688 (W)	Cuba
11	GUAT1	Guatemala	66	CUB893 (C)	Cuba	121	TZ177	Tanzania
12	CUB640	Cuba (C)	67	CUB968 (U)	Cuba	122	CUB693 (C)	Cuba
13	CUB642**	Cuba (E)	68	CUB970 (C)	Cuba	123	CUB703 (C)	Cuba
14	TME 981	Nigeria	69	CUBY-26 (B)	Cuba	124	TZ192	Tanzania
15	CUB651	Cuba (C)	70	CUB1114** (E)	Cuba	125	CUB706 (C)	Cuba
16	CUB655	Cuba (C)	71	CUB1129 (E)	Cuba	126	CUB723 (C)	Cuba
17	TME 1040	Nigeria	72	CUB1134 (E)	Cuba	127	ARG 8	Argentina
18	CUB660 (W)	Cuba	73	CUB1149 (E)	Cuba	128	CUB727 (C)	Cuba
19	CUB662 (C)	Cuba	74	CUB1152 (E)	Cuba	129	CUB733 (U)	Cuba
20	TME 1115	Nigeria	75	CUB1155 (E)	Cuba	130	BRA 191	Brazil
21	CUB665 (W)	Cuba	76	CUB1162 (E)	Cuba	131	CUB741 (E)	Cuba
22	CUB670 (E)	Cuba)	77	CUB1167 (E)	Cuba	132	CUB742 (U)	Cuba
23	TME 1153	Nigeria	78	CUB1177 (E)	Cuba	133	COL 965	Colombia
24	CUB677 (C)	Cuba	79	CUB1186 (E)	Cuba	134	CUB745 (W)	Cuba
25	TZ10	Tanzania	80	CUB1207 (E)	Cuba	135	CUB764 (U)	Cuba
26	CUB679 (C)	Cuba	81	CUB1227 (E)	Cuba	136	COL 1818	Colombia
27	TZ45	Tanzania	82	CUB1324 (U)	Cuba	137	CUB767 (C)	Cuba
28	TZ50	Tanzania	83	CUB1509 (U)	Cuba	138	CUB769** (E)	Cuba
29	CUB691 (U)	Cuba	84	CUB1510 (U)	Cuba	139	MEX 104	Mexico
30	TZ71	Tanzania	85	CUB1513 (U)	Cuba	140	CUB781 (U)	Cuba
31	TZ72	Tanzania	86	CUB1605 (C)	Cuba	141	PER 534	Peru
32	CUB705 (C)	Cuba	87	CUB1908 (E)	Cuba	142	CUB796 (C)	Cuba
33	TZ141	Tanzania	88	CUB1909** (B)	Cuba	143	VEN 45-A	Venezuela
34	TZ167	Tanzania	89	CUB1911 (W)	Cuba	144	VEN 183	Venezuela
35	CUB725** (B)	Cuba	90	CUBY-1 (B)	Cuba	145	CUB812 (U)	Cuba
36	ARG 6	Argentina	91	CUBY-24 (B)	Cuba	146	TME8	Nigeria
37	BRA 134	Brazil	92	NGA-19	Nigeria	147	TAI8	Thailand
38	CUB734 (U)	Cuba	93	MCOL273-6	Colombia	148	CUB828 (C)	Cuba
39	BRA 522	Brazil	94	MCOL273-7	Colombia	149	CM527-7	Colombia

**Table 1 (continued).** List of 163 cassava (*Manihot esculenta*) accessions from different geographic origins used for genetic diversity in this study with 36 SSR markers — *Liste de 163 accessions de manioc (Manihot esculenta) d'origines géographiques différentes utilisées dans cette étude de la diversité génétique au moyen de 36 marqueurs SSR.*

No.	Accession	Origin	No.	Accession	Origin	No.	Accession	Origin
40	COL 638	Colombia	95	GUA246	Guatemala	150	TME7	Nigeria
41	CUB744 (C)	Cuba	96	GUA1030	Guatemala	151	CUB863 (E)	Cuba
42	COL 976	Colombia	97	CUB625** (B)	Cuba	152	MCOL273-8	Colombia
43	COL 1634	Colombia	98	51 (2002)	Guatemala	153	MCOL273-9	Colombia
44	CUB766 (U)	Cuba	99	GUA853	Guatemala	154	4 (2002)	Guatemala
45	COL1999	Colombia	100	CUB633 (W)	Cuba	155	CUB899 (C)	Cuba
46	MEX 12	Mexico	101	GUA839	Guatemala	156	9 (2002)	Guatemala
47	CUB775 (U)	Cuba	102	GUA559	Guatemala	157	PER 183	Peru
48	PER 259	Peru	103	CUB637 (C)	Cuba	158	CUB902 (U)	Cuba
49	CUB787 (W)	Cuba	104	CUB630 (C)	Guatemala	159	ECU72	Ecuador
50	PER 569	Peru	105	TME 931	Nigeria	160	TME4	Nigeria
51	CUB811 (E)	Cuba	106	CUB647 (C)	Cuba	161	CUB1109 (E)	Cuba
52	CUB817** (B)	Cuba	107	TME 1003	Nigeria	162	CUB620 (C)	Cuba
53	TME41	Nigeria	108	TME 1016	Nigeria	163	CUB894 (C)	Cuba
54	CUB818 (E)	Cuba	109	CUB657 (C)	Cuba			
55	CUB821 (B)	Cuba	110	TME 1052	Nigeria			

\*: genotypes used as positive control for its well-known genetic pattern for SSR markers evaluated — *genotypes utilisés comme témoin positif pour leur modèle génétique bien connu pour évaluer les marqueurs SSR*; \*\*: official Cuban cassava commercial cultivars — *cultivars commerciaux officiels cubains* (MINAG, 2008); For Cuban accessions — *pour les accessions cubaines*: W: collected from west region — *collecté dans des régions à l'Ouest*; C: collected from central region — *collecté dans les régions centrales*; E: collected from east region — *collecté dans des régions à l'Est*; B: obtained by genetic breeding at INIVIT — *obtenu par amélioration génétique à l'INIVIT*; U: unknown origin — *origine inconnue*.

grown under greenhouse conditions, were used. The total genomic DNA of each cultivar was carried out according to Dellaporta et al. (1983). Quality of total DNA was checked by electrophoresis in 0.8% agarose gel (Amresco, Solon, OH, USA) using 0.55 M Tris-EDTA buffer. Each DNA was visualized by staining in 0.5 µg·mg<sup>-1</sup> Ethidium bromide; dilutions of phage lambda DNA of known concentration were used as control. Finally, DNA obtained from each accession was quantified by fluorometry (DyNA Quant TM. 200 Fluorometer, Hoefer Pharmacia Biotech).

### 2.3. Polymerase Chain Reaction (PCR) with SSR markers

A dilution of each DNA sample was done to a final concentration of 10 ng·µl<sup>-1</sup> and was used for PCR analysis. The cocktail for PCR was prepared as follows: 50 ng DNA, 0.2 mM dNTPs, 2.5 buffering solution 10 X, 2 mM MgCl<sub>2</sub>, 1.25 µM primer F, 1 U of DNA polymerase *Taq* and H<sub>2</sub>O (quality HPLC) until 25 µl final volume. Amplification was carried

out in a thermocycler PTC-100, the thermocycling profile consisted of an initial denaturation step for 1 min at 94 °C, followed by 30 cycles of denaturation at 94 °C for 30 s, annealing at 55 °C or 45 °C for 1 min and primer extension at 72 °C for 1 min, and a final extension cycle of 5 min at 72 °C. The PCR product stayed at 4 °C during an indefinite time. The PCR product was denatured and electrophoresed on 4% polyacrylamide gels, using an automated DNA sequencer ABI model 377 (Pelkin Elmer Inc.).

Gels were fixed in 10% HAc and the bands were revealed by staining with silver nitrate (SilverSequence TM DNA Sequencing System, Promega) according to Fregene et al. (2002).

The raw gel data were extracted using the GENESCAN package of ABI PRISM for Windows and GENOTYPER software (Pelkin Elmer Inc.). The extracted data were exported as allele sizes to Microsoft Excel (Microsoft Inc.) for further formatting as input files for statistical analysis. A strictly diallelic model of inheritance was applied; in other words, markers having three or more alleles were eliminated according to Fregene et al. (2003).

## 2.4. Gene diversity and genetic differentiation studies

SSR markers used in this study are listed in **table 2** and were supplied by CIAT, and consisted of 36 primer pairs, which were selected from a subset of

SSR markers developed at CIAT (Mba et al., 2001). The 36 primer pairs were selected based on being polymorphic, having clear reproducible allele patterns, high polymorphic information content (PIC), evenly spread across the cassava genome (Mba et al., 2001; Hurtado et al., 2008).

**Table 2.** List of the 36 SSR markers used to assess genetic diversity in 163 cassava accessions from different geographic origins — *Liste des 36 marqueurs SSR employés pour l'étude de la diversité génétique dans 163 accessions de manioc d'origines géographiques différentes.*

SSR locus	Type of repeat	Right primer	Left primer	Product size (pb)
SSRY4	GA(16) GA(3)	ATAGAGCAGAAGTGCAGGCG	CTAACGCACACGACTACGGA	287
SSRY9	GT(15)	ACAATTCATCATGAGTCATCAACT	CCGTTATTGTTCTGGTCTCT	278
SSRY12	CA(19)	AACTGTCAAACCAATTCTACTTGC	GCCAGCAAGGTTTGCTACAT	266
SSRY19	CT(8)CA(18)	TGTAAGGCATTCCAAGAATTATCA	TCTCCTGTGAAAAGTGCATGA	214
SSRY20	GT(14)	CATTGGACTTCCTACAAATATGAAT	TGATGGAAAAGTGGTTATGTCCTT	143
SSRY21	GA(26)	CCTGCCACAATATTGAAATGG	CAACAATTGGACTAAGCAGCA	192
SSRY34	GGC(7)	TTCCAGACCTGTTCCACCAT	ATTGCAGGGATTATTGCTCG	279
SSRY38	CA(17)	GGCTGTTCGTGATCCTTATTAAC	GTAGTTGAGAAAACCTTTCATGAG	122
SSRY51	CT(11)CA(18)	AGGTTGGATGCTTGAAGGAA	GGATGCAGGAGTGCTCAACT	298
SSRY59	CA(20)	GCAATGCAGTGAACCACTCTTT	CGTTTGTCTTTCTGATGTTT	158
SSRY63	GA(16)	TCAGAATCATCTACCTTGGCA	AAGACAATCATTTTGTGCTCCA	290
SSRY64	CT(13)CT(6)	CGACAAGTCGTATATGTAGTATTC	GCAGAGGTGGCTAACGAGAC	194
SSRY69	CT(18)	CGATCTCAGTCGATACCCAAG	CACTCCGTTGCAGGCATTA	239
SSRY82	GA(24)	TGTGACAATTTTCAGATAGCTTCA	CACCATCGGCATTAACCTTTG	211
SSRY100	GT(17)TTCT(7)	ATCCTTGCCTGACATTTTGC	TTCGCAGAGTCCAATTGTTG	210
SSRY102	GT(11)	TTGGCTGCTTTCACTAATGC	TTGAACACGTTGAACAACCA	179
SSRY103	GA(22)	TGAGAAGGAAACTGCTTGAC	CAGCAAGACCATCACCAGTTT	272
SSRY105	GT(6)GCGA(16)	CAAACATCTGCACTTTTGGC	TCGAGTGGCTTCTGGTCTTC	225
SSRY106	CT(24)	GGAAACTGCTTGACAAAAGA	CAGCAAGACCATCACCAGTTT	270
SSRY108	CT(24)CCT	ACGCTATGATGTCCAAAGGC	CATGCCACATAGTTCTGCT	203
SSRY110	GT(12)	TTGAGTGGTGAATGCGAAAAG	AGTGCCACCTTGAAAGAGCA	247
SSRY127	GT(8)GC(5)	GCTGAACTGCTTTGCCAACT	CTTCGGCCTCTACAAAAGGA	130
SSRY132	CA(8)	CTTTTGGCAGTCTTCTGTC	TGTCCAATGTCTTCTTTCTCT	196
SSRY135	CT(16)	CCAGAAACTGAAATGCATCG	AACATGTGCGACAGTGATTG	253
SSRY147		GTACATCACCACCAACGGGC	AGAGCGGTGGGGCGAAGAGC	113
SSRY148		GGCTTCATCATGGAAAAACC	CAATGCTTTACGGAAGAGCC	114
SSRY151		AGTGGAAATAAGCCATGTGATG	CCCATAATTGATGCCAGGTT	182
SSRY155		CGTTGATAAAGTGGAAAAGAGCA	ACTCCACTCCCAGTCTCGC	158
SSRY161	GT(11)GA(23)	AAGGAACACCTCTCCTAGAATCA	CCAGCTGTATGTTGAGTGAGC	220
SSRY164	GA(29)	TCAAACAAGAATTAGCAGAAGCTGG	TGAGATTTGTAATATTCATTTCACTT	187
SSRY169	GA(19)	ACAGCTCTAAAAACTGCAGCC	AACGTAGGCCCTAACTAACC	100
SSRY171	TA(5)GA(23)	ACTGTGCCAAAATAGCCAAATAGT	TCATGAGTGTGGGATGTTTTATG	291
SSRY177	CT(N)(65)CT(18)	ACCACAAACATAGGCACGAG	CACCCAATTCACCAATTACCA	268
SSRY179	GA(28)	CAGGCTCAGGTGAAGTAAAGG	GCGAAAGTAAGTCTACAACCTTTCTAA	226
SSRY180	GA(16)	CCTTGGCAGAGATGAATTAGAG	GGGGCATTCTACATGATCAATAA	163
SSRY181	GA(22)	GGTAGATCTGGATCGAGGAGG	CAATCGAAACCGACGATACA	199

Gene diversity and genetic differentiation analyses were carried out with data from the polymorphic SSR markers that showed distinct and scorable DNA bands, their diallelic nature and clear patterns.

All unambiguous SSR fragments from each of the polymorphic SSR markers were scored using binary numbers as presence (1) and absence (0) of an allele across all 163 genotypes. The genotypes TAI8 (Thailand) and ECU72 (Ecuador) were also eliminated because of too many missing data. From the obtained matrix, similarity degree within and between countries were calculated using Dice similarity coefficient (Dice, 1945). The following analysis related with diversity and differentiation studies only included the information related with cultivars from the most widely represented countries in this investigation: Cuba, Colombia, Guatemala, Nigeria and Tanzania (hereinafter, consider all accessions from each country as a 'group' or 'population').

Genetic diversity within and between countries was estimated using the software package GENSURVEY (Vekemans et al., 1997), with the following statistics: percentage of polymorphic loci, mean number of alleles per polymorphic locus, average observed heterozygosity ( $H_o$ ) and average gene diversity ( $H_e$ ) (Nei, 1978). For all loci and for all accessions, total heterozygosity ( $H_T$ ) and the proportion of among-accession differentiation ( $G_{ST}$ ) were estimated according to Nei (1978). Standard deviations for the above parameters were estimated over loci and sampled by jackknifing (Quenoille, 1956) using 200 replications. Given the small evolutionary divergence times for the accessions, the infinite allele model (Kimura et al., 1964) was assumed.

Genetic differentiation was quantified by the F-statistic estimator ( $F_{ST}$ , theta) (Wright, 1951), as described by Weir et al. (1984), using FSTAT 2.9 (Goudet, 2001), and by  $G_{ST}$  (Nei, 1978); FSTAT performs bootstrapping over loci and, given the large number of unlinked SSR loci employed in this study, it provides rigorous testing of hypotheses of genetic differentiation.  $F_{ST}$  gives a similar estimate of genetic differentiation as  $G_{ST}$  but  $G_{ST}$  takes into account variation in samples (accessions groups = accessions group by country) sizes, which is relevant in this study.  $F_{ST}$  values were estimated per allele, per locus and overall. Confidence intervals were calculated per locus over samples, and over loci by jackknifing (200 replicates), and by bootstrapping (1,000 bootstraps), over loci.

Pairwise genetic distances between individual accessions were calculated from the raw allele size-data based on the 1-proportion of shared alleles (1-PSA) (Bowcock et al., 1994), and estimated using the computer program "MICROSAT" (Minch, 1996). Distances based on 1-PSA give a more exact representation of genetic relationships when using

closely related genotypes and SSR markers, compared with other distance estimates (Bertin et al., 2001). The distance matrix generated was subjected to principal component analysis (PCA), using the program JMP (SAS Institute, 1999) with the purpose to obtain a structure of relationships among cassava accessions; a scatter plot of all cassava accessions and, another independent scatter plot of only Cuban accessions in two dimensions were produced.

A comparison matrix between countries was created with  $F_{ST}$  values using FSTAT (Goudet, 2001), and the relationship was analyzed by cluster analysis, using the unweighted pair group method using arithmetic averages (UPGMA). UPGMA clustering was performed using the SAHN programme of NTSYS-PC (Rohlf, 1998) to construct the dendrogram with a view to analyze the relationship between the Cuban landraces and the other cultivars by selecting the tree plot option in the same software package.

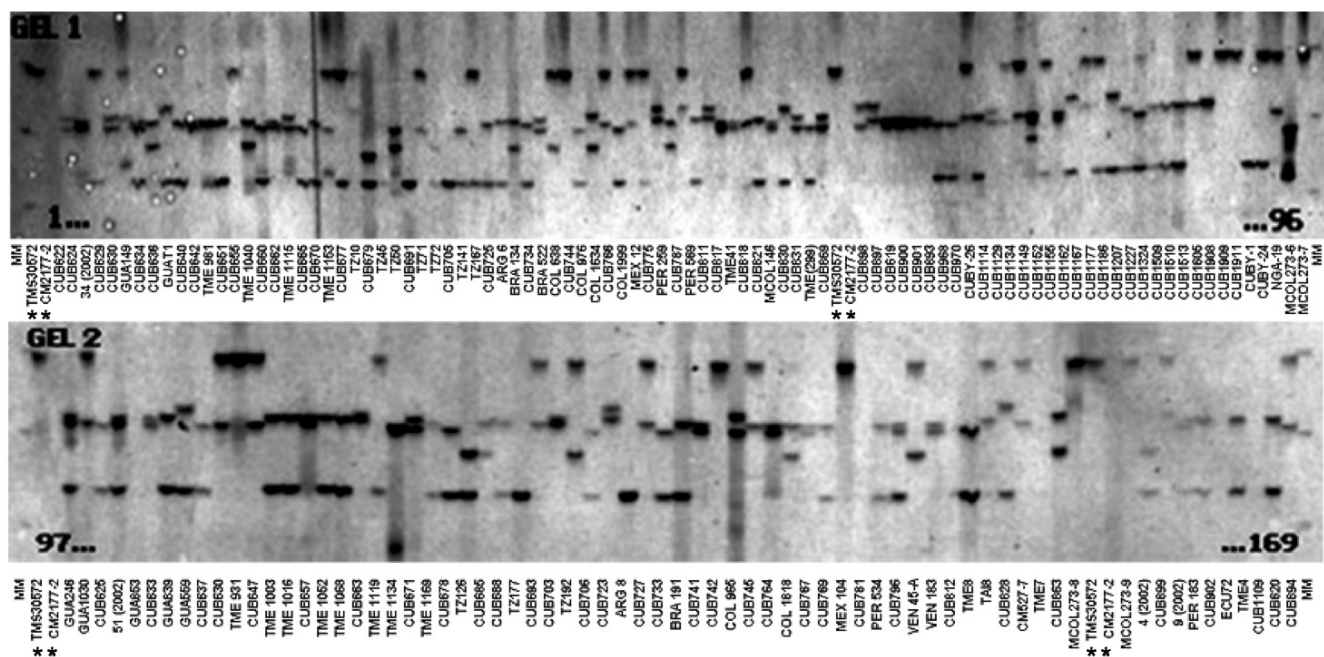
### 3. RESULTS AND DISCUSSION

Genetic diversity in cassava was evaluated in 163 genotypes from different geographic origins with 36 SSR primers to obtain a quantitative estimate of diversity in local cassava varieties; only two primers were monomorphic (SSRY-127 and SSRY-132) and they were excluded of the final analysis. Thirty-four polymorphic SSR primers that produced clear and scorable bands were analyzed.

The results confirmed the high polymorphism of SSR markers and the **figure 1** shows the observed polymorphism with the primer SSRY51 in the studied accessions from the five countries with the highest number of evaluated genotypes.

Data from the genetic distance analysis based on 34 polymorphic microsatellites were sufficient to distinguish all genotypes tested, and all images of the electrophoretic separations for those SSRs are available on the following web site: <http://isa.ciat.cgiar.org/molcas/estudios.jsp?code=7&pais=Cuba>. All those markers have shown adequate values in their polymorphism rates (PIC) in several previous diversity studies with other cassava genotypes (Mba et al., 2001; Hurtado et al., 2008; Lekha et al., 2011).

The number of SSR primer loci detected among all cassava landraces ranged from 2 to 10 alleles per locus with a 4.64 average, which is comparable with the values obtained by Raji et al. (2009) in a research on African cassava cultivars and elite germplasm from Asia and Latin America. However, Siqueira et al. (2009) only found an average of 3.3 alleles per locus in a similar genetic diversity study between Brazilian cassava landraces assessed with the same SSRs markers. Kawuki et al. (2009), working with



**Figure 1.** Polyacrylamide gel of PCR amplification of cassava accessions from Cuba, Guatemala, Colombia, Nigeria and Tanzania with the SSR51 microsatellite — *Gel de polyacrylamide des produits PCR obtenus au moyen du microsatellite SSR51 pour des accessions issues de Cuba, Guatemala, Colombie, Nigéria et Tanzanie.*

MM: molecular weight marker — *marqueur de poids moléculaire*; \*: the two controls have been used in more than one position — *les deux témoins ont été utilisés dans plusieurs positions.*

three different populations, from Asian, African and American cassava reported numbers of alleles ranging between 3 and 11.

Seventeen unique alleles with low appearance frequency (in all cases less than 25%) were found in Cuban accessions, in 12 of the 34 polymorphic markers evaluated: SSR54 (0.04), SSR520 (0.006), SSR538 (0.005), SSR559 (0.006 and 0.079), SSR563 (0.033), SSR569 (0.023), SSR5100 (0.011), SSR5103 (0.052, 0.012, 0.012 and 0.006), SSR5135 (0.005), SSR5151 (0.05), SSR5171 (0.012 and 0.036) and SSR5177 (0.014). Unique alleles were also found in genotypes from Colombia (6), Nigeria (3), Tanzania (3) and Guatemala (1).

The high polymorphism showed for this primers set also confirms the appropriateness of SSR markers for use in cassava genome analysis (Mba et al., 2001). The utility of SSR has been widely confirmed by recent studies on cassava diversity (Hurtado et al., 2008; Kunkeaw et al., 2010; Lekha et al., 2011).

The unique and broad diversity of cassava landraces found in Cuba reveals an invaluable germplasm resource for cassava breeding targeted to the region. The unique diversity suggests that the Cuban germplasm might have particular genetic information that will be more widely studied in the future.

### 3.1. Estimation of genetic diversity

In the genetic diversity indices calculated from the SSR data within and between countries, a high genetic polymorphism was observed for all microsatellites tested. On an average, there were more than 98% of polymorphic loci across all accessions and countries (see <http://isa.ciat.cgiar.org/molcas/estudios.jsp?code=7&pais=Cuba>). As the best represented countries in this research were Cuba, Colombia, Guatemala, Nigeria and Tanzania, hereinafter other studies were developed to deepen into their diversity and genetic differentiation.

Cuba and Guatemala genotypes showed 100% of polymorphic loci (**Table 3**); a high number of alleles was found per locus, with an average of  $4.64 \pm 0.0351$ ; Cuban landraces revealed the highest value (5.8 alleles per polymorphic locus), followed by cultivars from Colombia and Nigeria with 4.6 alleles per locus. Cassava accessions showed a mean of 4.72 alleles per polymorphic locus across the accessions.

Fregene et al. (2003) reported an average number of alleles of 6.0 for cassava landraces from Colombia and 5.2 for Brazilian cassava. Turyagyenda et al. (2012) found an average number of 5.923 in fifty-four cassava accessions from Uganda. A study conducted

**Table 3.** Genetic diversity within groups of cassava landraces classified according to the country of origin. Standard deviations (SD) were estimated by jackknifing over loci (200 replications).  $H_T$ ,  $H_S$ ,  $D_{ST}$  and  $G_{ST}$  are given over loci and over groups — *Diversité génétique intra groupes de variétés locales de manioc classées selon leur pays d'origine. Les déviations standard (SD) ont été estimées par jackknifing sur les loci (200 répétitions).  $H_T$ ,  $H_S$ ,  $D_{ST}$  et  $G_{ST}$  sont donnés pour les loci et les groupes.*

Group	Sample size	No. of loci	No. of pol. loci	Percent of pol. loci	Mean No. alleles/locus	Mean No. pol. alleles/locus	$H_O$	$H_E$	$H_{E,P}$
Cuba	86	34	34	100.0	5.8	5.8	0.6016	0.6314	0.6351
Guatemala	10	34	34	100.0	4.2	4.2	0.5556	0.6063	0.6385
Colombia	11	34	33	97.1	4.5	4.6	0.5675	0.6087	0.6396
Nigeria	16	34	33	97.1	4.5	4.6	0.5885	0.5949	0.6136
Tanzania	10	34	31	91.2	4.2	4.5	0.6459	0.5869	0.6190
Mean				97.06	4.64	4.72	0.5918	0.6057	0.6292
SD				3.60	0.65	0.61	0.0351	0.0169	0.0120
	$H_T$	$H_S$	$D_{ST}$	$G_{ST}$					
Mean	0.6538	0.6057	0.0482	0.0740					
SD	0.1770	0.1682	0.0253	0.0377					
95% CI	0.5780	0.5341	0.0383	0.0618					
95% CI	0.7137	0.6663	0.0585	0.0878					

$H_T$ : total heterozygosity (total genetic diversity) in the entire data set — *hétérozygote totale (diversité génétique totale) de l'échantillon complet*;  $H_S$ : heterozygosity within the country (average genetic diversity) averaged over entire data set — *hétérozygote totale intra pays (diversité génétique moyenne) de l'échantillon complet*;  $D_{ST}$ : average genetic diversity between countries — *diversité génétique moyenne entre populations*;  $G_{ST}$ : coefficient of genetic differentiation — *coefficient de différenciation génétique*; pol.: polymorphic — *polymorphique*;  $H_O$ : average observed heterozygosity within country — *hétérozygote moyenne observée intra pays*;  $H_E$ : average expected heterozygosity within country — *hétérozygote moyenne attendue intra pays*;  $H_{E,P}$ : average expected heterozygosity corrected for small simple sizes — *hétérozygote moyenne observée corrigée pour des petits échantillons* (Nei, 1978).

on cassava in Puerto Rico by Montero-Rojas et al. (2011) reported an average of 7.15 alleles per locus ranging between 2 and 14 alleles per locus. Therefore, the number of alleles among cassava accessions from Cuba is comparable to those previous results from Colombia, Brazil and Uganda.

The probability that two randomly selected alleles in a given accession are different, estimated by the average expected heterozygosity ( $H_E$ ), on average, could be considered high ( $0.6292 \pm 0.0120$ ), corrected for small simple sizes according to Nei (1978), with the highest  $H_E$  being found for Cuban cultivars (0.6314), followed by Colombian (0.6087) and Guatemalan (0.6063) accessions.

The average observed heterozygosity ( $H_O$ ) was also high ( $0.5918 \pm 0.0351$ ), which confirms cassava's outcrossing and highly heterozygous nature (Fregene et al., 2003). Cultivars from Tanzania and Cuba show the highest  $H_O$  indexes, 0.6459 and 0.6016, respectively.

These results reveal a comparable level of polymorphism in all groups from Africa (Nigeria and Tanzania) and Latin America (Colombia and Guatemala), and show a comparable polymorphism

level in Cuba with regard to the level found previously in other populations studied by Fregene et al. (2003). It is interesting that although the overall heterozygosity ( $H_T$ ) in all genotypes was high ( $0.6538 \pm 0.1770$ ), only 7.4% ( $G_{ST} = 0.0740 \pm 0.0377$ ) of this value was due to differentiation among cassava genotypes within countries, the rest is due to variations between countries ( $H_S$ ).

Several studies have indicated that the number of DNA markers used for genetic studies in plants varies with the total number of accessions assessed (Thottapilly et al., 1996). The reliability of estimates for genetic variation, such as  $H_E$ ,  $H_O$ ,  $F_{ST}$  and genetic distances, depends more on the number of loci than the number of individuals sampled (Baverstock et al., 1996). In cassava, Fregene et al. (2003) demonstrated that a minimum of 30 unlinked markers is required to obtain the maximum amount of information on allelic diversity; Hurtado et al. (2008) confirmed that 36 SSR are sufficient for cassava genetic studies.

The results indicate an important genetic variability in the populations studied, which represents a potential that can be useful for improving national



germplasm conservation strategies and especially for Cuba. This significant genetic diversity should be adequately preserved in germplasm collection and also usefully exploited in breeding programs. These results also point out the necessity to deepen into the knowledge of the genetic potentialities of the studied Cuban genotypes.

### 3.2. Genetic differentiation

For genetic differentiation, pairwise calculations of  $F_{ST}$  (theta) over all loci between pairs of country accessions groups provide a picture of germplasm exchange, based on their genetic similarity coefficient, between countries. All  $F_{ST}$  (theta) estimators were low to moderate (**Table 4**), superior to 0.04 as estimated by jackknifing and as calculated by bootstrapping, with a 99% confidence interval ( $F_{ST}$  values ranged from 0.0439 to 0.0605). The highest  $F_{ST}$  value was reached between Colombia and Tanzania (0.0605), followed by Guatemala and Colombia (0.0561) and between Nigeria and Tanzania (0.0560).

According to stadigraph-F calculated, the average value of  $F_{ST}$  ( $0.05 \pm 0.006$ ) is close to that obtained for  $G_{ST}$  ( $0.0427 \pm 0.0385$ ), estimated by FSTAT on all accessions from the five countries and for all loci, and indicated a differentiation from low to moderate within the five collections (countries) studied.

On the other hand, although the overall heterozygosity ( $H_T$ ) considering all the genotypes was high ( $0.6538 \pm 0.1770$ ), only 7.4% ( $G_{ST} = 0.0740 \pm 0.0377$ ) of this was due to differentiation among cassava accessions within the same country, the rest is due to variations between countries ( $H_S = 0.6057 \pm 0.1682$ ). Similar results were observed by Fregene et al. (2003) and Montero-Rojas et al. (2011) during two previous genetic diversity studies with cassava accessions from different African and Latin American countries.

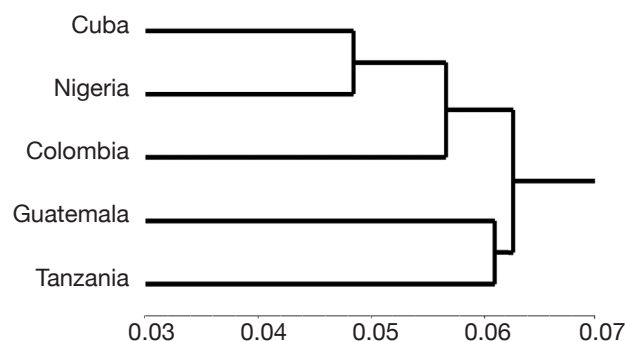
Estimates of  $F_{ST}$  over all loci between pairs of countries provide an idea on germplasm flow between them, which has been limited in Cuba due to its geographic isolation.

The pairwise  $F_{ST}$  (theta) estimates resulted in a dendrogram of the landraces that shows the relationships between African and Latin American genotypes (**Figure 2**). In general, a low differentiation among countries was observed or at least those accessions are as different as equals.

The close association between cultivars from Cuba and Nigeria results to be interesting if the geographical remoteness of countries and diversity indicators and differentiation observed are considered during this investigation. This result may be due to several reasons

**Table 4.** Genetic differentiations according to F-statistics estimator ( $F_{ST}$ ) in cassava accessions from Cuba, Guatemala, Colombia, Nigeria and Tanzania, estimated by jackknifing over country samples for all loci (200 replications) and as calculated by bootstrapping over loci (1,000 bootstraps) — *Les différenciations génétiques d'après l'estimateur F-statistique ( $F_{ST}$ ) ont été estimées par jackknifing (200 répétitions) sur tous les loci et par groupes de pays et comme calculées par bootstrapping sur loci (1000 bootstraps) pour des accessions du manioc de Cuba, Guatemala, Colombie, Nigéria et Tanzanie.*

	Cuba	Guatemala	Colombia	Nigeria	Tanzania
Cuba	0.0000	0.0527	0.0513	0.0439	0.0469
Guatemala	0.0527	0.0000	0.0561	0.0550	0.0533
Colombia	0.0513	0.0561	0.0000	0.0488	0.0605
Nigeria	0.0439	0.0550	0.0488	0.0000	0.0560
Tanzania	0.0469	0.0533	0.0605	0.0560	0.0000



**Figure 2.** Unweighted pair-group method with arithmetic averaging (UPGMA) dendrogram of the pairwise fixation index ( $F_{ST}$ ) between cassava landraces grouped per country — *Dendrogramme établi au moyen de l'algorithme UPGMA pour l'indice de fixation par couples ( $F_{ST}$ ) entre les variétés locales de manioc groupées par pays.*

and references found to explain them are limited. However, this is valuable information on the studied germplasm for cassava genetic improvement in Cuba.

Two previous studies could provide some elements although in both, other accessions not included in this study were studied. Cruz et al. (2011), during a diversity study using microsatellites markers, found a close relationship between the genotypes from Cuba and Colombia in the group of accessions that presented the highest heterozygosity (0,592). Also, Whankaew et al. (2012) during the study of 33 cassava accessions from 17 countries observed similar close relationships between groups of accessions from Cuba, Nigeria and Colombia.

During many years, common projects were developed between the cassava breeding programs from Cuba and CIAT in Colombia (S. Rodríguez, personal communication), from where every year

new improved lines were delivered to be evaluated by growers in Cuba. At the same time, similar convergence points exists between CIAT and other countries from Africa (H. Ceballos, personal communication); so there is a contact point between the cultivated cassava raised by growers in all these countries because of the germplasm exchange.

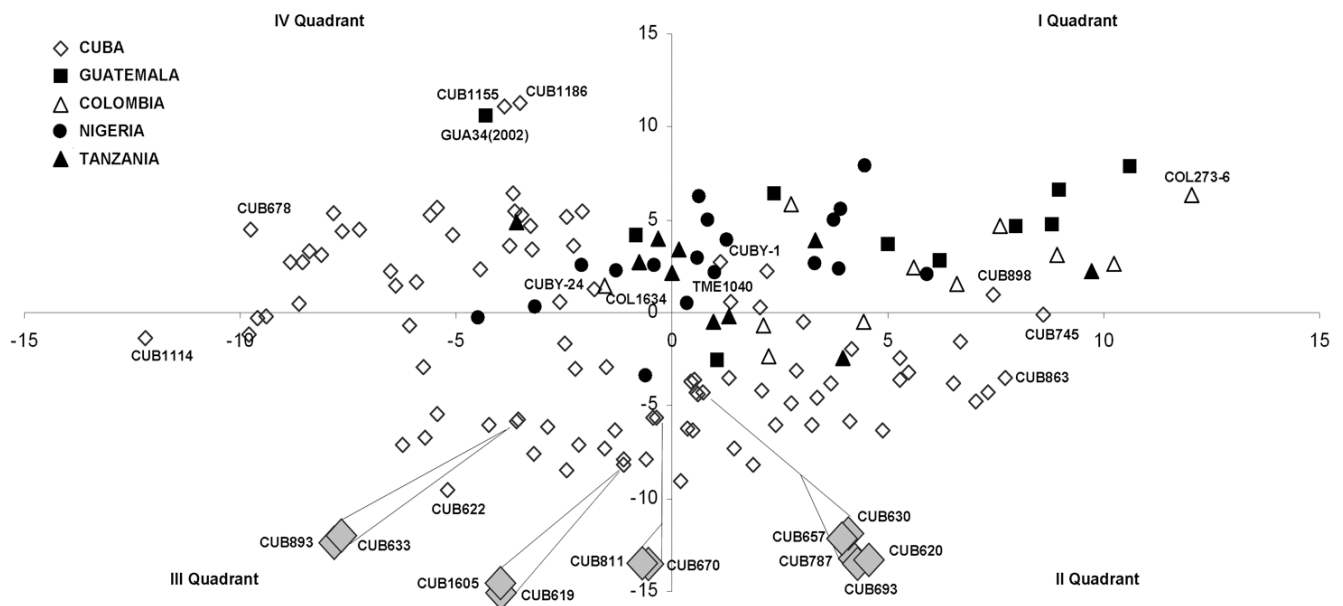
A graphical representation of the genetic relationships between all genotypes studied was obtained by a principal component analysis (PCA) performed on the genetic distance matrix. The accessions distribution based on genetic distances between all pairs of individual landraces and calculated by the 1-proportion of shared alleles (1-PSA) statistic (Bowcock et al., 1994) is showed in **figure 3**. The first two principal components explained 84.6% of the total variation. The genetic distance estimator 1-PSA has been found to be the most reliable for SSR data, assuming the infinite allele model (Kimura et al., 1964), and close relationships (Bertin et al., 2001).

Principal component analysis represents graphically the relationships among accessions from different geographic areas with the peculiarity that the largest number of accessions corresponds to Cuba. This analysis showed a wide dispersion of the genotypes, evidencing a rich genetic variability, especially those coming from Cuba. The presence of some Cuban genotypes in particular positions and outside of any possible cluster is noteworthy (see *e.g.* CUB1114,

CUB1155 and CUB1186). Most of them were collected from the east of the island and presumably share less genetic information with the other accessions studied. In general terms, this genetic structure is comparable with the one observed by Fregene et al. (2003) in which Cuban genotypes were inserted now. Well differentiated diversity groups were not observed indicating that it is very likely that dispersion is due to differential bands generated by SSR evaluated and shown in various combinations and frequencies.

The results showed that most of studied accessions share an important part of their genetic information. In case of Cuban germplasm, the accessions were collected many years ago in production fields in several regions of the country where edaphoclimatic conditions are different. The six Cuban accessions that are farthest away from PCA distribution ends (see CUB622, CUB678, CUB745, CUB1114, CUB1155 and CUB1186) were collected in savanna soils; three were from the east region of the country (Holguín) and two from the west (Matanzas). In the center of the figure, two promising cultivars from Cuba with two representing Colombia (CIAT international collection) and Nigeria (IITA collection) (see CUBY-24 and TME1040; CUBY-24 and COL1634) are located. The four accessions are the result of national breeding schemes but CIAT has been a common point of origin.

The principal components analysis based on genetic distances showed several Cuban accessions closely



**Figure 3.** Distribution of cassava accessions evaluated from Cuba, Colombia, Guatemala, Nigeria and Tanzania according to the principal components analysis (PCA) based on the genetic distances between all pairs of individual accessions and calculated by 1-PSA (1-proportion of shared alleles) with the use of 34 SSR markers — *Distribution d'accessions de manioc de Cuba, Colombie, Guatemala, Nigéria et Tanzanie obtenue par analyse en composantes principales (PCA) basée sur les distances génétiques entre toutes les paires d'accessions individuelles et calculée par 1-PSA (1-proportion d'allèles partagés) en employant 34 marqueurs SSR.*

related, for example, in the second quadrant: CUB620, CUB630, CUB657, CUB693 and CUB787. The first four were collected in the municipality of Santo Domingo (Villa Clara) in savanna conditions from the central region of Cuba, while CUB787 came from Pinar del Río province in the west of the island. On this point, it is important to carry out a depth study in the future on morphological and agronomic characteristics of each.

The lack of previous researches concerning such relation in accessions from the Cuban cassava germplasm, limits the scientific arguments that can explain this result. However, this suggests that they might possess important distinctive features from the genetic point of view, and therefore, those genotypes could have special characteristics for cassava genetic breeding.

According to experiences of similar researches in other countries, these results may be related to the role of seedlings, human and natural selection, the heterozygous nature of cassava (Ojulong et al., 2010) and also with the influence of geographic isolation imposed by an insular environment during thousands of years.

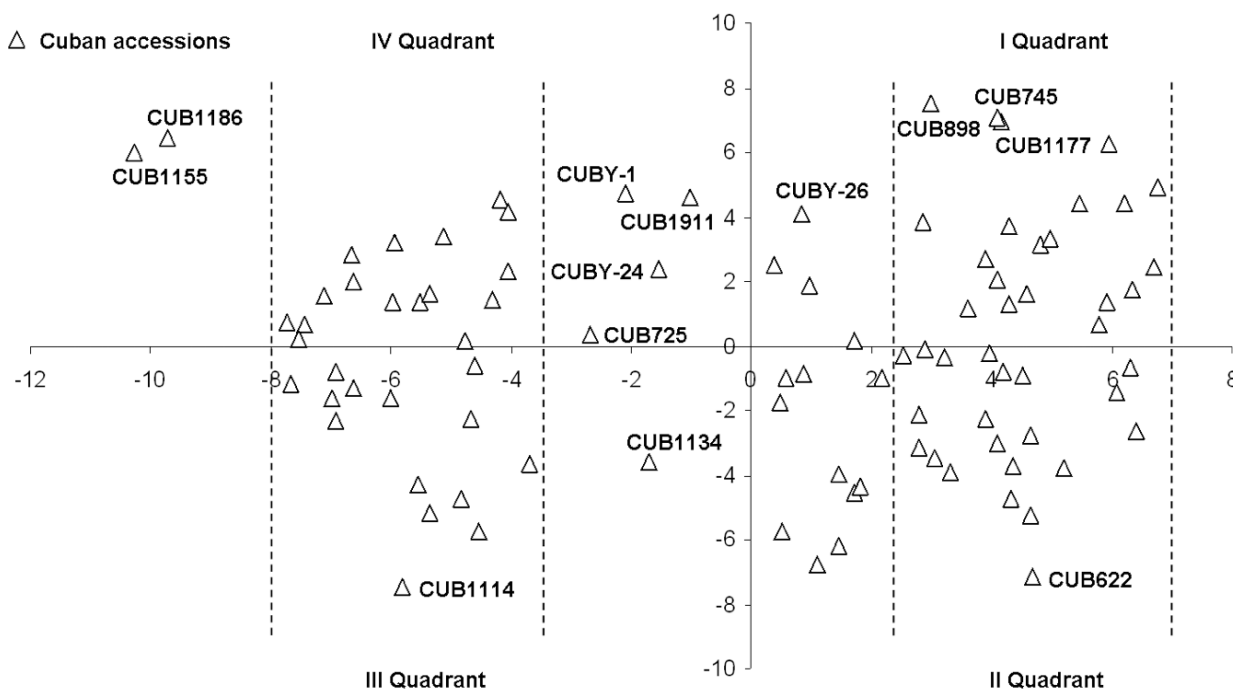
Ojulong et al. (2010) consider that agricultural practices of slash and burn by Amerindian farmers and the allogamous nature of cassava have been

demonstrated to produce a large pool of volunteer seedlings that may be a valuable source to produce new varieties through natural and human selection. The production of new varieties not only maintains a high level of genetic diversity but also serves as insurance against crop failure due to biotic and abiotic stresses, once the adaptation to the new environmental conditions of each country or region takes place in new genotypes.

Taking into consideration the behavior of genotypes from Cuba in comparison with the rest of African and Latin American countries studied, a principal components analysis (PCA) to assess genetic variability within all Cuban cassava accessions was conducted (**Figure 4**). The first two principal components account 86.2% of the total variation.

A wide variability among Cuban accessions could be noticed but no type of grouping is clearly appreciated between them. However, a trend toward a higher concentration of accessions into two main sectors was noted.

Interestingly, almost all accessions that were located in distant positions now and before the study with five countries (**Figure 3**) were collected in the east of Cuba. The eastern region shows contrasting climatic conditions such as: high temperatures, savanna and high mountains, very rainy climate or semi-desert areas.



**Figure 4.** Distribution of cassava accessions evaluated from Cuba according to the principal components analysis (PCA) based on the genetic distances between all pairs of individual accessions and calculated by 1-PSA (1-proportion of shared alleles) with the use of 34 SSR markers — *Distribution des accessions de manioc de Cuba d'après l'analyse en composantes principales (PCA) basée sur les distances génétiques entre toutes les paires d'accessions individuelles et calculée par 1-PSA (1-proportion d'allèles partagés) en employant 34 marqueurs SSR.*

From a historical-cultural perspective, that area has been considered the territory where the first inhabitants settled in Cuba and probably the cultivated cassava in the area came, like them, from South America (Valdés et al., 1991).

Therefore, interesting findings have been obtained in this study. Cultivars from the eastern region of the country may have distinctive features or peculiarities to differentiate them from the rest of the germplasm conserved in Cuba. Even these characteristics may be different from those found in continents. For such reasons, further researches on genetic diversity are being developed (in progress).

Cuba is a long and narrow island where cassava has been cultivated for thousands of years (Valdés et al., 1991; Vásquez, 2010), because of this, probably the geographic isolation and limited germplasm exchange had a direct impact on the occurrence of a peculiar differentiation process in accessions coming from the east of the island.

This finding is important and perhaps, a hypothesis could be stated considering that Cuban genotypes after many years of cultivation under insular conditions could set spontaneous deleterious mutations as a result of genetic drift in time. Fregene et al. (2003) consider that in these conditions also act the traditional cassava propagation system (vegetative or asexual), the natural and artificial selection and the allogamous nature of the crop. Likewise, cassava diversity is influenced significantly by growers during planting material exchange and in the selection and extension process of the best cultivars (Montero-Rojas et al., 2011). This is a hypothesis that, in the absence of previous studies that can be used as reference, rests on the existence of several unique and rare alleles in these accessions and suggests the convenience of further studies aimed at evaluating the specific response to adverse factors or expression of traits of interest.

At the same time, this study confirms one important thing: the Cuban germplasm studied is an important reservoir of genetic variation that will be basic for future food security actions. There is evidence that, the level of genetic diversity is fundamental for world food security and its maintenance is considered crucial in preventing the loss of evolutionary potential and ensuring the survival of given species because of its association with fitness traits and increased biological success. Notably, the estimation of genetic variation in the Cuban germplasm was performed on traditional cultivars, which according to Vimala et al. (2011), are usually characterized by their dynamics in crop fields and are genetically more heterogeneous, and may also provide an excellent source of genes for adaptive traits and pest resistance.

Recent morphological and agronomic studies on a significant group of these Cuban accessions also

showed the existence of wide genetic variability in the native germplasm, as well as their potentialities for breeding and human and animal consumption and for their high yields and high dry matter percentage in roots (Beovides et al., 2013; Beovides et al., 2014).

In general, molecular analysis applied to genotypes from the Cuban cassava germplasm, support the usefulness of this type of studies to better understand the genetic variability in cassava producing countries and in the case of Cuba, allowed to estimate parameters of genetic variation and to evaluate possible genetic relations of Cuban accessions with some other genetic relatives in America and Africa. All this contributes to better use and conservation of diversity, and generates wide useful information for possible use in better strategies for crop genetic breeding.

#### 4. CONCLUSIONS

Thirty-four polymorphic microsatellite markers used were useful to differentiate Cuban cassava accessions, and allowed to estimate the genetic relation between them and the evaluated genotypes from other geographical origins.

The use of studied microsatellite markers on Cuban cassava accessions showed a high genetic differentiation between them and although their diversity was high, the relation with cultivated genotypes from Guatemala, Colombia, Nigeria and Tanzania was evident.

#### Acknowledgements

We would like to thank all of the technical staff from INIVIT and CIAT due to their contribution, interest and support shown to this project. We are very grateful to the Cassava Biotechnology Network (CBN) due to the financial support for conducting researches.

#### Bibliography

- Bang T., Raji A. & Ingelbrecht I., 2011. A multiplex microsatellite marker kit for diversity assessment of large cassava (*Manihot esculenta* Crantz) germplasm collections. *Plant Mol. Biol. Rep.*, **29**(3), 655-662.
- Baverstock P.R. & Moritz C., 1996. Project design. In: Hillis D.M., Moritz C. & Mable B.K., eds. *Molecular systematics*. Sunderland, MA, USA: Sinauer Associates, 17-27.
- Beovides Y. et al., 2013. Cultivares cubanos de yuca (*Manihot esculenta* Crantz) con rendimiento y potencial genético para la agroindustria. *Centro Agrícola*, **40**(3), 71-78.
- Beovides Y. et al., 2014. Caracterización morfológica y agronómica de cultivares cubanos de yuca (*Manihot esculenta* Crantz). *Cultivos Tropicales*, **35**(2), 43-50.

- Bertin P., Grégoire D., Massart S. & De Froimont D., 2001. Genetic diversity among European cultivated spelt revealed by microsatellites. *Theor. Appl. Genet.*, **102**, 148-156.
- Bowcock A.M. et al., 1994. High resolution of human evolution with polymorphic microsatellites. *Nature*, **368**, 455-457.
- Cruz A.C. et al., 2011. Diversidad genética y contenido de carotenos totales en accesiones del germoplasma de yuca (*Manihot esculenta* Crantz). *Acta Agron.*, **60**(2), 97-107.
- Dellaporta S.L., Wood J. & Hicks J.R., 1983. A plant DNA minipreparation: version II. *Plant Mol. Biol. Rep.*, **1**, 19-21.
- Dice I.R., 1945. Measures of the amount of ecologic association between species. *Ecology*, **26**, 297-302.
- Ferguson M.E. et al., 2012. Molecular markers and their application to cassava breeding: past, present and future. *Trop. Plant Biol.*, **5**, 95-109.
- Fregene M. et al., 1997. A molecular genetic map of cassava (*Manihot esculenta* Crantz). *Theor. Appl. Genet.*, **95**, 431-441.
- Fregene M., Gutiérrez J.P. & Buitrago C., 2002. *Protocols of cassava molecular genetic*. CIAT Work Document. Cali, Colombia: CIAT.
- Fregene M. et al., 2003. Simple Sequence Repeat marker diversity in cassava landraces: genetic diversity and differentiation in an asexually propagated crop. *Theor. Appl. Genet.*, **107**(7), 1083-1093.
- Goudet J., 2001. FSTAT, Version 1.2: A computer program to calculate F-statistics. *Heredity*, **86**, 485-486.
- Hurtado P. et al., 2008. Comparison of simple sequence repeats (SSR) and diversity array technology (DArT) markers for assessing genetic diversity in cassava. *Plant Genet. Resour.*, **6**(3), 208-214.
- Kawuki R.S. et al., 2009. Identification, characterization and application of single nucleotide polymorphisms for diversity assessment in cassava (*Manihot esculenta* Crantz). *Mol. Breed.*, **23**(4), 669-684.
- Kimura M. & Crow J.T., 1964. The number of alleles that can be maintained in a finite population. *Genetics*, **49**, 725-738.
- Kunkeaw S. et al., 2010. Genetic linkage map of cassava (*Manihot esculenta* Crantz) based on AFLP and SSR markers. *Plant Breed.*, **129**(1), 112-115.
- Kunkeaw S. et al., 2011. Construction of a genetic linkage map using simple sequence repeat markers from expressed sequence tags for cassava (*Manihot esculenta* Crantz). *Mol. Breed.*, **27**(1), 67-75.
- Lekha S.S., Teixeira da Silva J.A. & Pillai S.V., 2011. Genetic variability studies between released varieties of cassava and central Kerala cassava collections using SSR markers. *J. Stored Prod. Postharvest Res.*, **2**(4), 79-92.
- Mba R.E.C. et al., 2001. Simple Sequence Repeats (SSR) markers survey of the cassava (*Manihot esculenta* Crantz) genome: towards an SSR-based molecular genetic map of cassava. *Theor. Appl. Genet.*, **102**, 21-31.
- Milián M.D. et al., 2000. Caracterización, evaluación y conservación de la colección cubana de germoplasma de yuca (*Manihot esculenta* Crantz). In: Carvalho L.J.C.B., Thro A.M. & Vilarinhos A.D., eds. *Proceeding of the 4<sup>th</sup> international scientific meeting cassava biotechnology network*. Brasilia: Cassava Biotechnology Network.
- MINAG, 2008. *Instructivo Técnico sobre el Cultivo de la Yuca*. La Habana: SEDGRI/AGRINFOR.
- Minch E., 1996. *MICROSAT. Version 1.4*. Stanford, CA, USA: Stanford University Medical Centre.
- Montero-Rojas M., Correa A.M. & Siritunga D., 2011. Molecular differentiation and diversity of cassava (*Manihot esculenta*) taken from 162 locations across Puerto Rico and assessed with microsatellite markers. *AoB PLANTS*, doi:10.1093/aobpla/plr010, (23/10/12).
- Nei M., 1978. Estimation of average heterozygosity and genetics distances from a small number of individuals. *Genetics*, **89**, 583-590.
- Ojulong H., Labuschagne M.T., Herselman L. & Fregene M.A., 2010. Yield traits as selection indices in seedling populations of cassava. *Crop Breed. Appl. Biotechnol.*, **10**(3), 191-196.
- Quenoille M.H., 1956. Notes on bias in estimation. *Biometrika*, **43**, 353-360.
- Raghu D. et al., 2007. Morphological and simple sequence repeats (SSR) based finger printing of South Indian cassava germplasm. *Int. J. Integr. Biol.*, **2**, 141-149.
- Raji A.A. et al., 2009. Gene-based microsatellites for cassava (*Manihot esculenta* Crantz): prevalence, polymorphisms, and cross-taxa utility. *BMC Plant Biol.*, **9**, 118.
- Rohlf F.J., 1998. *NTSYS-PC numerical taxonomy and multivariate analysis system. Version 2.02*. New York, NY, USA: Applied Biostatistics Inc.
- Roubus K., Moustakas M. & Aravanopoulos F.A., 2010. Molecular identification of Greek olive (*Olea europaea*) cultivars based on microsatellite loci. *Genet. Mol. Res.*, **9**, 1865-1876.
- SAS Institute Inc., 1999. *Statistic Analysis System. Version 8.0*. Cary, NC, USA: SAS Institute Inc.
- Siqueira M.V.B.M. et al., 2009. Genetic characterization of cassava (*Manihot esculenta*) landraces in Brazil assessed with simple sequence repeats. *Genet. Mol. Biol.*, **32**(1), 104-110.
- Thottapilly G., Crouch J.H. & Quin F.M., 1996. Overview of DNA marker research at IITA. DNA marker-assisted improvement of the staple crops of the sub-Saharan Africa. In: Jonathan H.C. & Abdou T., eds. *Proceedings of the Workshop on DNA markers*. Ibadan, Nigeria: IITA.
- Turyagyenda L.F. et al., 2012. Genetic diversity among farmer-preferred cassava landraces in Uganda. *Afr. Crop Sci. J.*, **20**(S1), 15-30.
- Valdés M.M., Albelo R.M. & Gallo G., 1991. La comunidad primitiva en Cuba. In: Ramos M.G. & Galán D.J.,

- eds. *Historia de Cuba*. La Habana: Editorial Pueblo y Educación, 4-14.
- Vázquez M., 2010. *Alimentos esenciales en la cocina cubana: boniato, yuca y maíz*, <http://www.cubasolar.cu/biblioteca/energia/Energia43/HTML/Articulo11.htm>, (21/04/2011).
- Vásquez A. & López C., 2012. Identificación de polimorfismos en genes candidatos de resistencia en yuca (*Manihot esculenta* Crantz). *Acta Agron.*, **61**(2), 133-142.
- Vekemans X. & Lefevre C., 1997. On the evolution of heavy metal tolerant populations in *Armeria maritima*: evidence from allozyme variation and reproductive barriers. *J. Evol. Biol.*, **10**, 175-199.
- Vieira E.A. et al., 2011. Caracterização molecular de acessos de mandioca açucarados e não açucarados. *Ciênc Agrotec*, **35**(3), 455-461.
- Vimala B., Thushara R., Nambisan B. & Sreekumar J., 2011. Effect of processing on the retention of carotenoids in yellow-fleshed cassava (*Manihot esculenta* Crantz) roots. *Int. J. Food Sci. Technol.*, **46**, 166-169.
- Weir B.S. & Cockerham C.C., 1984. Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358-1370.
- Whankaew S. et al., 2012. Characterization of microsatellite markers in cassava based on microsatellite-AFLP technique. *Genet. Mol. Res.*, **11**(2), 1319-1326.
- Wright S., 1951. The genetical structure of populations. *Ann. Eugenics*, **15**, 323-354.
- (45 ref.)