American Journal of Botany 87(11): 1647–1655. 2000.

Cross-species amplification of cassava (*Manihot esculenta*) (Euphorbiaceae) microsatellites: allelic polymorphism and degree of relationship[1]

Ana C. Roa,[2] Paul Chavarriaga-Aguirre, Myriam C. Duque, María M. Maya,[3] Merideth W. Bonierbale,[4] Carlos Iglesias,[5] and Joe Tohme[6]

Cassava Program and Biotechnology Research Unit, Centro Internacional de Agricultura Tropical (CIAT), A.A. 6713, Cali, Colombia

Microsatellite amplification was performed on cassava (*Manihot esculenta*) and six other different species (all wild) of the *Manihot* genus. We used ten pairs of microsatellite primers previously developed from cassava, detecting 124 alleles in a sample of 121 accessions of the seven species. The number of alleles per locus ranged from four to 21 alleles, and allelic diversity was greater in the wild species than in cassava. Seventy-nine alleles, including unique ones, were detected in the wild species but were not found in the crop. The lower level of heterozygosity in some wild species probably resulted from a combination of fine-scale differentiation within the species and the presence of null alleles. Overall, microsatellite primers worked across the genus, but, with increasing genetic distance, success in amplifying loci tended to decrease. No accession of *M. aesculifolia, M. carthaginensis,* and *M. brachyloba* presented a banding pattern at locus *Ga*-140; neither did one appear for *M. aesculifolia* at locus *Ga*-13. Previous work with amplified fragment length polymorphism (AFLP) markers and this microsatellite analysis show that these three wild taxa are the most distant relatives of the crop, whereas the wild forms *M. esculenta* subsp. *flabellifolia* and *M. esculenta* subsp. *peruviana* appear to be the closest.

Key words: cassava; cross-amplification; Euphorbiaceae genetic diversity; heterozygote deficiency; *Manihot* species; microsatellite loci.

Cassava (*Manihot esculenta* Crantz) is a root crop produced by small farmers in tropical Africa, Asia, and Latin America. A staple food, it occupies fourth place after rice, sugar cane, and maize as a source of calories in the human diet (Cock, 1985; Best and Henry, 1994). It is also used by small-scale industries. Based on different kinds of evidence, the crop may have been first cultivated in one of two possible places: southern Mexico and Guatemala (Rogers, 1963; Renvoize, 1972) or the Amazon Basin in Brazil (Nassar, 1978; Allem, 1987, 1994).

Several wild species have been proposed as being the closest related taxa to the crop: *M. aesculifolia,* widely distributed in Central America (Rogers and Appan, 1973; Bertram, 1993); *M. carthaginensis,* which originated in the Caribbean Coastal regions of Colombia and Venezuela (Bertram, 1993); *M. esculenta* subsp. *flabellifolia* and *M. esculenta* subsp. *peruviana* (Allem, 1994); and *M. tristis* (Allem, 1987).

Molecular markers such as restriction fragment length polymorphisms (RFLPs) (Bertram, 1993; Fregene et al., 1994) and amplified fragment length polymorphisms (AFLPs) (Roa et al., 1997) have been used to search for cassava's closest relatives. Chloroplast and ribosomal DNA RFLPs analyses suggested two pairs of probable ancestors of cassava: *M. aesculifolia* and *M. carthaginensis* (Bertram, 1993), and *M. tristis* and *M. esculenta* subsp. *flabellifolia* (Fregene et al., 1994). Recent analysis carried out on seven taxa with AFLPs showed that *M. esculenta* subsp. *flabellifolia* and *M. esculenta* subsp. *peruviana* are the closest relatives to cassava (Roa et al., 1997).

Microsatellites or simple sequence repeats (SSRs) are tandem repetitive DNA sequences of two to five nucleotides long (Akagi et al., 1997). Accumulated evidence shows that they are widely dispersed in all eukaryotic genomes (Tautz, Trick, and Dover, 1986; Morgante and Olivieri, 1993; Dow, Ashley, and Howe, 1995; Byrne et al., 1996), for example, the dinucleotide repeats (AT/TA)n and (GA/CT)n are the most common found in higher plants (Morgante and Olivieri, 1993; Dow, Ashley, and Howe, 1995; Akagi et al., 1997; Steinkellner et al., 1997). To amplify microsatellites, specific primers are designed on the flanking regions of the SSRs. These markers are abundant, codominant, highly polymorphic, even within populations, spread throughout the genome, easily amplified by polymerase chain reaction (PCR), and the great majority are probably selectively neutral (Ashley and Dow, 1994). They offer great potential for studies on parentage (Dow and Ashley, 1996), gene flow within and between populations (Innan, Terauchi, and Miyashita, 1997), mapping and breeding (McCouch et al., 1997), and extent and maintenance of genetic diversity

[1] Manuscript received 25 August 2000; revision accepted 10 February 2000.

The authors thank Dr. Morry Levy, Purdue University, for his statistical advice on microsatellite analysis, Dr. Gary Kochert, University of Georgia, for generously providing the cassava microsatellite primers, and Elizabeth de Paéz for editorial assistance. This research was supported in part by a grant from the Instituto Colombiano para el Desarrollo de la Ciencia y la Tecnología ''Francisco José de Caldas'' (COLCIENCIAS).

[2] Current address: Center for the Application of Molecular Biology to International Agriculture-CAMBIA GPO Box 3200, Canberra, ACT 2601, Australia.

[3] Current address: Purdue University, Department of Biological Sciences, Lilly Hall, Room G-420A, West Lafayette, Indiana 47907-1392 USA.

[4] Current address: Centro Internacional de la Papa-CIP, Apartado 1558, Lima 12, Peru.

[5] Current address: Weaver Popcorn NC, 1000 North 325 W, PO Box 207, New Richmond, Indiana 47967 USA.

[6] Auther for reprint requests (Tel.: ++57-2-4450000 ext. 3265; fax: ++57-2-4450073; e-mail: j.tohme@cgiar.org).

(Byrne et al., 1996; Witsenboer, Vogel, and Michelmore, 1997).

The widespread use of microsatellites has been limited by the fact that PCR primers require a high degree of homology to work, implying that novel species-specific markers would have to be isolated when starting the analysis of a new species (Steinkellner et al., 1997). Success in the cross-species amplification of any DNA sequence is inversely related to the evolutionary distance between the two species (Steinkellner et al., 1997). Hence, research on species relationships has increasingly focused on assessing the ability of SSR primers to amplify the same loci across different species and genera (Byrne et al., 1996; Katzir et al., 1996; Isagi and Suhandono, 1997; Smulders et al., 1997; Steinkellner et al., 1997; Witsenboer et al., 1997).

For cassava, 14 different primer sequences were designed to amplify SSRs containing mostly perfect or imperfect GA repeats. The primers were tested on 522 accessions of the cultivated cassava core collection conserved at CIAT and showed heterozygosity values between 0.00 and 0.88, with as many as 15 different alleles at one locus (Chavarriaga-Aguirre et al., 1998).

To evaluate the ability of these primers to amplify SSR loci in congeneric species, a set of six wild *Manihot* species and a diverse sample of cassava were assessed with ten pairs of primers. The sequence conservation of the flanking primer regions is discussed, together with the utility of using a codominant marker, such as microsatellites, to assess the genetic diversity and degree of relationship between cassava and its wild relatives.

MATERIALS AND METHODS

Plant material—A sample of 121 accessions was selected from seven species. For cassava (referred as ESC in tables and figures), 38 accessions were selected from the cassava core collection held in trust at CIAT. The core collection was previously set according to geographic, morphological, and biochemical parameters as described by Hershey et al. (1994). The remaining 83 accessions were chosen, based on previous analysis with AFLPs (Roa et al., 1997), to represent the wild taxa *M. aesculifolia* (AES), *M. brachyloba* (BLO), *M. carthaginensis* (CTH), *M. esculenta* subsp. *flabellifolia* (ESC-FLA), *M. esculenta* subsp. *peruviana* (ESC-PER), and *M. tristis* (TST) (Table 1).

DNA extraction—DNA samples were prepared by first drying young leaves at 48°C for 20 h, then grinding them, using a mortar and pestle, until a fine powder was obtained. The DNA was then extracted in eppendorf tubes, according to the following protocol (C. Colombo, 1995, personal communication, Montpellier, France): 0.1 g of ground tissue was resuspended in 1.5 mL of hexadecyltrimethyl-ammonium bromide (CTAB) buffer (0.1 mol/L Tris-HCl pH 8.0; 1.25 mol/L NaCl; 0.02 mol/L ethylenediaminetetra acetic acid (EDTA); 4% CTAB, and 0.1% 2-mercaptoethanol) and incubated at 65°C for 1 h 30 min. After two chloroform-isoamylic extractions, the aqueous phase was collected and the nucleic acids precipitated with cold isopropanol. The pellet collected by centrifugation was washed with 75% ethanol and dried at room temperature. The nucleic acids were dissolved in $T_{10}E_1$ (10 mmol/L Tris-HCl (pH 8.0), 1 mmol/L EDTA) and treated with RNAse A (10 μg/mL). Electrophoresis and fluorometry determined DNA quality and concentration.

Primers and PCR conditions—DNA from each sample was amplified, using phosphoramidite-labeled primers (Perkin Elmer/Applied Biosystems, Foster City, California, USA, or Research Genetics, Huntsville, Alabama, USA). Multiplex PCR reactions were set according to the method developed by Chavarriaga-Aguirre et al. (1998). Primer pairs were combined as follows: one quadruple with primers for the loci *Ga*-21, *Ga*-126, *Ga*-134, and *Ga*-136; and three duplex for the loci *Ga*-16/*Ga*-140; *Ga*-13/*Ga*gg-5, and *Ga*-12/*Ga*-131. The optimal annealing temperature was the same for all primer sets. PCR reactions and amplification profiles were carried out according to Chavarriaga-Aguirre et al. (1998), except that no polymerase activation cycle at 94°C for 10 min was carried out before amplification. In this study we used Ampli*Taq* polymerase (Perkin Elmer/Applied Biosystems), while Chavarriaga et al. (1998) used Ampli*Taq*-Gold, which required an activation cycle as mentioned above.

Electrophoresis conditions and allele sizing—A volume of 1.5 μL of the PCR product, combined with 0.5 μL of the internal size standard (GeneScan TAMRA 500, Perkin Elmer/Applied Biosystems) and 2 μL of deionized formamide, mixed with loading buffer (5:1), was denatured at 98°C for 2 min. This mixture was loaded on 4% polyacrylamide denaturing gels, containing 6 mol/L urea. Electrophoresis was carried out with 1× Tris-borate-EDTA (TBA) buffer on an ABI Prism 377 automatic DNA sequencer (Perkin Elmer/Applied Biosystems) using the Gene Scan Analysis Software version 2.0.2. The GeneScan module used to run the gels was GS 36C–2400 [36 cm well-to-read (WTR), virtual filter C, 3000 V/2400 scans/h in 2.5 h].

Allele sizing was performed, using the third-order least-square algorithm, which relies on regression analysis to build a best-fit size-calling curve. It compensates for any fragments that may run anomalously (Perkin-Elmer, 1995). Allele sizes were determined for all 121 individuals and compared, using Genotyper Software version 1.1 (Perkin-Elmer, 1995).

Data analysis—Presence and absence of alleles were registered for all 121 individuals. Only the strongest bands were considered alleles because lighter bands may have been stutter bands that resulted from slippage of the Taq polymerase during PCR (Lagercrantz, Ellengran, and Anderson, 1993; Wu and Tanksley, 1993). The number of alleles and percentage of polymorphism were calculated for each locus and species.

Levels of differentiation among taxa and degrees of nonrandom association of alleles within species, allele frequencies, heterozygosity, and F_{is} values were estimated for each locus–species combination, using the PopGene software (version 1.2; Yeh, Rongcai, and Boyle, 1997). F_{is}, called the fixation index or inbreeding coefficient, is the correlation between the number of homologous alleles within individuals, with reference to the local population. It is also a measure of heterozygote deficiency or excess (Avise, 1994). All individuals were assumed to be diploid. Therefore, when only one allele was observed at a locus, the individual was considered to be a homozygote.

At microsatellite loci, species heterozygote deficiencies were observed and may be explained by the presence of null alleles (Brookfield, 1996). To estimate how the frequency of null alleles (r_b) affects the observed heterozygosity (H_{obs}), r_b was calculated according to Brookfield (1996),

$$r_b = H_{exp} - H_{obs}/1 + H_{exp} \quad (1)$$

where H_{exp} is the expected heterozygosity.

A similarity matrix, calculated by using the Nei–Li index (Nei and Li, 1979), was used to construct a dendrogram to show the degree of genetic similarity within and among taxa. The dendrogram was constructed by employing the option TREE and the unweighted pair group method (UPGMA) with the NTSYS program (version 1.8; Rohlf, 1994). The dendrograms and similarity indexes obtained with AFLPs (Roa et al., 1997) were compared with those generated by the SSRs. Cophenetic values for each dendrogram and the assembly of a cophenetic matrix for each marker type were calculated, using the MXCOMP option of the NTSYS program. The Mantel matrix test was used to compare cophenetic matrices (Sokal and Rohlf, 1995).

RESULTS

Cross-species amplification—Crucial to the use of SSRs for assessing the genetic variation within and among species is the ability of primers from one species to amplify homologous loci in related species (Kijas, Fowler, and Thomas, 1995). At eight loci, primers amplified the target sequence across all species tested (Table 2), but in the wild species *M. aesculifolia,*

TABLE 1. Germplasm of *Manihot* species selected for microsatellite analysis.

Wild species	Accession code[a]	No. of indiv.	Country of origin	*M. esculenta* accession code[b]	Country of origin	Common name
M. aesculifolia	AES 002	1	Mexico	HCM 1	Colombia	-[c]
	AES 404-004	1	Mexico	M ARG 11	Argentina	Duro do Valle 30
	WU 26 AES 2	1	Mexico	M BOL 3	Bolivia	Rosada de Bolivia
	WU 32 AES 4	1	Mexico	M BRA 12	Brazil	-[c]
	WU 27 AES 5	1	Mexico	M BRA 97	Brazil	Saracura II
	WU 25 AES 6	1	Mexico	M BRA 110	Brazil	Pangola
	WU 24 AES 7	1	Mexico	M BRA 383	Brazil	Vassourão
	WU 82 AES 8	1	Mexico	M BRA 881	Brazil	Branca de Sta. Catarina
M. brachyloba	BLO 001	1	-[c]	M BRA 885	Brazil	Jabarú
	BLO 401 A	1	Colombia	M BRA 900	Brazil	Mandim Branca
	BLO 402	7	Colombia	M BRA 931	Brazil	Enganha Ladrão
M. carthaginensis	CTH 005	1	Colombia	M COL 1468	Brazil	Sip24-2 Mantiqueira
	CTH 012	1	Colombia	M COL 22	Colombia	Uvita
	CTH 106	1	Colombia	M COL 1438	Colombia	Llanera
	CTH 121	1	Colombia	M COL 1522	Colombia	Algodonera Amarilla
	CTH 246	1	Colombia	M COL 1684	Colombia	-[c]
	CTH Z	1	-[c]	M COL 2061	Colombia	Regional Morada
	CTH 59	1	-[c]	M COL 2066	Colombia	Chiroza Gallinaza
	CTH 409	2	Colombia	M COL 2215	Colombia	Venezolana 1
	CTH 411	1	Colombia	M CR 32	Costa Rica	Yuca Mangi
	CTH 413	5	Colombia	M CUB 51	Cuba	Pinera
	CTH 414	2	Colombia	M CUB 74	Cuba	Señorita
	CTH 415	4	Colombia	M ECU 41	Ecuador	De Tres Meses
	CTH 416	1	Colombia	M ECU 82	Ecuador	Blanca
	CTH 417	4	Colombia	M IND 33	Indonesia	No. 734-5
M. esculenta subsp. *flabellifolia*	ESC-FLA 423	2	Brazil	M MAL 2	Malaysia	Black Twig
	ESC-FLA 427	2	Brazil	M MAL 48	Malaysia	Red Twig
	ESC-FLA 428	2	Brazil	M MEX 59	Mexico	-[c]
	ESC-FLA 430	2	Brazil	M PAN 51	Panama	-[c]
	ESC-FLA 438	2	Brazil	M PAR 110	Paraguay	Tacuara Sayyu
	ESC-FLA 439	2	Brazil	M PTR 19	Puerto Rico	No. 9588
	ESC-FLA 443	2	Brazil	M TAL 1	Thailand	Rayong 1
	ESC-FLA 444	2	Brazil	M VEN 25	Venezuela	Querepa Amarga
M. esculenta subsp. *peruviana*	ESC-PER 406	2	Brazil	M VEN 45 A	Venezuela	-[c]
	ESC-PER 407	2	Brazil	M COL 1505	Venezuela	-[c]
	ESC-PER 411	2	Brazil	CM 2177-2	CIAT[d]	Hybrid
	ESC-PER 412	2	Brazil	CM 3306-9	CIAT[d]	Hybrid
	ESC-PER 413	2	Brazil	M NGA-2	IITA[e]	Hybrid
	ESC-PER 414	2	Brazil			
	ESC-PER 415	2	Brazil			
	ESC-PER 417	2	Brazil			
M. tristis	TST 001	1	Brazil			
	TST 002	1	Brazil			
	TST 007	1	Brazil			
	TST 008	1	Brazil			
	TST 009	1	Brazil			
	TST 012	1	Brazil			
	TST 018	1	Brazil			
	TST 026	1	Brazil			

[a] *Manihot* species were collected as true-seed populations.
[b] Cassava accessions are maintained as vegetative clones.
[c] No passport data were available.
[d] Centro Internacional de Agricultura Tropical, Colombia.
[e] International Institute of Tropical Agriculture, Nigeria.

M. brachyloba, and *M. carthaginensis,* no products were amplified at locus *Ga*-140. Similarly, with the locus *Ga*-13, no PCR products were observed for the first species (Table 2). At four of the eight loci, amplification products were absent in some accessions.

Number of alleles and polymorphism in Manihot—A total of 124 alleles were found in 121 individuals, using ten microsatellite loci. The number of alleles detected in all species varied greatly among loci. The highest number, 21 alleles, was found at locus *Ga*-131, and was followed by locus *Ga*-126 with 20 alleles. Five SSR loci showed between ten and 20 alleles, and fewer than ten alleles were detected in three loci, four alleles being the lowest number found (for locus *Ga*-13) (Fig. 1).

At all loci, a higher number of alleles were detected in wild species than in the cultivated cassava (Fig. 1). Considering the entire set of germplasm, 79 alleles were present in the wild

TABLE 2. Cross-species amplification of ten microsatellite loci in seven *Manihot* taxa. "+" denotes successful amplification, "−" denotes no amplification.

Species	Locus										Polymorphic loci (%)
	Gagg-5	*Ga*-13	*Ga*-12	*Ga*-16	*Ga*-21	*Ga*-126	*Ga*-131	*Ga*-134	*Ga*136	*Ga*140	
M. esculenta	+	+	+	+	+	+	+	+	+	+	100
M. esculenta subsp. flabellifolia	+	+	+	+	+	+	+	+	+	+	90
M. esculenta subsp. *peruviana*	+	+	+	+	+	+	+	+	+	+	100
M. aesculifolia	+	−	+	+	+	+	+	+	+	−	80
M. brachyloba	+	+	+	+	+	+	+	+	+	−	60
M. carthaginensis	+	+	+	+	+	+	+	+	+	−	90
M. tristis	+	+	+	+	+	+	+	+	+	+	60

species but absent in cassava. Forty-two alleles were shared between cassava and the wild conspecific taxa and just three SSR alleles were found only in the crop, suggesting a larger pool of SSR alleles for the wild species. Analyzing only cassava and the wild forms of *M. esculenta,* 47 additional alleles were present in the subspecific taxa. Thus, 32 unique alleles were found in species outside the *M. esculenta* complex (Fig. 1).

The percentage of polymorphic loci in the seven *Manihot* species ranged from 60 to 100% (Table 2), corresponding to a mean polymorphism of 83%. The taxa with the highest level of polymorphic loci (100%) were the cultivated species *M. esculenta* and the wild form *M. esculenta* subsp. *peruviana,* followed by *M. esculenta* subsp. *flabellifolia* with 90% polymorphism. *Manihot carthaginensis* and *M aesculifolia* exhibited 90 and 80% polymorphism, respectively, and 60% of the SSR loci were polymorphic in *M. brachyloba* and *M. tristis* (Table 2).

Allelic distribution and frequencies—Figure 2 shows the different allele sizes scored among the *Manihot* species. At most of the loci, allele sizes varied from 100 to 175, except for loci *Ga*-134 and *Ga*-126 where sizes were from >200 base pairs (bp) to 405 bp. The smallest difference between the highest and lowest values of allele size length was 6 bp at locus *Ga*-13, and the largest difference (225 bp) was detected at locus *Ga*-126. The allele sizes scored at the other remaining loci presented differences between 14 and 53 bp (Fig. 2). In most cases, the number of unique alleles (i.e., amplified products in just one individual or one species) was positively correlated with the total number of alleles per locus and their size differences. For instance, 13 was the maximum number of unique alleles detected at locus *Ga*-126. *Manihot aesculifolia* presented seven of the 13 unique alleles detected at this locus (Fig. 2). At locus *Ga*-131, with 21 alleles and 53 bp of difference between the shortest and largest alleles, a total of seven unique alleles were scored (Fig. 2).

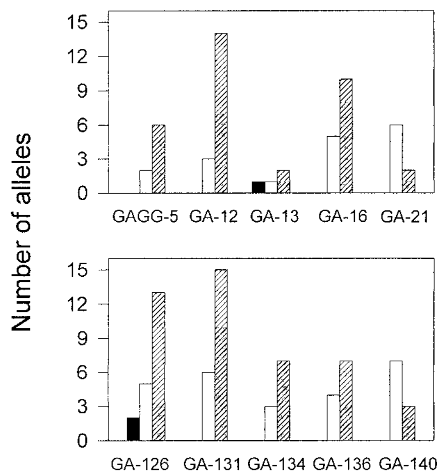

Fig. 1. Number and distribution of alleles at microsatellite loci (*Gagg*-5 to *Ga*-140) in *Manihot* species. Black bars = alleles unique to cassava; white bars = alleles shared between cassava and wild taxa; hatched bars = alleles unique to wild taxa (gray: present in the two subspecies of *M. esculenta*; white: absent in the two subspecies of *M. esculenta*).

At most of the SSR loci, alleles were detected at frequencies <0.33, both within and among the seven *Manihot* taxa (Fig. 2). Nevertheless, at the loci *Gagg*-5 and *Ga*-13, some alleles showed frequencies >0.66 in more than two species (Fig. 2). Unique alleles were also observed, with few exceptions, at low frequencies. Only for *M. brachyloba* and *M. carthaginensis* were alleles detected at frequencies of 0.7–1 at four SSR loci and at locus *Gagg*-5, respectively (Fig. 2). For the locus *Ga*-140 where no amplification was obtained for *M. aesculifolia, M. Brachyloba,* and *M. carthaginensis,* the alleles were evenly distributed among the remaining species and the frequencies varied mainly from 0 to 0.33. At the intraspecific level, in species such as *M. esculenta* subsp. *peruviana,* a larger percentage (60%) of the alleles was present at very low frequency (<0.1) (Fig. 2).

Level of heterozygosity in Manihot—Outcrossing plants with dioecious floral morphology are expected to have high genetic heterozygosity within populations (Avise, 1994). The *Manihot* species evaluated fit into this group of plants, but, in contrast to the theory, in 42 of 70 (60%) locus-species comparisons, the observed number of heterozygotes (H_{obs}) was less than expected (H_{exp}) (Table 3). Pooling across all ten loci, a heterozygote deficit was found in five of the seven taxa, all being wild species. The only species with the same overall observed (mean H_{obs}) and expected heterozygosity (mean H_{exp}) figures was *M. esculenta* (Table 3).

Forty-six of 70 (65.7%) F_{is} values were greater than zero (Table 3), indicating that nonrandom association of alleles can occur in the species tested (Gibbs et al., 1997). According to Gibbs et al., (1997), F_{is} values cannot be calculated (referred as NA in Table 3) when only one allele is present in the species or a second allele is present in only one individual.

The r_b estimator, which estimates the frequency of null al-

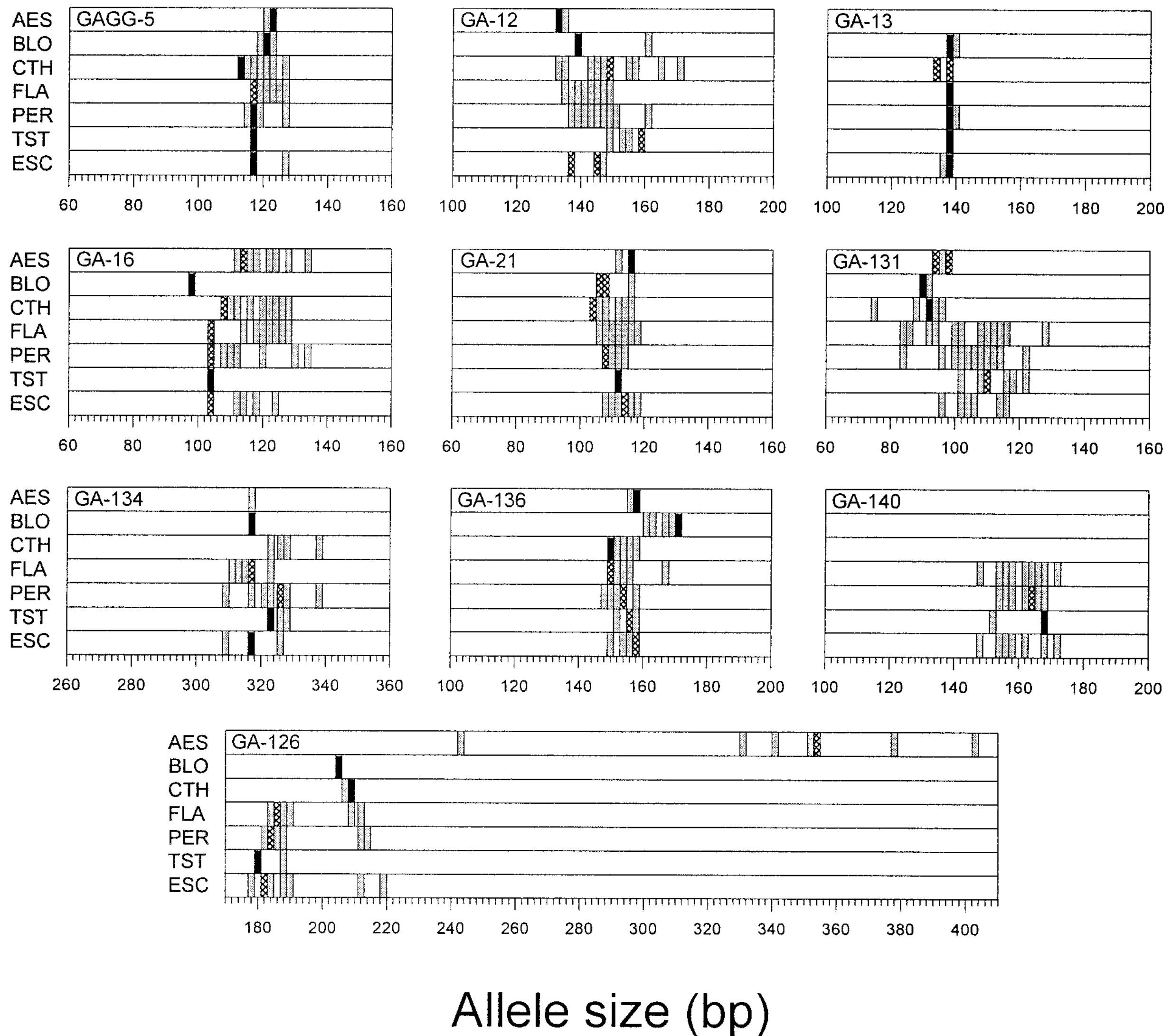

Fig. 2. Allelic distribution and frequencies in each of the *Manihot* taxa at different microsatellite loci. Bar colors: gray = frequency 0 to ≤0.33; cross-hatched = frequency >0.33 to ≤0.66; black = frequency >0.66 to 1.00. AES = *M. aesculifolia*; BLO = *M. brachyloba*; CTH = *M. carthaginensis*; FLA = *M. esculenta* subsp. *flabellifolia*; PER = *M. esculenta* subsp. *peruviana*; TST = *M. tristis*; ESC = *M. esculenta* (i.e., cassava).

leles in each species at each locus, varied from −0.14 in *M. esculenta* (locus *Gagg*-5) up to 0.33 in *M. esculenta* subsp. *peruviana* (locus *Ga*-134). When overall r_b was calculated for all loci in each species, the lowest (0.0) and highest (0.20) values were also obtained at *M. esculenta* and *M. esculenta* subsp. *peruviana,* respectively (Table 3).

***Cluster analysis*—**Figure 3 depicts the clustering of *Manihot* accessions into groups that corresponded, in most cases, with their taxonomic classification. With these ten SSR loci, the differentiation of every genotype was not possible in 12 cases, corresponding to accessions of the same populations in *M. brachyloba, M. carthaginensis, M. esculenta* subsp. *peruviana,* and *M. tristis.* Notwithstanding, the species *M. aesculifolia, M. brachyloba,* and *M. carthaginensis* formed discrete groups (clusters 7, 3, and 5, respectively) with <30% of similarity to the cultivated species (cluster 1). Most accessions from *M. esculenta* subsp. *flabellifolia, M. esculenta* subsp. *peruviana,* and *M. tristis* formed a mixed cluster (2), which was closer to cassava than the remaining species. Within this mixed cluster, the accessions of *M. esculenta* subsp. *flabellifolia* and some of *M. esculenta* subsp. *peruviana* were closest to cassava. As was shown by Roa et al. (1997), using AFLP markers on these species, the most similar group to cassava is the mixed cluster formed by the taxa proposed by Allem (1994) as the wild forms of the crop, followed by *M. tristis.*

With respect to the intraspecific variability, microsatellites showed a high heterogeneity within all taxa studied, even in cassava. The range varied from <10% within some populations of species such *M. brachyloba,* 50% within the sample of cassava and *M. aesculifolia,* and even further in the case of *M. carthaginensis* (70% of variability). One population of *M. esculenta* subsp. *flabellifolia* (cluster 4) and two accessions classified as *M. carthaginensis* (cluster 6) appeared to be very distant from their conspecifics (Fig. 3).

***Comparison between AFLPs and microsatellites*—**To compare the results obtained with AFLP and microsatellite markers

TABLE 3. Heterozygosity values (H_{exp} and H_{obs}), frequency of null alleles (r_b), and fixation index (F_{is}) at microsatellite loci of different *Manihot* species.

Locus	Species[a] (number of accessions)						
	ESC (38)	AES (8)	BLO (9)	CTH (26)	ESC-FLA (16)	ESC-PER (16)	TST (8)
Ga-21							
H_{exp} (H_{obs})	0.65 (0.50)	0.12 (0.13)	0.62 (0.56)	0.62 (0.23)	0.82 (0.31)	0.64 (0.13)	0 (0)
r_b	0.09	−0.007	0.04	0.24	0.28	0.31	0
F_{is}	0.25	NA	0.17	0.64	0.64	0.82	NA
Ga-136							
H_{exp} (H_{obs})	0.66 (0.76)	0.12 (0.13)	0.38 (0.22)	0.47 (0.15)	0.71 (0.38)	0.72 (0.50)	0.71 (0.38)
r_b	−0.06	−0.007	0.12	0.21	0.20	0.13	0.20
F_{is}	−0.14	NA	0.47	0.68	0.50	0.34	0.52
Ga-126							
H_{exp} (H_{obs})	0.73 (0.79)	0.75 (0.63)	0 (0)	0.07 (0.08)	0.77 (0.38)	0.73 (0.31)	0.43 (0.38)
r_b	−0.03	0.07	0	−0.003	0.22	0.24	0.04
F_{is}	−0.06	0.23	NA	−0.02	0.53	0.59	0.19
Ga-134							
H_{exp} (H_{obs})	0.39 (0.42)	0.22 (0)	0 (0)	0.33 (0.04)	0.75 (0.44)	0.76 (0.19)	0.32 (0.25)
r_b	−0.02	0.18	0	0.22	0.18	0.33	0.05
F_{is}	−0.06	1.00	NA	0.89	0.44	0.77	0.28
Ga-16							
H_{exp} (H_{obs})	0.52 (0.47)	0.66 (0.63)	0 (0)	0.70 (0.35)	0.68 (0.50)	0.71 (0.50)	0 (0)
r_b	0.03	0.02	0	0.21	0.10	0.12	0
F_{is}	0.11	0.11	NA	0.52	0.29	0.33	NA
Ga-140							
H_{exp} (H_{obs})	0.79 (0.76)	0 (0)	0 (0)	0 (0)	0.83 (0.31)	0.79 (0.50)	0.30 (0.13)
r_b	0.02	0	0	0	0.28	0.16	0.14
F_{is}	0.05	NA	NA	NA	0.64	0.39	0.63
Ga-12							
H_{exp} (H_{obs})	0.64 (0.58)	0.12 (0.13)	0.10 (0.11)	0.62 (0.42)	0.85 (0.50)	0.82 (0.25)	0.72 (0.25)
r_b	0.04	−0.007	−0.006	0.12	0.19	0.31	0.27
F_{is}	0.11	NA	NA	0.34	0.44	0.71	0.69
Ga-131							
H_{exp} (H_{obs})	0.75 (0.74)	0.63 (0.38)	0.20 (0)	0.36 (0.35)	0.87 (0.75)	0.82 (0.31)	0.73 (0.88)
r_b	0.01	0.16	0.16	0.007	0.06	0.28	−0.08
F_{is}	0.04	0.46	1.00	0.07	0.17	0.64	−0.13
Ga-13							
H_{exp} (H_{obs})	0.03 (0.03)	0 (0)	0.20 (0)	0.50 (0.08)	0 (0)	0.06 (0.06)	0 (0)
r_b	0	0	0.16	0.28	0	0	0
F_{is}	NA	NA	1.00	0.85	NA	NA	NA
Gagg-5							
H_{exp} (H_{obs})	0.43 (0.63)	0.43 (0.38)	0.29 (0.11)	0.31 (0.12)	0.54 (0.44)	0.32 (0.31)	0 (0)
r_b	−0.14	0.04	0.14	0.15	0.07	0.007	0
F_{is}	−0.45	0.19	0.65	0.64	0.23	0.06	NA
Mean H_{exp}	0.56	0.30	0.18	0.40	0.68	0.64	0.32
Mean H_{obs}	0.57	0.24	0.10	0.18	0.40	0.31	0.23
Mean r_b	0	0.05	0.07	0.16	0.17	0.20	0.07

[a] NA = not applicable (see text for explanation); ESC = *M. esculenta*, i.e., cassava; AES = *M. aesculifolia*; BLO = *M. brachyloba*; CTH = *M. carthaginensis*; ESC-FLA = *M. esculenta* subsp. *flabellifolia*; ESC-PER = *M. esculenta* subsp. *peruviana*; TST = *M. tristis*.

in a common group of germplasm, the Mantel matrix correspondence test was used. Similarity matrices and UPGMA dendrograms with each marker class were generated with 96 selected *Manihot* materials (data not shown). All the cophenetic correlation coefficients between the similarity matrices and cophenetic matrices obtained with both marker types were high, 0.81 and 0.87, respectively, and statistically significant (probability of random $Z <$ obs. Z: $P = 1.00$).

DISCUSSION

The present work intended to determine the extent to which pairs of primers designed for the amplification of SSR loci in cassava can be used for assessing its wild relatives. The utility of the cassava primers to produce PCR -amplified products across the genus was demonstrated. Nevertheless, as it has been pointed out in other works on SSR loci conservation in plant species, a decline of amplification success was observed with increase of genetic distance (Isagi and Suhandono, 1997; Steinkellner et al., 1997; White and Powell, 1997; Witsenboer et al., 1997). In *M. aesculifolia, M. carthaginensis,* and *M. brachyloba,* the species most distant from cassava, no amplification products were detected. In species with low degrees of relationship, therefore, the same SSRs loci cannot be found. The result obtained with this codominant marker agrees with the findings reported for AFLP markers used with the same species (Roa et al., 1997): that *M. aesculifolia* and *M. carthaginensis* are the wild species most distantly related to cassava.

According to Smulders et al. (1997), the lack of amplification of an allele in certain accessions can be the result of divergence in the sequences flanking the microsatellite, creating a null allele. The production of an undetectable amount of PCR product is another explanation given by Smulders et

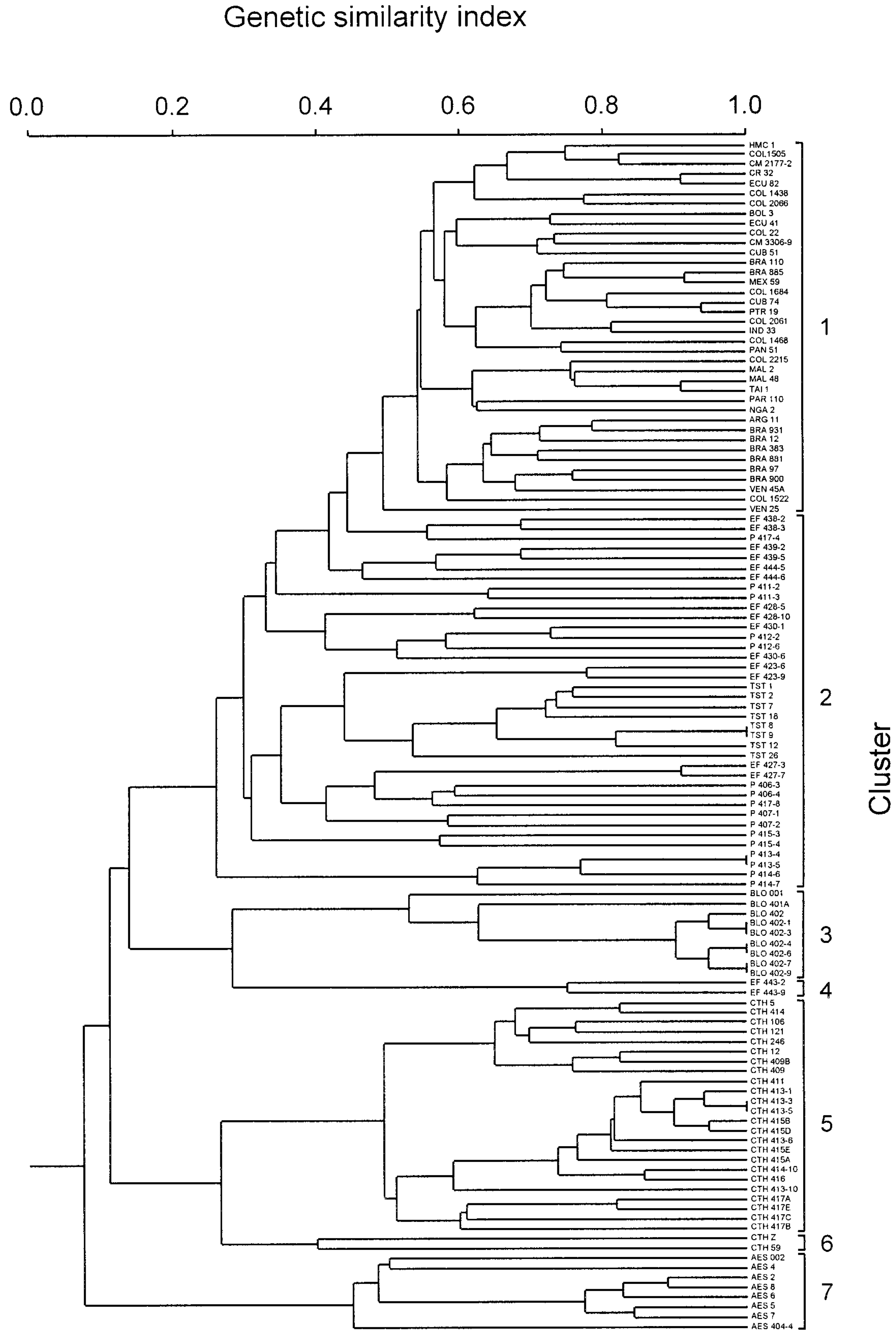

Fig. 3. Dendrogram of genetic similarity in *Manihot* species, drawn from a UPGMA analysis of Nei–Li estimates based on a microsatellite analysis. The nine clusters shown are discussed in the text. 1 = *M. esculenta,* that is, cassava; 2 = *M. esculenta* subspp. *flabellifolia* (EF) and *peruviana* (P), and *M. tristis* (TST); 3 = *M. brachyloba* (BLO); 4 = outlying population of *M. esculenta* subsp. *flabellifolia*; 5 = *M. carthaginensis* (CTH); 6 = outlying accessions of *M. carthaginensis*; 7 = *M. aesculifolia* (AES).

al. (1997) and Lavi et al. (1994). Nevertheless, these studies were carried out, using a detection method that is less sensitive than fluorescence. In our survey, null alleles were confirmed after several repetitions of the assay, using the same conditions and also different annealing temperatures, to ensure that no reaction failure existed. Nonamplifying or null alleles have not been highlighted in literature on plant molecular ecology, but, from human studies, two facts emerge. First, these alleles are common ; and, second, where flanking sequences were obtained for nonamplifying alleles, a mutation was found to have occurred in one of the priming sites (Pemberton et al., 1995). The exclusion of even a few nonamplifying homozygotes can have dramatic effects on the interpretation of genotype frequency distributions and could lead to mistaken interpretations about the level of inbreeding in a population (Pemberton et al., 1995).

Microsatellite variation was clearly detected in the set of germplasm studied. The number of alleles varied greatly among the SSR loci evaluated, but, in all, more alleles were found in the wild species than in cassava. Almost 64% of the scored alleles were present in the wild taxa while absent from cassava. Considering only cassava and the wild subspecies, 47 additional alleles were found in the second group, suggesting a larger pool of alleles for the *M. esculenta* subspecies than for cassava. This result shows that the primary gene pool of cassava, composed of the wild forms and a few other species (Allem, 1994), contains a valuable source of diversity that could be useful for improving cassava.

The high level of polymorphism (83%) obtained with this type of marker makes it a powerful tool for assessing genetic diversity in crop plants and its relatives. Ashley and Dow (1994) proposed that the high polymorphism found at microsatellite loci is related to the mechanism of mutation and the high rate at which this occurs. Nevertheless, the polymorphism rate reported here was lower than the AFLP polymorphism (98.6%) obtained by Roa et al. (1997) in the same species. This could be explained by the lack of amplification products in some *Manihot* species at certain SSR loci.

Currently, different models try to explain the mutational dynamics of the microsatellite loci. According to Slatkin's model (1995), they mutate in an unconstrained fashion, meaning that no limit exists on the number of possible allele sizes. However, evidence suggests that biases may exist in mutation direction and therefore a limit does exist to maximum allele size (Lehmann, Hawley, and Collins, 1996; Goodman, 1997). Considering the length differences that were found for most of the SSR loci evaluated, the amplification patterns obtained may correspond to one locus with several alleles. Nevertheless, the length difference between the shortest and longest allele for locus *Ga*-126 (>100 steps) may indicate the presence of other loci for these dinucleotide repeats in *M. aesculifolia.*

The substantial number of unique alleles (and null alleles) present in species such as *M. aesculifolia, M. carthaginensis,* and *M. brachyloba* suggests that their separation from cassava on a evolutionary scale took place long ago, allowing mutation to generate new alleles. Similar results have been reported in animal studies (Gibbs et al., 1997). The wild forms of *M. esculenta* (subspp. *flabellifolia* and *peruviana*) also presented a larger number of unique alleles than did the cultivated species, indicating a larger pool of SSR alleles in the wild relatives. However, substantial conclusion on *Manihot* evolution or cassava wild relatives awaits a full generic phylogenetic revision including both molecular and morphological data. This study, evaluating the ability of selected primers to amplify SSR loci at the generic level, is preliminary to such a definitive study.

The levels of heterozygosity found in *Manihot* were, in most cases, lower than expected. Species heterozygote deficiency can be explained as a result of different factors: (a) unrecognized genetic structure within populations, (b) inbreeding due to consanguineous mating, or (c) presence of null alleles, such that many apparent homozygotes are, in reality, heterozygotes between a visible and a null allele (Pemberton et al., 1995; Brookfield, 1996). Direct evidence for null alleles and genetic differentiation between populations was collected (r_b and F_{is} values). Assuming that the heterozygote deficit is totally due to null alleles, the r_b values suggest that such alleles are present, on the average, at relatively high frequencies (~11%) in *Manihot* species.

The other component of heterozygote deficiency implies a genetic differentiation, that is, in allele sizes and frequencies, within and between populations of a species. Similar results were obtained in the AFLP study (Roa et al., 1997). This fine-scale differentiation within the species could be reflected in a substantial difference between the observed and expected heterozygosity. In the crop sample, 60% of the observed alleles in all loci were at frequencies between 0.1 and 0.99 (Fig. 2). This feature indicates a low differentiation among accessions and could explain why the difference between observed and expected heterozygosity was not significant. However, in species such as *M. esculenta* subsp. *peruviana,* a larger percentage (60%) of alleles was present at low frequencies (<0.1) (Fig. 2), indicating a high differentiation among the accessions at the SSR loci. A combination of the presence of null alleles and fine-scale genetic differentiation may have been the cause of the heterozygote deficiency in *Manihot* species, but we could not quantify, using our data, the importance of each factor as a cause of the heterozygote deficit.

The cluster analysis of the seven *Manihot* species with SSR markers showed that the so-called wild forms of *M. esculenta* are the most closely related taxa to cassava among those studied here. The high correlation indexes obtained between the SSR and AFLP markers mean that a similar genetic structure of *Manihot* is depicted by both marker types. Dominant (AFLPs) (Roa et al. 1997) and codominant (SSR) markers demonstrated that cassava probably arose from the wild Brazilian forms *M. esculenta* subsp. *flabellifolia* and *M. esculenta* subsp. *peruviana,* as described by Allem (1994).

The utility of a codominant marker such as microsatellites for accurately assessing the occurrence of heterozygote genotypes in species with an outbreeding mode of reproduction such as cassava is clear. High levels of intraspecific variation were detected in most of the *Manihot* species surveyed, making SSR markers suitable for studies of population differentiation. This factor is essential for establishing the criteria for conservation of populations of a species under in situ and ex situ conditions.

LITERATURE CITED

AKAGI, H., Y. YOKOZEKI, A. INAGAKI, AND T. FUJIMURA. 1997. Highly polymorphic microsatellites of rice consist of AT repeats, and a classification of closely related cultivars with these microsatellite loci. *Theoretical and Applied Genetics* 94: 61–67.

ALLEM, A. C. 1987. *Manihot esculenta* is a native of the neotropics. *Plant Genetic Resources Newsletter (International Board of Plant Genetic Resources)* 71: 22–24.

———. 1994. The origin of *Manihot esculenta* Crantz (Euphorbiaceae). *Genetic Resources and Crop Evolution* 41: 133–150.

ASHLEY, M. V., AND B. D. DOW. 1994. The use of microsatellite analysis in population biology: background, methods and potential applications. *In* B. Schierwater, B. Streit, G. P. Wagner, and R. Desalle [eds.], Molecular Ecology and Evolution: Approaches and Applications, 185–201. Birkhäuser Verlag, Basel, Switzerland.

AVISE, J. C. 1994. Molecular markers, natural history and evolution. Chapman and Hall, New York, New York, USA.

BERTRAM, R. B. 1993. Application of molecular techniques to genetic resources of cassava (*Manihot esculenta* Crantz, Euphorbiaceae): interspecific evolutionary relationships and intraspecific characterisation. Ph.D. dissertation, University of Maryland, College Park, Maryland, USA.

BEST, R., AND G. HENRY. 1994. Cassava: towards the year 2000. *In* International network for cassava genetic resources, 3–11. Report of the First Meeting of the International Network for Cassava Genetics Resources, CIAT, Cali, Colombia, 18–23 August. International Crop Network Series number 10. International Plant Genetic Resources Institute (IPGRI), Rome, Italy.

BROOKFIELD, J. F. Y. 1996. A simple new method for estimating null frequency from heterozygote deficiency. *Molecular Ecology* 5: 453–455.

BYRNE, M., M. I. MARQUEZ-GARCÍA, T. UREN, D. S. SMITH, AND G. F. MORAN. 1996. Conservation and genetic diversity of microsatellite loci in the genus *Eucalyptus. Australian Journal of Botany* 44: 331–341.

CHAVARRIAGA-AGUIRRE, P., M. M. MAYA, M. W. BONIERBALE, S. KRESOVICH, M. FREGENE, J. TOHME, AND G. KOCHERT. 1998. Microsatellites in cassava (*Manihot esculenta* Crantz): discovery, inheritance, and variability. *Theoretical and Applied Genetics* 97: 493–501

COCK, J. H. 1985. Cassava: new potential for a neglected crop. Westview Press, Boulder, Colorado, USA.

DOW, B. D., AND M. V. ASHLEY. 1996. Microsatellite analysis of seed dispersal and parentage of saplings in bur oak, *Quercus macrocarpa. Molecular Ecology* 5: 615–627.

———, ———, AND H. F. HOWE. 1995. Characterization of highly variable (GA/CT)-n microsatellites in the bur oak, *Quercus macrocarpa. Theoretical and Applied Genetics* 91: 137–141.

FREGENE, M. A., J. VARGAS, F. ANGEL, J. TOHME, R. A. ASIEDU, M. O. AKORADA, AND W. M. ROCA. 1994. Chloroplast DNA and nuclear ribosomal DNA variability in cassava (*Manihot esculenta* Crantz) and its wild relatives. *Theoretical and Applied Genetics* 89: 719–727.

GIBBS, H. L., K. A. PRIOR, P. J. WEATHERHEAD, AND G. JOHNSON. 1997. Genetic structure of the threatened eastern massasauga rattlesnake, *Sistrurus c. catenatus*: evidence from microsatellite DNA markers. *Molecular Ecology* 6: 1123–1132.

GOODMAN, S. J. 1997. R_{ST} Calc: a collection of computer programs for calculating estimates of genetic differentiation from microsatellite data and determining their significance. *Molecular Ecology* 6: 881–885.

HERSHEY, C., C. IGLESIAS, M. IWANAGA, AND J. TOHME. 1994. Definition of a core collection for cassava. *In* International network for cassava genetic resources, 145–156. Report of the First Meeting of the International Network for Cassava Genetic Resources, CIAT, Cali, Colombia, 18–23 August. International Crop Network Series number 10. International Plant Genetic Resources Institute (IPGRI), Rome, Italy.

INNAN, H., R. TERAUCHI, AND N. T. MIYASHITA. 1997. Microsatellite polymorphism in natural populations of the wild plant *Arabidopsis thaliana. Genetics* 146: 1441–1452.

ISAGI, Y., AND S. SUHANDONO. 1997. PCR primers amplifying microsatellite loci of *Quercus myrsinifolia* Blume and their conservation between oak species. *Molecular Ecology* 6: 897–899.

KATZIR, N., Y. DANIN-POLEG, G. TZURI, Z., KARCHI, U., LAVI, AND P. B. CREGAN. 1996. Length polymorphism and homologies of microsatellites in several Cucurbitaceae species. *Theoretical and Applied Genetics* 93: 1282–1290.

KIJAS, J. M. H., J. C. S. FOWLER, AND M. R. THOMAS. 1995. An evaluation of sequenced tagged microsatellite site markers for genetic analysis within *Citrus* and related species. *Genome* 38: 349–355.

LAGERCRANTZ, U., H. ELLENGRAN, AND L. ANDERSON. 1993. The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic Acids Research.* 21: 1111–1115.

LAVI, U., M. AKKAYA, A. BHAGWAT, E. LAHAV, AND P. B. CREGAN. 1994. Methodology of generation and characteristics of simple sequence repeat DNA markers in avocado (*Persea americana* M.). *Euphytica* 80: 171–177.

LEHMANN, T., W. A. HAWLEY, AND F. H. COLLINS. 1996. An evaluation of evolutionary constraints on microsatellite loci using null alleles. *Genetics* 144: 1155–1163.

MCCOUCH, S. R., X. CHEN, O. PANAUD, S. TEMNYKH, Y. XU, Y. G. CHO, N. HUANG, T. ISHII, AND M. BLAIR. 1997. Microsatellite marker development, mapping and applications in rice genetics and breeding. *Plant Molecular Biology* 35: 89–99.

MORGANTE, M., AND A. M. OLIVIERI. 1993. PCR-amplified microsatellites as markers in plant genetics. *Plant Journal* 3: 175–182.

NASSAR, N. M. A. 1978. Conservation of the genetic resources of cassava (*Manihot esculenta* Crantz). Determination of wild species localities with emphasis on probable origin. *Economic Botany* 32: 311–320.

NEI, M., AND W.-H. LI. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences, USA* 76: 5269–5273.

PEMBERTON, J. M., J. SLATE, D. R. BANCROFT, AND J. A. BARRETT. 1995. Nonamplyfying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology* 4: 249–252.

PERKIN-ELMER. 1995. ABI Prism GenScan Analysis Version 2.0.2. Applied Biosystems Division. Foster City, California, USA.

———. 1995. Genotyper Software Version 1.1. Applied Biosystems Division. Foster City, California, USA.

RENVOIZE, B. S. 1972. The area of origin of *Manihot esculenta* as a crop plant-a review of the evidence. *Economic Botany* 26: 352–360.

ROA, A. C., M. M. MAYA, M. C. DUQUE, J. TOHME, A. C., ALLEM, AND M. W. BONIERBALE. 1997. AFLP analysis of relationships among cassava and other *Manihot* species. *Theoretical and Applied Genetics* 95: 741–750.

ROGERS, D. J. 1963. Studies of *Manihot esculenta* and related species. *Bulletin of the Torrey Botanical Club* 90: 43–54.

———, AND S. G. APPAN. 1973. *Manihot, Manihotoides* (Euphorbiaceae). Flora Neotropica Monograph number 13. Hafner Press, New York, New York, USA.

ROHLF, F. J. 1994. NTSYS-pc: numerical taxonomy and multivariate system, version 1.80. Exeter Software, Setauket, New York, USA.

SLATKIN, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139: 457–462.

SMULDERS, M. J. M., G. BREDEMEIJER, W. RUS-KORTEKASS, P. ARENS, AND B. VOSMAN. 1997. Use of short microsatellites from database sequences to generate polymorphisms among *Lycopersicon esculentum* cultivars and accessions of other *Lycopersicon* species. *Theoretical and Applied Genetics* 97: 264–272.

SOKAL, R. R., AND J. ROHLF. 1995. Biometry: the principles and practice of statistics in biological research, 3rd. ed. W.H. Freeman, New York, New York, USA.

STEINKELLNER, H., C. LEXER, E. TURETSCHEK, AND J. GLÖSSL. 1997. Conservation of (GA)n microsatellite between *Quercus* species. *Molecular Ecology* 6: 1189–1194.

TAUTZ, D., M. TRICK, AND G. A. DOVER. 1986. Cryptic simplicity in DNA is a major source of genetic variation. *Nature (London)* 322: 652–656.

WHITE, G., AND W. POWELL. 1997. Cross-species amplification of SSR loci in the Meliaceae family. *Molecular Ecology* 6: 1195–1197.

WITSENBOER, H., J. VOGEL, AND R. W. MICHELMORE. 1997. Identification, genetic localisation, and allelic diversity of selectively amplified microsatellite polymorphic loci in lettuce and wild relatives (*Lactuca spp.*). *Genome* 40: 923–936.

WU, K. S., AND S. D. TANKSLEY. 1993. Abundance, polymorphism and genetic mapping of microsatellites in rice. *Molecular and General Genetics* 241: 225–235.

YEH, F. C., Y. RONGCAI, AND T. BOYLE. 1997. PopGene Version 1.2, Microsoft Window- based software for population genetics analysis: a quick user's guide. University Alberta/Center for International Forestry Research, Alberta, Canada.