Theor Appl Genet (1998) 97: 493–501

P. Chavarriaga-Aguirre · M. M. Maya
M. W. Bonierbale · S. Kresovich · M. A. Fregene
J. Tohme · G. Kochert

Microsatellites in Cassava (*Manihot esculenta* Crantz): discovery, inheritance and variability

Received: 4 September 1997 / Accepted 16 March 1998

Abstract Fourteen microsatellites containing GA-repeats were isolated and characterized in cassava (*Manihot esculenta* Crantz, Euphorbiaceae). Microsatellite heterozygosity (*h*) was estimated in 48 accessions using (^{32}P)-end-labeled primers and in more than 500 accessions using fluorescence-based genotyping. Heterozygosity values ranged from 0.00 to 0.88 and the number of alleles detected varied from 1 to 15. The reproducibility of allele sizing was also assessed using fluorescence-based genotyping. The average inter-gel size difference was 1.03 nucleotides. Chi-square tests (χ^2) were performed to analyse segregation distortion and the linkage between alleles segregating from either or both parents in an F_1 mapping population. Most microsatellite loci segregated in the expected 1:1, 1:2:1 or 1:1:1:1 ratio. Linkage was detected between loci segregating from either parent, and segregation distortion from the male parent was detected for locus GA-131. Approximately 80% of the microsatellites detected one or two alleles per accession, suggesting a low degree of microsatellite locus duplication, an unexpected finding for a putative allopolyploid, highly heterozygous species. The high *h* values of most microsatellites, their amplification in other *Manihot* taxa and their suitability for high-throughput, fluorescence-based genotyping, make microsatellites the marker of choice for germplasm characterization and saturation of the cassava map.

Key words Cassava · Microsatellites · Fluorescence-based genotyping · Heterozygosity · Linkage

Communicated by M. A. Saghai Maroof

P. Chavarriaga · G. Kochert (✉)
2502 Plant Sciences, Department of Botany, University of Georgia, Athens GA 30605, USA

M. M. Maya · M. W. Bonierbale[1]
Cassava Program, CIAT, A.A. 6713, Cali, Colombia

S. Kresovich
USDA-ARS, Plant Genetic Resources Conservation Unit, University of Georgia, Griffin, GA, 30223-1797, USA

M. A. Fregene · J. Tohme
Biotechnology Research Unit, CIAT, A.A. 6713, Cali, Colombia

Present address:
[1] Merideth Bonierbale, International Potato Center (CIP), Apartado 1558, Lima 12, Perú

Introduction

Cassava (*Manihot esculenta* Crantz, Euphorbiaceae) is a root crop that originated in the new world tropics. Although the exact place of cassava domestication is still a matter of debate (Roa et al. 1997; Bertram 1993; Fregene et al. 1994; Schaal et al. 1995), it was taken from South America and spread throughout the Old World tropics by Portuguese sailors in the 16th century (Rogers 1963). Since then, cassava has gradually become an economically important crop in developing countries. In Africa, the starchy roots of cassava constitute a staple food. In Thailand, Indonesia and Brazil, cassava is an important export commodity (FAO 1992; CIAT 1993). Today cassava ranks fourth among the major sources of carbohydrates in the tropics, where it is an important food for over 500 million people.

DNA-based molecular markers such as RAPDs, nuclear RFLPs and microsatellites were used to develop the cassava molecular genetic map (Fregene et al. 1997). Similarly, RAPDs, AFLPs and rDNA, cpDNA and cDNA RFLPs have been used to assess the genetic variability of small sets of cassava germplasm and to establish relationships between cassava and its wild relatives (Beeching et al. 1993; Bertram 1993; Fregene et al. 1994; Marmey et al. 1994; Schaal et al. 1995; Roa et al. 1997). Assessing the genetic variability of larger

germplasm collections, such as the cassava core collection that contains over 600 accessions (Hershey et al. 1994), would require highly polymorphic markers as well as high-throughput genotyping systems (Ziegle et al. 1992). Such systems have been employed for plant germplasm characterization (Mitchell et al. 1997) and for the large-scale genotyping of humans (Ghosh et al. 1997; for a review see Hall et al. 1996). In the case of cassava, germplasm characterization will contribute to the establishment of relationships between accessions of the cultivated genepool and, consequently, facilitate the development of cassava cultivars to satisfy market standards and to respond to diverse biotic and abiotic challenges. One type of molecular marker that may be suitable for cassava germplasm characterization is the microsatellite.

Microsatellites are short stretches of tandemly repeated, 1–5 nucleotide sequences, such as $(GA)_n$. They are ubiquitously present in eukaryotic genomes and are highly polymorphic (Tautz 1989; Beckman and Soller 1990; Weber 1990). Conservation of the microsatellite flanking sequences allows the designing of primers for PCR amplification. Microsatellites, like RFLPs, are considered codominant markers. However, insertions or base substitutions in priming sites may convert them to dominant markers. Their high polymorphism makes microsatellites suitable markers for population genetics, the development of linkage maps and for phylogenetic studies (Zhao and Kochert 1993; Saghai Maroof et al. 1994; Broun and Tanksley 1996). They have been useful for the germplasm characterization of grapes, soybean, Brassica and wheat (Thomas et al. 1994; Kresovich et al.1995; Roder et al. 1995; Rongwen et al. 1995). The objective of the present work was to clone cassava microsatellites suitable for mapping and for the characterization of the genetic variability of the cassava core collection. Fourteen cassava microsatellite loci containing GA-repeats are described, along with heterozygosity estimates, the mode of inheritance and linkage analysis of 11 loci, and relevant aspects of microsatellite allele sizing using fluorescence-based, semi-automated genotyping systems.

Materials and methods

Plant material and DNA extraction

A segregating cassava F_1 population of 83 individuals (cross CM7857), produced by crossing cultivars TMS30572 and CM2177-2, was received in vitro, or as DNA samples, from CIAT (Cali, Colombia). In vitro material was transferred to the greenhouse after a subculture period of 6 weeks. Leaf tissue was collected for DNA extraction from the population mentioned above and from 522 accessions of the cassava core collection grown in the field at CIAT. For passport data of all accessions consult the SINGER database at web site htpp://noc1.cgiar.org/seartype.htm.

Genomic DNA was isolated from fully expanded young leaves of greenhouse-grown plants by a modification of the Kochert et al. (1991) procedure. Basically, 6–8 g of fresh tissue were powdered in liquid nitrogen and the DNA extracted by adding 5 ml of lysis buffer, avoiding the isolation of a crude nuclear preparation. All DNA solutions were treated with 1 µl of 10 µg/ml RNAse for 1 h at 37°C. DNA extractions for field-grown plants were carried out according to Dellaporta et al. (1983), with the modifications reported by González et al. (1995).

Microsatellite cloning and sequencing

Approximately 100 µg of DNA from cassava accession MCol22 was double-digested with *Eco*RI and *Xho*I, according to instructions provided by the manufacturer (New England Biolabs). The restricted DNA was run on 0.5% low-melting-point agarose gels until a good separation of fragments between 0.3 and 0.8 kb was achieved. DNA fragments within this size range were isolated from the gel, purified using a phenol-chloroform procedure (Sambrook et al. 1989), and cloned into the vector λ ZAP II (Stratagene). The total number of recombinant clones obtained was 8×10^5, as assessed with the iPTG X-Gal screening procedure. Plaque lifts (10^4 phages/plate) were screened with $(GA)_{15}$, $(GT)_{15}$, $(AT)_{15}$, $(TTAG)_8$, $(TCT)_{10}$, $(GTAG)_7$ and $(GAGG)_7$ (^{32}P)-end-labeled synthetic oligonucleotides. Hybridizations were conducted overnight at 42–50°C in 6 × SSC, 0.5% SDS, 1 mM EDTA pH 8.0, 0.1% powdered milk and 10 mM of monobasic/dibasic potassium phosphate. One to three filter washes were performed at 42–50°C for 15 min each in 6 × SSC and 0.1% SDS. Positive clones were sequenced after at least one more round of screening. DNA sequencing was performed on an automated sequencer (Perkin Elmer/Applied Biosystems model 373 A) using T3 and/or T7 primers.

Primer design, PCR conditions and electrophoresis

Primers were designed for microsatellites containing at least ten perfect or imperfect repeats, using PRIMER version 0.5 software (Whitehead Institute for Biomedical Research) and/or by visual inspection of the sequences. Nonradioactive PCR reactions (100 µl) contained 0.1–0.6 µg of genomic DNA, 0.2 µM of each primer in 10 mM TRIS-HCl, 50 mM KCl, 1.5 mM $MgCl_2$, 0.01% gelatin, 200 mM of each dNTP and 0.5–2.5 U of Ampli*Taq* DNA polymerase (Perkin Elmer/Cetus). Microsatellites were amplified with a Perkin-Elmer (Cetus) 480 or a PTC-100 Hot Bonnet (MJ-Research) thermal cycler. An initial denaturation cycle was done at 94°C for 2 min, followed by 30 cycles of amplification by denaturing at 94°C for 1 min, annealing at 56 or 45°C for 1 min, and extension at 72°C for 2 min. The final step was a single extension cycle at 72°C for 5 min. Nonradioactive PCR products were run on 5% metaphor agarose gels following recommendations from the supplier (FMC Bioproducts).

Radioactive PCR reactions (10 µl) followed the protocol described above with 0.1–0.2 µg of genomic DNA/reaction and (^{32}P)-end-labeled primers. PCR products were run on 6% polyacrylamide sequencing gels for a minimum of 3 h at 80 W. Since all microsatellites isolated contained dinucleotide GA-repeats (see Results), microsatellite allele sizing was visually estimated based on the sequence-determined size of the original clone from cassava accession Mcol22, and assuming a 2-bp difference between the most intense band and stutter bands.

Semi-automated genotyping using fluorescent primers

Primers labeled with fluorescent dyes were commercially synthesized by Perkin Elmer/Applied Biosystems or Research Genetics (Foster City, Calif.; Huntsville, Ala.). To confirm the accuracy of semi-automated allele sizing, four primer sets (GA-21, GA-126, GA-134, GA-136) were fluorescently labeled to perform quadriplex PCR amplifications. Data from ten independent gel runs were

collected from PCR-amplified products of cassava accession MCol22. The range, mean and standard deviation of inter-gel allele sizes were determined for each locus. Allele sizes were then compared to those obtained by sequencing the original microsatellite clones from the same accession.

Multiplex PCR amplifications were carried out in 25-µl volumes containing 25 ng of target DNA, 5 pmol of each primer, 1 × Perkin Elmer PCR buffer II, 1.5 mM of $MgCl_2$, 0.25 mM of dNTPs and 0.5 U of Ampli*Taq*-Gold polymerase (Perkin Elmer). Temperature cycling was done on a GeneAmp PCR System 9600 (Perkin Elmer) using 1 s ramp times. The amplification profile for all PCR reactions using fluorescently labeled primers consisted of: one polymerase activation cycle at 94°C for 10 min; one DNA denaturation cycle at 95°C for 4 min; 25 amplification cycles at 95°C for 1 min, 55°C for 2 min and 72°C for 2 min; and a final elongation cycle at 72°C for 10 min.

After PCR amplification, 1.5 µl of PCR product were mixed with 0.5 µl of GeneScan 500 internal lane standard labeled with TAMRA (Perkin Elmer/Applied Biosystems) and 2.0 µl of 50% formamide. The mixture was incubated at 95°C for 5 min, and 1–2 µl were loaded on 6% denaturing (7 M urea) acrylamide : bisacrylamide (19 : 1) gels (12 cm well-to-read). The samples were separated by electrophoresis in 1 × TBE at 29 W for a minimum of 3 h on an automatic DNA sequencer (Perkin Elmer/Applied Biosystems model 373 A with GeneScan 672 software version 1.2.1). Fragment sizes were calculated using the "local Southern" algorithm (Elder and Southern 1987).

Heterozygosity

To estimate the informativeness of the 14 microsatellites in cassava, heterozygosity (h) values were calculated from single gel runs of PCR-amplified products from accessions of the cassava core collection. If (^{32}P)-end-labeled primers were used to PCR-amplify microsatellites, 48 accessions were analyzed for h determination. With fluorescently labeled primers, 522 accessions were analyzed to estimate h (see Table 3). Heterozygosity was calculated according to the formula: $h = 1 - \Sigma(p_i)^2$ (Nei 1978; see also Rongwen et al. 1995 for application), where h represents the probability that two alleles from the same locus would be different if chosen at random in the population (Li and Graur 1991), and (p_i) is the frequency of the i^{th} allele at a locus. h gives an indication of the polymorphism of a marker in a given population.

Segregation and linkage analysis

Chi-square tests ($\alpha = 0.01$, with $n - 1$ degrees of freedom, where n = number of phenotypic classes) were performed to ascertain the significance of segregation distortion and linkage of 11 microsatellite loci. The tests were performed using 83 F_1 individuals of the cassava mapping population. For segregation analysis, two kinds of segregation ratios were expected: a backcross-like or an F_2-like ratio. In the former, microsatellite alleles segregate from only one of the parents in a 1 : 1 ratio, while in the latter both parents are heterozygous for the microsatellite and segregation results in a 1 : 2 : 1 or 1 : 1 : 1 : 1 ratio. Linkage analysis was performed using the gamete segregation ratios for pairs of microsatellite loci for each parent in the case of backross-like expected ratios, and for both parents in the case of F_2-like expected ratios.

Results and discussion

Microsatellite characterization and primer design

A size-fractionated cassava DNA library was screened with seven different oligonucleotide probes to detect microsatellites. Sixty positive clones were purified and sequenced to confirm the expected repeats. Thirty-two clones had the expected repeats, while the rest contained unexpected or no repeats. The majority of the microsatellites detected contained GA-repeats, which were variable in length and composition, some having up to 19 perfect repeats. Microsatellites containing GA-repeats are common in many plants (Lagercrantz et al. 1993; Terauchi 1994; Wang et al. 1994; Depeiges et al. 1995), which may also be the case for cassava. The isolation of other types of microsatellites may require the screening of libraries enriched for specific repeats.

Primers were designed for 22 microsatellites, although only those primers successfully amplifying the expected products are reported in Table 1. Unless otherwise indicated, all primers were used to amplify cassava DNA at an annealing temperature of 56°C.

Table 1 Forward and reverse primer sequences for 14 cassava microsatellites

Microsatellite name	5′ to 3′ primer sequences
GAGG-5	TAATGTCATCGTCGGCTTCG GCTGATAGCACAGAACACAG
GA-12	GATTCCTCTAGCAGTTAAGC CGATGATGCTCTTCGGAGGG
GA-13	TTCCCTCGCTAGAACTTGTC CTATTTGACCGTCTTCGCCG
GA-16	GTACATCACCACCAACGGGC AGAGCGGTGGGGCGAAGAGC
GA-21	GGCTTCATCATGGAAAAACC CAATGCTTTACGGAAGAGCC
GA-57	AGCAGAGCATTTACAGCAAGG TGTGGAGTTAAAGGTGTGAATG
GA-123	CAATGCAGGTGAAGTGAATACC AGGGTGCTCTTCAGAGAAAGG
GA-126	AGTGGAAATAAGCCATGTGATG CCCATAATTGATGCCAGGTT
GA-127[a]	CTCTAGCTATGGATTAGATCT GTAGCTTCGAGTCGTGGGAGA
GA-131	TTCCAGAAAGACTTCCGTTCA CTCAACTACTGCACTGCACTC
GA-134	ACAATGTCCCAATTGGAGGA ACCATGGATAGAGCTCACCG
GA-136	CGTTGATAAAGTGGAAAGAGCA ACTCCACTCCCGATGCTCGC
GA-140	TTCAAGGAAGCCTTCAGCTC GAGCCACATCTACTCGACACC
GA-161	TGTTCTTGATCTTCTGCTGCA TGATTGTGGACGTGGGTAGA

[a] Annealing temperature 45°C

Table 2 Sizing of four fluorescently labeled microsatellite loci amplified in a quadriplex PCR reaction from cassava cultivar MCol22 and run on 12-cm denaturing gels using a 373A DNA sequencer (Perkin Elmer/Applied Biosystems). Sizing is relative to the GeneScan 500 internal standard (TAMRA) using the "local Southern" algorithm. Inter-gel allele size differences are summarized for ten independent gel runs

Locus	Size range) (nt)	Inter-gel size differences (nt)	Mean size (nt)	Standard deviation (nt)	Size in nt of the sequenced clone[a]
GA-21					
Allele #1	106.23–107.79	1.56	107.19	0.4162	114
Allele #2	114.02–115.63	1.61	114.98	0.4764	
Allele #3	119.65–121.82	2.17	120.79	0.6329	
GA-126					
Allele #1	182.28–183.25	0.97	182.68	0.3001	182
Allele #2	188.13–188.76	0.63	188.49	0.2094	
Allele #3	189.78–190.51	0.73	190.26	0.2467	
GA-134					
Allele #1	307.82–308.56	0.74	308.22	0.3343	318
Allele #2	317.65–318.89	1.24	318.52	0.3946	
GA-136					
Allele #1	151.18–151.95	0.77	151.69	0.2296	158
Allele #2	155.42–155.84	0.42	155.66	0.1535	
Allele #3	159.12–159.68	0.56	159.50	0.1994	
		Mean = 1.0363		Mean = 0.3267	

[a] Sequence of original clone from accession Mcol22. (nt) nucleotides

Semi-automated genotyping

The results obtained for inter-gel allele size comparisons are summarized in Table 2. Only inter-gel allele size differences are reported since they are usually higher than intra-gel size differences according to Smith (1995) and Mitchell et al. (1997). Mitchell et al. (1997) report that the maximum inter-gel allele size difference should not exceed 0.5 nucleotides, since alleles can be classified incorrectly if sizes are rounded off to the next integer. This problem is accentuated when DNA fragment sizes vary by one nucleotide, due either to real size differences or to the terminal transfer activity of *Taq* DNA polymerase. The literature reports inter-gel size differences of 0.7 to 1.3 nt (see Mitchell et al. 1997 and references therein), which are similar to the ranges reported here, 0.42 to 1.61 nt, except for one allele of locus GA-21, where successive determinations varied up to 2.17 nt. This discrepancy may be due to the terminal transfer activity of *Taq* DNA polymerase, and gel or electrophoresis detection conditions that make the sizing of alleles less precise on 373 A sequencers (Ghosh et al. 1997). Deletion and insertion events that occur in the clone once the cloning vector is introduced into *Escherichia coli* for replication may also explain large size differences between the sequenced clone and the PCR-amplified product (Freund et al. 1989, in Mitchell et al. 1997). Dinucleotide repeats are difficult to score due to overlapping alleles. However, the use of external adjustment and binning algorithms, like those reported by Ghosh et al. (1997), may be used to reduce inaccuracies in allele sizing.

Cassava germplasm characterization is often limited by budget constraints. Thus, unequivocal identification of each accession is essential to avoid duplicate samples, which increases germplasm maintenance costs. The set of microsatellites described here, combined with appropriate allele-sizing algorithms and multiplex, semi-automated genotyping systems, could be used to characterize cassava accessions from the core collection or from other collections. This would expedite the process of selecting representative sets of germplasm for conservation, characterization or breeding.

Heterozygosity

To determine how useful microsatellites would be for the fingerprinting or characterization of cassava germplasm and for mapping purposes, heterozygosity (h) estimates were made using (^{32}P)-end-labeled or fluorescently labeled primers. The number of alleles detected for microsatellite loci ranged from 1 to 15 (see example in Fig. 1), and h varied from 0.00 to 0.88 (Table 3). An example of a low h value was microsatellite GA-123, for which only one fluorescently labeled PCR product of equal size was detected in 3 out of 39 accessions analyzed; the other 36 accessions did not show detectable PCR products (data not shown), probably due to mutations in priming sites. However, more than 75% of the microsatellites analyzed had h values above 0.5, which makes them useful markers for cassava germplasm characterization and mapping. A study of the genetic diversity of cultivated and wild *Manihot* germplasm using microsatellites GA-12, GA-21 and GAGG-5, revealed between 8 and 18 alleles for each locus (Roa 1997) and detected 18 alleles for locus GA-12 alone, ten more than the eight reported here, probably due to the inclusion of seven *Manihot* taxa. Nevertheless, the

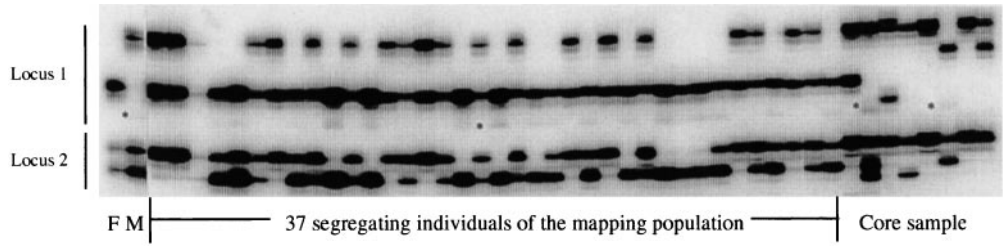

Fig. 1 Segregation and polymorphism of microsatellite GA-161. *F* and *M*, female and male parents respectively; Core sample: (left to; right) MBra71, MBra73, MBra356, MCol191, MCol474, MCR18, MEcu50 and MMex2. PCR amplifications were done using (^{32}P)-end labeled primers. (*) indicates a weakly amplified segregating allele from the male parent, which is also present in MBra73 and MCR18

number of alleles observed in Roa's study is comparable with the ones reported here and suggest microsatellite priming-site conservation among cassava and its wild relatives.

Assessment of microsatellite heterozygosity was done using small or large sets of cassava germplasm, though the number of alleles was naturally higher in the larger sample sets (Table 3). A 27% increase in the number of alleles was observed for microsatellite GA-134 when the number of accessions screened increased from 38 to more than 500 individuals. However, Table 3 shows that *h* values for microsatellites GA-21, GA-126 and GA-134 remained around 0.6, 0.8 and 0.5, respectively, regardless of sample size. This indicated that the small core subsamples fairly represented the variability existing in the larger core collection.

The set of cassava microsatellites reported here thus have heterozygosity values that may be useful for the germplasm characterization of cultivated cassava and its wild related species. This will undoubtedly contribute to the understanding of genetic relationships among cassava accessions, wild taxa, or both.

Microsatellite segregation and linkage analysis

Chi-square tests were performed on segregation ratios of 11 microsatellite loci that segregated in the mapping population from the female, the male, or both parents. Three microsatellites, GA-13, GA-57 and GA-123, were not polymorphic in the parents, yet the former two were polymorphic in other cassava accessions (data not shown). The results of the chi-square tests are summarized in Table 4. All microsatellite loci, except GA-131, fit expected 1 : 1, 1 : 2 : 1 or 1 : 1 : 1 : 1 disomic segregation ratios for one or two loci, a finding previously reported for isozyme and RFLP loci (Roca et al. 1992; Lefevre and Charrier 1993; Sarria et al. 1993; Fregene et al. 1997) and of significance to cassava breeding programs. Microsatellite GA-131 did not fit an expected 1 : 1 : 1 : 1 ratio, and exhibited an excess of alleles derived from the female parent (Table 4), although it has been mapped to linkage group G of the male- and

Table 3 Heterozygosity estimates for 14 cassava microsatellites using (^{32}P)-end-labeled and/or fluorescently labeled primers for the detection of PCR amplification products. A maximum of two alleles per accession was the general observation, except for loci GA-161 and GA-127 for which up to five alleles per accession were detected. No null alleles were considered for *h* estimations, which may result in underestimation of *h*

Microsatellite	# Accessions scored for *h*	# Alleles detected	Size range (nt)	Heterozygosity (*h*)
GA-12	39	6	131–157	0.54
GA-13	42	2	137–139	0.07
GA-16	41	6	89–129	0.63
GA-21	42	6	104–126	0.61
GA-21[a]	522	8[c]	107–121	0.65
GA-57	28	6[c]	153–183	0.67
GA-123	3[b]	1[c]	175	0.00
GA-126	34	8	178–214	0.85
GA-126[a]	521	12[c]	178–220	0.79
GA-127	34	10	203–239	0.74
GA-131	45	12	75–119	0.88
GA-134	38	4[c]	309–337	0.46
GA-134[a]	512	15[c]	308–338	0.55
GA-136[a]	522	9[c]	145–161	0.67
GA-140	35	5	154–164	0.76
GA-161	39	12	64–140	0.85
GAGG-5	28	6[c]	109–127	0.46

[a] Microsatellites amplified in quadriplex reactions
[b] Thirtynine accessions screened but only three showed a PCR product (see text)
[c] Alleles detected using fluorescently labeled primers, all others using (^{32}P)-labeled primers

Table 4 Goodness-of-fit χ^2 test of expected and observed disomic segregation ratios for 11 microsatellite loci in the F_1 cassava mapping population CM7857. Segregation analysis was done using (^{32}P)-labeled primers and sequencing gels ([a]), or by running nonradioactive PCR products on 5% Metaphor agarose gels stained with ethidium Bromide ([b]). The null hypothesis that there were no significant differences between observed and expected segregation ratios was tested at $\alpha = 0.01$ and $n - 1$ degrees of freedom, where n = number of phenotypic classes (if $n = 2$, 3 or 4, the critical regions are = 6.63, 9.21 or 11.34 respectively). GA-127 and GA-134 had two linked loci in each parent but only one locus is shown here[c]. GA-13, GA-57 and GA-123 did not segregate in the mapping population. (+ +) segregating, weakly-amplified allele from male parent for locus GA-161a

Microsatellite	Parental genotype TMS30572 (left) and CM2177-2	Genotypes classes: expected (top) and observed ratios	Total individuals scored	Calculated χ^2 and significance at a = 0.01
GAGG-5[b]	– – – –	– – – 1 : 1 33 : 42	75	1.08 NS
GA-12[a]	– – – –	– – – – – – – 1 : 1 : 1 : 1 20 : 30 : 19 : 14	83	6.49 NS
GA-16[a,b]	– – – –	– – – – 1 : 2 : 1 21 : 43 : 19	83	0.20 NS
GA-21[b]	– – –	– – – – 1 : 1 34 : 43	77	1.05 NS
GA-126[a]	– – –	– – – – 1 : 1 35 : 35	70	0.00 NS
GA-127[a,b]	– – – –	– – – – – – – 1 : 1 : 1 : 1 17 : 20 : 17 : 14	68	1.06 NS
GA-131[a,b]	– – – –	– – – – – – – 1 : 1 : 1 : 1 22 : 8 : 11 : 31	72	18.55 Significant
GA-134[c]	– – –	– – – 1 : 1 41 : 38	79	0.11 NS
GA-136[a]	– – –	– – – 1 : 1 36 : 43	79	0.62 NS
GA-140[a]	– – – –	– – – – 1 : 2 : 1 21 : 39 : 16	76	0.71 NS
GA-161a[a]	– – + +	– – – + + 1 : 1 31 : 48	79	4.01 NS
GA-161b[a]	– – – –	– – – – 1 : 2 : 1 14 : 40 : 25	79	3.21 NS

Table 5 Goodness-of-fit χ^2 test to estimate the significance of linkage between pairs of microsatellite loci that segregate from the female (top table), the male (middle table), or both parents (lower table). Critical values for $\chi^2_{0.01}$ with 2, 3 or 5 degrees of freedom are 9.21, 11.34 and 15.08 respectively. Bold numbers indicate values too close to, or larger than the critical values. Also shown are the linkage groups to which loci have been assigned

Microsatellite	GA-12	GA-126	GA-127	GA-131	GA-136	GA-161b	Linkage group
GAGG-5	1.74	1.62	4.44	2.57	3.94	6.61	Q
GA-12		0.35	2.06	0.66	0.87	4.08	nd
GA-126			**11.33**	4.20	0.80	5.56	K
GA-127				0.57	3.02	6.15	K
GA-131					1.79	4.81	G
GA-136						7.73	nd

Microsatellite	GA-21	GA-127	GA-131	GA-134	GA-161a	GA-161b	Linkage group
GA-12	4.284	7.623	**15.94**	3.98	8.94	6.51	Unnamed
GA-21		0.632	**11.74**	0.25	5.79	5.59	nd
GA-127			**14.63**	0.70	5.19	5.95	K
GA-131				**13.9**	**13.69**	13.32	G
GA-134					2.86	5.84	nd
GA-161a						**49.05**	nd
GA-161b							nd

Microsatellite	GA-140	GA-161b	Linkage group
GA-16	2.662	3.451	nd
GA-140		0.794	nd

[source: Fregene et al. (1997); this report]
(nd) not yet determined

female-derived cassava maps (Fregene et al. 1997; Table 5).

The cassava genetic linkage maps reported by Fregene et al. (1997) were developed based on the segregation of unique alleles in the gametes of both parents, which resulted in one linkage map for each parent. A total of five microsatellite loci, GA-12, GA-126, GA-127, GA-131 and GAGG-5, were assigned to at least three linkage groups in both maps (Fregene et al. 1997; Table 5). Here we tested for linkage between pairs of all segregating microsatellite loci using chi-square tests performed on the segregation ratios of gametes from both parents (Table 5). As reported by Fregene et al. (1997), linkage between GA-126 and GA-127 was also detected, although the observed chi-square value (11.33) was slightly lower than the critical value (11.34). These two loci have been assigned to group K of the female-derived cassava map approximately 30 cM from each other. In our study, no linkage was detected for any other pair of microsatellite loci in the female parent. However, in the male parent, microsatellite GA-131 appeared to be linked to five other loci, including GA-12 and GA-127, with observed chi-square values larger than the critical value. Yet, these three microsatellite loci have been mapped to three different linkage groups of the male-derived cassava map (Table 5). We assume that the segregation distortion observed for this microsatellite on the male-derived gametes may explain the discrepancy between our results and those of Fregene et al. (1997).

Mapping microsatellites in cassava helped to identify common linkage groups in the female- and male-derived maps for microsatellites GA-12, GA-127 and GA-131 (Table 5; Fregene et al. 1997). Microsatellite GA-12 revealed heterozygosity in both parents, although it has been mapped from the male parent only (segregation data was not available for the female parent), falling into a linkage group that has not been reconciled with the female-derived map. Success in mapping microsatellites like GA-12, with unique alleles segregating from both parents, would facilitate the recognition of common linkage groups from among those based on female- or male-derived segregation data.

Duplicated loci

Barring priming-site mutations, PCR amplification of highly polymorphic microsatellites should detect all alleles from duplicated loci in highly heterozygous, true auto- or allo-tetraploid species. However, in the case of the cassava, 78% (11/14) of the microsatellites detected only one or two alleles in most accessions, and segregation analysis of 11 microsatellites showed only one locus or two linked loci (Table 4). Microsatellite GA-161 was an exception, detecting up to five alleles per genotype (Fig. 1), though segregation analysis uncovered two linked loci. These results were unexpected considering that cassava is an outcrossing, very heterozygous crop of presumable allotetraploid origin (Magoon et al. 1969; Umannah and Hartman 1973). Although knowledge on the evolution of microsatellite loci in plant genomes is limited, these findings may suggest a low degree of duplication of microsatellite loci in cassava, which agrees with the low number of duplicated RFLP loci (less than 5%) found by Fregene et al. (1997). Duplicated microsatellite and RFLP loci are present in the cassava genome, though their origin

as random genomic duplications or polyplodization events remains unknown.

Acknowledgments This research was supported by funds from the USAID and the Rockefeller Foundation, for which the authors are most grateful. We are also grateful for the excellent technical assistance provided by the late Christopher A. Jester and by Dr. Sharon E. Mitchell, both from USDA-ARS, Plant Genetic Resources Conservation Unit of the University of Georgia, Griffin, Ga. We also thank Perkin Elmer/Applied Biosystems for providing fluorescently labelled primers.

References

Beckman JS, Soller M (1990) Toward a unified approach to genetic mapping of eukaryotes based on sequence-tagged microsatellite sites. Biotechnology 8 : 930–932

Beeching JR, Marmey P, Gavalda MC, Noirot M, Haysom HR, Hughes MA, Charrier A (1993) An assessment of genetic diversity within a collection of cassava (*Manihot esculenta* Crantz) germplasm using molecular markers. Ann Bot 72 : 515–520

Bertram RB (1993) Application of molecular techniques to genetic resources of cassava (*Manihot esculenta* Crantz – Euphorbiaceae): interspecific evolutionary relationships and intraspecific characterization. PhD thesis, University of Maryland

Broun P, Tanksley SD (1996) Characterization and genetic mapping of simple repeat sequences in the tomato genome. Mol Gen Genet 250 : 39–40

CIAT (1993) Cassava: the latest facts about an ancient crop. CIAT communication unit. AA 6713, Cali, Colombia

Dellaporta SL Wood J, Hicks JR (1983) A plant DNA minipreparation: version II. Plant Mol Biol Rep 1 : 19–21

Depeiges A, Goubely C, Lenior A, Cocherel S, Picard G, Raynal M, Grellet F, Delseny M (1995) Identification of the most represented repeated motifs in *Arabidopsis thaliana* microsatellite loci. Theor Appl Genet 91 : 160–168

Elder JK, Southern EM (1987) Computer-aided analysis of one-dimensional restriction fragment gels. In: Bishop MJ, Rawlings CJ (eds) Nucleic acid and protein sequence analysis – a practical approach. IRL Press, Oxford, UK, pp 165–172

FAO (1992) Agrostat database 1992

Fregene M, Vargas J, Ikea J, Angel F, Tohme J, Asiedu RA, Akoroda MO, Roca WM (1994) Variability of chloroplast DNA and nuclear ribosomal DNA in cassava (*Manihot esculenta* Crantz). Theor Appl Genet 89 : 719–727

Fregene M, Angel F, Gomez R, Rodriguez F, Chavarriaga P, Roca WM, Tohme J, Bonierbale MW (1997) A molecular genetic map for cassava (*Manihot esculenta* Crantz). Theor Appl Genet 95 : 431441

Freund A, Bichara M, Fuchs RPP (1989) Z DNA-forming sequences are spontaneous deletion hot spots. Proc Natl Acad Sci USA 86 : 7465–7496

Ghosh S, Karanjawala ZE, Hauser ER, Ally D, Knapp JI, Rayman JB, Musick A, Tannenbaum J, Te C, Shapiro S, Eldridge W, Musick T, Martin C, Smith JR, Carpten JD, Brownstein MJ, Powell JI, Whiten R, Chines P, Nylund SJ, Magnuson VL, Boehnke M, Collins FS, FUSION Study Group (1997) Methods for precise sizing, automated binning of alleles and reduction of error rates in large-scale genotyping using fluorescently labeled dinucleotide markers. Genome Res 7 : 165–178

González DO, Palacios N, Tohme J (eds) (1995) Protocolos para Marcadores Moleculares. Unidad de Investigación en Biotecnología, Centro Internacional de Agricultura Tropical (CIAT), Cali-Colombia, Publicación CIAT No. 258

Hall JM, LeDuc CA, Watson AR, Roter AH (1996) An approach to high-throughput genotyping. Genome Res 6 : 781–790

Hershey C, Iglesias C, Iwanaga M, Tohme J (1994) Definition of a core collection for cassava. In: International network for cassava genetic resources. Report of the first meeting of the international network for cassava genetic resources, CIAT, Cali, Colombia, 18–23 August 1992. International Crop Network Series No. 10. International Plant Genetic Resources, Rome, Italy, pp 145–156

Kochert G, Halward T, Branch DW and Simpson CE (1991) RFLP variability in peanut (*Arachis hypogaea* L.) cultivars and wild species. Theor Appl Genet 81 : 565–570

Kresovich S, Szewc-McFadden AK, Bliek SM, McFerson JR (1995) Abundance and characterization of simple sequence repeats (SSRs) isolated from a size-fractionated genomic library of *Brassica napus* L. (rapeseed). Theor Appl Genet 91 : 206–211

Lagercrantz U, Ellengren H, Anderson L (1993) The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. Nucleic Acid Res 21 : 1111–1115

Lefevre F, Charrier A (1993) Heredity of 17 isozyme loci in cassava (*Manihot esculenta* Crantz) germplasm. Euphytica 66 : 171–178

Li WH, Graur D (1991) Fundamentals of molecular evolution. Sinauer Assoc Inc, Sunderland, Massachusetts, USA

Magoon ML, Krishnan R, Bai KV (1969) Morphology of the pachytene chromosomes and meiosis in *Manihot esculenta* crantz. Cytologia 34 : 612–626

Marmey P, Beeching JR, Hamon S, Charrier A (1994) Evaluation of cassava (*Manihot esculenta* Crantz) germplasm collections using RAPD markers. Euphytica 74 : 203–209

Mitchell SE, Kresovich S, Jester CA, Hernandez J, Szewc-McFadden AK (1997) Application of multiplex and fluorescence-based, semi-automated allele sizing technology for genotyping plant genetic resources. Crop Sci 37 : 617–624

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

Roa AC (1997) Estimación de la diversidad genética en *Manihot* spp. mediante morfología y marcadores moleculares. Tesis de Maestría, Universidad del Valle, Facultad de Ciencias, Programa Magister en Ciencias-Biología. Cali, Valle, Colombia

Roa AC, Maya MM, Duque MC, Tohme J, Allem AC, Bonierbale MW (1997) AFLP analysis of relationships among cassava and other *Manihot* species. Theor Appl Genet 95 : 741–750

Roca WM, Henry G, Angel F, Sarria R (1992) Biotechnology research applied to cassava improvement at the International Center of Tropical Agriculture (CIAT). Agric Biotechnol News Inf 4 : 303N–308N

Roder MS, Plaschke J, König SU, Börner A, Sorrells ME, Tanksley SD, Ganal MW (1995) Abundance, variability and chromosomal location of microsatellites in wheat. Mol Gen Genet 246: 327–333

Rogers DJ (1963) Studies on *M. esculenta* Crantz cassava and related species. Bull Torrey Bot Club 9 : 43–54

Rongwen J, Akkaya MS, Bhagwat AA, Lavi U, Cregan PB (1995) The use of microsatellite DNA markers for soybean genotype identification. Theor Appl Genet 90 : 43–48

Saghai Maroof MA, Biyashev RM, Yang GP, Zhang Q, Allard RW (1994) Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal locations, and population dynamics. Proc Natl Acad Sci USA 91 : 5466–5470

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Sarria R, Ocampo C, Ramirez H, Hershey C, Roca WM (1993) Genetics of esterase and glutamate oxaloacetate transaminase isozymes in cassava (*Manihot esculenta* Crantz). In: Roca WM, Thro AM. Proc 1st Int Sci Meeting of the Cassava Biotechnology Network, Cartagena, Colombia, 25–28 August 1992. Centro Internacional de Agricultura Tropical, Cali, Colombia, Working document, no. 123

Schaal B, Olson P, Prinzie T, Carvalho LJCB, Tonukari NJ, Hayworth D (1995) Phylogentic analysis of the genus *Manihot* based on molecular markers. In: The cassava biotechnology network:

Proc 2nd Int Sci Meeting, Bogor, Indonesia, 22–26 August 1994. Centro Internacional de Agricultura Tropical, Cali, Colombia, Working document, no. 150

Smith RN (1995) Accurate size comparison of short tandem repeat alleles amplified by PCR. BioTechniques 18 : 122–128

Tautz D (1989) Hypervariability of simple sequences as a general source of polymorphic DNA markers. Nucleic Acids Res 17 : 6463–6471

Terauchi R (1994) A polymorphic microsatellite marker from the tropical tree *Dryobalanops lanceolata* (Dipterocarpaceae). Jpn J Genet 69 : 567–576

Thomas MR, Cain P, Scott NS (1994) DNA typing of grapevines: a universal methodology and database for describing cultivars and evaluating genetic relatedness. Plant Mol Biol 25 : 939–949

Umannah EE, Hartman RW (1973) Chromosome numbers and karyotypes of some *Manihot* species. J Am Soc Hort Sci 98 : 272–274

Wang Z, Weber JL, Zhong G, Tanksley SD (1994) A survey of plant short tandem repeats. Theor Appl Genet 88 : 1–6

Weber JL (1990) Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. Genomics 7 : 524–530

Zhao X, Kochert G (1993) Phylogenetic distribution and genetic mapping of a $(GGC)_n$ microsatellite from rice (*Oryza sativa* L.) Plant Mol Biol 21 : 607–614

Ziegle JS, Su Y, Corcoran KP, Nie L, Mayrand PE, Hoff LB, McBride LJ, Kronick MN, Diehl SR (1992) Application of automated DNA-sizing technology for genotyping microsatellite loci. Genomics 14 : 1026–1031