

Development and characterization of microsatellite markers for genetic analysis of a Brazilian endangered tree species *Caryocar brasiliense*

ROSANE GARCIA COLLEVATTI†§, ROSANA VIANELLO BRONDANI‡§ & DARIO GRATTAPAGLIA*§

†Departamento de Ecologia, Universidade de Brasília, Brasília 70910-900, D.F. Brasil, ‡Departamento de Biologia Celular, Universidade de Brasília, Brasília 70910-900, D.F. Brasil, and §Laboratório de Genética de Plantas, Recursos Genéticos e Biotecnologia—EMBRAPA, C.P. 02372, Brasília, D.F. Brasil

In this work we report the development and characterization of 10 microsatellite loci for the endangered tree species *Caryocar brasiliense*. Using genomic library enrichment, the efficiency of SSR marker development was 14.4% from sequencing data to operationally useful loci. Primer sequences for this set of 10 loci are made available together with their estimates of expected heterozygosity, probability of paternity exclusion and probability of identity. Mendelian inheritance and segregation was confirmed for all 10 loci in open-pollinated half-sib families as well as the absolute transferability of these 10 loci to five other species of the same genus. Number of alleles per locus ranged from 10 to 22 with a mean value of 16 and expected heterozygosity varying from 0.84 to 0.94. The combined probability of genetic identity was on the order of 10^{-17} clearly demonstrating that SSR multilocus genotypes are likely to be unique and capable of readily discriminating individuals of *C. brasiliense*. The very high combined probability of paternity exclusion (0.99999995) also indicates that these markers will permit detailed parentage studies in natural populations even in situations where both maternity and paternity are unknown. The battery of microsatellite markers developed and characterized in this study opens a new perspective for the generation of fundamental population genetic data for devising sound collection and conservation procedures for *C. brasiliense* and related species of the genus.

Keywords: *Caryocar brasiliense*, Caryocaraceae, Cerrado, microsatellites, tropical tree.

Introduction

Caryocar brasiliense Camb. (Caryocaraceae) is a widespread but endangered Brazilian Cerrado tree species, pollinated by small-sized nectarivorous bats. Seeds are surrounded by a woody endocarp coated by a yellow fleshy mesocarp rich in oil and vitamin A, which is eaten by several wild animals, such as greater rhea, macaws, parrots, pampas deer and paca (Gribel & Hay, 1993). Additionally, this species plays a significant role in the local economy of the inhabitants of central Brazil, who use the yellow mesocarp as an important source of oil for cooking and for home-made recipes for candies, ice-cream and liqueur.

The fragmentation of the 'Cerrado' vegetation in Brazil and the higher frequency of fire caused by

agricultural practices have been affecting recruitment and ultimately population size and dynamics of *C. brasiliense*, augmented by the intense commerce of its fruits. Habitat fragmentation may reduce genetic variability through genetic bottlenecks. Founder effect, genetic drift and restriction of gene flow, added to the enhancement of inbreeding, may increase population genetic isolation and divergence. Additionally, these genetic hazards may lead to fixation of deleterious alleles, endangering species persistence in habitat fragments (Gilpin & Soulé, 1986; Young *et al.*, 1996). It is therefore fully recognized that a number of genetic parameters such as inbreeding and outbreeding depression, genetic bottlenecks, loss of heterozygosity and adaptability have to be considered together with population demography in 'population vulnerability analysis' (PVA) and the estimation of 'minimum viable population' (MVP) (Gilpin & Soulé, 1986).

*Correspondence. E-mail: dario@cenargen.embrapa.br

Molecular markers have been increasingly used as a very effective tool for the understanding of population genetic structure, gene flow, parentage, population viability and, ultimately, to quantify the effects of habitat fragmentation and to guide conservation strategies (Young *et al.*, 1996; Parker *et al.*, 1998). DNA polymorphisms based on SSR (Simple Sequence Repeats or microsatellites) are one of the most powerful molecular markers to estimate genetic parameters of populations and understand detailed patterns of gene flow and parentage composition. Microsatellites display a very high content of genetic information, as they are codominant and highly multiallelic, with expected heterozygosity values typically above 0.7. Furthermore they are abundant, uniformly distributed in plant genomes, and typically transportable among closely related species because of genome sequence homology (Morgante & Olivieri, 1993). Additionally, it has been demonstrated that some classes of SSR constitute an important source of quantitative genetic variation, coding for functional elements of protein and acting as regulatory elements of transcription (Kashi *et al.*, 1997). In plants, SSR motifs of many types are as frequently and widely distributed as in human and other mammalian genomes (Tautz & Renz, 1984; Weber, 1990; Morgante & Olivieri, 1993; Wang *et al.*, 1994). Despite the usefulness of microsatellite makers for the investigation of population genetics and conservation, reports on the development, characterization and use of SSR loci in tropical tree species are still scarce (but see Condit & Hubbell, 1991; Chase *et al.*, 1996; White & Powell, 1997a; Aldrich *et al.*, 1998). This is partly because of the tradition that isozyme polymorphisms still hold and the often not so accessible molecular technologies needed to develop a battery of polymorphic microsatellite markers.

We are interested in understanding the population genetic structure, patterns of gene flow and mating system of *C. brasiliense*, in order to generate useful information for conservation strategies. As part of this project we report here the development, characterization and inheritance of a battery of highly polymorphic SSR loci in *C. brasiliense*. Besides estimating the genetic information content of this set of markers for the study of genetic structure and parentage analysis, we investigated the transferability of these loci to other species of the same genus.

Materials and methods

Plant material and DNA extraction

For SSR development total genomic DNA was extracted from expanded leaves of a single individual tree of *C. brasiliense*, sampled at Água Limpa Forestry Park

(15°57'12''S, 47°56'35''W), Brasília, Brazil. For SSR loci characterization, at least 30 individuals per population from four populations 200–1000 km apart (a total of 123 individuals) were used. These four populations were: (1) Campus of the Federal University of Mato Grosso do Sul (20°30'24''S, 54°36'53''W), Campo Grande; (2) Itirapina Ecological Reserve (22°13'13''S, 47°51'03''W), São Paulo; (3) Brasília National Park (15°44'26''S, 47°59'19''W), Brasília; (4) Grandes Sertões Veredas National Park (15°13'29''S, 45°49'12''W), Minas Gerais. Genomic DNA extraction from expanded leaves followed standard CTAB procedure (Doyle & Doyle, 1987) both for SSR development and genotyping experiments.

Construction of SSR-enriched genomic libraries

Protocols described by Rafalski *et al.* (1996) and optimized for tropical tree genomes at Embrapa — Genetic Resource and Biotechnology (Brondani *et al.*, 1998) were used. DNA from an individual of *C. brasiliense* was digested with three different restriction enzymes, *Mse*I, *Tsp*509 and *Sau*3A, according to manufacturer's instructions, in order to select one that would produce a larger amount of fractionated DNA in the range of 280–600 bp. Approximately 50 µg of genomic DNA was digested with *Mse*I (TTAA), and fragments between 280 and 600 bp were recovered by DEAE-cellulose NA-45 membrane (Schleicher and Schuell, NY) via electrophoresis on 2% agarose gel. Around 30 µg of DNA fragments were ligated to adaptors to the *Mse*I restriction site. Fragments containing SSR sequences were selected by hybridization with biotinylated oligonucleotides complementary to the repetitive sequence AG/CT, and recovered by magnetic beads linked to streptavidine. Fragments were amplified by PCR and cloned in the plasmid vector pGEM-T (Stratagene, CA) and then transformed by electroporation into *E. coli* strain XL1-Blue and grown on ampicillin and tetracycline containing agar plates. Transformants were picked, streaked on 132-mm plates (100 per plate) and regrown at 37°C for 12 h. Duplicate plates containing colonies from these transformants were stamped onto positively charged nylon membranes (Hybond N, Amersham Pharmacia), grown, lysed, denatured, neutralized and UV cross-linked.

Selection of recombinants for repeat sequences

Recombinant colonies having SSR were identified by hybridization with a poly (dA-dG) probe labelled with Digoxigenin-11-ddUTP using a DIG oligonucleotide 3'-end labelling Kit (Boehringer Mannheim) according to the manufacturer's instructions. The temperature

used for prehybridization and hybridization was 65°C for the poly AG/TC oligonucleotide. Processed membranes were exposed to X-ray film for 2–3 h at 37°C.

Sequencing of positive clones and primer design

Positive clones were picked and grown overnight in liquid ampicillin LB media. Plasmid DNA was extracted with Wizard Minipreps (Promega Co., WI). DNA inserts were sequenced on an Applied Biosystems 377 (Perkin-Elmer, CA) instrument using dye-terminator fluorescent chemistry. Oligonucleotides complementary to the repeats were designed using the software 'PRIMER' (Lincoln *et al.*, 1991). To reduce problems with spurious banding patterns generated during amplification and to allow later development of single-reaction multiplex PCR, some stringent criteria in primer sequence design were applied: (i) primer T_m of 72°C; (ii) 3°C difference in T_m between primer pairs; (iii) GC content ranging from 40% to 60%; and (iv) absence of complementarity between primers. Primers were synthesized by Operon Technologies Inc. (Alameda, CA).

Primer screening and PCR amplification

Thirty-two individuals of *C. brasiliense*, randomly chosen from the 123 individuals (eight individuals per population), were used for primer screening. Microsatellite amplification was performed in a 13 μ L reaction mix containing 0.9 μ M of each primer, 1 unit Taq DNA polymerase, 200 μ M of each dNTP, 1% reaction buffer (10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM MgCl₂), DMSO 50% and 10.0 ng of template DNA. Amplifications were performed using a PT-100 thermal controller (MJ Research) with the following conditions: 96°C for 2 min (1 cycle), 94°C for 1 min, 54 or 56°C for 1 min (according to each primer pair annealing temperature), 72°C for 1 min (30 cycles); and 72°C for 7 min (1 cycle). Each primer pair was initially screened for product polymorphism, and the annealing temperature was later optimized to produce clear and robust DNA band amplification in all loci. Analysis of amplified fragments produced by these amplifications was carried out in 3.5% Metaphor gels stained with ethidium bromide (1.0 mg mL⁻¹) and sized by comparison to a 1 kb DNA ladder standard (Gibco, MD). For genotype determination and precise estimates of allele sizes, the amplified products were separated on 4% denaturing polyacrylamide gels stained with silver nitrate (Bassam *et al.*, 1991) and sized by comparison to a 10-bp DNA ladder standard (Gibco, MD) on a computer screen. Allele sizes were estimated using the software SEQAID II (Rhoads & Roufa, 1990) taking into consideration the

expected allelic series in base pairs from the primer designed and the original DNA clone from which the SSR locus was developed.

Analysis of inheritance and transferability of SSR loci

To verify the inheritance of the microsatellites developed, we examined the segregation in an open-pollinated half-sib family. Sixteen seeds were collected from a mother tree and DNA was extracted directly from the embryo because of the low germination of dormant seeds. For DNA extraction from embryos we used the Fast DNATM Kit H, and FP120 FastPrep Cell DisruptorTM (BIO101/Savant Instruments Inc., CA), according to manufacturer's instructions. PCR amplification and visualization of allele segregation followed the same protocols used for leaves.

To test the transferability of SSR loci we extracted DNA from leaves of eight individuals of five species of the same genus: *C. coriaceum*, from Riachão das Neves, Bahia, *C. edule*, Porto Seguro, Bahia, and *C. glabrum*, *C. pallidum* and *C. villosum*, from Manaus, Amazônia. PCR amplification followed the same protocol used for *C. brasiliense*. Transferability was visualized both in 3.5% agarose gels and on silver-stained 4% denaturing polyacrylamide gels.

SSR loci characterization

Ten selected SSR loci were characterized for number of alleles per locus, allelic frequency and observed and expected heterozygosities under Hardy–Weinberg (Nei, 1978), using the 123 individuals of *C. brasiliense*. Genetic analyses were carried out using the software Genetic Data Analysis (GDA; Lewis & Zaykin, 1998). Based on estimated allele frequencies, two parameters of genetic information content for parentage studies were estimated for each locus: (i) probability of genetic identity (I) (Paetkau *et al.*, 1995), which corresponds to the probability of two random individuals displaying the same genotype; and (ii) paternity exclusion probability (Q) (Weir, 1996), which corresponds to the power with which a locus excludes an individual tree of being the parent of an offspring. The combined probability of paternity exclusion, $QC = 1 - \prod(1 - Q_i)$ and the combined probability of genetic identity $IC = \prod I_i$ were also estimated for the combined battery of loci.

Results

SSR development

Digestion of the *C. brasiliense* genome with three different enzymes revealed that *Mse*I produced the most

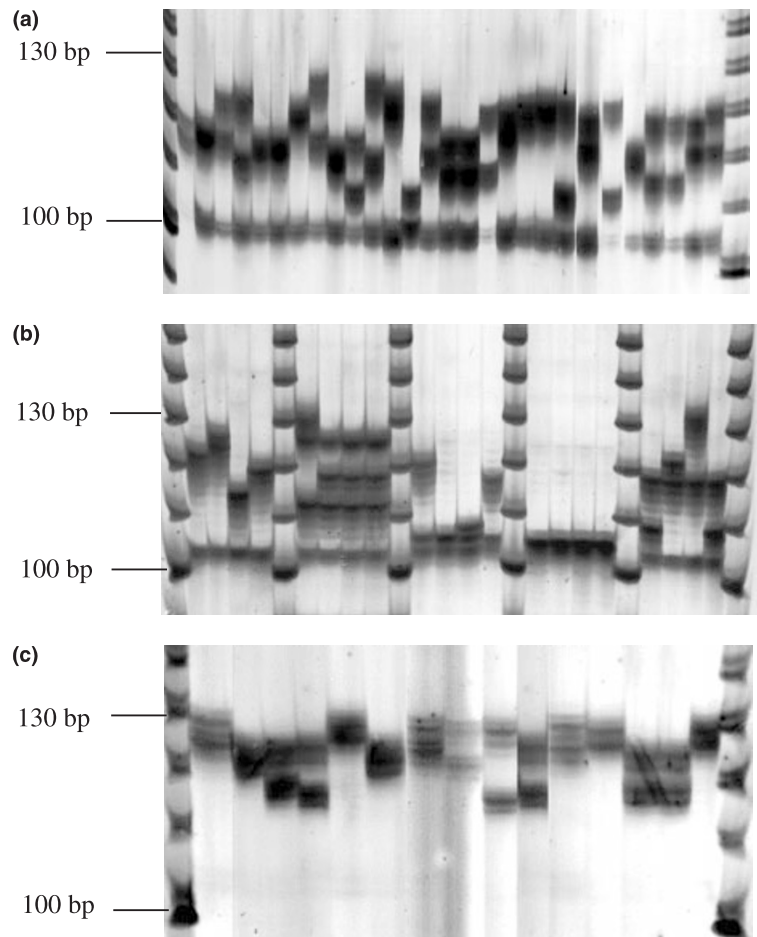


Fig. 1 Microsatellite polymorphisms in *Caryocar brasiliense* for locus *cb6* visualized in silver-stained denaturing polyacrylamide gels. First and last lanes are 10 bp ladder (Gibco). (a) DNA polymorphism for 28 unrelated individuals; (b) transferability of the locus to *C. coriaceum* (lanes 2–5), *C. edule* (lanes 7–10), *C. glabrum* (lanes 11–14), *C. pallidum* (lanes 16–19) and *C. villosum* (lanes 21–24); (c) inheritance and segregation in an open-pollinated half-sib family, lane 2 maternal tree followed by 14 progeny individuals.

adequate digestion profile for library development, with fragments ranging from 200 to 800 bp. One library enriched for AG microsatellite repeats was constructed. After the enrichment step, screening of 1000 recombinant colonies with the AG/TC probe detected 195 positive clones (19.5%). From the 195 positive clones sequenced, 28 were useful sequences (14.4%) from which primer pairs were designed. In 19 clones no SSRs were found in the sequences, suggesting that positive clones were misidentified. In 78 clones the repeated motif localization was too close to the end of the insert with not enough sequence for primer design. In 51 clones a high quality DNA sequence could not be obtained for both regions flanking the SSR. Finally, adequate primer sequences could not be designed from the DNA sequences of 19 clones.

Sequences of the DNA inserts containing microsatellites showed three categories of repeats as classified by Weber (1990). Twenty-two inserts contained single 'perfect' repeats (with no interruption in the repeat sequence) and four microsatellites were imperfect (with interruption in repeat sequence), but none of them was a clearly interpretable locus. Two microsatellites were

compound repeats (with different repeats in tandem), one of them was monomorphic, and the other was polymorphic. Primers were named using the prefix *cb* (from *Caryocar brasiliense*).

Screening of SSR and allele size determination

From 28 primer pairs developed for *C. brasiliense* 15 (54%) were amplified using a single PCR protocol and generated clearly interpretable products. Of these, five (33%) were monomorphic in a sample of 32 individuals and the other 10 were polymorphic showing clear allele size variation (Fig. 1a). From the 10 loci characterized, only *cb12* was a compound microsatellite (Table 1). The other nine were perfect microsatellites, with repeat motif size ranging from 18 to 28 (Table 1).

Inheritance, transferability and characterization

Inheritance was verified for all 10 SSR loci by analysing a heterozygous mother tree for the locus and its

Table 1 Primer sequences, repeat motifs, expected fragment size from sequencing data and observed size range in detected alleles, annealing temperature (T_a) and total number of alleles (A) for the 10 SSR loci developed for *Caryocar brasiliense*

Locus	Repeat motif	Primer pair sequences (5'–3')	Fragment size and range (bp)	T_a (°C)	A
<i>cb1</i>	(AG) ₂₇	ggTgTgAgCTTAgAgCTgAA gTCCAgCTTAATgTCCgACT	189 150–195	54	18
<i>cb3</i>	(AG) ₂₅	CAGCCATggTTCACgTTAgT CgCACATggAAACgCTTA	158 130–175	56	19
<i>cb5</i>	(AG) ₁₉	gTCAgAATgAAggCAGCTTg ATAgAATCCAaggCCACACCA	153 130–180	56	16
<i>cb6</i>	(AG) ₁₉	CTACCACAACCTCggAgACAA gACACTCCTgCAACTCCATT	123 105–160	56	14
<i>cb9</i>	(AG) ₂₁	ATCgAgATgAgCCAACCgAC ggAAggTgTTgCAGCACTgA	92 55–95	56	13
<i>cb11</i>	(AG) ₂₈	ggTCgTTATTgCTgTggT gTgAACATgAgCATCggT	176 135–185	56	15
<i>cb12</i>	(AG) ₂₁ (AC) ₉	GACATgTggCAATAggCggT TTgTgTgTgAAggTgTgTTgTgTT	179 150–210	56	20
<i>cb13</i>	(AG) ₁₅	AgCAgTTAgATgATgAAgTg CTATTgCCATATgTCgTAgT	145 125–170	54	15
<i>cb20</i>	(AG) ₂₂	TgACACAACCATCACATTCT gCAACTgTCgCAATAAACAA	164 140–185	56	17
<i>cb23</i>	(AG) ₁₈	ATACCAgCTCTgACAgAA AAgCCTgAgAgTAgAgAA	153 110–185	56	23

open-pollinated half-sib family (Fig. 1b). All sibs displayed one of the maternal alleles, confirming Mendelian inheritance and suggesting no seed contamination. Loci *cb6* and *cb12*, however, showed more than two alleles in the profile suggesting locus duplication possibly because of the ancient polyploid nature of the *Caryocar* genome (see below) or DNA or chromosome duplication. At locus *cb12*, interpretation of polymorphisms was not a problem as the second locus was monomorphic. However at *cb6* interpretation of genotypes for some individuals was impossible as alleles at both loci co-migrated to the same position in the gel. All 10 loci were fully transferable to the five species surveyed, displaying clear genotypes using the same protocols for PCR amplification (Fig. 1c).

Considering the 10 SSR loci analysed in the present work, a total of 123 individuals genotyped and the opened-pollinated half-sib family, the least and the most variable loci displayed 13 (*cb9*) and 23 alleles (*cb23*), respectively (Table 1). Progeny individuals genotyped in the inheritance analysis displayed alleles that were not detected in individuals genotyped in the characterization study (*cb1*, *cb6* and *cb23* — one allele, *cb9* — three, *cb13* — four). All loci presented three or four more frequent alleles (Fig. 2), except the most polymorphic loci (*cb12*,

cb20, *cb23*), which had a more uniform frequency distribution. For *cb9*, the locus with the lowest number of alleles, one allele (allele 60) represented more than 30.0% of the total. Paternity exclusion probabilities ranged from 0.69 to 0.95 with a combined value (over all loci) of 0.99999995 (Table 2). As expected, the two loci with the lowest number of alleles, *cb9* with 10 alleles and *cb13* with 11, displayed the lowest values of paternity exclusion probability and the highest values of probability of genetic identity (Table 2). Probability of genetic identity ranged from 0.01 to 0.4 with a combined value (over all loci) of 3.1×10^{-17} .

Discussion

Our results show that AG sequence repeats in the *C. brasiliense* genome are relatively abundant and therefore amenable to isolation for the development of microsatellite markers. Primer pairs that amplified easily-interpretable markers were developed for 14.4% of the sequenced plasmid clones from an enriched library. An anchor-PCR screening prior to sequencing could have significantly improved the yield of useful sequences, by eliminating false-positives and positive sequences with repeated motifs positioned too close to the vector

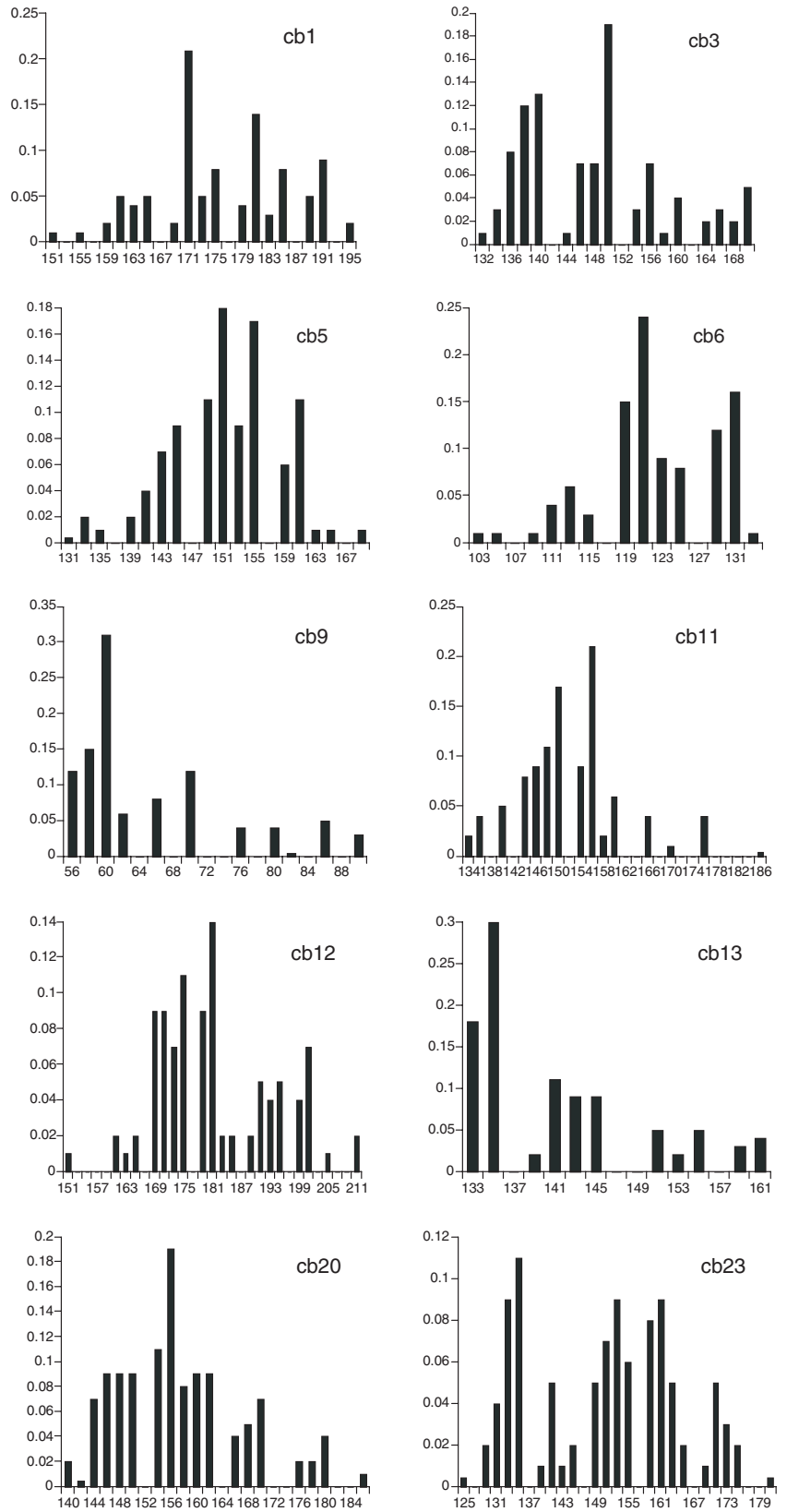


Fig. 2 Allele frequency distribution for the 10 SSR loci. *x*-axis, allele size in base pairs; *y*-axis, allele frequency.

Table 2 Characterization of 10 SSR loci of *Caryocar brasiliense*, based on a sample of 123 unrelated individuals

Locus	<i>A</i>	<i>H_e</i>	<i>H_o</i>	<i>Q</i>	<i>I</i>
<i>cb1</i>	17	0.90	0.83	0.79	0.02
<i>cb3</i>	19	0.91	0.95	0.79	0.02
<i>cb5</i>	16	0.89	0.71	0.79	0.02
<i>cb6</i>	13	0.86	0.81	0.74	0.04
<i>cb9</i>	10	0.84	0.72	0.70	0.04
<i>cb11</i>	15	0.88	0.77	0.80	0.03
<i>cb12</i>	20	0.94	0.56	0.84	0.01
<i>cb13</i>	11	0.84	0.59	0.69	0.04
<i>cb20</i>	17	0.91	0.64	0.95	0.02
<i>cb23</i>	22	0.94	0.69	0.85	0.01
Mean	16.0	0.89	0.73	<i>QC</i> = 0.99999995	<i>IC</i> = 3.1×10^{-17}

A, number of alleles; *H_o*, observed heterozygosity; *H_e*, expected heterozygosity; *Q*, probability of paternity exclusion; *QC*, combined probability of paternity exclusion; *I*, probability of genetic identity; *IC*, combined probability of genetic identity.

(Brondani *et al.*, 1998). However, given the likelihood of observing high expected heterozygosity values at the SSR loci and considering the research objectives contemplated in this programme — population genetic structure — only a relatively small number of loci was needed, so that an intense screening of positive clones prior to sequencing was deemed unnecessary.

Microsatellite polymorphisms in *C. brasiliense* were detected by silver staining. Attempts were made to use fluorescence labelling and semiautomated detection of SSR loci. Both methods offer advantages and limitations. Silver staining detection on polyacrylamide gels is an accessible technology and allows the genetic analysis of species that show nonspecific fragment amplification because of low quality of DNA, as these products usually migrate off the allele size range. On the other hand, rapid data generation by multiplexing of several loci in a single gel lane is limited. Fluorescence-based DNA detection offers the potential of multifluorescence multiplex genotyping ability and precise allele size determination (Mitchell *et al.*, 1997). On the other hand, fluorescence detection is not so accessible as silver staining because of the high cost of equipment for detection and the high quality of DNA required. This last aspect was the limiting factor for employing fluorescence labelling and detection of SSR loci in *C. brasiliense*. The DNA obtained from leaves and seeds of *C. brasiliense* typically was contaminated with polysaccharides and polyphenols that were hard to remove and seriously affected PCR. As several attempts were made to optimize conditions for fluorescence detection and yielded unsatisfactory results, we chose to perform the genotyping work on silver-stained gels, because this method was found to

yield robust data and to be significantly less influenced by DNA contamination. Moreover, adoption of this marker technology by other research groups in tropical countries should be significantly more straightforward if based on silver staining.

The most variable locus in *C. brasiliense* (*cb23* with 23 alleles) was the shortest one in number of repeat units, displaying only 18 AG repeats. Although the number of microsatellite loci surveyed in this study was limited, this observation does not support the view that the number of alleles per locus is positively correlated with the number of repeat motifs (Weber, 1990; Taramino & Tingey, 1996). In fact, this relationship is controversial because the size of a nonrepeat portion of the amplified fragment may be different among loci (Valdes *et al.*, 1993; Goldstein & Pollock, 1997).

Transferability of microsatellite loci between closely related species is a consequence of the homology of flanking regions of simple sequence repeats. Other studies in tropical trees have already demonstrated the high rate of transferability of SSR loci among taxonomically related tree species, such as in the Leguminosae (Dayanandan *et al.*, 1997), Meliaceae (White & Powell, 1997b) and among *Eucalyptus* species (Brondani *et al.*, 1998). The absolute transferability (100%) of the microsatellite loci developed for *C. brasiliense* to five other species of the genus (*C. coriaceum*, *C. edule*, *C. glabrum*, *C. pallidum* and *C. villosum*) indicates a high level of genome homology and will allow comparative studies of population genetic structure in all these species. *Caryocar brasiliense* and *C. villosum* have been described as having a high similarity in chromosome number and karyotype. Both are polyploid, as are most of the species of the order Theales, with $2n = 46$ (Ehrendorfer *et al.*, 1984). Despite polyploidy and therefore potential locus duplication, most of the microsatellite loci showed amplification from a unique site with the exception of loci *cb6* and *cb12*. This result suggests that the polyploidization event is a relatively ancient one and that sufficient time has passed to allow sequence divergence of the duplicated genomes. Alternatively, allopolyploidy might have occurred between species with disparate genomes such that only one set of homeologues contains a site that can be amplified.

Relatively high levels of multiallelism were observed at all 10 SSR loci developed. Mean number of alleles per locus (16.0, for 10 loci) and expected heterozygosity range (0.84–0.94) were higher than those found by White & Powell (1997a) for *Swietenia humilis*, an endangered tropical hardwood species in Central America (9.7 alleles per locus for 10 loci). The broad range observed in expected heterozygosity values results from the broad variation in number of alleles per locus, and allele frequency distribution within populations.

Loci with smaller numbers of alleles or with a skewed frequency distribution such as *cb9* and *cb13*, tend to have lower heterozygosity values and consequently lower probability of paternity exclusion, and higher probability of genetic identity. The number of alleles per locus reported in this study is most likely a minimum value. Because of the widespread distribution of the species in Brazil the number of alleles should increase when new populations are sampled. Indeed, new alleles previously undetected in the adults were seen when analysing the progeny individuals.

The combined probability of genetic identity, i.e. the probability that two individuals drawn at random from a population have identical multilocus genotypes at all 10 loci, was on the order of 10^{-17} . This clearly demonstrates that SSR multilocus genotypes will be unique and capable of readily discriminating individuals of *C. brasiliense*. This excellent power of discrimination is a very useful tool to identify precisely clonality in natural populations. In *C. brasiliense* this battery of microsatellite markers should therefore allow the precise identification of clonal regeneration arising by root sprouting, a common event in some ecosystems and with important implications for conservation strategies. The very high combined power of paternity exclusion (0.99999995) also indicates that these markers will permit detailed parentage studies in natural populations, even in situations where both maternity and paternity are unknown *a priori*. Furthermore, the exact determination of parentage of regenerants and seeds will allow a precise understanding of reproductive success of adults and the dynamics of genetic structure of natural populations. In conclusion, the microsatellite markers developed and characterized in this study open a new perspective for generating fundamental data to devise sound conservation procedures for *C. brasiliense* and related species of the genus.

Acknowledgements

This study is part of the Doctoral Thesis of the first author, and has been supported by the Universidade de Brasília, WWF (World Wildlife Fund), CNPq/PADCT (Brazilian Ministry of Science and Technology) competitive grant 62.00059/97-4 to D.G. and CAPES (National Council for Research and Technology) whose assistance we gratefully acknowledge. We wish to thank John Hay for continued support, Paul Lewis and Dmitri Zaykin for the authorization for the use of the GDA software as well as Antônio Dimas Costa Jr, Fabian Borghetti, Flavia Pinto, Frederico S. Lopes, Galdino R. Santos, Maura C. Almeida and Neilton R. Oliveira for their kind help in field work. We also acknowledge the support of IBAMA (Brazilian Ministry of Environment) and IEF-SP (State

Forest Institute of São Paulo) for authorizing plant collections in Brasília and Grandes Sertões Veredas National Parks and in Itirapina Ecological Station.

References

- ALDRICH, P., HAMRICK, J. L., CHAVARRIAGA, P. AND KOCHERT, G. 1998. Microsatellite analysis of demographic genetic structure in fragmented populations of the tropical tree *Symphonia globulifera*. *Mol. Ecol.*, **7**, 933–944.
- BASSAM, B. J., CAETANO-ANOLLES, G. AND GRESSHOFF, P. M. 1991. Fast and sensitive silver staining of DNA in polyacrylamide gels. *Analyt. Biochem.*, **196**, 80–83.
- BRONDANI, R. P. V., BRONDANI, C., TARCHINI, R. AND GRATTA-PAGLIA, D. 1998. Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. *Theor. Appl. Genet.*, **97**, 816–827.
- CHASE, M., KESSELL, R. AND BAWA, K. 1996. Microsatellite markers for population and conservation genetics of tropical trees. *Am. J. Bot.*, **83**, 51–57.
- CONDIT, R. AND HUBBELL, H. 1991. Abundance and DNA sequence of two-base repeat regions in tropical tree genomes. *Genome*, **34**, 66–71.
- DAYANANDAN, S., BAWA, K. S. AND KESSELL, R. 1997. Conservation of microsatellites among tropical trees (Leguminosae). *Am. J. Bot.*, **84**, 1658–1663.
- DOYLE, J. J. AND DOYLE, J. L. 1987. Isolation of plant DNA from fresh tissue. *Focus*, **12**, 13–15.
- EHRENDORFER, F., MORAWETZ, W. AND DAWE, J. 1984. The neotropical angiosperm families Brunelliaceae and Caryocaraceae: first karyosystematical data and affinities. *Pl. Syst. Evol.*, **145**, 183–191.
- GILPIN, M. E. AND SOULÉ, M. E. 1986. Minimum viable populations: process of species extinction. In: Soulé, M. E. (ed.) *Conservation Biology, the Science of Scarcity and Diversity*, pp. 19–34. Sinauer Associates, Sunderland, MA.
- GOLDSTEIN, D. B. AND POLLOCK, D. D. 1997. Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. *J. Hered.*, **88**, 335–342.
- GRIBEL, R. AND HAY, J. D. 1993. Pollination ecology of *Caryocar brasiliense* (Caryocaraceae) in Central Brazil cerrado vegetation. *J. Trop. Ecol.*, **9**, 199–211.
- KASHI, Y., KING, D. AND SOLLER, M. 1997. Simple sequence repeats as a source of quantitative genetic variation. *Trends Genet.*, **13**, 74–78.
- LEWIS, P. AND ZAYKIN, D. 1998. *Genetic data analysis: computer program for the analysis of allelic data*. Version 1.0 (d12). Free program distributed by the authors over the internet from the GDA Home Page at <http://chee.unm.edu/gda/>.
- LINCOLN, S. E., DALY, M. J. AND LANDER, E. S. 1991. *'Primer' software*. Available from E. Lander, Whitehead Institute, Cambridge, MA.
- MITCHELL, S. E., KRESOVICH, S., JESTER, C. A., HERNANDEZ, C. J. AND SZEWC-McFADDEN, A. K. 1997. Application of multiplex PCR and fluorescence-based, semi-automated allele sizing technology for genotyping plant genetic resources. *Crop Sci.*, **37**, 617–624.

- MORGANTE, M. AND OLIVIERI, A. M. 1993. PCR-amplified microsatellites as markers in plant genetics. *Plant J.*, **3**, 175–182.
- NEI, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individual. *Genetics*, **89**, 583–590.
- PAETKAU, D., CALVERT, W., STIRLING, I. AND STROBECK, C. 1995. Microsatellite analysis of population structure in Canadian polar bears. *Mol. Ecol.*, **4**, 347–354.
- PARKER, P. G., SNOW, A. A., SCHUG, M. D., BOOTON, G. C. AND FUERST, P. A. 1998. What molecules can tell us about populations: choosing and using a molecular marker. *Ecology*, **79**, 361–382.
- RAFALSKI, J. A., MORGANTE, M., POWELL, W., VOGEL, J. M. AND TINGEY, S. V. 1996. Generating and using DNA markers in plants. In: Birren, B. and Lai, E. (eds) *Analysis of Non-Mammalian Genomes – A Practical Guide*, pp. 75–134. Academic Press, New York.
- RHOADS, D. D. AND ROUFA, D. J. 1990. *SEQAID II 3.80*. Molecular Genetics Laboratory, Kansas State University, New York.
- TARAMINO, G. AND TINGEY, S. 1996. Simple sequence repeats for germplasm analysis and mapping in maize. *Genome*, **39**, 277–287.
- TAUTZ, D. AND RENZ, M. 1984. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucl. Acids Res.*, **12**, 4127–4138.
- VALDES, A. M., SLATKIN, M. AND FREIMER, N. B. 1993. Allele frequencies at microsatellite loci: the stepwise mutation model revisited. *Genetics*, **133**, 737–749.
- WANG, Z., WEBER, J. L., ZHONG, G. AND TANKSLEY, S. D. 1994. Survey of plant short tandem repeats. *Theor. Appl. Genet.*, **88**, 1–6.
- WEBER, J. L. 1990. Informativeness of human (dC-dA)_n (dG-dT)_n polymorphisms. *Genomics*, **7**, 524–530.
- WEIR, B. S. 1996. *Genetic Data Analysis II*. Sinauer Associates, Sunderland, MA.
- WHITE, G. AND POWELL, W. 1997a. Isolation and characterization of microsatellite loci in *Swietenia humilis* (Meliaceae): an endangered tropical hardwood species. *Mol. Ecol.*, **6**, 851–860.
- WHITE, G. AND POWELL, W. 1997b. Cross-species amplification of SSR loci in the Meliaceae family. *Mol. Ecol.*, **6**, 1195–1197.
- YOUNG, A., BOYLE, T. AND BROWN, T. 1996. The population genetic consequences of habitat fragmentation for plants. *Trends Ecol. Evol.*, **11**, 413–418.