Genet Resour Crop Evol (2007) 54:1653–1657
DOI 10.1007/s10722-007-9277-7

SHORT COMMUNICATION

Cross-amplification and characterization of microsatellite loci for three species of *Theobroma* (Sterculiaceae) from the Brazilian Amazon

Maristerra R. Lemes · Thieme M. Martiniano · Vanessa M. Reis · Camila P. Faria · Rogério Gribel

Received: 9 March 2007 / Accepted: 3 September 2007 / Published online: 19 October 2007

Abstract This study reports on the cross-species amplification of 23 microsatellite markers previously developed for *Theobroma cacao* L. (Sterculiaceae), source of chocolate in three economically important Amazonian species of *Theobroma* (*T. grandiflorum, T. subincanum, T. sylvestre*). Thirteen of the 23 microsatellite loci tested were polymorphic across the three species at 2–13 alleles per locus. The observed heterozygosity per locus varied from 0.18 to 0.84 and expected heterozygosity ranged from 0.28 to 0.87. The high level of transferability and genetic information content of these microsatellite loci indicate their usefulness for population genetic, mating system and breeding studies of these economically important Amazonian fruit trees.

Keywords Amazon · Cocoa · Genetic diversity · SSR · Transferability · Tropical tree

Introduction

The genus *Theobroma* (Sterculiaceae) is exclusively Neotropical and comprises 22 species, ten of which occur in the Brazilian Amazon (Cuatrecasas 1964). The most widely cultivated and economically important species in the genus is *Theobroma cacao* L., source of cacao or chocolate. The chocolate industry is based on the selection of *T. cacao* cultivars, some of them polyembryonic and or originated by crossing with other species of the genus. In the Amazonia region *T. cacao* occurs spontaneously in the wild (Cheesman 1944); pulp from its fruits is used in juices and ice cream by local populations.

Other species of the genus are cultivated in home gardens or exploited from wild sources in the Amazonia. Fruit pulp from cupuaçu (*Theobroma grandiflorum* (Willdenow ex Sprengel) Schumann) is used in juices, desserts, ice cream, and cosmetics, and is considered one of the most profitable crops of the Amazonia region with excellent potential for agroforestry cultivation (Smith et al. 1998; Browder and Pedlowski 2000). Cupuaçu seeds can also be used for the production of high-quality chocolate. Other under-exploited cocoa species occurring in the Brazilian Amazon and cultivated in home gardens include *Theobroma sylvestre* Martius and *Theobroma subincanum* Martius (Schmidt 2003).

Microsatellites or SSRs (Simple Sequence Repeats) are polymorphic DNA loci that are one to six nucleotide repeat sequences wide and randomly dispersed in the genomes of all prokaryotes and eucaryotes (Litt and Lutty 1989; Tautz 1989). Owing to their high variability, these markers have been widely employed as a powerful tool in fields such as

M. R. Lemes (✉) · T. M. Martiniano · V. M. Reis · C. P. Faria · R. Gribel
Laboratório de Genética e Biologia Reprodutiva de Plantas, Instituto Nacional de Pesquisas da Amazônia, Avenida André Araújo, 2936, Manaus, AM 69083-000, Brazil
e-mail: mlemes@inpa.gov.br

genetic mapping, breeding, analyses of parentage, mating systems, gene flow, and conservation (Morgante and Olivieri 1993; Chase et al. 1996, Jarne and Lagoda 1996; Zane et al. 2002; Ellegren et al. 2004; Varshney et al. 2005). The main limitations to widespread use of microsatellites are the expensive and laborious cloning and screening procedures involved in their development. If phylogenetically related species present homologous SSR flanking sequences, it is possible that SSR primers developed for one species could be used to detect polymorphism in others. Several studies of tropical tree species have reported the utility of heterologous primers for the identification of SSR loci in closely related taxa (Dayanandan et al. 1997; White and Powell 1997; Collevatti et al. 1999; Roa et al. 2000; Zucchi et al. 2002; Braga et al. 2007).

Lanaud et al. (1999) reported on the isolation and characterization of 23 microsatellite loci in *Theobroma cacao* and tested their amplification for eight other species of the Sterculiaceae. Here we report on the ability of these markers to amplify SSR loci in three congeneric species (*Theobroma grandiflorum, T. subincanum*, and *T. sylvestre)* with the goal of making a set of polymorphic markers available to researchers for investigations of genetic diversity, breeding, and mating systems of these species.

Material and methods

Plant material

For the genetic analysis we collected leaves from 33 adult trees of *T. grandiflorum*, 31 of *T. subincanum*, and 33 of *T. sylvestre* from cultivated (*T. grandiflorum*) and natural (*T. subincanum* and *T. sylvestre*) populations located in the Manaus region in Central Amazonia. The leaves were dried in silica gel and stored at –20°C until DNA extraction.

DNA extraction

Total genomic DNA was extracted following the standard CTAB protocol (Doyle and Doyle 1987). DNA quantification was performed by comparison with known concentrations of a DNA standard (Lambda DNA) in 1% agarose gel stained with ethidium bromide.

Microsatellite analysis

Twenty-three pairs of microsatellite primers developed for *Theobroma cacao* (Lanaud et al. 1999) were tested in *T. grandiflorum*, *T. subincanum* and *T. sylvestre*. PCR amplification was carried out in a final reaction volume of 13 µl containing 0.9 µM of each primer, 1 U *Taq* DNA polymerase, 200 µM of each dNTP, 10 mM Tris–HCl pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, BSA (2.5 mg/ml), and 7.5 ng of genomic DNA. PCR amplifications proceeded according to the following protocol: 94°C for 5 min followed by 30 cycles of 94°C for 1 min; annealing temperature (°C) for each locus/species for 1 min (Table 1); 72°C for 1 min; and a final elongation step at 72°C for 7 min. Amplified fragments were visualized in 3.5% agarose gel and sized with 1 kb DNA ladder.

After PCR optimization, the loci that showed clear and robust band amplification in agarose gels were selected for analysis of polymorphisms in 4% polyacrylamide gel (PAGE) stained with silver nitrate (Creste et al. 2001). DNA ladder (10 bp) was used to estimate the allele sizes.

Data analysis

Polymorphic loci were characterized for each species considering the number of alleles per locus, expected and observed heterozygosities for each locus and averaged over all loci, using the software GDA (Lewis and Zaykin 2001). Hardy-Weinberg equilibrium was tested by Fisher's exact test. Two parameters of genetic information content were estimated for each locus and over all loci: (1) probability of genetic identity (I) (Paetkau et al. 1995), and (2) paternity exclusion probability (Q) (Weir 1996).

Results and discussion

From 23 microsatellite markers developed for *T. cacao*, 17 loci (74%) were successfully amplified for

Table 1 Characteristics of microsatellite loci for three *Theobroma* species from the Brazilian Amazon. (A) number of alleles per locus, (He) expected heterozygosity, (Ho) observed heterozygosity, (I) probability of genetic identity, and (Q) paternity exclusion probability. Ta = annealing temperature; n = number of adult individuals genotyped; np = no polymorphism

Locus	Alelle size range (bp)	*T. grandiflorum* (n = 33)						*T. subincanum* (n = 33)						*T. sylvestris* (n = 31)					
		Ta (°C)	A	He	Ho	I	Q	Ta (°C)	A	He	Ho	I	Q	Ta(°C)	A	He	Ho	I	Q
mTcCIR02	260–292	63	8	0.81	0.69	0.05	0.63	57	9	0.77	0.51	0.09	0.58	np	np	np	np	np	np
mTcCIR03	174–194	48	7	0.78	0.72	0.09	0.57	np	np	np	np	np	np	np	np	np	np	np	np
mTcCIR04	232–264	57	6	0.71	0.62	0.14	0.48	np	np	np	np	np	np	52	6	0.75	0.77	0.10	0.54
mTcCIR09	246–290	np	np	np	np	np	np	np	np	np	np	np	np	51	9	0.74	0.52	0.10	0.55
mTcCIR11	320–360	np	np	np	np	np	np	51	7	0.75	0.61	0.10	0.55	51	5	0.73	0.84	0.13	0.50
mTcCIR12	196–282	np	np	np	np	np	np	56	11	0.80	0.78	0.09	0.62	52	7	0.69	0.55	0.14	0.48
mTcCIR13	240–254	np	np	np	np	np	np	57	6	0.68	0.79	0.16	0.44	np	np	np	np	np	np
mTcCIR15	228–250	np	np	np	np	np	np	46	10	0.81	0.78	0.07	0.63	np	np	np	np	np	np
mTcCIR17	248–300	55	7	0.78	0.67	0.09	0.58	51	13	0.87	0.76	0.04	0.73	51	6	0.74	0.52	0.12	0.51
mTcCIR19	344–376	54	6	0.28	0.18	0.54	0.16	np	np	np	np	np	np	52	7	0.80	0.48	0.07	0.61
mTcCIR22	260–310	51	6	0.68	0.51	0.14	0.47	46	8	0.78	0.56	0.08	0.58	49	10	0.77	0.58	0.08	0.59
mTcCIR25	120–148	59	5	0.71	0.69	0.14	0.48	60	8	0.77	0.57	0.09	0.58	46	2	0.43	0.61	0.42	0.19
mTcCIR26	240–286	46	10	0.71	0.73	0.11	0.52	np	np	np	np	np	np	np	np	np	np	np	np
Mean			6.9	0.68	0.60	6.6×10^{-8}	0.9982		9	0.78	0.67	2.6×10^{-9}	0.9862		6.5	0.71	0.61	5.1×10^{-8}	0.9974

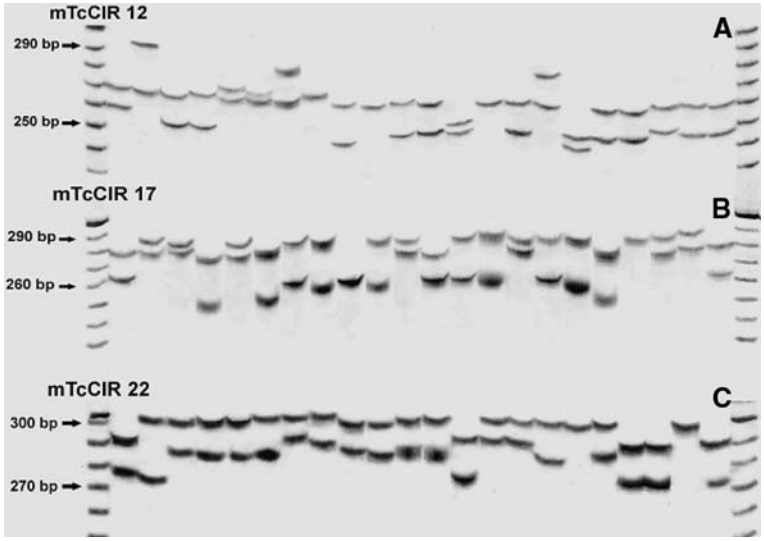

Fig. 1 Allelic variation for three microsatellite loci in *Theobroma subincanum* (**A**), *T. grandiflorum* (**B**) and *T. sylvestre* (**C**). First and last lanes in the gel show a 10 bp ladder

the three species (*T. grandiflorum, T. subincanum*, *T. sylvestre)*. Of these, 13 loci showed polymorphisms across the three species. For each species we found eight informative loci (47%). According to Peakall et al. (1998), studies on cross-species transferability of SSR loci within genera in plants have shown that the proportion of putative loci yielding amplified products that are informative (polymorphic) range from 20 to 100%.

Figure 1 shows allelic variation in polyacrylamide silver stained gels for the loci mTcCIR02, mTcCIR17 and mTcCIR22 for *T. subincanum*, *T. grandiflorum* and *T. sylvestre* respectively. The number of alleles per locus varied from 5 to 10 (mean 6.9) for *T. grandiflorum*, 6 to 13 (mean 9.0) for *T. subincanum* and 2 to 10 (mean 6.5) for *T. sylvestre*. The allelic diversity found here for *T. grandiflorum* was higher than that registered by Alves et al. (2006) for the same species analyzing 21 SSR loci (mean = 5.38, ranging from 2 to 11) over 214 individuals from seven populations.

The values for observed and expected heterozygosities across the three species ranged from 0.18 to 0.84 and 0.28 to 0.87, respectively (Table 1). The values for observed heterozygosities were higher in our study compared to Lanaud et al. (1999) and Alves et al. (2006). The mean observed heterozygosities per locus were slightly lower than mean expected heterozygosities in the three species, suggesting predominantly allogamous reproductive systems with low levels of inbreeding. Most loci deviated from Hardy–Weinberg proportions ($p < 0.05$) in all three species with the exception of loci mTcCIR25 for *T. grandiflorum* and mTcCIR15 for *T. subincanum*.

The combined probability of paternity exclusion (*QC*) indicates the probability of correctly excluding a random nonparent individual in the population over all loci, and is estimated based on allele frequencies (Weir et al. 1996). *QC* values were high for these three species (0.9982, 0.9862, 0.9974 for *T. grandiflorum, T. subincanum and T. sylvestre*, respectively). Values determined for the combined probability of genetic identity (*IC*)—the probability that two individuals selected at random from a population will have identical genotypes—were extremely low (6.6×10^{-8}, 2.6×10^{-9}, 5.1×10^{-8} for *T. grandiflorum, T. subincanum*, and *T. sylvestre*, respectively).

Results reported here demonstrate that these loci allow for precise discrimination among species and even individual trees, making these markers extremely useful for studies of breeding and mating systems, parentage, and gene flow of the study species.

Acknowledgements Financial support was provided by CNPq-Brazil (Grant no. 469926/2000-7 to RG), PROBIO/MMA-Brazil (Project contract # 0114-00/04), and European Union (GENEO-TROPECO Project contract # ERBIC18CT970149). We would like to thank Chris Dick and Jimmy Grogan for revision of the manuscript.

References

Alves RM, Sebbenn AM, Artero AS, Figueira A (2006) Microsatellite loci transferability from *Theobroma cacao* to *Theobroma grandiflorum*. Mol Ecol Notes 6: 1219–1221

Braga AC, Reis AMM, Leoi LT, Pereira RW, Collevatti RG (2007) Development and characterization of microsatellite markers for the tropical tree species *Tabebuia aurea* (Bignoniaceae). Mol Ecol Notes 7:53–56

Browder JO, Pedlowski MA (2000) Agroforestry performance on small farms in Amazonia: findings from the Rondonia Agroforestry Pilot Project. J Agroforestry Syst 49:63–83

Chase M, Kesseli R, Bawa K (1996) Microsatellite markers for population and conservation genetics of tropical trees. Am J Bot 83:51–57

Cheesman EE (1944) Notes on the nomenclature, classification and possible relationships of cocoa populations. Trop Agric 21:144–159

Collevatti RG, Brondani RV, Grattapaglia D (1999) Development and characterization of microsatellite markers for genetic analysis of a Brazilian endangered tree species *Caryocar brasiliense*. Heredity 83:748–756

Creste S, Tulmann-Neto A, Figueira A (2001) Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. Plant Mol Biol Rep 19:299–306

Cuatrecasas J (1964) Cacao and its allies: a taxonomic revision of the genus *Theobroma*. Contr US Natl Herb 35:379–614

Dayanandan S, Bawa K, Kesseli R (1997) Conservation of microsatellites among tropical trees (Leguminosae). Am J Bot 84:1658–1663

Doyle JJ, Doyle JL (1987) Isolation of plant DNA from fresh tissue. Focus 12:13–15

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. Nature Reviews Genetics 5:435–445

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. Trends Ecol Evol 11:424–429

Lanaud C, Risterucci AM, Pieretti I, Falque M, Bouet A, Lagoda PJL (1999) Isolation and characterization of microsatellites in *Theobroma cacao* L. Mol Ecol 8: 2141–2152

Lewis PO, Zaykin D (2001) Genetic data analysis: computer program for the analysis of allelic data. Version 1.0 (d16c). Free program distributed by the authors over the Internet from http://lewis.eeb.uconn.edu/lewishome/software.html

Litt M, Luty JA (1989) A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle action gene. Am J Hum Genet 444:397–401

Morgante M, Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. Plant J 3:175–182

Paetkau D, Calvert W, Stirling I, Strobeck C (1995) Microsatellite analysis of population structure in Canadian polar bears. Mol Ecol 4:347–354

Peakall R, Gilmore S, Keys W, Morgante M, Rafalski A (1998) Cross-species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants. Mol Biol Evol 15(10):1275–1278

Roa AC, Chavarriaga-Aguirre P, Duque MC, Maya MM, Bonierbale MW, Iglesias C, Tohme J (2000) Cross-species amplification of cassava (*Manihot esculenta*) microsatellites: allelic polymorphism and degree of relationship. Am J Bot 87:1647–1655

Schmidt MJ (2003) Farming and patterns of agrobiodiversity on the Amazon floodplain in the vicinity of Mamirauá, Amazonas, Brazil. MSc thesis, University of Florida, Gainesville, USA

Smith N, Dubois J, Current D, Lutz E, Clement C (1998) Agroforestry experiences in the Brazilian Amazon: constraints and opportunities. Pilot program to conserve the Brazilian Rain Forest, Ministério do Meio Ambiente, Recursos Hídricos e Amazônia Legal & World Bank, Brasília, DF, Brazil

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

Varshney RK, Graner A, Sorrells ME (2005) Genic microsatellite markers in plants: features and applications. Trends Biotechnol 23:48–55

Weir BS (1996) Genetic data analysis II. Sinauer Associates, Sunderland, MA, USA, pp 209–212

White GM, Powell W (1997). Cross-species amplification of SSR loci in the Meliaceae family. Mol Ecol 6:1195–1197

Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. Mol Ecol 11:1–16

Zucchi MI, Brondani RPV, Pinheiro JB, Brondani C, Vencovski R (2002) Transferability of microsatellite markers from *Eucalyptus* spp. to *Eugenia dysenterica* (Myrtaceae family). Mol Ecol Notes 2:512–513