Molecular Ecology Resources (2008) **8**, 802–804 doi: 10.1111/j.1755-0998.2007.02069.x

PERMANENT GENETIC RESOURCES

Isolation and characterization of microsatellite markers for *Casearia sylvestris* Sw. (Salicaceae), a neotropical medicinal tree

M. M. CAVALLARI,*†‡ C. BILLOT,† J.-M. BOUVET,‡ B. FAVREAU,‡ M. I. ZUCCHI,§ D. A. PALMIERI¶ and M. A. GIMENES§

**UNESP-Universidade Estadual Paulista 'Julio de Mesquita Filho', Departamento de Genética, Distrito de Rubião Junior s/n, Botucatu 18618–000, SP, Brazil, †CIRAD, Unité Mixte de Recherche 'Développement et Amélioration des plantes', TA A96-03, 34398 Montpellier cedex 5, France, ‡CIRAD, Unité Propre de Recherche de Génétique Forestière, Campus International de Baillarguet TA10/C, 34398 Montpellier cedex 5, France, §IAC – Instituto Agronômico de Campinas, Centro de Recursos Genéticos, Avenida. Barão de Itapura, 1481, Campinas 13020–902, SP, Brazil, ¶UCSal – Universidade Católica do Salvador, Laboratório de Estudos em Meio Ambiente, Avenida. Prof. Pinto de Aguiar, 2589, Salvador 41740–090, BA, Brazil, **Embrapa Recursos Genéticos e Biotecnologia, Avenida W5 Norte, CP 02372, Brasília 70770–900, DF, Brazil*

Abstract

Casearia sylvestris **Sw. is a widespread neotropical tree utilized in popular medicine. Recent research ranked** ***Casearia*** **as one of the most promising genus in the search of drugs against cancer. Despite its wide distribution and pharmacological importance, no microsatellite markers have yet been developed for this genus. In this study, we provide 10 polymorphic microsatellite loci specifically designed for** ***C. sylvestris*****, used to analyse 90 individuals distributed in two populations from São Paulo state, Brazil. On average, 12.3 alleles per locus were identified, showing the ability of the markers to detect microsatellite polymorphism in this species.**

Keywords: *Casearia sylvestris*, Flacourtiaceae, genetic diversity, microsatellites, molecular markers, Salicaceae

Received 29 September 2007; revision accepted 20 November 2007

The genus *Casearia* comprises about 180 species distributed in the tropics (Sleumer 1980). Traditionally considered belonging to the Flacourtiaceae (Cronquist 1981), the taxonomic position of *Casearia* has been recently suggested to fit among the Salicaceae (APG II 2003). *Casearia sylvestris* Sw. is a widespread neotropical tree, ranging from Mexico to Argentina (Sleumer 1980) and occurring abundantly in various habitats. Its fruits, small and copious, are an important resource for several bird species (R. Torres, personal communication).

Recent research ranked *Casearia* as one of the most promising genus in the search of drugs against cancer (Balunas *et al*. 2006). Antitumoral properties of *C. sylvestris* were found by Itokawa *et al*. (1990). Population genetic structure of *C. sylvestris* was studied by Silva *et al*. (2006) through random amplified polymorphic DNA (RAPD) markers.

Although *Casearia* is ecologically and pharmacologically important, no highly polymorphic and codominant markers, such as microsatellites, have yet been developed for this genus. To our knowledge, there are no specific microsatellite markers for any of the tropical genus of Salicaceae (*sensu* APG II 2003), the same applying to the Flacourtiaceae (*sensu* Cronquist 1981) genus.

In this study, we provide 10 highly polymorphic microsatellite (simple sequence repeat) loci specifically designed for *C. sylvestris* Sw. These markers may cross-amplify in other *Casearia* species, encouraging additional genetic diversity studies on this large genus.

Marker isolation involved construction of a genomic library enriched for (CT)n and (GT)n repeats. The methodology was based on biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles, as

Correspondence: Marcelo Mattos Cavallari, Fax: +55 14 38153744; E-mail: mmcavall@ibb.unesp.br

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd

Table 1 Characteristics of 10 microsatellite markers of *Casearia sylvestris* Sw.

Locus name	GenBank	Motif repeat	Primers sequence	T_a	Allele size	n_a	N	n_a SV	n_a CB	H_O SV	H_O CB	H_E SV	H_E CB	P – HW SV	P – HW CB	Null alleles freq. SV	Null alleles freq. CB
Csy04	EU179316	$(CT)_{17}$	F: *CATAACTCTTTGCTTGCCCC R: TTGAACTCACATCTGCTGCC	67.0	117–167	15	88	10	13	0.737	0.880	0.730	0.888	0.5498ns	0.4095ns	0.0000	0.0043
Csy06	EU179317	$(CT)_{17}$	F: *TACCGTCCCTAGGACATTCG R: GCAATGCAGGTGATTTCTGA	65.0	280–322	20	82	18	19	0.628	0.659	0.909	0.849	0.0001†	0.0001†	0.1469	0.1028
Csy07	EU179318	$(TC)_{19}$	F* TCATGAACAATGCCAAGCTC R: TGAAGCGAAATTCTGCCTTT	62.5	246–268	6	86	4	4	0.057	0.157	0.109	0.182	0.0172ns	0.1624ns	0.0474	0.0219
Csy08	EU179319	$(TC)_{16}$	F: *GCCTTTAATTTCCTTTCGCC R: AAAGAGGTGATGTGCTGCTC	65.0	145–167	11	66	9	7	0.206	0.156	0.738	0.467	0.0001†	0.0001†	0.3065	0.2120
Csy09	EU179320	$(AG)_{12}$	F: *GGTTCAATCCTCTTCCAGCA R: CGGCCTAATTCCTATTGTGG	65.0	187–201	8	84	6	8	0.486	0.680	0.615	0.683	0.0348ns	0.4936ns	0.0800	0.0000
Csy11	EU179321	$(AG)_{18}$	F: *TTGTAGCACCACTTTGGCCT R: GGTCACTGTTGAAGTTTCTGGA	65.0	140–180	19	89	15	13	0.897	0.820	0.876	0.813	0.5600ns	0.5679ns	0.0000	0.0000
Csy14	EU179322	$(CA)_{10}$	F: *CTTTACATGGAAGGGCAACC R: TTTTTCCCTCACTGCATTCAT	65.0	218–244	9	87	6	9	0.564	0.583	0.474	0.765	0.0568ns	0.0012*	0.0000	0.1034
Csy15	EU179323	$(TC)_{11}$	F: *GATGGGTCAATTTCAGGAAC R: TGTTCGCTCCTAAATGCAAA	65.0	251–285	16	82	12	11	0.473	0.545	0.718	0.746	0.0001†	0.0005*	0.1416	0.1151
Csy16	EU179324	$(TG)_{10}$	F: *GCATCTGGTTGGCTCAAAAT R: TGGCAAAGAGATCGAGGATT	65.0	272–284	5	90	4	4	0.461	0.196	0.440	0.243	0.6671ns	0.0978ns	0.0000	0.0382
Csy18	EU179325	$(CT)_{11}$	F: *CCTAGTCGGTGGCAAACATT R: GCAAAGGAGCTTGAATCTGG	65.0	273–303	14	89	9	6	0.512	0.440	0.553	0.488	0.2130ns	0.1692ns	0.0258	0.0325

*5′ M13 tail: CACGACGTTGTAAAACGAC; F, forward sequence; R, reverse sequence; T_a, optimized annealing temperature; allele size range expressed in base pairs not taking into account the 19 bp of the M13 tail; n_a, number of observed alleles per locus; N, number of individuals successfully genotyped; H_O, heterozygosity observed; H_E, heterozygosity expected; SV, Santa Virginia population; CB, Carlos Botelho population; P HW, P values for the HWE test, significance threshold adjusted using Bonferroni correction: P (5%) = 0.0025 corresponding to ns: $P > 0.05$; *: $0.01 < P < 0.05$; †: $0.001 < P < 0.01$; null alleles frequencies estimated by the software MICRO-CHECKER 2.2.3 ('Brokfield 1' method).

described by Billotte *et al*. (1999). For library construction, total genomic DNA was extracted from leaves of one individual (voucher IAC 48.646). The purified total DNA (6.5 μg) was digested with *Rsa*I and enriched in (CT)n and (GT)n repeats. Microsatellite-enriched DNA fragments were ligated into pGEM-T Easy vector (Promega) and used to transform DH5α competent cells. Positive colonies (blue/white β-galactosidase selection) were tested by polymerase chain reaction (PCR) to check for the presence of inserts. A total of 162 recombinant colonies were obtained and sequenced from the SP6 primer using the BigDye terminator version 3.1 Cycle Sequencing Kit (Applied Biosystems) and a 3100 DNA Analyser (Applied Biosystems). Sequences obtained were analysed with the softwares PHRED, PHRAP, CONSED and CROSS_MATCH (Laboratory of Phil Green, Genome Sciences Department, University of Washington, www.phrap.org/index.html). Fifty-two clones containing a simple sequence repeat motif were identified by the Microsatellite Identification Tool (available at http://pgrc.ipk-gatersleben.de/misa/). Primers pairs were designed for 48 putative loci using the PRIMER 3 program (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi).

Ten microsatellite markers (Table 1) were tested using 90 samples collected from two natural populations (PESM-Santa Virginia and SM-Carlos Botelho) from São Paulo State Atlantic Forest, Brazil. Fragments were amplified by PCR on a Mastercycler (Eppendorf) thermocycler in 15-μL reaction volume, according to the following conditions: 1× PCR buffer [16.6 mM $(NH_4)_2SO_4$, 67 mM Tris-HCl pH 8.0, 0.2 mM dNTP, 0.4 mg/mL BSA, 3 mM $MgCl_2$ and 0.05% W-1 detergent solution Invitrogen, which is provided with the *Taq* DNA polymerase], 0.16 μM forward M13 (CACGACGTTGTAAAACGAC) 5′-tail-end primer and 0.2 μM of reverse primer, 0.30 μM IRDye fluorescent-labelled M13-primer (700 or 800 nM), 0.12 U/μL *Taq* DNA polymerase (Invitrogen) and 5.0 ng of genomic template DNA. A 'touchdown' cycling programme was used: 95 °C for 5 min, then 7 cycles of 94 °C for 30 s, 67 °C decreasing to 55 °C at 2 °C per cycle for 1 min 30 s, 72 °C for 1 min followed by 24 cycles of 94 °C for 30 s, 53 °C for 1 min 30 s, 72 °C for 1 min and a final extension of 60 °C for 30 min.

Amplified fragments were analysed at 700 and 800 nm by electrophoresis on an IR2-DNA analyser (LI-COR 4200 sequencer) at the Montpellier Languedoc-Roussillon Genopole genotyping platform.

Allele scoring was performed with the SAGA software (LI-COR). Standard genetic diversity parameters were determined with FSTAT (Goudet 1995), which also performed

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd

the tests of Hardy–Weinberg equilibrium (HWE) and of linkage disequilibrium between pairs of loci. The number of alleles per locus ranged from five (Csy16) to 20 (Csy06), and the mean number of alleles per locus was 12.3. The heterozygosity ranged from 0.109 (Csy07) to 0.909 (Csy06) in the Santa Virginia population, and from 0.182 (Csy07) to 0.888 (Csy04) in the Carlos Botelho population (Table 1). No significant linkage disequilibrium was observed for any pair of loci after Bonferroni correction. Three loci (Csy06, Csy08 and Csy15) in Santa Virginia population and four loci (Csy06, Csy08, Csy14 and Csy15) in Carlos Botelho population showed a significant deviation from HWE after Bonferroni correction (Table 1). The observed departures from HWE can be interpreted as the result of a Wahlund effect (substructuring of a population) and/or of the presence of null alleles. Presence of null alleles was verified with MICRO-CHECKER 2.2.3 (van Oosterhout *et al*. 2004). Null alleles were suggested to occur over most of loci, however, only the locus Csy08 presented systematic excess of homozygotes in seven other natural populations (data not shown).

The microsatellite markers presented here were also tested using eight individuals of *Casearia decandra* Jacq., one individual of *Casearia gossypiosperma* Briq. and one individual of *Casearia obliqua* Spreng. Amplification products were obtained only with Csy18 (for all three species). All other markers failed to generate amplification products in any of the three species. Additional tests should be performed with a larger number of samples to confirm these results.

The microsatellites selected in this study constitute an efficient tool to investigate genetic diversity and structure of *C. sylvestris* populations. They will be used to assess the genetic population structure of this species and to implement a strategy for its conservation and management.

Acknowledgements

We thank Dr Roseli B. Torres (IAC herbarium, Campinas, Brazil), for the identification of individuals sampled, Ronan Rivallan (Montpellier Languedoc-Roussillon Genopole – UMR DAP of CIRAD) for technical assistance, and the CAPES and FAPESP foundations, Brazil, for partial funding for the research.

References

APG II. (2003) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants. *Botanical Journal of the Linnean Society*, **141**, 399–436.

Balunas MJ, Jones WP, Chin YW *et al*. (2006) Relationships between inhibitory activity against a cancer cell line panel, profiles of plants collected, and compound classes isolated in an anticancer drug discovery project. *Chemistry and Biodiversity*, **3**, 897–915.

Billotte N, Lagoda PJL, Risterucci A-M, Baurens F-C (1999) Microsatellite enriched libraries: applied methodology for the development of SSR markers in tropical crops. *Fruits*, **54**, 277–288.

Cronquist A (1981) *An Integrated System of Classification of Flowering Plants*. Columbia University Press, New York.

Goudet J (1995) FSTAT, version 1.2: a computer program to calculate *F*-statistics. *Journal of Heredity*, **86**, 485–486.

Itokawa H, Totsuka N, Morita H *et al*. (1990) New antitumor principles, casearins A-F, for *Casearia sylvestris* Sw. (Flacourtiaceae). *Chemical and Pharmacological Bulletin*, **38**, 3384–3388.

van Oosterhout C, Hutchinson WF, Wills DP, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.

Silva MAS, Ming LC, Pereira AMS *et al*. (2006) Phytochemical and genetic variability of *Casearia sylvestris* Sw. from São Paulo State Atlantic Forest and Cerrado populations. *Revista Brasileira de Plantas Medicinais*, **8**, 159–166.

Sleumer HO (1980) *Flora Neotropica Monograph n.22 (FLACOURTIACEAE)*. The New York Botanical Garden, New York.

© 2008 The Authors
Journal compilation © 2008 Blackwell Publishing Ltd