

PERMANENT GENETIC RESOURCES

Isolation and characterization of microsatellite loci in the Guanacaste tree, *Enterolobium cyclocarpum*

MAUREEN B. PETERS,* CRIS HAGEN,* DORSET W. TRAPNELL,† J. L. HAMRICK,‡
OSCAR ROCHA,§ PETER E. SMOUSE¶ and TRAVIS C. GLENN*

*Savannah River Ecology Laboratory, University of Georgia, PO Drawer E, Aiken, SC 29802, USA, †Institute of Ecology, University of Georgia, Athens, GA 30602, USA, ‡Department of Plant Biology, University of Georgia, Athens, GA 30602, USA, §Department of Biological Sciences, Kent State University, Kent, OH 44242, USA, ¶Ecology, Evolution and Natural Resources, Rutgers University, New Brunswick, NJ 08901, USA

Abstract

We isolated nine microsatellite loci from the Guanacaste tree (*Enterolobium cyclocarpum*) and optimized them for future research on breeding populations of this species. Loci were screened across 53 individuals from one population and were shown to be variable with the number of alleles per locus ranging from five to 15. Polymorphic information content ranged from 0.420 to 0.900 and observed heterozygosity ranged from 0.547 to 0.906.

Keywords: *Enterolobium cyclocarpum*, Guanacaste tree, microsatellites, parentage analysis, PCR, pollen dispersal, population structure

Received 18 May 2007; revision accepted 11 June 2007

Enterolobium cyclocarpum (Jacq.) Griseb., the Guanacaste tree, is a dominant canopy tree of tropical dry forests (Janzen 1983) ranging from southern Mexico to northern South America (Janzen 1983; Zamora 1991). This moth-pollinated tree is a common element of primary forests but is also found at relatively high densities in disturbed landscapes such as pastures. Because of its relatively high densities, detailed studies of pollination movement and breeding patterns require the development of several hypervariable loci. In particular, we will use the simple sequence repeat (SSR) markers we have developed in paternity analyses designed to determine if the effective number of pollen parents and pollen movement distances are significantly affected by habitat disturbances (e.g. forest fragmentation).

We extracted total genomic DNA from seedling leaf tissue using the QIAGEN DNeasy Plant Kit. DNA was then serially enriched twice for microsatellites using three probe mixes [mix 2 = (AG)₁₂ (TG)₁₂ (AAC)₆ (AAG)₈ (AAT)₁₂ (ACT)₁₂ (ATC)₈; mix 3 = (AAAC)₆ (AAAG)₆ (AATC)₆ (AATG)₆ (ACAG)₆ (ACCT)₆ (ACTC)₆ (ACTG)₆; mix 4 = (AAAT)₈ (AACT)₈ (AAGT)₈ (ACAT)₈ (AGAT)₈]

following Glenn & Schable (2005). Briefly, the DNA was digested with restriction enzyme *RsaI* (New England Biolabs) and simultaneously ligated to double-stranded SuperSNX linkers (SuperSNX-24 Forward 5'-GTTTAAGGCCTA-GCTAGCAGCAGAATC and SuperSNX-24 Reverse 5'-GATTCTGCTAGCTAGGCCTTAAACAAAA). Linker-ligated DNA was denatured and hybridized to biotinylated microsatellite oligonucleotide mixes, which were then captured on magnetic streptavidin beads (Dynal). Unhybridized DNA was washed away and remaining DNA was eluted from the beads, amplified in polymerase chain reactions (PCR) using the forward SuperSNX-24 as a primer, and cloned with TOPO TA Cloning Kits (Invitrogen). We isolated and purified a total of 192 clones with inserts and sequenced 96 plasmids with M13 forward and reverse primers using the BigDye version 3.1 (Applied Biosystems) on an ABI PRISM 3130xl sequencer. Sequences from both strands were assembled and edited in SEQUENCHER 4.2 (Gene Codes) and exported to EPHEMERIS 1.0 (available at www.uga.edu/srel/DNA_Lab/programs.htm) for microsatellite searching. We designed PCR primers using OLIGO 6.67 (Molecular Biology Insights). One primer in each pair was modified on the 5' end with an engineered sequence [(CAG tag 5'-CAGTCGGGCGTCATCA-3' or M13R tag 5'-GGAAACAGCTATGACCATG-3'); see www.uga.edu/srel/DNA_Lab/protocols.htm to allow use of a

Correspondence: Maureen B. Peters, Fax: 803 7253309; E-mail: peters@srel.edu

Table 1 Characterization of nine polymorphic microsatellite loci for *Enterolobium cyclocarpum*. T_a corresponds to highest annealing temperature in a touchdown thermal cycling programme; size indicates the range of observed alleles in bp; k is number of alleles; PIC is polymorphic information content, H_O and H_E are observed and expected heterozygosities, respectively; P value indicates significance of deviation from Hardy–Weinberg equilibrium. A total of 53 individuals from a single population were genotyped

Locus accession no.	Primer sequence 5'–3'	Repeat motif	N	T_a (°C)	Size (bp)	k	PIC	H_O	H_E	P
Ency-4 EF492440	U: GTTTCACCCAGAAGACTATAT L: FAM [†] ACTCCGTCGAAGATTAG	(AC) ₉ ... (ACAT) ₄ ... (AAAC) ₅	53	60	194–206	6	0.645	0.736	0.691	0.668
Ency-8 EF492434	U: FAMGACAAACAAACAGAAAGGTAAT L: GGGACGCTAAAAGAAAACATA	(AAAG) ₄	53	60	208–224	5	0.420	0.566	0.492	0.934
Ency-9 EF492433	U: FAMAGACCATTCCACACTTAGT L: GTTCTCACAAAAGGCTTTATC	(AG) ₁₆	53	60	182–222	15	0.839	0.774	0.861	0.004
Ency-13 EF492439	U: FAMGAGAGGGGAAAAAAATTTCTA L: GTTTGAGCAGCAATGGAACCTCTAC	(ATC) ₁₀	53	60	223–241	6	0.652	0.698	0.709	0.049
Ency-17 EF492438	U: GTTTATTAGGAGCCTCGACTGTTA L: NED [‡] ACCTGCACCTTCCAACATAGT	(GTTT) ₆	53	60	262–296	6	0.599	0.547	0.668*	0.004
Ency-21 EF492432	U: GTTTGTCCAAAAGGTAGTA L: NED [‡] TAGGCTCATGTTTCAGATA	(AC) ₈ ... (ACAT) ₁₂	53	60	285–321	15	0.900	0.906	0.916	0.283
Ency-22 EF492436	U: TCCAGGCAGGGCAGTAGTA L: NED [‡] AGCAATGGGCAAAAAGAGT	(CT) ₁₄ ... (ACTC) ₄	53	60	125–149	6	0.693	0.774	0.743	0.429
Ency-24 EF492435	U: GTTTCATGCAACGGAATATAC L: NED [‡] AGACCAGGAACTTCACTA	(CT) ₁₅	53	60	128–154	9	0.575	0.604	0.619	0.125
Ency-33 EF492437	U: NED [‡] GATTTCTTTCTTCAGATTAT L: GTTTACAGAAGTGAAGATTGATAT	(ACAT) ₈	53	60	156–192	10	0.690	0.717	0.735	0.403

Primers with CAG tag (5'-CAGTCGGGCGTCATCA-3') are indicated with superscript FAM or NED (except see below), which was used for genotyping; *significant deviation from Hardy–Weinberg equilibrium is indicated at $P < 0.05$; †primer tagged with M13R tag (5'-GGAAACAGCTATGACCATG-3') and fluorescently labelled with FAM.

third primer in the PCR (identical to either the CAG or M13 tag) that is fluorescently labelled for detection on the ABI 3130xl].

We tested 24 primer pairs for amplification and polymorphism using DNA obtained from seedling leaf tissue (using the same protocol as above) of seven trees originating from seed collections made at Hacienda Solimar, southern Guanacaste Province, Costa Rica. PCR amplifications were performed in a 12.5- μ L volume [10 mM Tris pH 8.4, 50 mM KCl, 25.0 μ g/mL BSA, 0.4 μ M unlabelled primer, 0.08 μ M tag-labelled primer, 0.36 μ M universal dye-labelled primer, 2 mM MgCl₂, 0.15 mM dNTPs, 0.5 U JumpStart *Taq* DNA Polymerase (Sigma), and 20–40 ng DNA template] using an ABI thermal cycler. Touchdown thermal cycling programmes (Don *et al.* 1991) encompassing a 10 °C span of annealing temperatures ranging between 65 °C and 55 °C, 60 °C and 50 °C or 55 °C and 45 °C were used for the amplification (see Table 1). Cycling parameters were 21 cycles of 96 °C for 20 s, highest annealing temperature (decreased 0.5 °C per cycle) for 20 s, and 72 °C for 30 s; and 15 cycles of 96 °C for 20 s, lowest annealing temperature for 20 s, and 72 °C for 30 s. PCR products were run on an ABI 3130xl sequencer and sized with Naurox size standard prepared as described in DeWoody *et al.* (2004), except that

unlabelled primers started with GTTT. Results were analysed using GENEMAPPER version 4.0 (Applied Biosystems). Nine of the tested primer pairs amplified high-quality PCR product showing polymorphism across seven individuals.

We further assessed variability of these nine loci on 46 additional Guanacaste seedlings for a total of 53 analyzed individuals from the Hacienda Solimar population. Each seedling originated from a different maternal individual. Conditions and characteristics of the nine loci are given in Table 1. We used CERVUS version 2.0 (Marshall *et al.* 1998) to estimate the number of alleles per locus (k), observed and expected heterozygosities, and polymorphic information content (PIC). Deviations from Hardy–Weinberg equilibrium and linkage disequilibrium were determined using GENEPOP version 3.4 (Raymond & Rousset 1995). One locus, Ency-17, deviated from expectations of Hardy–Weinberg equilibrium ($P = 0.0042$) after Bonferroni correction for multiple comparisons. Linkage was indicated in three paired loci comparisons, Ency-4 and Ency-21, Ency-21 and Ency-22, and Ency-21 and Ency-33 ($P < 0.05$), but there is no significant indication of linkage among Ency-4, 22, and 33. These loci should now be tested with samples from known parent–offspring pairs to ensure markers behave as expected.

Acknowledgements

This work was supported by the U.S. Department of Energy, through Financial Assistance Award No. DE-FC09-07SR22506 to the University of Georgia Research Foundation; National Science Foundation Grant No. DEB-0211526 to J.L. Hamrick; National Science Foundation Grant No. DEB-0211430 to P. Smouse, and CR-USA Foundation Grant to O.J. Rocha.

Disclaimer: This report was prepared as an account of work sponsored by an agency of the US Government. Neither the US Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favouring by the US Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the US Government or any agency thereof.

References

- DeWoody AJ, Schupp J, Kenefic LP *et al.* (2004) Universal method for producing ROX-labeled size standards suitable for automated genotyping. *BioTechniques*, **37**, 348–350.
- Don RH, Cox PT, Wainwright BJ, Baker K, Mattick JS (1991) 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Research*, **19**, 4008.
- Glenn TC, Schable NA (2005) Isolating microsatellite DNA loci. *Methods in Enzymology*, **395**, 202–222.
- Janzen DH (1983) *Costa Rican Natural History*. University of Chicago Press, Chicago, Illinois.
- Marshall TC, Slate J, Kruuk L, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology*, **7**, 639–655.
- Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.
- Zamora N (1991) Tratamiento de la familia Mimosaceae (Fabales) de Costa Rica. *Brenesia*, **36**, 63–149.