Molecular Ecology Notes (2002), **2**, 363–365 doi:10.1046/j.1471-8278 .2002.00252.x

PRIMER NOTE

Isolation and characterization of nuclear microsatellite loci in the tropical tree *Corythophora alta* (Lecythidaceae)

M. B. HAMILTON*

Georgetown University, Department of Biology, Reiss Sciences Building 406, Washington, D.C. 20057–1229, USA, Biological Dynamics of Forest Fragments Project, National Institute for Research in the Amazon, Manaus, AM 69011–970, Brazil

Abstract

Studies of tropical plant population structure and gene flow are needed to estimate genetic repercussions of tropical deforestation. Twelve microsatellite loci were isolated for *Corythophora alta* (Lecythidaceae), a canopy tree at the Biological Dynamics of Forest Fragments Project, near Manaus, Brazil. These microsatellites will be used to estimate population structure and infer seed paternity. Loci were isolated with a subtractive hybridization method. Loci averaged 10.7 alleles per locus with high levels of heterozygosity. Four loci had significant deficits and one locus had a significant excess of heterozygotes compared to Hardy–Weinberg expectations. Many loci had high average probabilities of paternity exclusion.

Keywords: *Corythorphora alta*, forest fragmentation, Lecythidaceae, microsatellite, seed paternity, tropical tree

Received 1 April 2002; revision received 9 May 2002; accepted 9 May 2002

Studies of tropical plant population structure and gene flow have become an important paradigm in efforts to understand basic population processes in tropical forests and to estimate future repercussions of tropical deforestation (e.g. Nason & Hamrick 1997; Aldrich & Hamrick 1998; Dick 2001). The primers described here were designed to detect population variation in *Corythophora alta*, a bee-pollinated canopy tree in the Brazil nut family (Lecythidaceae) that flowers and fruits annually. *Corythophora alta* are frequent in the experimental forest fragments and areas of continuous forest located at the Biological Dynamics of Forest Fragments Project (BDFFP), near Manaus, Brazil (Gascon and Bierregaard 2001). These microsatellite markers were isolated with two goals. First, the loci will be used to estimate population structure among trees in several forest fragments and in continuous forest areas. Second, these loci will be used to infer the paternity of seeds and thus estimate the distribution of pollen dispersal distance as well as the proportion of seed paternity as a result of trees within 1- or 10-hectare forest sampling areas.

Correspondence: M. B. Hamilton. *Present address: Georgetown University, Department of Biology, Reiss Sciences Building 406, 37th and O Streets NW, Washington, DC 20057–1229 USA. Fax: (202) 687 5662; E-mail: hamiltm1@georgetown.edu

The loci described here were isolated following the subtractive hybridization method of Hamilton *et al*. (1999); see detailed protocol at http://www.georgetown.edu/faculty/hamiltm1. Genomic DNA was obtained from frozen leaf tissue ground in liquid nitrogen and extracted with a DNeasy plant kit (QiaGen, Valencia, CA) following the manufacturer's instructions. Genomic DNA was digested with *Hae*IIII, *Alu*I and *Rsa*I (all from New England Biolabs, Cambridge, MA) to obtain a majority of DNA in the 200–1200 base pair size range. After treating digests with mung bean nuclease to create blunt ends and ligation of the SNX linker, streptavidin-bead subtractive hybridization used 30-mer biotinylated oligonucleotides with GT, CT, CAC, CAT, GGA, CCG, ACG, CAA, CTC, CAG, AGG, AAT and GGAT repeat motifs. Products of the subtractive hybridizations were amplified by polymerase chain reaction (PCR), dot blotted onto a nylon membrane and hybridized to these same oligonucleotides. Products with strong hybridization signals (CT, GA, CTC, CAC, GAA, CCG) were ligated into vector and transformed into competent cells. Hybridization-positive colonies were picked and the insert was PCR amplified with T7 and T3 primers. PCR products were purified with QiaQuick spin columns (QiaGen, Valencia, CA), sequenced with both T7 and T3 primers in reactions containing 4.6 μL water, 2 μL template, 2.4 μL primer (1 μM) and 6 μL dRhodamine terminator reaction

Table 1 Annealing temperature (T_A), number of PCR cycles, PCR 72 °C extension time and thermal cycler

Locus	T_A (°C)	No. PCR cycles	Extension time (s)	Thermal cycler*
CCG25–3#11	59	28	15	AB
CCG25–3#6	57	28	15	AB
CCG25–4#20	63	28	20	RC
CCG25–6#7	62	28	20	RC
CT50–2#3	62	28	20	RC
CT50–2#4	57	28	22	RC
CTC40–1#5	57	28	22	RC
CTC40–3#12	59	30	15	RC
CTC40–3#5	57	28	30	RC
GGA25–1#19	59	27	30	RC
GGA50–2#3	57	28	20	RC
GGA50–2#7	57	25	15	AB

*Applied Biosystems GeneAmp 9700 (AB) or Stratagene Robocycler Gradient 96 (RC).

ready mix (Applied Biosystems, Foster City, CA) and electrophoresed on a model 377 sequencer (Applied Biosystems, Foster City, CA). The resulting sequences were aligned into contigs and edited using SEQUENCHER 3.1.1 (GeneCodes, Ann Arbor, MI). Potential primers were identified manually and tested with AMPLIFY 1.2 (Engels 1993).

Primers were synthesized (Operon Technologies, Alameda, CA) and after some experimentation with PCR constituents and cycling parameters, all primer pairs amplified a single band in the expected size range. Clone sequencing templates diluted to 10 000 : 1 were used as positive controls to verify the PCR product molecular weight. PCR reactions contained 1 μL of DNA template (DNA concentration was not determined), 2 μL of 10× Thermopol buffer (containing 20 mM $MgSO_4$), 0.2 mM each dNTP, 0.4 μM of each primer, 1.25 M betaine (Aldrich, cat. num. 21,906–1, Milwaukee, WI) and 0.4 units of Vent exo-polymerase (New England Biolabs, Cambridge, MA) in a total volume of 20 μL. The thermal cycling profiles were 5 min at 96° followed by cycles of 96° for 45 s, annealing temperature for 45 s and 72° extension for 15–30 s (Table 1).

Loci were initially tested for polymorphism by running PCR products from 10 to 12 individuals on Spreadex EL 400 precast gels (Elchrom Scientific, Jamaica Estates, NY) or 4% NuSieve 3 : 1 precast gels (BioWhittaker, Rockland, ME). For those loci that showed multiple alleles, high-performance liquid chromatography purified primers sets were ordered with a forward primer label of 6-FAM, HEX (Operon Technologies, Alameda, CA), or NED (Applied Biosystems, Foster City, CA). To score genotypes, PCR products were electrophoresed along with GeneScan ROX 400 molecular marker on a 377 or 3100 sequencer and sized using GENESCAN

Table 2 Locus name, primer sequences, repeat motif, observed size range, number of individuals genotyped (*N*), number of alleles (*k*), observed heterozygosity (H_O), expected heterozygosity (H_E), average probability of paternity exclusion (P_E) and GenBank accession number

Locus	Primer sequences (5′–3′)	Repeat	Size range (bp)	*N*	*k*	H_O	H_E	P_E	Accession number
CCG25–3#11	F: GAC CGG CGT CAA GAT AGT GG R: GCC CAA TAT AGC TAT TTG CAG G	GT	170–212	60	15	0.767	0.857	0.708	AF484236
CCG25–3#6	F: CCA TTA TCT AGG AGA GGT TGG R: GTA GAG AAT TGA TGC GTG AGG	GT	182–238	56	18	0.607*	0.904	0.791	AF484237
CCG25–4#20	F: GAC CAG AGT TCT GAG CAT ACC R: CGG CTT CAC TAT CTT CAG C	AG	126–142	57	7	0.772	0.768	0.551	AF484238
CCG25–6#7	F: TAC CAA CGG TGT TAG CGA CG R: TTC CTC CTG CAA CAG TCA CC	GT	145–187	52	12	0.769	0.795	0.610	AF484239
CT50–2#3	F: TCC CAA CAT TCA ATC TTC GC R: TGT GTG CTG GCT GAG TGC	CT	138–158	37	8	0.919	0.840	0.662	AF484240
CT50–2#4	F: TTC TGA ACA ACA TAG GTG GTC R: CAA ATA GTC CAT CAA AGG GAG	CT	166–194	47	14	0.851	0.882	0.746	AF484241
CTC40–1#5	F: TAC GAC GGT AAC ACA GCT C R: CTG AGA ACT AGC TAG TGC C	CTC	276–306	52	9	0.673*	0.776	0.568	AF484242
CTC40–3#12	F: CAT AGT CTT AAT TCG GGT CCA C R: TAA CTC TTG CAG ACT AAA ACC TC	CTT /CTC	137–167	51	11	0.431*	0.825	0.657	AF484243
CTC40–3#5	F: TTA GTG CCA TTA TGC CTT ACC R: GTG GTT ACT TTG CTG TGA GC	CTC	93–114	60	7	0.850*	0.771	0.541	AF484244
GGA25–1#19	F: CTG GTA GTC TGC CTT CGA CTC R: TAG GAC AAC ATT AAC GTG TCA GC	CCT /TTC	314–344	56	10	0.911	0.838	0.666	AF484245
GGA50–2#3	F: CAA TCT CAT CTT CTT CCC AAG R: TCA GGC TGA TAT TCA CAC CG	CCT /TTC	179–212	47	10	0.489*	0.858	0.703	AF484246
GGA50–2#7	F: TCA CCA CCA GAA AGC TGG C R: AGT CAT GGA TTG GAC TCG G	CAG	231–258	56	7	0.625	0.580	0.371	AF484247

*Observed and Hardy–Weinberg expected homozygote frequencies different ($P > 0.05$).

3.1 or 3.7, respectively (all from Applied Biosystems, Foster City, CA). Genotype data for each locus were collected from individuals in the 1202 population at the BDFFP, a 10-hectare forest fragment containing 63 known *C. alta* individuals with diameter at breast height ≥ 10 cm.

CERVUS 2.0 (Marshall *et al*. 1998) was used to summarize the allele and genotype data (Table 2). These loci were highly polymorphic with between seven and 18 alleles, an average of 10.7 alleles per locus and high levels of heterozygosity. GENEPOP web version 3.1c (http://wbiomed.curtin.edu.au/genepop; Raymond & Rousset 1995) was used to test individual loci for deviation from Hardy–Weinberg expected heterozygote frequency. The probability test was employed (option 1, suboption 3, equivalent to a two-tailed hypothesis test) using default values for the Markov chain parameters. Five loci did not fit Hardy–Weinberg expectations and four of these loci showed significant deficits of heterozygous genotypes consistent with null alleles. Many of the loci also had high average probabilities of paternity exclusion, which should provide powerful markers for paternity exclusion when seed genotype data from several loci are combined. For example, the combined average probability of paternity exclusion for five loci (CCG25–3#11, CCG25–6#7, CT50–2#3, CT50–2#4 and GGA50–2#7) is 0.99673 based on allele frequencies in the 1202 population. These loci will be of great utility for estimating population structure and conducting paternity analyses in the BDFFP populations of *C. alta*.

Acknowledgements

This work was supported by the Smithsonian Institution, the National Geographic Society, the Biological Dynamics of Forest Fragments Project, Georgetown University and the National Science Foundation (DEB9983014). I thank R. Fleischer for supporting this project while I was a postdoctoral fellow at the Smithsonian Institution. N. Zhu, M. Venkatesan, G. Baltazar and C. Lund provided valuable technical assistance. I thank C. Hurley for generous sharing of two ABI 377 instruments.

References

Aldrich PR, Hamrick JL (1998) Reproductive dominance of pasture trees in a fragmented tropical forest mosaic. *Science*, **281**, 103–105.

Dick C (2001) Genetic rescue of remnant tropical trees by an alien pollinator. *Proceedings of the Royal Society of London Series B*, **268**, 2391–2396.

Engels WR (1993) Contributing software to the internet: the Amplify program. *Trends in Biochemical Sciences*, **18**, 448–450.

Gascon C, Bierregaard RO (2001) The biological dynamics of forest fragments project: the study site, experimental design, and research activity. In: *Lessons from Amazonia: the Ecology and Conservation of a Fragmented Forest* (eds Bierregaard RO, Gascon C, Lovejoy TE, Mesquita R), pp. 31–42. Yale University Press, New Haven.

Hamilton MB, Pincus EL, DiFiore A, Fleischer RC (1999) Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. *Biotechniques*, **27**, 500–507.

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. *Molecular Ecology*, **7**, 639–655.

Nason JD, Hamrick JL (1997) Reproductive and genetic consequences of tropical forest fragmentation: two case studies of neotropical canopy trees. *Journal of Heredity*, **88**, 264–276.

Raymond M, Rousset F (1995) GENEPOP (Version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.