Conservation Genet Resour (2014) 6:987–989
DOI 10.1007/s12686-014-0265-x

MICROSATELLITE LETTERS

Characterization of twenty-six microsatellite markers for the tropical pioneer tree species *Cecropia insignis* Liebm (Urticaceae)

Na Wei · Christopher W. Dick

Received: 13 June 2014 / Accepted: 7 July 2014 / Published online: 17 July 2014

Abstract *Cecropia insignis* is an ecologically important Neotropical pioneer tree and major vertebrate food source. Although this species is relatively common in faunally intact tropical rainforests, its population dynamics may be negatively impacted by hunting of seed-dispersing animals. To better understand gene flow and regeneration dynamics in *C. insignis*, we characterized 26 microsatellite markers in a population sampled from Barro Colorado Island, Panama. Eleven loci of ≥3 alleles were tested on 48 individuals, whereas the remaining 15 loci of two alleles were tested on 12 individuals. Allelic richness ranged from 2 to 9 per locus. Observed and expected heterozygosity averaged 0.478 and 0.440 respectively. Polymorphism information content was between 0.141 and 0.757. Only two loci exhibited deviation from Hardy–Weinberg proportions.

Keywords *Cecropia insignis* · Microsatellite markers · Tropical tree · Seed dispersal

Cecropia insignis is a dioecious, gap-dependent canopy tree species distributed broadly in lowland moist forests of Central and northern South America. It provides important food resources (e.g., leaves, nectar, fruits) for forest-dwelling animals. Although *Cecropia* trees represent one of the few primarily wind-pollinated taxa, its seed dispersal is mediated by vertebrates including large birds and monkeys. Hunting pressures are increasingly threatening the persistence of such seed-dispersing vertebrates, and have measurably altered tropical forest dynamics (Terborgh et al. 2008). We developed twenty-six polymorphic microsatellite markers for *C. insignis* to evaluate the impact of hunting and other anthropogenic changes on gene flow and regeneration in this species.

Methods used to obtain genomic data using circular consensus sequencing of Pacific Biosciences (PacBio) are described by Wei et al. (in press). Briefly, a PacBio 500-bp SMRTbell library was established from the genomic DNA of one *C. insignis* tree, and then sequenced using four SMRT cells with C2 chemistry. In total, 198,989 circular consensus reads were generated. A quality-control step (for details, see Wei et al. in press) was performed before searching for microsatellite loci and designing primers in QDD v2.1 (Meglécz et al. 2010). In total, 512 microsatellites loci were retrieved. From the pure (non-interrupted) microsatellite loci ($n = 404$), we synthesized 69 primer pairs (38 di-, 30 tri-, and 1 tetra-nucleotide motifs).

For marker validation, we isolated genomic DNA from 48 reproductive-sized trees of *C. insignis* growing on Barro Colorado Island, Panama. We adjusted the use of DNeasy Plant Mini Kit (QIAGEN, Valencia, California, USA) for high-throughput DNA isolation by replacing DNA binding columns with E-Z® 96 DNA Plates (Omega Bio-Tek, Norcross, Georgia, USA). After an initial screening of primer amplification on three individuals, polymorphic loci were tested on another 9 samples. Then microsatellite loci showing ≥3 alleles based on these 12 samples were scored on an additional 36 individuals. PCRs were carried out as follows: 94 °C for 4 min; 28 cycles of 94 °C for 30 s, 59 °C (decreasing 0.2 °C per cycle) for 40 s and 72 °C for 60 s; 10 cycles of 94 °C for 30 s, 53 °C for 40 s and 72 °C

N. Wei (✉) · C. W. Dick
Department of Ecology and Evolutionary Biology, University of Michigan, 830 North University Avenue, Ann Arbor, MI 48109-1048, USA
e-mail: weina@umich.edu

C. W. Dick
Smithsonian Tropical Research Institute, PO Box 0843-03092, Balboa, Ancón, Republic of Panama

Table 1 Characteristics of 26 microsatellite markers developed in *C. insignis*

Locus	Primer sequence (5′–3′)[a]	Motif	Size range	*A*	H_O	H_E	PIC	Sample size	Accession no.
CEC_08	F: CTGCAATTGACTTGCCACAC	$(AAG)_{11}$	149–206	5	0.771	0.642	0.593	48	KF680367
	R: GGTGTGAAATGAAAGTGACCC								
CEC_10	F: ATTGCTCGTGCAACCAAAG	$(AAT)_8$	258–285	5	0.596	0.565	0.523	48	KF680369
	R: TTGTGCCATGTTAATAGCCC								
CEC_12	F: TTCCAATCCGGAGATAAACG	$(AAG)_{10}$	110–128	4	0.708	0.581	0.524	48	KF680371
	R: AAGCAAGAATCTCAAAGCCG								
CEC_17	F: TTCTTGATCGTGTTTGCTGC	$(AAT)_7$	115–127	4	0.458	0.425	0.364	48	KF680376
	R: AAATGTTCAAGGCATTGGTTC								
CEC_37	F: CAAGAGATGCGTCGAGAGTG	$(AG)_{16}$	151–157	4	0.479	0.545	0.466	48	KF680388
	R: GGCAATCAATTTGCGTAACC								
CEC_43	F: TTCGTGTATGAGGACAACGAG	$(AG)_{14}$	293–317	5	0.583	0.688	0.624	48	KF680393
	R: AATTCCACGAGGAAGCAGAG								
CEC_45	F: TTTACCAAACCCAATTCCC	$(AG)_{13}$	118–152	9	0.875	0.785	0.757	48	KF680394
	R: ATTCTCAGCAAGTTCCCAGC								
CEC_46	F: AGTACAACACCCGGATCGAC	$(AG)_{13}$	112–136	8	0.604	0.528	0.503	48	KF680395
	R: TCGAATATAACGCCTCTCGC								
CEC_56	F: TGGCCTTCTTGAGTTGTTTG	$(AC)_{10}$	193–201	3	0.625	0.539	0.447	48	KF680402
	R: TCAGCCACTCTCACTCTTCG								
CEC_61	F: TCCAAGTAACATCCTCTCCCTC	$(AG)_{10}$	115–121	3	0.188	0.205	0.188	48	KF680406
	R: TCCCTCAGAAAGCGAAGAAC								
CEC_64	F: TTTGTCTTTGGCTTTGGACC	(AC)9	145–155	4	0.542	0.536	0.497	48	KF680408
	R: CAACCTTTGCAAATTGGTCTAC								
CEC_15	F: ACCAGAGCCTTGAACAATCC	$(AAG)_7$	119–122	2	0.167	0.278	0.239	12	KF680374
	R: TTCTTTGGACGAGAAATCGG								
CEC_22	F: CCGCATGGATAATTTCTCTTC	$(AAT)_8$	204–207	2	0.333	0.375	0.305	12	KF680381
	R: ACATCGTTGCATGAGCTTTG								
CEC_31	F: GGGTGTATGCTCTCACACTTG	$(AAT)_7$	129–138	2	0.333	0.278	0.239	12	KF680386
	R: TCCATGATATGGTTTGGGTG								
CEC_34	F: TTAGGACTACTGCCTTCGCAC	$(AC)_{19}$	153–163	2	0.417	0.330	0.276	12	KF680387
	R: TATTGAGGCATGGAGGCTTG								
CEC_38	F: TTACAGAGCATTGTGACCCG	$(AG)_{15}$	159–161	2	0.500	0.486	0.368	12	KF680389
	R: TGATGGAAGCTCTGAAGCAC								
CEC_40	F: TTATGGGCAACTACGGCTTC	$(AG)_{15}$	121–125	2	0.500	0.375	0.305	12	KF680390
	R: CCATGTTCTAAACAATGTGTCC								
CEC_41	F: TGAGCAAGCTGGAAAGGAAG	$(AG)_{15}$	156–166	2	0.583	0.413	0.328	12	KF680391
	R: TGCAAACCCAGCTATAAATGC								
CEC_49	F: GAATTGCACATTGCCCTCTC	$(AG)_{12}$	116–118	2	0.417	0.330	0.276	12	KF680397
	R: CTCCGGTCTCTTCCTTCCC								
CEC_52	F: ACCTTTGACCGTGGGATTC	$(AC)_{10}$	126–132	2	1.000***	0.500	0.375	12	KF680398
	R: TGGTTGTCAAACTGTAAGGCAG								
CEC_53	F: GGCTGAGAGCTTTGGAGATG	$(AG)_{10}$	142–150	2	0.250	0.330	0.276	12	KF680399
	R: ACACTGTAGCAGAGCGGAGC								
CEC_59	F: CCTCGGTGACCTTGAACTTG	$(AG)_{10}$	154–156	2	0.167	0.153	0.141	12	KF680404
	R: AAGAAACCCTTCAATCTCTGC								
CEC_60	F: CTCAGCATAGATCTCGTTGCC	$(AG)_{10}$	184–186	2	0.250	0.413	0.328	12	KF680405
	R: TCTACTCAACAACCCGACCC								

Table 1 continued

Locus	Primer sequence (5′–3′)[a]	Motif	Size range	*A*	H_O	H_E	PIC	Sample size	Accession no.
CEC_62	F: GTTTGGTGGGTTCACATGG	$(AG)_{10}$	115–117	2	0.583	0.413	0.328	12	KF680407
	R: CGATGTGTCACACTTGGGTC								
CEC_65	F: TGAGGAATCTCCAAGGGAAG	$(AC)_9$	117–121	2	0.167*	0.444	0.346	12	KF680409
	R: TCAGTGATTGGACTTCTGTTCC								
CEC_67	F: CTTGAAACCGGCTCCTGAAC	$(AG)_9$	157–163	2	0.333	0.278	0.239	12	KF680411
	R: TCGGGAATGGAAATAAATATGC								

A = number of alleles per locus; H_O = observed heterozygosity; H_E = expected heterozygosity; PIC = polymorphism information content

Significant deviation from Hardy–Weinberg proportions at $P < 0.05$ (*) and $P < 0.001$ (***)

[a] M13 tail (TGTAAAACGACGGCCAGT) attached to the 5′ end of individual forward primers

for 60 s; and 72 °C for 10 min. Each 8-μL PCR contained 1 μL of 4 ng/μL DNA, 0.05 μL of 1 μM HEX-labeled or 1.5 μM FAM-labeled M13 primer (TGTAAAACGACGGCCAGT), 0.12 μL of 5 μM M13-tagged forward primer, 0.48 μL of 5 μM reverse primer, 0.8 μL of 25 mM $MgCl_2$, 4 μL of GoTaq Colorless Master Mix (Promega, Madison, Wisconsin, USA), and 1.55 μL H_2O. PCR products of two loci labeled by different dyes were sized in a single lane on an ABI 3730 DNA Analyzer (Life Technologies, Carlsbad, California, USA). Alleles were then scored using GeneMarker v2.4.1 (SoftGenetics, State College, Pennsylvania, USA). Allelic richness, observed and expected heterozygosity, and Hardy–Weinberg equilibrium (HWE) were estimated using GenAlEx v6.5 (Peakall and Smouse 2012). Polymorphism information content (PIC) was assessed in PowerMarker v3.25 (Liu and Muse 2005).

We described here only the 26 polymorphic microsatellite loci. For the eleven markers screened on 48 individuals (Table 1), allelic richness averaged 5 per locus (range 3–9). H_O ranged from 0.188 to 0.875; H_E varied between 0.205 and 0.785. PIC was between 0.188 and 0.757 (mean = 0.499). All of these eleven loci conformed to HWE. For the fifteen markers showing two alleles and tested on 12 individuals (Table 1), observed and expected heterozygosity averaged 0.400 and 0.360 respectively. PIC was between 0.141 and 0.375. Two of the 15 loci (CEC_52 and CEC_65) deviated from Hardy–Weinberg expectations.

Acknowledgments This work was supported by a CTFS-Forest-GEO grant from Smithsonian Tropical Research Institute and Center for Tropical Forest Science.

References

Liu KJ, Muse SV (2005) PowerMarker: an integrated analysis environment for genetic marker analysis. Bioinformatics 21(9):2128–2129. doi:10.1093/bioinformatics/bti282

Meglécz E, Costedoat C, Dubut V, Gilles A, Malausa T, Pech N, Martin JF (2010) QDD: a user-friendly program to select microsatellite markers and design primers from large sequencing projects. Bioinformatics 26(3):403–404. doi:10.1093/bioinformatics/btp670

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28(19):2537–2539. doi:10.1093/bioinformatics/bts460

Terborgh J, Nun ez-Iturri G, Pitman NCA, Valverde FHC, Alvarez P, Swamy V, Pringle EG, Paine CET (2008) Tree recruitment in an empty forest. Ecology 89(6):1757–1768

Wei N, Bemmels JB, Dick CW (in press) The effects of read length, quality and quantity on microsatellite discovery and primer development: from Illumina to PacBio. Mol Ecol Resour. doi:10.1111/1755-0998.12245