 doi: 10.1111/j.1471-8286.2007.01808.x

PRIMER NOTE

Isolation and characterization of microsatellite loci from the endangered highlands scrub hypericum (*Hypericum cumulicola*)

C. E. EDWARDS,*† M. ARAKAKI,*† P. F. QUINTANA-ASCENCIO,‡ D. E. SOLTIS* and P. S. SOLTIS†
*Department of Botany, University of Florida, Gainesville, FL 32611, USA, †Florida Museum of Natural History, Gainesville, FL 32611, USA, ‡Department of Biology, University of Central Florida, 4000 Central Boulevard, Orlando, FL 32816, USA

Abstract

We report the isolation of 19 primer pairs for amplification of polymorphic microsatellite loci for *Hypericum cumulicola*. These markers were evaluated in 24 individuals from one population; two to four alleles were detected per locus, and observed heterozygosity ranged from 0 to 0.5. Two loci demonstrated significant heterozygote deficiencies, possibly due to null alleles, and significant linkage disequilibrium was found between six pairs of loci. The remaining microsatellite loci will help determine if genetic differentiation is responsible for life-history differences between natural and anthropogenically disturbed populations of *H. cumulicola*.

Keywords: highlands scrub hypericum, Hypericaceae, *Hypericum cumulicola*, microsatellite

Received 25 February 2007; revision accepted 20 March 2007

Hypericum cumulicola (highlands scrub hypericum) is a small, federally endangered, perennial herb endemic to the Lake Wales Ridge, Florida (Dolan *et al*. 1999). It generally occurs in rosemary scrub, a highly xeric, open subtype of Florida scrub occupying well-drained, white sand soils; however, the species has also colonized open-sand areas created by anthropogenic disturbances, such as fire lanes and roadsides (Quintana-Ascencio *et al*. 2007). *Hypericum cumulicola* life-history differs between scrub and road populations: individuals in road populations grow faster, have more variable life spans, earlier reproduction, and higher fecundity (Quintana-Ascencio *et al*. 2007). Genetic differentiation and phenotypic plasticity were suggested as possible causes of these differences. To assess whether genetic differentiation is responsible for these differences, highly variable genetic markers are needed to distinguish among individuals. Allozyme variation is extremely low in *H. cumulicola* (Dolan *et al*. 1999), and because microsatellite markers often have much higher mutation rates than allozymes (Jarne & Lagoda 1996), microsatellites may be better suited for differentiating individuals within a population. Thus, our goal was to develop highly variable microsatellite loci to differentiate among individuals within and among populations of *H. cumulicola.*

We used a microsatellite enrichment procedure based on Kandpal *et al*. (1994). Genomic DNA was extracted using the DNeasy Mini Kit (QIAGEN) and digested with *Sau*3AI. DNA was purified and fractionated using Chroma Spin columns (Clontech Laboratories) to remove fragments < 400 base pairs (bp). Remaining DNA was ligated to *Sau*3AI linkers and amplified by polymerase chain reaction (PCR). The fragment library was enriched for $(CA)_n$ and $(GT)_n$ repeats by hybridizing DNA fragments to a biotinylated nucleic acid probe. Probe-target fragments were captured using VECTREX® Avidin D (Vector Laboratories), amplified by PCR, ligated into a TOPO TA pCR vector, transformed into One Shot® *Escherichia coli* competent cells (Invitrogen), and selected by growing them overnight on Luria broth (LB)/kanamycin/agar plates. Transformed colonies were cultured in LB overnight and screened for the repeat using PCRs with the CA repeat primer and M13F/M13R. PCRs were visualized on 2% agarose gels, and colonies that produced a band of the appropriate size (> 200 bp) were amplified by rolling circle amplification (RCA) and sequenced in one direction using an ABI 3730xl DNA Analyser (Applied Biosystems) with the T7 primer and BigDye Terminator Cycle Sequencing chemistry (Applied Biosystems).

Correspondence: C. E. Edwards, Fax: 352 846 2154; E-mail: edwardsc@ufl.edu

Table 1 Characteristics of 19 microsatellite loci developed for *Hypericum cumulicola*, based on a sample of 24 individuals from one population (all individuals were successfully genotyped at all loci). N_A, number of alleles found in all individuals; H_O, observed heterozygosity; H_E, expected heterozygosity; P, probability of a heterozygote deficiency under the assumption of Hardy–Weinberg equilibrium (significant P values at the 5% nominal level are indicated in bold using Bonferroni corrected significance level = 0.0026)

Locus	Primer sequence (5′–3′)	Repeat array	N_A	Size (bp)	H_E	H_O	P	GenBank Accession no.
Hyp1	F: *TTTGACAACAAGTCAACATAGTATAGC R: GCACCATTGTCATTTTGTTCC	$(GA)_6$GTGATAAA$(CA)_8$	2	263–285	0.120	0.125	1.0000	EF440625
Hyp2	F: *TGACTGTGCATCAGTGCAAC R: AAGCAGGACTGACAATTCAGG	$(TC)_8$ACGTGCTC$(GT)_5$	2	243–245	0.510	0.208	0.0053	EF440626
Hyp7	F: *GACCCAAGTGTGTTCTGTCG R: GAGCGGTGGTAGAAACCAGA	$(CT)_8$GTAT$(CT)_8(GT)_6$	2	237–239	0.311	0.208	0.1868	EF440627
Hyp9	F: *TGGTCGTCCATTTCACGATA R: CTCCCGCATCTCCTTTACCT	$(GA)_{12}(GT)_2$GCGTCGGATGTTGGAG$(TTGG)_2(TG)_9$	4	199–207	0.563	0.500	0.2421	EF440628
Hyp12	F: *GCGCATGTCTGGAACAGTAA R: CGCAGGAGACCATCTACCTC	$(GC)_6(AC)_{10}$	2	225–227	0.510	0.208	0.0079	EF440629
Hyp21	F: *TCAGTCCAGAGCTTCCTCAG R: AAGGCTTAAGAAATCACC	$(GT)_4$GGGTGTCTGT$(GA)_{13}$	2	203–211	0.337	0.333	0.7395	EF440630
Hyp23	F: *ATAGAGTTATGTTTATATGGCCC F: TCAGACTGGGTTACATAGTATCGG	$(TC)_{11}(AC)_{21}$	3	231–235	0.670	0.416	0.0053	EF440631
Hyp24	F2: *TGGGAATTGGGACGAGGCGGC R2: AGGCTGAGAAAGCAATGTTGG	$(GT)_8$	2	271–273	0.507	0.00	**0.0026**	EF440632
Hyp27	F: *ATCAGTCACTCTGTGGTATTCCC R: TCCCTTTGCTTTCTTATTTAATGC	$(CT)_7(CA)_9$	2	221–223	0.284	0.25	0.4927	EF440633
Hyp28	F: *TTCCTCTGGTCAGGATGGGCGC R: ATGTCCTGATGTTTGGCCTGG	$(CA)_8$	2	235–237	0.439	0.375	0.3974	EF440634
Hyp30	F2: *AATCTACTTCCGCGGATACATACTCC R: TCGATGTAAATAACTAACTCATCTGAA	$(CA)_{11}$A$(CA)_6$	2	229–301	0.042	0.042	1.0000	EF440635
Hyp31	F: *AACAGCACGGACGCGTCACAACCC R: AGCCTCCTCAGGCTCCACTTCC	$(CA)_8$	2	157–159	0.156	0.083	0.1316	EF440636
Hyp32	F: *TACCATGAAGAACCCATCGATAGCC R: AGACGACAATTGCATGCATGAACG	$(CA)_{11}$	2	197–205	0.503	0.042	**0.0026**	EF440637
Hyp33	F: *AAGAACATTATAGAAGTGGGTATGGG R: ATGCTCACAATGACGTAACGAGC	$(TC)_{15}$	2	201–203	0.383	0.25	0.1211	EF440638
Hyp38	F: *TTTATTCTCACTGTTGGGGTTTAGG R: TTGGTCGTCCGCGAACGGCG	$(GT)_{16}$	2	171–181	0.667	0.422	0.2711	EF440639
Hyp40	F: *TGAAGGAGAATGATGCCTACTG R: CAGCTTTAGCAAGGACTGGAA	$(CA)_{10}$	2	249–253	0.667	0.422	0.3026	EF440640
Hyp41	F2: *TCACATTCGTTGCTTGTTCC R: ACCATTAGCTGGGACAAAAGG	$(CT)_7(CA)_7$	3	207–211	0.595	0.333	0.0053	EF440641
Hyp43	F: *GGTGCAGGTCTATGCTTTGC R: TTTCCCACAGAATTCATCTCC	$(GT)_{28}$	2	391–393	0.563	0.242	0.0079	EF440642
Hyp46	F: *GAGCACAAAAGCCACATCC R: TTTAGGTGTAAATTATTGGTTAGGG	$(TG)_2(TC)_2(GT)_{10}$	2	307–313	0.439	0.208	0.0158	EF440643

*M13 tag (CACGACGTTGTAAAACGAC) added to 3′ end of primer for amplification with fluorescently labelled M13.

Sequences were edited and checked for redundancies using SEQUENCHER 4.2 (Gene Codes Corp.). Primers were designed manually or by using PRIMER3 (Rozen & Skaletsky 2000). An M13 tail (CACGACGTTGTAAAAC) was added to the 5′ end of each forward primer for amplification using a universal dye-labelling method (Boutin-Ganache *et al*. 2001). Of 267 colonies sequenced, 88 contained microsatellites, and we designed primers for 46 unique loci. We initially screened primers for amplification in eight *H. cumulicola* individuals following Symonds *et al*. (in prep.); PCR amplifications were performed in 10 μL reactions containing 0.5 U of GoTaq Flexi DNA polymerase (Promega) or *Taq* DNA polymerase (New England BioLabs), 1× Promega Colorless GoTaq Flexi Buffer or New England BioLabs ThermoPol Buffer [20 mM Tris-HCl, 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 2 mM $MgSO_4$, 0.1% Triton X-100, pH 8.8 at 25 °C], 1.5 mM $MgCl_2$, 0.15 μM each of the reverse primer and D4-labelled M13 primer (Beckman Coulter), 0.01 μM of the extended forward primer, and 50 μM of each dNTP. PCR temperature cycling conditions were as follows: (i) 3 min at 94 °C, (ii) denaturation for 30 s at 94 °C, (iii) annealing for 30 s at 52 °C, (iv) extension for 45 s at 72 °C, (v) 35 repetitions of steps 2–4, and (vi) a final elongation at 72 °C for 20 min. Labelled PCR products were run on a CEQ 8000 capillary sequencer (Beckman Coulter) by loading 1 μL of PCR product, 0.35 μL of CEQ TM DNA Size Standard Kit-400 (Beckman Coulter) and 25 μL of formamide. Resulting electropherograms were analysed using CEQ GENETIC ANALYSIS SYSTEM software (Beckman Coulter) and scored manually.

Of 46 original loci, 17 demonstrated poor amplification in the initial screening and were not tested further. The remaining 29 loci were tested for variability in 24 *H. cumulicola* individuals; 10 loci were monomorphic. The remaining 19 loci (Table 1) were successfully genotyped in all 24 individuals, and these results were analysed to assess deviation from Hardy–Weinberg equilibrium and linkage disequilibrium among loci using Fisher's exact tests in FSTAT version 2.9.3.2 (Goudet 2002); significance was assessed using permutation procedures. Descriptive statistics (Table 1) were measured using GENEPOP version 3.4 (Raymond & Rousset 1995). After Bonferroni correction for multiple comparisons, significant linkage disequilibrium was detected between six pairs of loci (Hyp23 × Hyp24, Hyp23 × Hyp32, Hyp24 × Hyp32, Hyp2 × Hyp43, and Hyp33 × Hyp40). No locus demonstrated an excess of heterozygotes, but after Bonferroni correction, Hyp24 and Hyp32 showed significant heterozygote deficiency. The cause of the deficit is unknown but may be due to inbreeding or the presence of null alleles. The loci described here will allow us to determine if life-history differences between scrub and road populations of *H. cumulicola* are caused by genetic differentiation, and to help monitor gene flow between populations.

Acknowledgements

We thank Vaughan Symonds, Matt Gitzendanner, and Ashley Morris for guidance with laboratory techniques.

References

Boutin-Ganache I, Raposo M, Raymond M, Deschepper CF (2001) M13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allele sizing methods. *BioTechniques*, **31**, 24–28.

Dolan RW, Yahr R, Menges ES, Halfhill MD (1999) Conservation implications of genetic variation in three rare species endemic to Florida rosemary scrub. *American Journal of Botany*, **86**, 1556–1562.

Goudet J (2002) *FSTAT (version 2.9.3.2): a computer program to calculate F-statistics*. URL: http://www2.unil.ch/popgen.softwares.fstat.htm.

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. *Trends in Ecology & Evolution*, **11**, 424–429.

Kandpal RP, Kandpal G, Weissman SM (1994) Construction of libraries enriched for sequence repeats and jumping clones, and hybridization selection for region-specific markers. *Proceedings of the National Academy of Sciences USA*, **91**, 88–92.

Quintana-Ascencio P, Weekly CW, Menges ES (2007) Comparative demography of a rare species in Florida scrub and road habitats. *Biological Conservation*, in press.

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

Rozen S, Skaletsky HJ (2000) PRIMER 3 on the WWW for general users and for biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds Krawetz S, Misener S), pp. 365–386. Humana Press, Totowa, New Jersey.