Molecular Ecology Notes (2002) **2**, 589–592 doi: 10.1046/j.1471-8278 .2002.00329.x

PRIMER NOTE

Isolation of polymorphic microsatellites in the stemless thistle (*Cirsium acaule*) and their utility in other *Cirsium* species

ALISTAIR S. JUMP, DEBORAH A. DAWSON, CELIA M. JAMES,* F. IAN. WOODWARD and TERRY BURKE

*Department of Animal and Plant Sciences, University of Sheffield, Sheffield, S10 2TN, UK, *Horticultural Research International East Malling, West Malling, Kent, ME19 6BJ, UK*

Abstract

The genus *Cirsium* includes species with both widespread and restricted geographical distributions, several of which are serious weeds. Nine polymorphic microsatellite loci were isolated from the stemless thistle *Cirsium acaule*. Eight were polymorphic in *C. acaule*, six in *C. arvense* and seven in *C. heterophyllum*. One locus monomorphic in *C. acaule* showed polymorphism in *C. heterophyllum*. The mean number of alleles per locus was 4.1 in *C. acaule*, 6.2 in *C. arvense* and 2.9 in *C. heterophyllum*. These nine loci were also amplified in *C. eriophorum* and *C. vulgare*, suggesting that these markers may be of use throughout the genus.

Keywords: Asteraceae, *Cirsium*, cross-species amplification, microsatellite, thistle, weed

Received 27 July 2002; revision received 3 September 2002; accepted 3 September 2002

The genus *Cirsium* occurs throughout the Northern Hemisphere with approximately 60 species native to Europe. *Cirsium arvense* and *C. vulgare* have become serious weeds throughout temperate zones across the globe. Several species reach geographical limits within the British Isles (*C. acaule*, *C. dissectum*, *C. eriophorum* and *C. heterophyllum*). We developed microsatellite markers in *Cirsium* towards a study investigating differences in population genetic structure between core and peripheral regions of a species' geographical range.

DNA extracted from fresh *C. acaule* leaves following the protocol of Doyle & Doyle (1987) was used to create a genomic library enriched for the dinucleotide sequences $[GC]_n$, $[AC]_n$, $[GA]_n$ and $[GT]_n$, and the trinucleotide sequences $[CAA]_n$, $[GCC]_n$, $[CTG]_n$ and $[CAG]_n$. The protocol was essentially that of Edwards *et al*. (1996). In brief, 200 ng DNA was digested with *Rsa*I (Amersham Pharmacia Biotech, Little Chalfont, Buckinghamshire, UK), ligated to *Mlu*I linkers (Edwards *et al*. 1996) and amplified by the polymerase chain reaction (PCR) using one linker sequence as a primer. The denatured fraction was hybridized with the dinucleotide and trinucleotide target sequences bound in combination (10 ng per oligonucleotide, one filter for trinucleotides, one for dinucleotides) to 0.5 cm^2 pieces of nylon membrane (Hybond N^+, Amersham Pharmacia Biotech). After extensive washing, the bound *C. acaule* DNA was recovered and amplified as before. The linkers were removed by digestion with *Mlu*I (Amersham Pharmacia Biotech) and the DNA ligated into a pJV1 plasmid (a pUC19 plasmid containing a *Bss*HII site, modified and supplied by K. Edwards, IACR, Long Ashton, UK). The plasmids were transformed into XL1 blue competent cells (Stratagene, La Jolla, CA, USA) and plated onto LB-agar plates containing ampicillin, IPTG and Xgal for selection. Colonies containing microsatellite sequences were identified by blotting onto nylon membrane and hybridizing against the target sequences radiolabelled with ^{32}P-dCTP. Positive clones were sequenced in both directions using *Mlu*I linker sequence as primer and BIG DYE terminators on an ABI 377 Sequencer (Applied Biosystems, Foster City, CA, USA). PCR primers were designed from unique sequences with the assistance of Primer 3 (Rozen & Skaletsky 1996/1997). Primers from each pair were modified at the 5′ end to include either a fluorescent phosphoramidite (TET, HEX or FAM) or the sequence GTTTCTT, added to reduce noise from variable adenylation during the PCR (Brownstein *et al*. 1996; Table 1).

Correspondence: AS Jump. Fax: +44 0 114 2220002; E-mail: a.s.jump@sheffield.ac.uk

Table 1 Characteristics of nine *Cirsium acaule* microsatellite loci

Locus (EMBL no.)	Repeat in cloned allele	Modification at 5′ primer end (5′–3′)	Primer sequence (5′–3′)	Allele size range (bp)	Cycling conditions: Program	Cycling conditions: $MgCl_2$ (mM)	No. alleles	H_O	H_E
Caca01	$(CA)_{10}$	F: HEX-	TTT GAA GTG GAT CTT CGC ACG	239–233	64	2.5	4	0.64	0.63
(AJ457836)		R: GTTTCTT-	CAT GGG AGA CGA ACT AAC AGA TGC						
Caca04	$(CA)_{12}$	F: GTTTCTT-	ATC ACC GCT TCC ACC GTC TC	103–122	65/55TD	2.5	3	0.52	0.42
(AJ457839)		R: FAM-	GCT TAT TAG AAC CGC CAT TGA AAG C						
Caca05	$(CA)_{12}$	F: TET-	ACC CAA CCC TCG ATC TGA A	164–166	62/52TD	1.5	2	0.52	0.52
(AJ457840)		R: GTTTCTT-	GAG GAT ACC GGC GAT TGT TA						
Caca07	$(GT)_9(GA)_{11}$	F: GTTTCTT-	CCC AAA CTC CCA CCT TCA TTT G	148–160	64	2.5	5	0.72	0.71
(AJ457842)		R: HEX-	GTC GGA GAT GCT CCG GTG AC						
*Caca10**	$(TTC)_8$-$(TC)_{20}$	F: GTTTCTT-	GAA TTC GCG ACA ACA CAC GC	196	65/55TD	1.5	4	NCA	
(AJ457845)		R: FAM-	GGT AAG GAA TGA ATG ATT GGG CTC						
Caca16	$(GT)_{10}$	F: TET-	TCG TGC TCT TCG ATT GAT TG	117–135	60	2.5	4	0.48	0.52
(AJ457851)		R: GTTTCTT-	CAG AAA ACC GCT CCA TTG C						
*Caca17**	$(CT)_{12}(GT)_9$-	F: TET-	GGC ATA CTG ACA TTC TCA AAC GC	328–332	55	1.5	2	NCA	
(AJ457852)	$(GT)_4$-$(GT)_8$	R: GTTTCTT-	CGT GAT GTG ATG GCA TGT TC						
*Caca22***	$(TC)_9$-$(TC)_9$	F: TET-	GGC TCT GCC TCA CCC ATC TC	190	65/55TD	1.5	1	0	0
(AJ457856)		R: GTTTCTT-	AGG TGT TCA GCA CGG TTC GG						
Caca24	$(CA)_{10}$	F: HEX-	TGG ATA ACG CGC TAG ATC AC	232–244	62/52TD	2.5	7	0.71	0.73
(AJ457858)		R: GTTTCTT-	AAG AAC TCA ATT AGT AGG AAG TGG						

PCR programs
55: (95 °C 3 min): 35 cycles of (94 °C 30 s, 55 °C 30 s, 72 °C 30 s): (72 °C 10 m)
60: (95 °C 3 min): 35 cycles of (94 °C 30 s, 60 °C 30 s, 72 °C 30 s): (72 °C 10 m)
64: (95 °C 3 min): 35 cycles of (94 °C 30 s, 64 °C 30 s, 72 °C 30 s): (72 °C 10 m)
62/52TD: (95 °C 3 min): 5 cycles of (94 °C 30 s, T °C 30 s, 72 °C 30 s, where T drops from 62 to 54 in 2 °C steps): 15 cycles of (94 °C 30 s, 52 °C 30 s, 72 °C 30 s): (72 °C 10 m)
65/55TD: (95 °C 3 min): 5 cycles of (94 °C 30 s, T °C 30 s, 72 °C 30 s, where T drops from 65 to 57 in 2 °C steps): 15 cycles of (94 °C 30 s, 55 °C 30 s, 72 °C 30 s): (72 °C 10 m)

*Locus not amplifying consistently, across all individuals (NCA) in *C. acaule* but of utility in *C. heterophyllum* and *C. arvense*; **monomorphic in *C. acaule* but polymorphic in *C. heterophyllum*. See Table 2. n = 25 individuals tested.

Table 2 Amplification and polymorphism of *Cirsium acaule* microsatellite loci in other *Cirsium* species

	Cycling conditions		*Cirsium arvense* (n = 25)				*Cirsium heterophyllum* (n = 25)				*Cirsium vulgare* (n = 5)	*Cirsium eriophorum* (n = 5)
Locus	Program	$MgCl_2$ (mM)	Allele size range (bp)	No. alleles	H_O	H_E	Allele size range (bp)	No. alleles	H_O	H_E		
Caca01	64	2.5	233–242	5	0.96	0.66	235–237	2	0.28	0.31	+	+ +
Caca04	65/55TD	2.5	103–119	5	0.60	0.55	112–126	3	0.16	0.34	+	+
Caca05	62/52TD	1.5	144–174	7	0.92	0.72	144	1	0	0	+ +	+ +
Caca07	64	2.5	142–168	7	NCA		—	NP	0	0	+	+
Caca10	65/55TD	1.5	167–183	6	0.48	0.60	168–192	2	0.76	0.63	+	+ +
Caca16	60	2.5	—	NP	0	0	118–143	3	0.96	0.60	+ +	+ +
Caca17	55	1.5	290	1	0	0	305–317	3	0.12	0.11	+	+
Caca22	65/55TD	1.5	192–218	7	NCA		210–227	3	0.08	0.21	+	+
Caca24	62/52TD	2.5	220	1	0	0	222–266	4	0.48	0.53	+ +	+ +

NP = no product of the expected size, NCA = not amplifying consistently across all individuals. + indicates PCR amplification observed, + + indicates possible polymorphism. PCR programs are detailed in Table 1.

Twenty-five individuals from each of three *Cirsium* species (*C. acaule*, *C. arvense* and *C. heterophyllum*) were used to assess polymorphism. PCR reactions were carried out in 10 μL volumes containing 50 ng genomic DNA, 1 μM of each forward and reverse primer, 0.25 units Thermoprime Plus DNA polymerase (ABGene, Epsom, Surrey, UK) in the manufacturer's buffer (final concentrations 20 mM $(NH4)_2SO_4$, 75 mM Tris-HCl pH 9.0, 0.01% Tween), 1.5–2.5 mM $MgCl_2$ (see Table 1), and 200 μM of each dNTP, with the addition of 0.5% Tween 20 (final concentration). PCR amplification was performed in a Hybaid Touchdown thermal cycler (Thermo Hybaid, Ashford, Middlesex, UK). PCR products were visualized initially on 2% agarose gels stained with ethidium bromide; primer pairs amplifying a consistent product were then screened for polymorphism on 6% polyacrylamide gels stained with silver (Promega, Southampton, UK) (Bassam *et al*. 1991). Genotypes were assigned using an ABI 377 Sequencer and GeneScan v3.1 software (Applied Biosystems). Expected heterozygosities were calculated using GENETIX v4.02 (Belkhir *et al*. 2001). Reaction profiles are given in Tables 1 and 2.

Of the 110 clones sequenced, 44 contained dinucleotide repeats, three contained trinucleotide repeats and one contained a tetranucleotide repeat. Primers were designed from 24 unique sequences with a minimum of nine uninterrupted repeat units. Seven did not produce a product of the expected size and eight did not amplify consistently across all individuals of at least one species. All 24 sequences were submitted to the EMBL database (accession nos AJ457836–AJ457859). Eight loci detected polymorphism in *C. acaule*, six in *C. arvense* and seven in *C. heterophyllum*. These revealed two to seven alleles per locus with a mean observed heterozygosity of 0.60 in *C. acaule*, 0.74 in *C. arvense* and 0.41 in *C. heterophyllum* (Tables 1 and 2). PCR amplification of all nine loci was also observed in the more distant species *C. eriophorum* and *C. vulgare*; polymorphism in these species was not assessed (Table 2). A search of the EMBL sequence database identified a 47-base pair (bp) region of similarity between sequence *Caca20* (AJ457855) and the NADH dehydrogenase gene from a range of plant species including *Arabidopsis thaliana* and *Oryza sativa*.

Microsatellite repeats are found less frequently in plants than animals (Lagercrantz *et al*. 1993). There is generally a lower level of cross-utility of loci among different plant taxa and when amplification occurs, fewer of the loci are usually found to be polymorphic (Whitton *et al*. 1997; Peakall *et al*. 1998). Screening loci across several congeneric species, as reported here, demonstrates that some loci that show poor amplification, or appear monomorphic in the species from which the markers are developed, may be of utility and detect polymorphism in related species. Cross-amplification of the *C. acaule* markers in the other species described here suggests that these markers are likely to be of wider application within the genus *Cirsium*.

Acknowledgements

This work was supported by a Hossein Farmy scholarship from the University of Sheffield and assisted by the Sheffield Molecular Genetics Facility (supported by the Natural Environmental Research Council). We thank Andy Krupa for help with genotyping and Robert Jehle and Lisa Pope for their comments on the manuscript.

References

Bassam BJ, Caetano-Anolles G, Gresshoff PM (1991) Fast and sensitive silver staining of DNA in polyacrylamide gels. *Analytical Biochemistry*, **196**, 80–83.

Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F (2001) GENETIX, *Version* 4.02. *Logiciel Sous Windows TM Pour la Génétique*

Des Populations. CNRS UPR. 9060. Laboratoire Génome et Populations, Université de Montpellier II, Montpellier, France.

Brownstein MJ, Carpten JD, Smith JR (1996) Modulation of non-templated nucleotide addition by *Taq* DNA polymerase: primer modifications that facilitate genotyping. *Biotechniques*, **20**, 1004–1010.

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, **19**, 11–15.

Edwards KJ, Barker JHA, Daly A, Jones C, Karp A (1996) Microsatellite libraries enriched for several microsatellite sequences in plants. *Biotechniques*, **20**, 758–760.

Lagercrantz U, Ellegren H, Andersson L (1993) The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. *Nucleic Acids Research*, **21**, 1111–1115.

Peakall RS, Gilmore S, Keys W, Morgante M, Rafalski A (1998) Cross-species amplification of soybean (*Glycine max*) simple sequence repeats (SSRs) within the genus and other legume genera: implications for the transferability of SSRs in plants. *Molecular Biology and Evolution*, **15**, 1275–1287.

Rozen S, Skaletsky HJ (1996/1997) Primer3. Code available at: http://www-genome.wi.mit.edu/genome_software/other/primer3.html

Whitton J, Rieseberg LH, Ungerer MC (1997) Microsatellite loci are not conserved across the Asteraceae. *Molecular Biology and Evolution*, **14**, 204–209.