

PRIMER NOTE

Nine polymorphic microsatellite markers in *Centaurea stoebe* L. [subspecies *C. s. stoebe* and *C. s. micranthos* (S. G. Gmelin ex Gugler) Hayek] and *C. diffusa* Lam. (Asteraceae)

ROBIN A. MARRS,*† RUTH A. HUFBAUER,* STEVEN M. BOGDANOWICZ‡ and RENÉ SFORZA§

*Department of Bioagricultural Science and Pest Management, Colorado State University, Fort Collins, CO 80523, USA, †Graduate Degree Program in Ecology, Colorado State University, USA, ‡Evolutionary Genetics Core Facility, Cornell University, USA, §European Biological Control Laboratory, USDA-ARS, Montpellier, France

Abstract

Centaurea stoebe (subspecies *C. s. stoebe* and *C. s. micranthos* [S. G. Gmelin ex Gugler] Hayek) and *Centaurea diffusa* are Eurasian plant species that have invaded much of North America. We isolated seven microsatellite loci from *C. stoebe* and two loci from *C. diffusa*. All loci described here amplify in both species and have between six and 25 alleles each. These markers will be useful in examining population structure and addressing questions regarding these invasions.

Keywords: *Centaurea diffusa*, *Centaurea maculosa*, *Centaurea stoebe micranthos*, *Centaurea stoebe stoebe*, invasive species, knapweed, microsatellites

Received 29 January 2006; revision accepted 17 March 2006

Spotted knapweed, *Centaurea stoebe* L. [including subspecies *C. s. stoebe* (diploid) and *C. s. micranthos* (S. G. Gmelin ex Gugler) Hayek (tetraploid), synonyms: *Centaurea maculosa* L. and *Centaurea biebersteinii* DC], and diffuse knapweed, *Centaurea diffusa* Lam., (Asteraceae) are two of the most important and pernicious rangeland weeds plaguing the western USA (Sheley *et al.* 1998). In their invasive range, these species can form monocultures (Sheley & Larson 1996), displace native grass species (Kedzie-Webb *et al.* 2001) and are able to invade natural systems in the absence of human disturbance (Tyser & Key 1988). Both species were introduced inadvertently to North America from Eurasia in the late 1800s or early 1900s, possibly multiple times each (Watson & Renney 1974). In the native range, tetraploid and diploid forms of both species are spatially isolated; tetraploids are found in the eastern Eurasian and Asian parts of their range, while diploids occur primarily in the western European portion of the range (Ochsmann 2000). Morphological intermediates between *C. stoebe* and *C. diffusa* were described as a hybrid species, *Centaurea xpsammogena*, in the native range (Gáyer 1909), and similar intermediates have been observed in the introduced range (Watson & Renney 1974;

Ochsmann 2000). In light of the economic and ecological importance of these species, as well as the interesting evolutionary questions raised by the potential for interspecific crossing in these taxa, we have developed nine polymorphic microsatellite markers that amplify in both *C. stoebe* and *C. diffusa*.

Microsatellite markers were developed from mixed genomic DNA from several introduced-range sources. Genomic DNA from *C. s. micranthos* collected in Montana and Colorado was used to develop markers CM26, 21CM36, 38CM22, 42CM27, CM15, 25CM6 and CM17. Markers CD37 and CD9 were developed from mixed *C. diffusa* genomic DNA extracted from plants found in Montana and Washington.

Microsatellite cloning and sequencing was performed at the Evolutionary Genetics Core Facility of Cornell University following a modification of the protocol of Hamilton *et al.* (1999). The genomic DNA was digested with *AluI* and *HaeIII*, a SNX linker was ligated and the DNA was enriched for repeats by hybridization to 3'-biotinylated (GT)₁₅ and (CT)₁₅ oligonucleotides followed by magnetic capture with streptavidin-coated magnetic beads. Enriched fragments were made double-stranded by polymerase chain reaction (PCR), digested with *NheI* and cloned into *XbaI*-digested, dephosphorylated pUC19. Colonies were grown on Luria-Bertani plates with ampicillin, and replicas were transferred

Correspondence: Robin A. Marrs, Fax: 1(970)491-3862; E-mail: rmarrs@lamar.colostate.edu

Table 1 Marker information for *Centaurea stoebe* and *Centaurea diffusa*; locus names, GenBank Accession nos, forward (F) and reverse (R) primer sequences [*indicates location of M13(-21) tail for attachment of fluorescent dye to forward primer, sequence = TGTAACACGACGGCCAGT], annealing temperatures (N_m), fluorescent dye, repeat motifs, size ranges of PCR products, number of alleles observed, and observed (H_O) and expected (H_E) heterozygosities for each locus, calculated by species and ploidy

Locus	GenBank ID	Primer Sequence (5' → 3')	N_m (°C)	Dye	Repeat Motif	Size Range (bp)	No. of alleles	2n C. s. stoebe		4n C. s. micranthos		4n C. diffusa	
								H_O	H_E	H_O	H_E	H_O	H_E
Primer set 1:													
CM26	DQ356425	F:* GAAGGGCTACGAGGGTGTTC R: GAAGTGTGTGCAATTTCAATCTATT	55	VIC	(TG) ₉ T(TA) ₃	133–172	22	0.455	0.510	0.815	0.852	0.737	0.829
CD37	DQ356426	F:* AGGTGCACCTTTCCTGTCAAC R: CAACCCAATAAGATTACTTCCACTTC	58	NED	(CA) ₉	154–176	7	monomorphic		0.646	0.703	0.301	0.652
21CM36	DQ356419	F:* GCTATTAAACAACCTCCAAAATGAACAG R: CCTGCTCCAACAAGTTTCCTC	45	6-FAM	(CA) ₆ (TA) ₅ (TG) ₁₆	187–244	6	0.250	0.228	0.536	0.488	0.565	0.423
38CM22	DQ356427	F:* GGCTACATTAAGCTTATCCATTC R: CTCGCATGTTATCCCTCCCTC	55	VIC	(GA) ₁₂ AA(GA) ₈	289–344	25	0.000	0.699	0.291	0.900	0.436	0.802
CD9	DQ356418	F:* GGTCCTCCATCTTCAAGCTAATAAC R: ATGCTTCCCTTCTCAATGTTTTCTCT	55	NED	(CA) ₁₇	301–325	9	0.100	0.700	0.718	0.863	0.295	0.808
Primer set 2:													
42CM27	DQ356422	F:* TGGGATATTCGTTGGTTTAGTTTT R: CCTCCCACTCCCGTTTGAC	58	6-FAM	(TG) ₁₄	155–175	11	0.205	0.298	0.718	0.753	0.600	0.733
CM15	DQ356424	F:* GGAGGGCATGGGATTAAGAGAT R: TGGATGCATCGGTCTGGAAATA	55	VIC	(GT) ₉	185–220	10	0.263	0.689	0.525	0.729	0.283	0.730
25CM6	DQ356420	F:* ATGGGACATAAGATCCACAACAG R: TAATTCAGCATTCAAAAATTTAGAAGAC	45	NED	(CA) ₉	255–272	7	0.190	0.251	0.200	0.824	0.458	0.864
CM17	DQ356421	F:* TACTTGGGCTTTTCGCTAATGAT R: ACAACGTGTTCCAGCAGCAG	60	VIC	(AC) ₉	379–430	16	0.063	0.502	0.398	0.782	0.706	0.775

to nylon membranes and probed with radiolabelled (GT)₁₅ and (CT)₁₅. Positive colonies were sequenced with a BigDye Terminator Cycle Sequencing Kit (Applied Biosystems) and a universal M13 primer that flanks the cloning site. Reactions cycled at 95 °C for 50 s, 50 °C for 20 s and 60 °C for 4 min (25 cycles total) then were analysed on an ABI 377 Automated Sequencer.

Primer pairs were designed for regions flanking the microsatellites using the program PRIMER SELECT (DNASar). Forward primers had universal M13(-21) tails added to their 5' ends following the protocol of Schuelke (2000). These primer pairs were tested for amplification with genomic DNA extracted from both species, *C. stoebe* (subspecies *C. s. stoebe* and *C. s. micranthos*) and *C. diffusa*. These DNAs were isolated with Plant Mini Kits (QIAGEN). PCR amplifications were performed with 10 µL total reaction volumes of 1× PCR buffer (20 mM Tris-HCL, pH 8.4, 50 mM KCl), 2 mM MgCl₂, 0.2 mM dNTPs, 0.5 U *Taq* polymerase (Novagen), 0.1 µL *TaqStart* antibody (Clontech), 0.4 pmol of forward primer with M13(-21) tail, 1.8 pmol each of reverse primer and fluorescently labelled (6-FAM, NED or VIC) M13(-21) universal primer and 1 µL of extracted genomic DNA. Samples were amplified using a Mastercycler or Mastercycler Gradient thermal cycler (Applied Biosystems) with an initial denaturation step of 94 °C for 1 min, then 36 cycles of 50 s at 94 °C, 1 min at annealing temperature (see Table 1) and 1 min at 72 °C for polymerase extension, followed by a final extension at 72 °C for 10 min. PCR products were stored at -20 °C until genotyping. For this initial marker development test, we genotyped 229 individuals from several sample locations in their native and introduced ranges.

Prior to genotyping, PCR products were multiplexed in two batches according to the scheme given in Table 1. Four (primer set 2) or five (primer set 1) of these loci can be genotyped in a single sequencer run without overlap of similarly fluorescently labelled alleles. Genotyping was performed on an ABI 3100-Avant capillary sequencer (Applied Biosystems) at Colorado State University. Reference DNA samples were run with each genotyping plate. Files were analysed with GENEMAPPER version 3.0 (Applied Biosystems).

The ploidy of our samples was unknown, and we used microsatellite banding patterns to infer ploidy. If a sample location contained an individual that showed three or more alleles at any of the nine loci, all individuals in that sample were analysed as tetraploids. This assumes uniform ploidy within populations, a hypothesis that has not yet been tested. All *C. diffusa* sample locations contained individuals with three or more alleles, so all *C. diffusa* were analysed as tetraploids. Samples from two *C. stoebe* locations in Western Europe appeared to be diploid (*C. s. stoebe*), while the rest appeared to be tetraploid (*C. s. micranthos*).

The number of alleles at each locus ranged from six to 25. Observed (H_O) and expected (H_E) heterozygosities at each locus were calculated using ARLEQUIN version 2.000

(Schneider *et al.* 2000) for diploid populations (Table 1). Observed heterozygosity was calculated by hand for tetraploids, and expected heterozygosity was calculated using SPAGED1 version 1.2 (Hardy & Vekemans 2002) (Table 1). Note that some of the assumptions of Hardy-Weinberg equilibrium may be violated in these invasive species, so H_E may not be as useful a metric for measuring allelic diversity as H_O . Observed heterozygosity was generally highest in *C. s. micranthos* (tetraploid), followed closely by tetraploid *C. diffusa*. *Centaurea stoebe stoebe* (diploid) had the lowest observed heterozygosity at each locus, which is to be expected as it has two possible alleles, while the tetraploids have four possible alleles and many more ways of being heterozygous.

Acknowledgements

We thank Aaron Jackson for laboratory assistance, Lincoln Smith for collecting samples and Susan Knudson for sequencer expertise. This work was supported by USDA ARS Cooperative Agreement 5840121141, USDA NRI 2002-35320-12137 and the Colorado Agricultural Experiment Station.

References

- Gáyer G (1909) Vier neue Centaureen der Flora von Ungarn. *Magyar Botanikai Lapok*, **8**, 58–61.
- Hamilton MB, Pincus EL, Di Fiore A, Fleischer RC (1999) Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. *BioTechniques*, **27**, 500–507.
- Hardy OJ, Vekemans X (2002) SPAGED1: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.
- Kedzie-Webb SA, Sheley RL, Borkowski JJ, Jacobs JS (2001) Relationships between *Centaurea maculosa* and indigenous plant assemblages. *Western North American Naturalist*, **61** (1), 43–49.
- Ochsmann J (2000) *Morphologische und molekularsystematische Untersuchungen an der Centaurea stoebe L.-Gruppe (Asteraceae-Cardueae) in Europa*. PhD Dissertation (ed. Cramer J), Stuttgart, Georg-August University, Göttingen.
- Schneider S, Roessli D, Excoffier L (2000) *ARLEQUIN: A Software for Population Genetics Data Analysis Version 2.000*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.
- Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments: a poor man's approach to genotyping for research and high-throughput diagnostics. *Nature Biotechnology*, **18**, 233–234.
- Sheley RL, Larson LL (1996) Emergence date effects on resource partitioning between diffuse knapweed seedlings. *Journal of Range Management*, **50**, 263–267.
- Sheley RL, Jacobs JS, Carpinelli MF (1998) Distribution, biology, and management of diffuse knapweed (*Centaurea diffusa*) and spotted knapweed (*Centaurea maculosa*). *Weed Technology*, **12**, 353–362.
- Tyser RW, Key CH (1988) Spotted knapweed in natural area fescue grasslands: an ecological assessment. *Northwest Science*, **62** (4), 151–159.
- Watson AK, Renney AJ (1974) The biology of Canadian weeds. 6. *Centaurea diffusa* and *C. maculosa*. *Canadian Journal of Plant Science*, **54**, 687–701.