American Journal of Botany: e4–e6. 2010.

AJB Primer Notes & Protocols in the Plant Sciences

Isolation and characterization of novel microsatellite markers for *Arctium minus* (Compositae)[1]

Sara López-Vinyallonga[2,6], Mónica Arakaki[3], Núria Garcia-Jacas[2], Alfonso Susanna[2], Matthew A. Gitzendanner[4], Douglas E. Soltis[4] and Pamela S. Soltis[5]

[2]Botanic Institute of Barcelona (CSIC-ICUB), Pg. del Migdia s. n. 08038 Barcelona, Spain; [3]Department of Ecology and Evolutionary Biology, Brown University, 80 Waterman St., Box G-W, Providence, Rhode Island 02912 USA; [4]Department of Biology, University of Florida, Gainesville, Florida 32611 USA; [5]Florida Museum of Natural History, University of Florida, Gainesville, Florida 32611 USA

- *Premise of the study:* Microsatellite primers were developed for the invasive plant *Arctium minus* to investigate the effects of facultative self-pollination and the biannual habit on population genetic structure, as well as the colonization of the Americas by this Eurasian species.
- *Methods and Results:* Sixteen di- and trinucleotide microsatellite loci were identified in six populations. The number of alleles per locus ranged from one to 10, observed heterozygosities ranged from 0.000 to 0.897, and the mean value of F_{IS} was 0.316.
- *Conclusions:* These results indicate the utility of these loci in future studies of population genetics in *A. minus*.

Key words: *Arctium minus*; enriched library; lesser burdock; microsatellites; population genetics; SSR.

Arctium minus (Hill) Bernh. is a diploid biennial plant, native to Eurasia. It is widespread and often naturalized as a weed in many parts of North and South America, and it also grows in Africa, where it is rare. *Arctium minus* is allogamous (outcrossing), although it can be self-pollinated if insect pollination fails. In previous phylogenetic studies (López-Vinyallonga et al., 2009), little sequence divergence among species of the *Arctium-Cousinia* complex was found, most likely due to rapid and recent divergence in the group. The high polymorphism of microsatellite markers can be helpful for studying the population genetics of this complex, as it has been for other recently diverged groups (e.g., Chirhart et al., 2005; Edwards et al., 2008). In addition, *A. minus* is considered an invasive plant in many countries such as the US, Canada, and New Zealand, and the study of its population genetics can be useful for developing management strategies. We report here a set of novel polymorphic microsatellites for *A. minus* that will be useful for assessing genetic variability and divergence as well as gene flow within and among populations.

METHODS AND RESULTS

Genomic DNA was extracted, using the NucleoSpin Plant Kit (Macherey-Nagel GmbH & Co. KG, Düren, Germany), from dried leaf tissue of specimens collected from six wild populations representing the native distribution of *A. minus*. Altogether, 134 individuals were sampled corresponding to six populations in central and southern Europe and eastern Asia with sample sizes ranging from 18 to 26 individuals per population (for details see figure caption in Table2). Two genomic libraries enriched for AAG and CA repeats were constructed from *A. minus* following Symonds et al. (2010). The genomic DNA was digested with *Sau*3AI and enriched for repeats by hybridization to 3′-biotinylated $(AAG)_{14}$TATAAGATA and $(CA)_{15}$TATAAGATA oligonucleotides, respectively, followed by magnetic capture with streptavidin-coated magnetic beads (Promega Corp, Madison, WI, USA). Enriched fragments were made double-stranded by polymerase chain reaction (PCR) and were ligated into a TOPO TA pPCR 4.0 vector, transformed into One Shot *Escherichia coli* competent cells (Invitrogen, Carlsbad, CA, USA) and grown on Luria Broth (LB) agar plates with Kanamycin. Colonies were picked and transferred to LB/Kanamycin 96-well plates and grown overnight. Screening for positives was PCR-based, involving two PCRs per sample, with a repeat primer (either AAG or CA) and a primer for the vector (M13F or M13R).

A total of 264 positive clones for AAG and 235 for CA was detected; 192 and 160 of these, respectively, were sequenced using an ABI 3730xl DNA Analyzer (Applied Biosystems, Foster City, CA, USA). Using the criterion of at least six repeat units in the target sequence, we designed primers from 28 (for AAG) and 14 (for CA) of these sequences using PRIMER3 version 0.4.0 (Rozen and Skaletsky, 2000) and Operon (Operon Biotechnologies, Huntsville, AL USA). Forward primers had universal M13 tails added to their 5′ ends following Boutin-Ganache et al. (2001).

Amplifications were performed in 10-µl reactions containing 0.5 U of Go*Taq* Flexi DNA polymerase (Promega), 1× Promega colorless Go*Taq* Flexi Buffer, 2.0 mM $MgCl_2$, 50 µM of each dNTP, 0.45 µM of the reverse primer, 0.012 µM of the extended forward primer, 0.45 µM of the labeled M13 primer (6-FAM, VIC, NED or PET, Applied Biosystems) and 25 ng/µl of template DNA. The profile used for amplifications consisted of a denaturation step of 94°C for 3 min, followed by 34 cycles of 94°C for 30 s, 52°C for 30 s, and 72°C for 45 s and a final elongation time of 20 min at 72°C. Labeled PCR products were diluted 1 : 20 and pool-plexed by using 1–4 µl of each PCR (up to four loci). One µl of the diluted PCR mixture, 9.9 µl of formamide, and 0.1 µl LIZ 600 size standard (Applied Biosystems) was loaded and run on an ABI 3730xl DNA Analyzer (Applied Biosystems) at the Interdisciplinary Center for Biotechnology Research (ICBR) facility at the University of Florida. Specific amplification was achieved for

[1] Manuscript received 7 December 2009; revision accepted 29 December 2009.
This work was supported by a grant from the Spanish Government through the Ministry of Education and Science (CGL 2004-04563-C02-01/BOS and CGL 2007-60781/BOS) to López-Vinyallonga. The authors thank Borşič, Montes-Moreno, Romashchenko, Romo and Vilatersana for their help with plant collection.

[6] Author for correspondence: slopez@ibb.csic.es

doi:10.3732/ajb.0900376

Table 1. Characteristics of 16 microsatellite primers developed in *Arctium minus*. The six populations tested are from France, Poland, Slovakia, Spain (2) and Turkey (N = 18–26 for each).

Locus	Primer Sequence (5′→3′)	T_a (°C)	Repeat Motif	Size Range (bp)	GenBank
Am2	F: *AGAAAGGAAAGGGGGAGCTT	50	$(AAG)_6$	212–247	GU354039
	R: TCTTCTGGATCTGCCTCGAT				
Am8	F: *ATCGCCATCGAAGACAAGAC	54	$(AAG)_9$	108–152	GU354040
	R: AATAAAATCCTGAGCCGGAAA				
Am17	F: *TCGTGGGACTCTACCACCTC	52	$(AAG)_6$	254–288	GU354041
	R: TCCTGGACCCAGATCGTACT				
Am25	F: *GGTTTGGTTCTCCCTCAGGT	45	$(TC)_8$	278	GU354042
	R: ATCAAGCCGGTGACCATATC				
Am26	F: *TTGGGGTGAGGTAGAAGCAG	52	$(AAG)_6$	260–265	GU354043
	R: TTCCACCAAGTTGGGTTAGC				
Am27	F: *CGGGTTCTGACTTAGCTTGC	50	$(AAG)_6$	361	GU354044
	R: GGTAACATCCGTTTCGTTCG				
Am30	F: *GCAAAGGGGTCCTAGAGCAT	52	$(CA)_8$	180–189	GU354045
	R: TCGAAGTGTTATCGGTTGCT				
Am31	F: *TGTGCAACTGCTCCTTCAGT	52	$(CA)_9$	182–272	GU354046
	R: CTCCAACAATGCAGAAACCA				
Am32	F: *GCTGTTGCCATGACTCTAAGG	52	$(CA)_8$	212–226	GU354047
	R: CGGAAAAGGACGACAAAGAA				
Am33	F: *TCCCTTGTGAAACGCAATTT	52	$(CA)_9$	197–203	GU354048
	R: CGGTCAGATCCATTTATCACG				
Am34	F: *CCATGCTCACCTCCATTTCT	52	$(GA)_9$	158–185	GU354049
	R: CAGCATAATGATACGGCAACA				
Am35	F: *AGTTAGTGTCATTGTTGAGAGAACTTA	50	$(TATG)_6(TG)_4[(TATG)_2(TG)_4]_4$	174–239	GU354050
	R: TGTGATAGCATCCAAACTCCA				
Am36	F: *TGTTATTTCAGCCCTGGATTG	54	$(CA)_{11}$-$(TA)_6$-$(GA)_5$	204	GU354051
	R: CAACTTACAATTTCAATGGTATCTCTC				
Am37	F: *TCTCACCGGCGATAGAAACT	50	$(TC)_9$	175–204	GU354052
	R: ATACCGGAAGACGGAATGTG				
Am39	F: *TGCAGACACGCATTACAACA	52	$(CA)_3CGC(CA)_5$	248–258	GU354053
	R: TGGCCCTAGAATGATGGAAA				
Am40	F: *CACTGTTGTGGTGGTGGTTT	52	$(CA)_3G(CA)_2N(CA)_5$	182–266	GU354054
	R: GTGGGTGGAGCTAATGTGGA				

Notes: * M13 tail (CACGACGTTGTAAAACGAC), optimized annealing temperature (T_a), repeat motifs, size ranges of PCR products, GenBank accession number.

Table 2. Results of initial primer screening in *Arctium minus*.

	France (N = 24)			**Poland (N = 20)**			**Slovakia (N = 18)**			**Spain-1 (N = 23)**			**Spain-2 (N = 23)**			**Turkey (N = 26)**		
	N_a	H_o	H_e	N_a	H_o	H_e	N_a	H_o	H_e	N_a	H_o	H_e	N_a	H_o	H_e	N_a	H_o	H_e
Locus	26	0,111	0,130	40	0,260	0,356	36	0,243	0,271	28	0,102	0,192	32	0,114	0,266	26	0,105	0,083
Am2	1	0,000	0,000	1	0,000	0,000	1	0,000	0,000	1	0,000	0,000	1	0,000	0,000	1	0,000	0,000
Am8	2	0,033	0,033	3	0,190	0,176	4	0,048	0,430*	2	0,034	0,034	1	0,000	0,000	4	0,138	0,131
Am17	1	0,000	0,000*	3	0,765	0,538*	1	0,000	0,000*	1	0,000	0,000	1	0,000	0,000*	1	0,000	0,000*
Am25	1	0,000	0,000	1	0,000	0,000	1	0,000	0,000	1	0,000	0,000	1	0,000	0,000	1	0,000	0,000
Am26	1	0,000	0,000	1	0,000	0,000	2	0,095	0,091	1	0,000	0,000	1	0,000	0,000	1	0,000	0,000
Am27	1	0,000	0,000	1	0,000	0,000	1	0,000	0,000	2	0,000	0,000	2	0,038	0,109	1	0,000	0,000
Am30	1	0,000	0,000*	1	0,000	0,000*	2	0,033	0,033	2	0,125	0,249*	4	0,000	0,525*	1	0,000	0,000
Am31	5	0,083	0,641*	4	0,037	0,628*	2	0,167	0,495*	4	0,000	0,432*	2	0,320	0,269*	4	0,897	0,528*
Am32	1	0,000	0,000*	3	0,444	0,580*	4	0,458	0,468	1	0,000	0,000	3	0,368	0,500	2	0,000	0,074*
Am33	2	0,632	0,432*	4	0,304	0,566*	3	0,069	0,131*	3	0,500	0,461*	2	0,167	0,495	2	0,000	0,064*
Am34	1	0,000	0,000*	6	0,417	0,785*	3	0,143	0,255*	2	0,333	0,420	3	0,333	0,431*	1	0,000	0,000*
Am35	3	0,000	0,131*	3	0,000	0,461*	5	0,500	0,689*	2	0,333	0,456	2	0,038	0,109	1	0,000	0,000*
Am36	1	0,000	0,000	1	0,000	0,000	1	0,000	0,000	1	0,000	0,000	1	0,000	0,000	1	0,000	0,000
Am37	1	0,000	0,000	4	0,577	0,454	2	0,640	0,435*	1	0,000	0,000	2	0,083	0,080	1	0,000	0,000*
Am39	1	0,000	0,000	1	0,000	0,000	1	0,000	0,000*	1	0,000	0,000	2	0,167	0,433*	1	0,000	0,000*
Am40	2	0,700	0,455	2	0,643	0,436	2	1,000	0,500	2	0,000	0,444	3	0,000	0,611*	2	0,333	0,278

Notes: Number of alleles (N_a), observed heterozygosity (H_o), expected heterozygosity (H_e). Significant departures from HWE: * $P < 0.005$. Voucher of populations: France, Hérault, Plateau de l'Escandorque, VII-2007, *Montes-Moreno & Vilatersana 1042* (BC); Poland, Kórnik, VI-2007, *Romo 14983* (BC); Slovakia, between Liesek and Hladovka, VI-2007, *Romo 14981* (BC); Spain-1, Barcelona, Dosrius, Canyamars, VIII-2007, *López-Vinyallonga & Vilatersana 1100* (BC); Spain-2, León, between Murias de Paredes and Montrondo, VIII-2007, *Borşič & Vilatersana 1074* (BC); Turkey, National Park of Uludag, near Bursa, VI-2007, *Romo 14639* (BC). The sample size for each population is shown in parentheses.

16 primer sets (Table 1). Fragment analysis was performed using GENEMARKER 1.5 (Soft Genetics, LLC) and Peak Scanner Software 1.0 (Applied Biosystems), with data scored manually in both cases. The data were analyzed with GenAlEx6 (Peakall and Smouse, 2006), and the results are shown in Table 2.

The number of alleles (*N*a) observed at each locus per population ranged from one to 10, with observed heterozygosities (H_o) ranging from 0.000 to 0.897, and expected heterozygosities (H_e) from 0.000 to 0.785 (Table 2). When tested the six populations across loci, estimates of the inbreeding coefficient F_{IS} (Weir and Cockerham, 1984) ranged from –0.752 to 0.890, with a mean value of 0.316 for the studied populations. GENEPOP version 3.4 (Raymond and Rousset, 1995) was used to test for departure from Hardy-Weinberg equilibrium and linkage disequilibrium in the six populations across loci. Six loci showed a significant departure from Hardy-Weinberg equilibrium ($P < 0.005$) due to heterozygote deficiency, and nine out of 78 pairwise locus comparisons showed significant linkage disequilibrium ($P < 0.005$). These results are consistent with the facultative ability of *A. minus* to self-pollinate.

Tests for the presence of null alleles were performed using MICRO-CHECKER 2.2.3 (van Oosterhout et al., 2004). There was no evidence for scoring errors due to stuttering and no evidence for large allele dropout at any of the 16 loci tested. In addition, null alleles were only detected at two loci, Am31 and Am34.

CONCLUSIONS

Microsatellite markers reported here for *A. minus* are suitable for population genetic studies, such as understanding past evolutionary and demographic events. Extensive sampling of more populations has been carried out, and genetic analyses of these populations are in progress to answer questions regarding the colonization of the Americas and other population processes in this species.

LITERATURE CITED

BOUTIN-GANACHE, I., M. RAPOSO, M. RAYMOND, AND C. F. DESCHEPPER. 2001. M-13-tailed primers improve the readability and usability of microsatellite analyses performed with two different allele sizing methods. *BioTechniques* 31: 24–28.

CHIRHART, S. E., R. L. HONEYCUTT, AND I. F. GREENBAUM. 2005. Microsatellite variation and evolution in the *Peromyscus maniculatus* species group. *Molecular Phylogenetics and Evolution* 34: 408–415.

EDWARDS, C. E., D. E. SOLTIS, AND P. S. SOLTIS. 2008. Using patterns of genetic structure based on microsatellite loci to test hypotheses of current hybridization, ancient hybridization and incomplete lineage sorting in *Conradina* (Lamiaceae). *Molecular Ecology* 17: 5157–5174.

LÓPEZ-VINYALLONGA, S., I. MEHREGAN, N. GARCIA-JACAS, O. TSCHERNEVA, A. SUSANNA, AND J. W. KADEREIT. 2009. Phylogeny and evolution of the *Arctium-Cousinia* complex (Compositae, Cardueae-Carduinae). *Taxon* 58: 153–171.

PEAKALL, R., AND P. E. SMOUSE. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288–295.

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

ROZEN, S., AND H. J. SKALETSKY. 2000. PRIMER3 on the WWW for general users and for biologist programmers. *In*: S. Krawetz and S. Misener [eds.], Bioinformatics Methods and Protocols: Methods in Molecular Biology, 365–386. Humana Press, Totowa, New Jersey, USA.

SYMONDS, V. V. 2010. Microsatellite marker development: a review. *Molecular Ecology Resources.* Invited review in press.

VAN OOSTERHOUT, C., W. F. HUTCHINSON, D. P. M. WILLS, AND P. SHIPLEY. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535–538.

WEIR, B. S., AND COCKERHAM, C. C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358–1370.