Development and Evaluation of Microsatellite Markers for a Native Prairie Perennial, Echinacea angustifolia (Asteraceae)

Authors: Ison, Jennifer L., Wagenius, Stuart, Reitz, Diedre, and Ashley, Mary V.

Source: Applications in Plant Sciences, 1(11)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1300049

Applications in Plant Sciences 2013 1(11): 1300049

PRIMER NOTE

DEVELOPMENT AND EVALUATION OF MICROSATELLITE MARKERS FOR A NATIVE PRAIRIE PERENNIAL, *ECHINACEA ANGUSTIFOLIA* (ASTERACEAE)[1]

JENNIFER L. ISON[2,3,5], STUART WAGENIUS[3], DIEDRE REITZ[4], AND MARY V. ASHLEY[2]

[2]Department of Biological Sciences, University of Illinois at Chicago, 845 West Taylor Street, Chicago, Illinois 60607 USA;
[3]Plant Conservation and Biology, Chicago Botanic Garden, 1000 Lake Cook Road, Glencoe, Illinois 60022 USA; and
[4]Department of Biology, Carleton College, One North College Street, Northfield, Minnesota 55057 USA

- *Premise of the study:* Microsatellite loci for the native prairie perennial *Echinacea angustifolia* were developed and evaluated for future use in population structure and paternity studies.
- *Methods and Results:* A total of 50 trinucleotide microsatellite regions were identified though an enrichment protocol that prescreens for microsatellite repeats before ligating into a vector. Of these, 11 loci were polymorphic and in Hardy–Weinberg equilibrium in three populations with varying numbers of plants. The loci had between three and 14 alleles and collectively provided high paternity exclusion probabilities.
- *Conclusions:* These sets of microsatellite primers will provide researchers and land managers with valuable information on the population genetic structure and gene flow between fragmented prairie populations.

Key words: Asteraceae; *Echinacea angustifolia*; fragmentation; microsatellites; perennial herb; tallgrass prairie.

The North American tallgrass prairie is highly fragmented, with less than 1% of the presettlement prairie in existence (Samson and Knopf, 1994). Due to this fragmentation and habitat destruction, native prairie plant populations could be suffering from a number of genetic consequences such as inbreeding depression, genetic drift, and reduced gene flow between fragmented populations (Young et al., 1996). Information on the population genetic structure and pollination patterns within these fragmented prairie plant remnants would help land managers determine conservation priorities and could be instrumental toward ensuring the population persistence of prairies.

Echinacea angustifolia DC. (Asteraceae) has characteristics that are common to many herbaceous prairie plants; it is long-lived, self-incompatible, and insect pollinated by a number of solitary bees (Wagenius and Lyon, 2010). Because of these characteristics, *E. angustifolia* has become a model prairie plant and has been used in a number of studies to assess the genetic and ecological consequences of prairie fragmentation (e.g., Wagenius, 2006; Wagenius and Lyon, 2010). Researchers have annual demographic data on plants from 27 *E. angustifolia* remnants in western Minnesota. In this study, we developed and evaluated 11 *E. angustifolia* microsatellite markers for use in population genetic and paternity studies.

METHODS AND RESULTS

We developed microsatellites for *E. angustifolia* using an enrichment protocol that prescreens for microsatellite repeats before ligating into a vector (Glenn and Schable, 2005). Approximately 10 µg of DNA from an *E. angustifolia* individual was cut with an *Rsa*I restriction enzyme. Double-stranded linkers were added to the ends of the DNA fragments to provide primer binding sites in future PCRs and to assist with ligating the fragment into a vector.

We prescreened for trinucleotide microsatellite regions using Dynabeads magnetic beads (Life Technologies, Grand Island, New York, USA) with a microsatellite oligonucleotide attached, either $(AAG)_8$, $(AAC)_6$, $(AAT)_{12}$, $(ACT)_{12}$, $(ATC)_8$, or $(AGC)_6$. The resulting fragments were ligated into a vector and cloned. We sequenced the clones on an ABI 3730 DNA Analyzer (Life Technologies) using Sanger sequencing. The development process resulted in sequences for 50 trinucleotide microsatellite regions ranging from six to 18 repeats.

We developed primers for all 50 microsatellite regions using Primer3 version 0.4.0 (Rozen and Skaletsky, 2000). The criteria inputted into Primer3 were optimal primer size of 20 bp (minimum 18 bp, maximum 27 bp); CG content of primer between 20% and 80%; optimal primer melting temperature of 60°C (minimum 57°C, maximum 63°C); a maximum self-complementarity score of 8; and a maximum 3′ self-complementarity score of 3. Of the 50 microsatellite regions, 20 loci were variable and had consistent amplification following optimization. We tested these 20 loci for allelic diversity, Hardy–Weinberg equilibrium (HWE), and repeatability; 11 performed well (Tables 1 and 2). None of these 11 loci showed evidence of null alleles (tested using MICRO-CHECKER version 2.2.3; van Oosterhout et al., 2004). Two loci were amplified by the same primer pair (Ech13 and Ech13Z) but had nonoverlapping alleles and were not linked (tested using FSTAT version 1.2; Goudet, 1995). These loci potentially represent a duplicated region in the genome with an insertion in one of the regions.

We collected and dried leaves (82 total) from adult *E. angustifolia* plants that were collected as seed in rural western Minnesota, USA, from three remnant populations: Staffanson Prairie Preserve (SPP; 45°48′58.09″N, 95°44′53.78″W),

[1]Manuscript received 3 June 2013; revision accepted 9 July 2013.

We would like to thank K. Feldheim of the Pritzker Laboratory, Field Museum of Natural History, for guidance and assistance in developing the microsatellite markers. This work was completed in partial fulfillment of the requirements for the doctoral degree (to J.L.I.) from the Graduate College, University of Illinois at Chicago.

[5]Author for correspondence: ison.jennifer@gmail.com. Current address: Department of Ecology and Evolutionary Biology, University of Toronto, 25 Willcocks Street, Toronto, Ontario M5S 3B2, Canada.

doi:10.3732/apps.1300049

TABLE 1. Characterization of 11 microsatellite loci in *Echinacea angustifolia*.

Locus	Primer sequences (5′–3′)	Repeat motif	Size range (bp)[a]	T_a (°C)	WellRED dye[b]	GenBank accession no.
Ech03	F: AGAAGAATTGCGTGCTACGG R: CAATAATTAGATCAAGCTCATCATCA	$(GAT)_{12}$	168–280	61	D3	KF150005
Ech07	F: GTTGCACCAGAAAATGATGG R: CGTCGTCTAGAAAGGTATCACCA	$(GAT)_{17}$	177–261	64	D2	KF150006
Ech05	F: TCCATTTGACCGCTTTATGTC R: TGCAAGTTGAAATGGGTCTG	$(GAT)_{10}$	170–220	54	D3	KF150007
Ech11	F: CCTGTCGCTGTAACCGAACT R: ACGCATCGTAATGTCCGTTT	$(GAA)_{11}$	205–250	57	D3	KF150008
Ech13	F: TTGCACCAGAGAATGATGGT R: CTTCACCTGCCTGCTTTTCT	$(GAT)_{18}$	155–180	55	D2	KF150009
Ech13Z*	F: TTGCACCAGAGAATGATGGT R: CTTCACCTGCCTGCTTTTCT	$(GAT)_{18}$	195–250	55	D2	KF150009
Ech15	F: CTTCCTTCACCTGCCTTTTT R: AAGGGCAAGAAGGCTATTCC	$(CAT)_{11}$	215–235	55	D4	KF150010
Ech28	F: GGTCCTTTCCCTAGCACCAT R: TAGCCACAACAACAGCCTCA	$(TTG)_6$	186–229	59.8	D4	KF150011
Ech36	F: GCTCATTTCCTCTCTTGGATCA R: GCTAGAGCCACAACCTCCAC	$(TCT)_8$	210–235	63	D4	KF150012
Ech37	F: TGGGGGTAAAAAGAACTTCAA R: CTGATTGGAGGATTGGCTTC	$(AAG)_8$	370–400	55	D4	KF150013
Ech47	F: GCAACCTAGACATGGCAAGTAA R: CGGGTTTGAAGTAGATGAGGTT	$(TTC)_7$	245–255	54	D3	KF150014

Note: F = forward; R = reverse; T_a = annealing temperature.

[a] Base pairs are calculated to the nearest five base pairs.

[b] Fluorescent label used for each locus.

* Ech13Z is amplified by the same primers as Ech13, but the alleles do not overlap and the loci are unlinked.

Landfill (LF; 45°50′17.99″N, 95°45′15.57″W), and Steven's Approach (SAP; 45°49′36.10″N, 95°40′11.75″W). A representative voucher has been deposited at the Chicago Botanic Garden's Nancy Poole Rich Herbarium. The three remnant populations were between 2.5 km and 6.7 km apart and varied in mean number of flowering plants per year from 25 to 2961 (Table 2). We extracted DNA from the leaf samples using either a modified cetyltrimethylammonium bromide (CTAB) method (Chaudhry et al., 1999) or a QIAGEN DNeasy Plant Kit (QIAGEN, Valencia, California, USA).

For eight loci (all but Ech03, Ech07, and Ech28), the PCR was conducted in 10-μL PCR volumes containing 10 ng of template DNA, 0.25 μm of tagged forward primer (labeled with WellRED fluorescent dye; D2, D3, or D4 [Sigma-Aldich, St. Louis, Missouri, USA]), 0.25 μm of reverse primer, 5 μL PCR MasterMix 2× (50 U/mL *Taq* DNA polymerase in a proprietary reaction buffer [pH 8.5], 400 μM of each dNTP, and 3 mM $MgCl_2$ [Promega Corporation, Madison, Wisconsin, USA]), and 5 μg/μL of bovine serum albumin (BSA); an additional 0.25 mM $MgCl_2$ was added to primer Ech37. The PCR profile consisted of an initial 3-min denaturing step at 94°C followed by 35 cycles of a 30-s denaturing step at 94°C, a 30-s annealing step, and a 1-min elongation step at 72°C, followed by a final 10-min elongation step at 72°C (Table 1). For the remaining three loci (Ech03, Ech07, and Ech28), an M13 sequence tail (5′-CACGACGTTGTAAAACGAC) was added to the forward primer. The tail was also attached to a fluorescent label (for methods see Boutin-Ganache et al., 2001). Two sequential PCRs were conducted with a final 15-μL PCR volume. The first PCR had the same reaction recipe and profile as the other eight primers except that it only consisted of 13 cycles. For the second PCR, an additional 2.5 μL of PCR MasterMix 2×, 2.5 μg/μL of BSA, 0.125 mM $MgCl_2$, and 2 μL of water were added to the product from the first run. In addition, 0.25 μm of tagged M13 primer (fluorescently labeled) was added to the reaction. The second PCR had the same profile as the first PCR but consisted of 27 cycles and used an annealing temperature of 55°C for all three primers. Amplified products were analyzed on a Beckman Coulter CEQ 8000 Genetic Analysis System version 9.0 (Beckman Coulter, Indianapolis, Indiana, USA).

The 11 microsatellite loci have between three and 14 alleles (Table 2). For each of the three populations, we calculated the number of alleles per locus, observed and expected heterozygosities, and deviation from HWE using GenAlEx version 6.5 (Peakall and Smouse, 2012; Table 2). The observed

TABLE 2. Summary information for 11 microsatellite loci in three *Echinacea angustifolia* populations.[a]

	SPP (N = 41)			LF (N = 20)			SAP (N = 21)			Total
Locus	A	H_o	H_e	A	H_o	H_e	A	H_o	H_e	A
Ech03	4	0.537	0.635	5	0.600	0.576	3	0.714	0.632	6
Ech05	6	0.683	0.678	4	0.700	0.700	6	0.727	0.665	8
Ech07	13	0.756	0.800	8	0.850	0.769	6	0.682	0.786	14
Ech11	12	0.829	0.819	9	0.600	0.775*	7	0.864	0.748	13
Ech13	9	0.854	0.835	10	0.750	0.845	7	0.955	0.811	10
Ech13Z	9	0.675	0.759	5	0.550	0.685	6	0.524	0.619	12
Ech15	5	0.659	0.663	4	0.400	0.404	4	0.682	0.649	5
Ech28	5	0.366	0.424*	4	0.550	0.488	3	0.714	0.625	6
Ech36	7	0.780	0.741	5	0.800	0.723	5	0.773	0.708	8
Ech37	5	0.659	0.732	5	0.750	0.734	6	0.864	0.778	6
Ech47	3	0.439	0.449	2	0.500	0.495	2	0.409	0.375	3

Note: A = number of alleles; H_e = mean expected heterozygosity; H_o = mean observed heterozygosity; LF = Landfill population; N = sample size of each population; SAP = Steven's Approach population; SPP = Staffanson Prairie Preserve population.

[a] The mean number of flowering plants for each population was determined by S. Wagenius from 1995–1997: SSP = 2961, LF = 101, SAP = 25.

* Indicates significant deviation from Hardy–Weinberg equilibrium ($P < 0.05$, after Bonferroni correction).

heterozygosity ranged from 0.366 to 0.955 (mean = 0.673, SE ± 0.26), while the expected heterozygosity ranged from 0.375 to 0.845 (mean = 0.670, SE ± 0.023). The samples from the largest population (SPP N = 41) were used to evaluate the loci's informative power for paternity studies. The polymorphic information content (PIC), nonexclusion probability for identity, first and second parents, and parent pairs were calculated using CERVUS 3.0 (Kalinowski et al., 2007; Appendix 1). The combined paternity exclusion probability across all loci was 0.9994.

CONCLUSIONS

In this study, we developed and evaluated 11 microsatellite markers for a native prairie plant, *E. angustifolia*. We demonstrated that the markers are polymorphic across three populations of varying sizes and in HWE. The markers also have a high paternity exclusion probability and should be informative in paternity assignment–based pollination studies. To the best of our knowledge, this is the first set of primers for a herbaceous tallgrass prairie perennial. We believe that these primers will provide researchers and land managers with vital information on the genetic consequences of prairie fragmentation, including assessment of genetic variability, genetic differentiation, and pollination patterns within and among remnant populations.

LITERATURE CITED

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

Chaudhry, B., A. Yasmin, T. Husnain, and S. Riazuddin. 1999. Mini-scale genomic DNA extraction from cotton. *Plant Molecular Biology Reporter* 17: 280.

Glenn, T. C., and N. A. Schable. 2005. Isolating microsatellite DNA loci. *In* E. A. Zimmer and E. H. Roalson [eds.], Methods in enzymology, vol. 395, Molecular evolution: Producing the biochemical data, Part B, 202–222. Academic Press, San Diego, California, USA.

Goudet, J. 1995. FSTAT (version 1.2): A computer program to calculate *F*-statistics. *Journal of Heredity* 86: 485–486.

Kalinowski, S. T., M. L. Taper, and T. C. Marshall. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* 16: 1099–1106.

Peakall, R., and P. E. Smouse. 2012. GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research: An update. *Bioinformatics (Oxford, England)* 28: 2537–2539.

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.], Methods in molecular biology, vol. 132: Bioinformatics methods and protocols, 365–386. Humana Press, Totowa, New Jersey, USA.

Samson, F., and F. Knopf. 1994. Prairie conservation in North America. *Bioscience* 44: 418–442.

van Oosterhout, C., W. F. Hutchinson, 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.

Wagenius, S. 2006. Scale dependence of reproductive failure in fragmented *Echinacea* populations. *Ecology* 87: 931–941.

Wagenius, S., and S. P. Lyon. 2010. Reproduction of *Echinacea angustifolia* in fragmented prairie is pollen-limited but not pollinator-limited. *Ecology* 91: 733–742.

Young, A., T. Boyle, and T. Brown. 1996. The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology & Evolution* 11: 413–418.

Appendix 1. Informative power for paternity studies of 11 *Echinacea angustifolia* microsatellite loci. Samples were taken from the Staffanson Prairie Preserve (SPP) population (N = 41).

Locus	PIC	NE-1P	NE-2P	NE-PP	NE-I
Ech03	0.567	0.794	0.644	0.488	0.201
Ech05	0.62	0.739	0.576	0.398	0.161
Ech07	0.781	0.541	0.364	0.17	0.059
Ech11	0.798	0.519	0.346	0.164	0.054
Ech13	0.816	0.494	0.324	0.149	0.047
Ech13Z	0.727	0.625	0.445	0.253	0.09
Ech15	0.608	0.756	0.591	0.414	0.169
Ech28	0.38	0.908	0.781	0.644	0.376
Ech36	0.697	0.668	0.492	0.311	0.111
Ech37	0.686	0.686	0.51	0.332	0.118
Ech47	0.359	0.899	0.815	0.715	0.394
Total	0.640 (mean)	0.0144	0.0006	4.05E-06	1.39E-10

Note: N = sample size of population; NE-1P = first parent nonexclusion probability; NE-2P = second parent (first parent known) nonexclusion probability; NE-I = identity nonexclusion probability; NE-PP = parent pair nonexclusion probability; PIC = polymorphic information content.