Development and Characterization of Microsatellite Primers in *Geranium carolinianum* (Geraniaceae) with 454 Sequencing

Author(s): Rebecca Y. Shirk, Travis C. Glenn, Shu-Mei Chang, and J. L. Hamrick
Source: Applications in Plant Sciences, 1(8) 2013.
Published By: Botanical Society of America
DOI: http://dx.doi.org/10.3732/apps.1300006
URL: http://www.bioone.org/doi/full/10.3732/apps.1300006

BioOne (www.bioone.org) is a nonprofit, online aggregation of core research in the biological, ecological, and environmental sciences. BioOne provides a sustainable online platform for over 170 journals and books published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Web site, and all posted and associated content indicates your acceptance of BioOne’s Terms of Use, available at www.bioone.org/page/terms_of_use.

Usage of BioOne content is strictly limited to personal, educational, and non-commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.
DEVELOPMENT AND CHARACTERIZATION OF MICROSATELLITE PRIMERS IN *Geranium carolinianum* (Geraniaceae) WITH 454 SEQUENCING¹

**REBECCA Y. SHIRK**²³⁵, TRAVIS C. GLENN⁴, SHU-MEI CHANG², and J. L. HAMRICK²

²Department of Plant Biology, University of Georgia, Athens, Georgia 30602 USA; ³Department of Botany, University of Wisconsin–Madison, Madison, Wisconsin 53706 USA; and ⁴Department of Environmental Health Science, University of Georgia, Athens, Georgia 30602 USA

- Premise of the study: Microsatellite primers were developed for *Geranium carolinianum*, a North American winter annual herb, for use in population genetic analyses.
- Methods and Results: Genomic DNA enriched for repeat-containing fragments was sequenced on the Roche 454 Titanium platform, resulting in 470 primer pairs developed from 1115 microsatellite-containing sequences. A subset of 37 primer pairs was screened for polymorphism across three native and three invasive populations. We identified four monomorphic and eight polymorphic loci. Polymorphic loci contained between two and seven alleles per locus, and mean within-population expected heterozygosity ranged from 0.100 to 0.290. Within populations, observed heterozygosity for individual loci ranged from zero to 0.857, and expected heterozygosity ranged from 0.046 to 0.559.
- Conclusions: These microsatellite markers will be useful for future studies of genetic diversity, structure, and mating systems across the geographic range of *G. carolinianum*, and may be transferable to other closely related species.

Key words: 454 sequencing; biological invasions; Geraniaceae; *Geranium carolinianum*.

*Geranium carolinianum* L. (Geraniaceae) is a weedy winter annual herb native to North America and naturalized in East Asia, South America, and the Caribbean (Aedo, 2000). In China, it is naturalized along roadsides and in agricultural fields, and is considered an invasive species with minor environmental impacts (Liu et al., 2006). The origin and colonization history of these introduced populations is unknown. Although it is not closely related to the horticultural geranium (genus *Pelargonium* L’Hér. ex Aiton), genetic resources developed for *G. carolinianum* are potentially transferable to other Geraniaceae L. species, which includes annuals and perennials that are both native and introduced in North America (Aedo, 2000). Here, we present 12 microsatellite loci developed using targeted enrichment and next-generation sequencing.

METHODS AND RESULTS

DNA was extracted from leaf tissue using a QIAGEN Plant Mini Kit (QIAGEN, Valencia, California, USA) and enriched for microsatellites using three probe mixes [Mix 2 = (AG)₁₂, (TG)₁₂, (AACC)₆, (AAGT)₆, (AAT)₁₂, (ACT)₁₂, (ATC)₆; Mix 3 = (AAACC)₆, (AAAGT)₆, (AATG)₆, (ACAG)₆, (ACCT)₆, (ACT)₆, (ACTG)₆; Mix 4 = (AAAT)₆, (AACT)₆, (AACT)₆, (ACAT)₆, (ACAT)₆, (AGAT)₆] as in Glenn and Schable (2005). DNA from one individual was digested separately with *RsaI* (New England Biolabs, Ipswich, Massachusetts, USA) and *AluI* (New England Biolabs). Digests were pooled, a single adenosine was added with exo-Klenow fragment (New England Biolabs), and fragments were ligated to CAG-SimpleXT-13 adapters (upper oligo = GTTTCAGTGCCGCGTCACTACA CCACGCCGGCAGACAT; lower oligo = /5Phos/ TGTTCCGGTCTCGTG). Linker-ligated DNA was pooled, denatured, and hybridized to biotinylated microsatellite oligonucleotide mixtures, which were subsequently captured on magnetic streptavidin beads (Dynabeads M-280; Invitrogen, Carlsbad, California, USA), and remaining DNA was washed away. The enriched DNA was eluted and amplified in PCR using Phos_Pig_CAG (5Phos/GTTTCAGTGCCGCGTCACTACA). PCR reactions were run in a 25 μL total volume with 2 μL eluted DNA, 0.2 μL Phusion Taq polymerase (5 units/μL; New England Biolabs), 1x Phusion reaction buffer (New England Biolabs), 0.025 mg/mL bovine serum albumin (BSA; New England Biolabs), 2.0 mM MgCl₂, 0.15 mM each dNTP (Sigma-Aldrich, St. Louis, Missouri, USA), and 0.5 μM Phos_Pig_CAG. Cycling conditions were 2 min at 95°C; 25 cycles of 20 s at 95°C, 20 s at 60°C, and 1.5 min at 72°C; and a final 30-min extension at 72°C. Multiple libraries with additional different SimpleX sequences were then pooled and ligated into a 454 RL MID-tagged library, pooled with additional 454 RL MID-tagged libraries, and sequenced on half a run of the Roche 454 FLX with Titanium chemistry using the manufacturer’s protocols.

Following demultiplexing, 1115 sequences with microsatellites were identified from 5586 sequence reads using MSATCOMMANDER version 0.8.1 (Faircloth, 2008). All microsatellite-containing sequences are provided in Appendix S1. MSATCOMMANDER also designs primers for microsatellite-containing sequences using Primer3 (Rozen and Skaltsky, 2000). Sequences were discarded if Primer3 could not design acceptable primers. All primers (excluding those tested as described below, a total of 433 pairs) are reported in Appendix S2. To avoid selecting primers within highly repetitive regions of the genome, sequences were assembled into contigs by automatic assembly using default parameters in Sequencher version 4.7 (Gene Codes Corporation, Ann Arbor, Michigan, USA).

¹Manuscript received 15 January 2013; revision accepted 25 February 2013.

The authors thank B. Faircloth for demultiplexing the 454 reads and revisions to MSATCOMMANDER, and an anonymous reviewer who provided helpful comments. This material is based upon work supported by the National Science Foundation’s Partnerships for International Research and Education (PIRE) program under grant no. OISE 0730218.

¹ doi:10.3732/apps.1300006

Applications in Plant Sciences 2013 1(8): 1300006; http://www.bioone.org/loi/apps © 2013 Shirk et al. Published by the Botanical Society of America. This work is licensed under a Creative Commons Attribution License (CC-BY-NC-SA).
Arbor, Michigan, USA). Microsatellite-containing sequences that were at high frequency (i.e., members of contigs with >10 sequences) were not considered for screening.

Of the remaining sequences, we selected 37 primer pairs for testing. The majority of primers tested were based on unique sequences in the data set and amplified tri- or tetranucleotide repeats. Forward primers were tagged on the 5′ end with the universal “CAG” sequence (CAGTCGGGCTGTACCATCA) to allow fluorescent labeling of PCR product using a 3-primer protocol as detailed below. Additionally, a 5′ GTTT “pigtail” was added to some reverse primers to ensure consistency in amplicon size (Brownstein et al., 1996). These 37 primer pairs were screened using DNA template from 12 samples across Geranium carolinianum’s native range. All primer pairs were initially tested using a “touchdown 58°C” PCR protocol. Cycling conditions were 3 min at 95°C; nine touchdown cycles of 30 s at 95°C, 30 s at 65–58°C (starting at 65°C and decreasing one degree each cycle), and 45 s at 72°C; 30 cycles of 30 s at 95°C, 30 s at 50°C, and 45 s at 72°C; and a final extension of 20 min at 72°C. Primer pairs that did not amplify consistently at touchdown 58°C were tested with a “touchdown 50°C” protocol, which was identical to touchdown 58°C except the annealing temperatures for the touchdown cycles decreased from 58°C to 50°C, followed by 30 cycles at an annealing temperature of 50°C. All PCR reactions were carried out in a 12.5 μL total volume with 10–200 ng DNA template, 0.1875 units JumpStart Taq DNA polymerase (Sigma-Aldrich), 1× JumpStart reaction buffer without magnesium (Sigma-Aldrich), 1× (0.1 mg/mL) BSA, 2 mM MgCl₂, 0.15 mM each dNTP, 0.025 μM forward (CAG-tagged) primer, 0.25 μM reverse (untagged) primer, and 0.25 μM universal CAG primer labeled with a 5′ FAM, HEX, or NED fluorescent dye. Primer pairs with consistent amplification at either touchdown 58°C or touchdown 50°C were screened for polymorphism by fragment analysis. PCR products from the initial 12 DNA samples were sized on an ABI 3730xl capillary sequencer (Applied Biosystems, Carlsbad, California, USA) using a ROX-labeled internal size standard (GGF500R; Georgia Genomics Facility, Athens, Georgia, USA) using GeneMapper version 3.7 (Applied Biosystems). Chromatograms were analyzed using GenAlEx version 6.2 (Peakall and Smouse, 2006). Results are given in Table 2. Across all samples, number of alleles per polymorphic locus ranged from two to seven, and five loci were monomorphic in at least one population. Mean population Hₑ (averaged across loci) ranged from 0.100 to 0.290. Within populations, Hₑ for individual loci ranged from zero to 0.857, and Hₑ ranged from 0.046 to 0.559. We did not test for Hardy–Weinberg equilibrium because this is a colonizing, mixed-mating species.

CONCLUSIONS

We report 12 microsatellite loci in Geranium carolinianum, eight of which are shown to be variable in native and invasive populations. Forward primer sequences are labeled with a 5′ universal sequence to allow fluorescent tagging of amplicons during PCR, and reverse primer sequences (with the exception of GC1-R, GC6-R, GC7-R, and GC10-R) are labeled with a 5′ GTTT pigtail sequence for consistent genotyping. Although not tested, addition of a 5′ pigtail to these primers should facilitate accurate genotyping. These microsatellite loci will be useful for studies of genetic variation, gene flow, and mating system in Geranium carolinianum and may be transferable to related species.

![Image](http://www.bioone.org/loi/apps)
Table 2. Results of primer screening of eight polymorphic microsatellite loci in three native (United States) and three invasive (China) populations of *Geranium carolinianum*.  

<table>
<thead>
<tr>
<th>Locus</th>
<th>A</th>
<th>H&lt;sub&gt;o&lt;/sub&gt;</th>
<th>H&lt;sub&gt;e&lt;/sub&gt;</th>
<th>A</th>
<th>H&lt;sub&gt;o&lt;/sub&gt;</th>
<th>H&lt;sub&gt;e&lt;/sub&gt;</th>
<th>A</th>
<th>H&lt;sub&gt;o&lt;/sub&gt;</th>
<th>H&lt;sub&gt;e&lt;/sub&gt;</th>
<th>A</th>
<th>H&lt;sub&gt;o&lt;/sub&gt;</th>
<th>H&lt;sub&gt;e&lt;/sub&gt;</th>
<th>A</th>
<th>H&lt;sub&gt;o&lt;/sub&gt;</th>
<th>H&lt;sub&gt;e&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>GC7</td>
<td>4</td>
<td>0.056</td>
<td>0.650</td>
<td>6</td>
<td>0.056</td>
<td>0.650</td>
<td>4</td>
<td>0.056</td>
<td>0.650</td>
<td>6</td>
<td>0.056</td>
<td>0.650</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GC10</td>
<td>2</td>
<td>0.056</td>
<td>0.153</td>
<td>1</td>
<td>0.056</td>
<td>0.153</td>
<td>4</td>
<td>0.056</td>
<td>0.153</td>
<td>4</td>
<td>0.056</td>
<td>0.153</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GC19</td>
<td>3</td>
<td>0.056</td>
<td>0.333</td>
<td>3</td>
<td>0.056</td>
<td>0.333</td>
<td>2</td>
<td>0.056</td>
<td>0.333</td>
<td>1</td>
<td>0.056</td>
<td>0.333</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GC35</td>
<td>2</td>
<td>0.000</td>
<td>0.091</td>
<td>4</td>
<td>0.000</td>
<td>0.091</td>
<td>4</td>
<td>0.000</td>
<td>0.091</td>
<td>4</td>
<td>0.000</td>
<td>0.091</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GC38</td>
<td>2</td>
<td>0.000</td>
<td>0.278</td>
<td>2</td>
<td>0.000</td>
<td>0.278</td>
<td>1</td>
<td>0.000</td>
<td>0.278</td>
<td>1</td>
<td>0.000</td>
<td>0.278</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note:  
- A = number of alleles;  
- H<sub>o</sub> = observed heterozygosity;  
- H<sub>e</sub> = expected heterozygosity;  
- n = sample size.

LITERATURE CITED


