Characterization of 13 Microsatellite Markers for Calochortus gunnisonii (Liliaceae) from Illumina MiSeq Sequencing

Authors: Ryan S. Fuller, Seth Frietze, and Mitchell E. McGlaughlin
Source: Applications in Plant Sciences,  3(8)
Published By: Botanical Society of America
URL: https://doi.org/10.3732/apps.1500051
CHARACTERIZATION OF 13 MICROSATELLITE MARKERS FOR *CALOCHORTUS GUNNISONII* (LILIACEAE) FROM ILLUMINA MiSeq SEQUENCING

RYAN S. FULLER\(^2,3\), SETH FRIETZE\(^2\), AND MITCHELL E. MCGLAUGHLIN\(^2\)

\(^2\)School of Biological Sciences, University of Northern Colorado, 501 20th Street, Greeley, Colorado 80639 USA

- **Premise of the study:** Microsatellite primers were designed for *Calochortus gunnisonii* (Liliaceae), a montane lily species of the central and southern Rocky Mountains, using next-generation DNA sequencing. The markers will be used to investigate population structure, genetic diversity, and demographic history.
- **Methods and Results:** Thirteen polymorphic microsatellite loci were isolated from *C. gunnisonii* using Illumina MiSeq next-generation DNA sequencing and bioinformatic screening. The mean number of alleles per locus ranged from 4.15 to 5.92 (avg. = 4.97). Observed and expected heterozygosity ranged from 0.077 to 0.871 and 0.213 to 0.782, respectively. The primers were also tested for cross-species amplification value with *C. flexuosus*, *C. nuttallii*, *C. kennedyi* var. *kennedyi*, and *C. subalpinus*.
- **Conclusions:** These primers will be useful for genetic and evolutionary studies across *C. gunnisonii*’s range within the southern and central Rocky Mountains. Furthermore, these markers have proven valuable for cross-species amplifications within *Calochortus*.

**Key words:** *Calochortus gunnisonii*; Illumina; Liliaceae; microsatellites; MiSeq.

---

*Calochortus* Pursh (Liliaceae) is a large genus of bulbous geophytes (ca. 70 spp.) originating in California ~7 million years ago (Patterson and Givnish, 2003). Its range includes a center of diversity in California that spreads north to British Columbia, east to the Dakotas, and south to Guatemala (Ownbey, 1940; Patterson and Givnish, 2003; Henss et al., 2013). The genus also occurs in a wide range of habitats including grasslands, deserts, vernal pools, woodland meadows, springs, montane woodlands, and forest understories, with most taxa occupying narrow geographic ranges (Ownbey, 1940; Patterson and Givnish, 2003; Fiedler and Zebell, 2012; Henss et al., 2013).

Gunnison’s mariposa lily, *C. gunnisonii* S. Watson, is a North American endemic populating portions of the central and southern Rocky Mountains. *Calochortus gunnisonii* has a broad distribution encompassing northeastern Arizona, northern New Mexico, much of Colorado, eastern Utah, large portions of Wyoming, southern Montana, and western South Dakota at elevations of 1200–3300 m (Fiedler and Zebell, 2012). This species achieves some of the highest elevations for the genus in the southern portion of the Rocky Mountains. Disjunctions in the northern portions of *C. gunnisonii*’s range exist across mountain ranges of the Big Horn Mountains, Black Hills, Absaroka Range, Sierra Madre, Medicine Bow, and Laramie Range. A population genetic study of *C. gunnisonii* across multiple, disjunct populations in the central and southern Rocky Mountains is currently being conducted.

However, previous genetic studies of members within the genus are limited to amplified fragment length polymorphisms (Henss et al., 2013) and chloroplast sequence comparisons (Patterson and Givnish, 2003). Here, we report the characterization of 13 microsatellite loci that will be used to investigate the role of glacial oscillatory demographic changes in shaping genetic structure of *C. gunnisonii* across multiple montane disjunctions of the central and southern Rocky Mountains.

**METHODOLOGY AND RESULTS**

Next-generation sequencing was used to acquire a large quantity of genomic sequence data in search of microsatellite repeats. Genomic DNA (gDNA) was extracted from leaf tissue using a modified cetyltrimethylammonium bromide (CTAB) protocol (Friar, 2005). Two individual DNA samples collected from two separate geographic populations (Hell Canyon Road, South Dakota: 43.724753°N, 103.854112°W; La Prele Reservoir, Wyoming: 42.706796°N, 105.578531°W [Appendix 1]) were pooled and sent to the Center for Genome Research and Biocomputing at Oregon State University. A total of ~400 ng of gDNA was used for library preparation and Illumina MiSeq sequencing (Illumina, San Diego, California, USA). A single, 300-bp paired-end MiSeq sequencing run resulted in 18,332,564 reads with an average length of 301 bases. Raw sequence reads were filtered, reformatted, and trimmed using the default commands of the Trimmomatic v.0.32 program (Bolger et al., 2014). Trimmomatic yielded 12,080,556 high-quality contigs (6,040,278 forward and reverse). Contigs were de novo assembled using Trinity (v. 07-04-2014) (Grabherr et al., 2011) producing 486,538 contigs, with an average size of 516 bp (N50 = 583). MSATCOMMANDER v.1.0.8 (Faircloth, 2008) found a total of 4118 perfect microsatellite repeats from the assembled contigs: 85 hexanucleotide with at least four repeat units, 585 pentanucleotide with at least four repeat units, 481 tetranucleotide with at least five repeat units, and 2967 trinucleotide with at least five repeat units. Dinucleotide repeats were excluded to avoid stutter and subsequent scoring problems in later analyses. The inset version of Primer3 (Rozen and Skaltsky, 2000) in MSATCOMMANDER designed 1153 primer pairs using default parameters (except that GC clamp = yes and repeat motif was ≥4 for penta- and hexanucleotide repeats). The PCR product size was set to 120–400 bp. One primer of each pair was designed with a common tag at the 5’ end following the procedure of Boutin-Ganache et al. (2001). Two common
Table 1. Characteristics of 13 microsatellite loci designed for *Calochortus gunnisonii*.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Primer sequences (5′–3′)</th>
<th>5′ Tag</th>
<th>Repeat motif</th>
<th>Fluorescent dye</th>
<th>Allele size range (bp)</th>
<th>$T_a$ (°C)</th>
<th>MgCl$_2$/MgSO$_4$</th>
<th>GenBank accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAGU_14</td>
<td>F: TTGTCAAGTGCGCAAGTGTC</td>
<td>M13</td>
<td>(ACACC)$_5$</td>
<td>FAM</td>
<td>246–266</td>
<td>62.4</td>
<td>MgCl$_2$</td>
<td>KR139841</td>
</tr>
<tr>
<td></td>
<td>R: ATCAAACCTGATCCTATACC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAGU_15</td>
<td>F: ATCCTCACTGCTCCATACC</td>
<td>M13</td>
<td>(AGAGG)$_4$</td>
<td>VIC</td>
<td>385–415</td>
<td>62.4</td>
<td>MgCl$_2$</td>
<td>KR139842</td>
</tr>
<tr>
<td></td>
<td>R: GTGCGAGATCCGCACTTCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAGU_22</td>
<td>F: CACATTGTGGTGTAGCGAG</td>
<td>CAGT</td>
<td>(ATCC)$_5$</td>
<td>PET</td>
<td>237–257</td>
<td>62.4</td>
<td>MgCl$_2$</td>
<td>KR139843</td>
</tr>
<tr>
<td></td>
<td>R: TTGTTACCTGCGACAGGCCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAGU_31</td>
<td>F: CACCCCAAGAGCGCTAAAGG</td>
<td>CAGT</td>
<td>(ACATC)$_4$</td>
<td>PET</td>
<td>310–340</td>
<td>62.4</td>
<td>MgCl$_2$</td>
<td>KR139844</td>
</tr>
<tr>
<td></td>
<td>R: TCCACCTGACTCCACAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAGU_35</td>
<td>F: TAATACCGGTGACTCCGCC</td>
<td>M13</td>
<td>(ATGC)$_6$</td>
<td>VIC</td>
<td>263–283</td>
<td>62.4</td>
<td>MgCl$_2$</td>
<td>KR139845</td>
</tr>
<tr>
<td></td>
<td>R: TTACCCAGCTGACGAGACC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAGU_36</td>
<td>F: CACCCAGCTCAGCACCAGAC</td>
<td>M13</td>
<td>(ATCG)$_5$</td>
<td>FAM</td>
<td>378–398</td>
<td>62.4</td>
<td>MgCl$_2$</td>
<td>KR139846</td>
</tr>
<tr>
<td></td>
<td>R: TAAGTTAGTAGAAGAGCACGC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAGU_39</td>
<td>F: TTACCCACAGCTCCAGAG</td>
<td>CAGT</td>
<td>(ACTG)$_5$</td>
<td>FAM</td>
<td>335–374</td>
<td>62.4</td>
<td>MgCl$_2$</td>
<td>KR139847</td>
</tr>
<tr>
<td></td>
<td>R: GTCTCTGCTGCTGCTCCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAGU_42</td>
<td>F: TCGTGGTTAAGTCTACATCG</td>
<td>M13</td>
<td>(AAGAC)$_5$</td>
<td>PET</td>
<td>380–415</td>
<td>66.9</td>
<td>MgCl$_2$</td>
<td>KR139848</td>
</tr>
<tr>
<td></td>
<td>R: GTCTCTGGAATACAGATCCAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAGU_45</td>
<td>F: TCAGTACGAAAACAGGAGGCC</td>
<td>M13</td>
<td>(ACGCG)$_5$</td>
<td>FAM</td>
<td>190–214</td>
<td>62.4</td>
<td>MgCl$_2$</td>
<td>KR139849</td>
</tr>
<tr>
<td></td>
<td>R: ACCCTACCTCTTTGCTTCGGT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAGU_46</td>
<td>F: GCACCTGACATCGATGGAC</td>
<td>T7</td>
<td>(ACTAT)$_4$</td>
<td>PET</td>
<td>164–184</td>
<td>59.8</td>
<td>MgSO$_4$</td>
<td>KR139850</td>
</tr>
<tr>
<td></td>
<td>R: TACCTGCAGTACAGTACG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAGU_47</td>
<td>F: TTCAAGGGATGGGATCGCC</td>
<td>CAGT</td>
<td>(AGCTCC)$_4$</td>
<td>VIC</td>
<td>345–381</td>
<td>61.1</td>
<td>MgCl$_2$</td>
<td>KR139851</td>
</tr>
<tr>
<td></td>
<td>R: GAACTCTCTCTGGCAGAC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAGU_48</td>
<td>F: TGCACCATAGAGACATG</td>
<td>CAGT</td>
<td>(ACAGAT)$_4$</td>
<td>FAM</td>
<td>406–448</td>
<td>61.1</td>
<td>MgSO$_4$</td>
<td>KR139852</td>
</tr>
<tr>
<td></td>
<td>R: GTCAAGGATCCAGGCTTCTC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CAGU_50</td>
<td>F: TAGGGAGGAGCTTCCAGAGAC</td>
<td>CAGT</td>
<td>(ACAGT)$_5$</td>
<td>PET</td>
<td>364–409</td>
<td>56.9</td>
<td>MgCl$_2$</td>
<td>KR139853</td>
</tr>
<tr>
<td></td>
<td>R: TAGGCTGCGGCACGTCATT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note:* $T_a$ = PCR annealing temperature.

*a The primer tagged with either M13R, CAGT, or T7 is indicated by an asterisk.

*b The 5′ tag used for incorporation of the fluorescent tag: M13R (AGGAAACGCTATGACC), T7 (GCTAGTTATTGCTCAGCG), or CAGT (ACAGTCCGGCGCTCATCA).

*c MgCl$_2$ or MgSO$_4$ used in PCR reactions.*
Table 2. Allelic diversity at 13 variable microsatellite loci for three populations of Calochortus gunnisonii.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Dixon</th>
<th>Sand Lake</th>
<th>Shavano Camp</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>A</td>
<td>H_o</td>
</tr>
<tr>
<td>CAGU_14</td>
<td>31</td>
<td>4</td>
<td>0.469</td>
</tr>
<tr>
<td>CAGU_15</td>
<td>31</td>
<td>8</td>
<td>0.563</td>
</tr>
<tr>
<td>CAGU_22</td>
<td>31</td>
<td>5</td>
<td>0.125</td>
</tr>
<tr>
<td>CAGU_31</td>
<td>31</td>
<td>6</td>
<td>0.531</td>
</tr>
<tr>
<td>CAGU_35</td>
<td>30</td>
<td>6</td>
<td>0.419</td>
</tr>
<tr>
<td>CAGU_36</td>
<td>30</td>
<td>5</td>
<td>0.387</td>
</tr>
<tr>
<td>CAGU_39</td>
<td>30</td>
<td>6</td>
<td>0.452</td>
</tr>
<tr>
<td>CAGU_42</td>
<td>31</td>
<td>7</td>
<td>0.250</td>
</tr>
<tr>
<td>CAGU_45</td>
<td>31</td>
<td>7</td>
<td>0.344</td>
</tr>
<tr>
<td>CAGU_46</td>
<td>28</td>
<td>5</td>
<td>0.517</td>
</tr>
<tr>
<td>CAGU_47</td>
<td>30</td>
<td>5</td>
<td>0.387</td>
</tr>
<tr>
<td>CAGU_48</td>
<td>29</td>
<td>5</td>
<td>0.763</td>
</tr>
<tr>
<td>CAGU_50</td>
<td>29</td>
<td>6</td>
<td>0.633</td>
</tr>
<tr>
<td>Mean</td>
<td>30.23</td>
<td>5.92</td>
<td>0.458</td>
</tr>
</tbody>
</table>

Note: A = number of alleles; H_o = observed heterozygosity; H_e = expected heterozygosity; N = sample size.

*Statistical significance associated with departure from Hardy–Weinberg equilibrium (HWE) is indicated with an asterisk (\( *P \leq 0.05 \)).

LITERATURE CITED


We identified 13 C. gunnisonii microsatellite loci that are variable and informative. These markers will be used to investigate the population genetic structure and levels of genetic variability of C. gunnisonii in the central and southern Rocky Mountains. Intra- and intermontane patterns of gene flow and divergence will be inferred within C. gunnisonii. Cross-species amplification was high in a closely related taxon, C. nuttallii, and decreased in more divergent sampled taxa.

CONCLUSIONS

We identified 13 C. gunnisonii microsatellite loci that are variable and informative. These markers will be used to investigate the population genetic structure and levels of genetic variability of C. gunnisonii in the central and southern Rocky Mountains. Intra- and intermontane patterns of gene flow and divergence will be inferred within C. gunnisonii. Cross-species amplification was high in a closely related taxon, C. nuttallii, and decreased in more divergent sampled taxa.

http://www.bioone.org/loi/apps

See the entire text for more details.
### APPENDIX 2. Cross-species amplification information for 13 microsatellite loci developed for *Calochortus gunnisonii* with four related *Calochortus* species.

<table>
<thead>
<tr>
<th>Locus</th>
<th><em>C. flexuosus</em></th>
<th><em>C. nuttallii</em></th>
<th><em>C. kennedyi var. kennedyi</em></th>
<th><em>C. subalpinus</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>CAGU_14</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>CAGU_15</td>
<td>—</td>
<td>+</td>
<td>+</td>
<td>—</td>
</tr>
<tr>
<td>CAGU_22</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>—</td>
</tr>
<tr>
<td>CAGU_31</td>
<td>—</td>
<td>+</td>
<td>+</td>
<td>—</td>
</tr>
<tr>
<td>CAGU_35</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>CAGU_36</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>CAGU_39</td>
<td>—</td>
<td>+</td>
<td>+</td>
<td>—</td>
</tr>
<tr>
<td>CAGU_42</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>—</td>
</tr>
<tr>
<td>CAGU_45</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>CAGU_46</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>CAGU_47</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>CAGU_48</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>—</td>
</tr>
<tr>
<td>CAGU_50</td>
<td>+</td>
<td>+</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

*Note:* + = amplification product fell within the product size range for the locus; — = amplification product was absent or did not fall within the product size range expected for the locus.