Isolation and characterization of polymorphic microsatellite markers for *Centaurea aspera* L. and *Centaurea seridis* L. (Asteraceae)

H. Merle\(^1\), I. Segura\(^2\), A. Garmendia\(^2\) and M. Ferriol\(^2\)

\(^1\) Departamento de Ecosistemas Agroforestales. Universidad Politécnica de Valencia. Camino de Vera s/n, E-46022 Valencia, Spain.

\(^2\) Instituto Agroforestal Mediterráneo. Universidad Politécnica de Valencia. Camino de Vera s/n, E-46022 Valencia, Spain.

Corresponding author: Hugo Merle
Fax: +34 96 387 92 69
Email address: humerfa@upvnet.upv.es

Keywords: *Centaurea aspera, Centaurea seridis*, genetic diversity, microsatellite marker, *Seridia* section

Running title: SSR markers for *Centaurea aspera* and *C. seridis*
Abstract

Eight polymorphic microsatellite loci were developed and characterized for the diploid Centaurea aspera L. and the tetraploid Centaurea seridis L., two species of the Seridia section of Asteraceae. We used 132 individuals collected from 5 locations. These markers provided high polymorphism ranging from 3 to 10 alleles per locus. These microsatellite loci will be useful tools to study polyploid complexes that include triploid individuals.

Main text

Centaurea L. is one of the widest and most complex genus inside Asteraceae, with 400 to 700 species depending on the author (Dittrich 1977; Bremer 1994; Wagenitz & Hellwig 1996; Hellwig 2004). New taxa and combinations are continuously published in many studies (Colas et al. 1997; Garcia-Jacas et al. 1997; Kalpoutzakis & Constantinidis 2004; Garcia-Jacas et al. 2006; Raimondo & Spadaro 2008; Trigas et al. 2008; Rahiminejad et al. 2010).

The diploid Centaurea aspera L. and the tetraploid C. seridis L. are closely related species from the section Seridia (Juss.) Czerep. Centaurea aspera grows in dry open habitats in S.W. Europe (Spain, France, Portugal and Italy), while C. seridis has a narrower distribution area in S.E. Spain (Tutin et al. 1976). When these two species grow side by side in sandy coastal dunes, a hybrid contact zone is found, and triploid forms with a confusing taxonomy have been identified (Garmendia et al. 2009; Mateo & Crespo 2009). Microsatellite markers have been described and published in Centaurea only for C. corymbosa (Freville et al. 2000), C. stoehbe, and C. diffusa (Marrs et al. 2006), and have been used to study the genetic structure of C. horrida (Mameli et al. 2008) and C. corymbosa (Hardy et al. 2004). We have tested most of these markers in C. aspera and C. seridis and only few of them were useful. In addition, amplification showed a very low polymorphism level. The need for using microsatellite
markers in \textit{C. aspera} and \textit{C. seridis} for genetic studies has leaded us to develop new ones. Here, we report the isolation and characterization of eight new polymorphic microsatellite loci that can be useful tools to study the polyploid complex.

Locus-specific primer pairs were designed to amplify SSRs and their flanking regions using the software EST2uni (Forment \textit{et al.} 2008). Public ESTs from \textit{C. solstitialis} L. and \textit{C. maculosa} Lam. were uploaded from Genbank. High-quality EST sequences were then assembled to obtain the unigene set using EST2uni. One hundred thirty-two individuals of \textit{C. aspera}, \textit{C. seridis}, and their hybrids, collected in five locations at S.E. of Spain, were examined for polymorphism using the developed microsatellites primers. Locations and their sample size were: South coast (Alicante), 54 individuals (20 \textit{C. aspera}, 22 \textit{C. seridis}, 12 hybrids); Central coast (Valencia), 63 individuals (21 \textit{C. aspera}, 25 \textit{C. seridis}, 17 hybrids); Soria, 3 individuals of \textit{C. aspera}; Sax (inner land, Alicante), 8 individuals (3 \textit{C. aspera}, 3 \textit{C. seridis}, 2 hybrids); and Montsant (Tarragona), 4 individuals of \textit{C. aspera}. Total genomic DNA was extracted from 0.05 g young buds using a modified cetyltrimethyl ammonium bromide (CTAB) protocol (Doyle & Doyle 1990). Polymerase chain reaction (PCR) was carried out in 25 µl final reaction volume containing 2.5 µl PCR buffer 10x, 0.75 to 1.25 µl MgCl\textsubscript{2} 50 mM (see Table 1), 1.25 µl dNTPs 10 mM, 0.5 µl forward primer 0,01 mM, 0.5 µl reverse primer 0,01 mM, 1 U Taq, 20 ng of template DNA and H\textsubscript{2}O miliQ to final volume (17 to 17.5 µl).

The reaction mixture was subjected to PCR amplification in a Peltier Thermal Cycler (Techne TC-412). We used a variable annealing temperature depending on the loci (Table 1). The cycling profile was 94°C for 5 min; 36 cycles: 94°C for 1 min, annealing temperature (see Table 1) for 1 min, 72°C for 1 min; and a final extension at 72°C for 10 min. Nine µl of loading mixture was added to the PCR products. Subsequently, the PCR products were examined by electrophoresis using 12% non-denaturing polyacrylamide gels stained with
silver nitrate. The sizes of the amplified DNA fragments were estimated by comparing them with a 100-bp DNA ladder standard.

We selected eight polymorphic SSR loci out of 26 initially tested loci. For each microsatellite locus the number of alleles per locus ($N_a$), the observed heterozygosity ($H_O$), the expected heterozygosity or gene diversity ($H_E$) and tests for deviation from the Hardy-Weinberg equilibrium (HWE) and linkage equilibrium were calculated using PopGene32 (Yeh & Boyle 1997). Allelic variation was estimated using the polymorphism information content (PIC) value (Weir 1990).

A total of 47 alleles were observed across the 8 loci (Table 1). The number of alleles per locus ranged from 2 to 10. The observed heterozygosity of the 8 loci ranged from 0.217 to 0.755 with a mean value of 0.498, while expected heterozygosity ranged from 0.254 to 0.866 with a mean value of 0.555. Four microsatellites showed a high polymorphic information content (PIC) of more than 0.5, indicating that the loci were very useful in assessing genetic diversity and population structure.

The eight polymorphic loci exhibited Hardy-Weinberg equilibrium after Bonferroni correction. The potential occurrence of null alleles at each locus was tested using Micro-Checker v. 2.2.3 (Van Oosterhout et al. 2004). Null allele frequency ranged from 0 to 0.071. No locus pairs were in gametic disequilibrium following sequential Bonferroni correction. These microsatellite markers will be useful tools for assessing the genetic relationships between $C. aspera$, $C. seridis$ and their hybrids, as well as genetic diversity and population structure among and within natural populations of each species ($C. aspera$ and $C. seridis$).

References


**Acknowledgements**

We thank COMA V (Valencian Institute of Conservation and Agrodiversity Improvement), Carmelo López Del Rincón, and José Miguel Blanca Postigo, for help in the lab. Financial support for this study came from GV (Generalitat Valenciana) and UPV (University Polytechnic of Valencia).
Table 1 8 microsatellite loci of *Centaurea aspera* and *C. seridis* and their amplification information (average across *C. aspera* populations)

<table>
<thead>
<tr>
<th>Locus</th>
<th>Primer sequence (5’-3’)</th>
<th>Repeat motif</th>
<th>Size (bp)</th>
<th>Ta (ºC)</th>
<th>Mg</th>
<th>Na</th>
<th>H₀</th>
<th>Hₑ</th>
<th>PIC</th>
<th>Pₑ</th>
<th>Null allele frequency</th>
<th>GenBank Accession No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA001</td>
<td>F: AACATGGGCGTTCTCATTCAC</td>
<td>(CAA)₃₂</td>
<td>281</td>
<td>65</td>
<td>1.5</td>
<td>8</td>
<td>0.674</td>
<td>0.797</td>
<td>0.79</td>
<td>0.075</td>
<td>0.045</td>
<td>EH716506.1</td>
</tr>
<tr>
<td></td>
<td>R: CGTAGCTAGCCGGAGTCTTTT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA002</td>
<td>F: ATCTCTCCATGACCACCTCTT</td>
<td>(TCT)₃₁₃</td>
<td>256</td>
<td>55.5</td>
<td>2</td>
<td>5</td>
<td>0.543</td>
<td>0.452</td>
<td>0.45</td>
<td>0.199</td>
<td>0.000</td>
<td>EH754166.1</td>
</tr>
<tr>
<td></td>
<td>R: TCTAGGGTACGAGGAGCCTGA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA003</td>
<td>F: GGGGGAGGAAAACAAACATAGA</td>
<td>(ACA)₄₁₄</td>
<td>229</td>
<td>57</td>
<td>1.5</td>
<td>3</td>
<td>0.286</td>
<td>0.254</td>
<td>0.25</td>
<td>1</td>
<td>-0.024</td>
<td>EH780862.1 EH772571.1</td>
</tr>
<tr>
<td></td>
<td>R: ATCCACCAAGAAACCTCAATCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA004</td>
<td>F: GTGGAAGTCTGTTGATGGAAGA</td>
<td>(CAA)₃₁₃</td>
<td>197</td>
<td>55</td>
<td>1.5</td>
<td>9</td>
<td>0.755</td>
<td>0.859</td>
<td>0.85</td>
<td>0.022</td>
<td>0.054</td>
<td>EH783651.1</td>
</tr>
<tr>
<td></td>
<td>R: TCTGTTCCCTTGGCATCCATCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA005</td>
<td>F: CAGTGTTGGATGATGGAACAC</td>
<td>(TTC)₆</td>
<td>299</td>
<td>54</td>
<td>2.5</td>
<td>4</td>
<td>0.486</td>
<td>0.552</td>
<td>0.54</td>
<td>0.260</td>
<td>0.057</td>
<td>EH789030.1</td>
</tr>
<tr>
<td></td>
<td>R: GAAGAAATTGCACAGGGACAGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA006</td>
<td>F: AGGACTATTACGCCCTTCTCA</td>
<td>(TCA)₁₂</td>
<td>208</td>
<td>55.5</td>
<td>2</td>
<td>3</td>
<td>0.217</td>
<td>0.335</td>
<td>0.33</td>
<td>0.243</td>
<td>0.000</td>
<td>EH751778.1</td>
</tr>
<tr>
<td></td>
<td>R: CAGGTCAACTGAGGTTTTG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA007</td>
<td>F: AAACCATGAATCAAGTGCAGC</td>
<td>(ATG)₁₄</td>
<td>200</td>
<td>57</td>
<td>2</td>
<td>10</td>
<td>0.704</td>
<td>0.866</td>
<td>0.86</td>
<td>0.015</td>
<td>0.071</td>
<td>EH777535.1</td>
</tr>
<tr>
<td></td>
<td>R: TGAATGCTTGGTTTCCCTCATC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA008</td>
<td>F: GGAGATGCGGATGTGAATTG</td>
<td>(CAT)₆</td>
<td>178</td>
<td>55.5</td>
<td>2</td>
<td>5</td>
<td>0.318</td>
<td>0.329</td>
<td>0.32</td>
<td>1</td>
<td>0.000</td>
<td>EH772209.1</td>
</tr>
<tr>
<td></td>
<td>R: ATGCAGGTGTCTCCTCAATTTG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
F, forward primer; R, reverse primer; Ta, annealing temperature °C; Mg, MgCl₂ 50 mM; Na, number of alleles for *C. aspera* and *C. seridis*; H₀, observed heterozygosity; Hₑ, expected heterozygosity; PIC, Polymorphism Information Content; PᵥHWE, P-value of Hardy-Weinberg exact test.; *, these p-values are non-significant (P > 0.05) after sequential Bonferroni corrections (□). □□□= □/n; \[ n = \frac{N_a^2 - N_a}{2} \]