Development and Characterization of Simple Sequence Repeat Loci for the Common Morning Glory, *Ipomoea purpurea* (Convolvulaceae)

Adam Kuester¹,⁴, Britnie Foltz², Grant Lor², Shu-Mei Chang², Jeffrey L. Bennetzen³, Regina Baucom¹

¹721 Rieveschl Hall, Biological Sciences Department, University of Cincinnati, Cincinnati, OH
²2502 Plant Sciences Building, Plant Biology Department, 120 Carlton St. University of Georgia, Athens, GA
³C426a Life Sciences Building, Department of Genetics, University of Georgia, Athens, GA
⁴Corresponding Author:

Adam Kuester, 721 Rieveschel Hall, Biological Sciences Department, University of Cincinnati, Cincinnati, OH, email: kuesteam@ucmail.uc.edu

Abstract

Twenty Simple Sequence Repeat (SSR) loci were identified and characterized from a 454 sequencing project of the common morning glory, *Ipomoea purpurea*, an important agricultural weed of North America. Eighteen of the twenty loci were identified as polymorphic, with 1 to 4 alleles in 21 individuals from a wild population from North Carolina. Observed and Expected diversity (*H₀* and *Hₑ*) ranged from 0.000-0.619 and 0.000-0.591, respectively. The resources reported here will enable additional population genetic study for the common morning glory.

Running title SSR Loci in the Common Morning Glory

Key words: morning glory, *Ipomoea purpurea*, simple sequence repeat loci, population genetics
Introduction

The common morning glory (*Ipomoea purpurea*; family Convolvulaceae) is a frequently cultivated ornamental plant with a global distribution. *I. purpurea* has been identified as an important agricultural weed in the United States (wssa.org), primarily affecting corn, soybean, cotton and peanut crops in the Southeastern and central United States (Webster and Colbe 1997). The morning glory genus, *Ipomoea*, has also been the subject of numerous investigations addressing species interactions, which has made it a valuable emerging model system (Baucom et al. 2011) for studies elucidating pollination biology (Epperson and Clegg 1987) and evolution of resistance traits against herbivores (Rausher et al. 1993), pathogens (Kniskern and Rausher 2007) and herbicides (Baucom and Mauricio 2010). Molecular genetic studies on morning glories have illuminated genetic pathways involved in extreme floral color variation, which has evident ecological implications on mating systems and interactions with associated pollinators (Clegg and Durbin 2000; Zufall and Rausher 2004).

Though morning glories have been widely studied for the past 80 years, the development of molecular genetic resources within the genus has only recently commenced (ex. Sweet potato (*I. batatas*: Hu et al. 2004, Schafleitner et al. 2010; *I. nil*: Tong et al. 2012), and few markers have been made freely available for the use in population genetic study of morning glories. In particular, development of neutral genetic markers will enable us to investigate population genetic structures across species ranges in areas where populations are recognized as agricultural pests and will also allow further study of plant mating systems. Here, we describe the utility of 20 Simple Sequence Repeat (SSR) loci developed from a 454 genomic sequencing project.

Materials and Methods
Marker Development

Genomic DNA was extracted from the leaf tissue of a single *I. purpurea* individual at the 4-leaf stage of growth with a DNeasy Plant Maxi Kit (QIAGEN). A random genomic DNA library was sequenced in a half-plate run on a 454 GS XLR70 Titanium genomic sequencer at the University of Florida ICBR Genomics Services Division following standard protocols of library preparation and sequencing (Roche, Inc.).

SSR Finder (Sanchez-Villeda et al 2003, available upon request) was used to identify putative microsatellite locus primer sets from a FASTA 454 sequencing dataset. Of the 3407 SSR loci identified by SSR Finder, we selected 50 primer pairs to screen for polymorphism that fit the suggested parameters of Temnykh et al (2001), which determined repeat motifs that were more likely to be polymorphic.

Genetic Material

Seed material from a natural population of *I. purpurea* near Duplin, NC (N 34°58'55.70", W 78°2'7.80") was collected in the fall of 2003. Seeds from twenty-one individuals were collected from a soy field. Seeds obtained from four additional southeastern natural populations were used for screening the SSR loci. For this initial screen, we used the following number of individuals per population: Bogart, GA: (N 33°52'28.27", W 83°31'34.14"): 4; Duplin West, NC (N 34°58'55.70", W 78°2'7.80"): 1, Old Kenley, NC (N 36°8'55.34", W 78°3'35.00"): 1, Willis, TN (N 35°18'45.60", W 85°56'10.40"): 1. Seeds were germinated in a top soil mixture in the University of Georgia Plant Biology Greenhouses and leaf tissue from expanding leaves was extracted for genetic analysis.
DNA Isolation

Fifteen milligrams of silica-dried leaf material were used for DNA isolation. The tissue was extracted using a 96 well plate CTAB extraction procedure developed by T. Culley (University of Cincinnati). DNA concentrations were determined by spectrophotometry, and dilutions of 20 ng/µL were prepared for use in PCR.

All forward primers were attached to an M13 tag (Schuelke 2000) with labeled dye (6-FAM; Applied Biosystems, Carlsbad, California, USA). Of the thirty loci that were not reported here, eight failed to amplify, nine were monomorphic and thirteen proved difficult to score due to non-specific binding or poor amplification. Information on loci that were monomorphic or failed to amplify in *I. purpurea* is available on request from the first author.

PCR Reaction and Multiplexing

To determine the level of marker polymorphism within *I. purpurea*, we genotyped 21 individuals from a natural population in Duplin County, North Carolina. All primers were obtained from Integrated DNA Technologies (Coralville, Iowa, USA). We determined multiplexed genotyping based on fragment size overlap size preclusions. The 20 SSRs were run in nine multiplexes, consisting of one, two or three loci per genotyping run. The PCR and genotyping multiplexes were as follows: Multiplex 1: IP6, IP26, IP47; Multiplex2: IP31, IP27; Multiplex 3: IP2, IP9; Multiplex4: IP12, IP21, IP42; Multiplex5: IP36, IP41; Multiplex6:IP8, IP24, IP44; Multiplex7: IP1, IP13, IP18; IP45 and IP34 were run singly.

For each reaction, 40 ng of genomic DNA was used in a 10 µL reaction, consisting of 5 µL of Taq PCR Master Mix [(2× concentrated) containing 0.5 unit Taq DNA Polymerase, Qiagen PCR Buffer (with 3 mmol MgCl2)] and 400 mmol of each dNTP (Taq PCR Master Mix Kit, Qiagen).
An Eppendorf (Hauppauge, NY) Mastercycler pro thermal cycler- was used for all PCR reactions. Thermocycler conditions followed recommendations of Schuelke (2000), consisting of 95°C for 3 min; 20 cycles of 94°C for 30 s, 58°C for 90 s, 72°C for 60 s; ten cycles of 94°C for 30 s, 56°C for 90 s, 72°C for 60 s and a final extension at 72°C for 30 min. Samples that failed to amplify were run a second time. One microliter of PCR product was used for detection using an Applied Biosystems 3730 DNA Analyzer at the Cornell University Life Sciences Core Facility (Ithaca, NY). An ABI GS500 Liz size standard was used for fragment length comparison. All sample genotypes were analyzed using Applied Biosystems PeakScanner 1.0 analytical software. A PP (Primer Peaks adjustment) sizing default was used for the analysis.

Population Genetic Analyses

Observed (H₀) and expected (Hₑ) heterozygosity within populations were estimated using GenAlEx version 6.1 (Peakall and Smouse 2006). Within the Duplin population we tested for deviation from Hardy–Weinberg equilibrium (HWE) and pairwise linkage disequilibrium (LD; 190 locus pairs) using GENEPOP version 4.0 (Raymond and Rousset 1995).

Results

All 20 primer pairs reported were polymorphic in a preliminary assessment of 7 samples from populations collected from North Carolina, Georgia and Tennessee in the southeastern US (unpublished data). In particular, primer sets that were monomorphic in the Duplin NC population were polymorphic from our preliminary screen from individuals collected across NC,
GA and TN (IP1: size range: 276-279; IP44 size range: 162-168). Eighteen of these loci produced polymorphic genotypes within the Duplin, NC population (Table 1). All loci for each individual amplified after two genotyping runs. Within-population observed heterozygosity (Hₒ) ranged from 0.000 to 0.619 (mean = 0.240) and expected (Hₑ) heterozygosity ranged from 0.000 to 0.591 (mean = 0.320) for the Duplin population (Table 1). The total number of alleles per locus (Nₐ) are presented in Table 1 and ranged from 1 to 4 (mean = 3).

We detected significant deviation from Hardy-Weinberg equilibrium at five of the twenty loci for the Duplin population (all P values < 0.05, Table 1). All loci were in linkage equilibrium.

Discussion

Simple sequence repeat loci for Ipomoea will be useful in quantifying genetic variation, in mating system analyses, and to assess gene flow within an important agricultural weed. We expect the loci described here will be of use in other closely related morning glory species, including other agricultural weeds. These SSR markers may also aid comparative studies between morning glories and the domesticated sweet potato, for which SSR loci have just recently have been developed (Schafleitner et al. 2010).

Acknowledgements

The authors thank T. Culley for her helpful procedural advice and contribution to this manuscript. Funding for this work was provided by USDA grant 1009672 to RSB and SMC, NSF grant 0640881 to SMC, and the Giles Endowment to JLB.

Baucom RS, Mauricio R Defense against the herbicide Round Up (R) predates its widespread use. Evolutionary Ecology Research 12, 131-141.


Table 1. Description of Simple Sequence Repeat loci developed in *Ipomoea purpurea*. Shown are the locus identification number (Locus ID), forward and reverse primer sequences, annealing temperature (\(T_a\)), repeat motif, number of individuals successfully amplified from the Duplin County, North Carolina population out of a total of 21 individuals (N), number of alleles observed (Na), allele size range, observed (\(H_o\)) and expected (\(H_e\)) heterozygosity, deviation of loci within the Duplin population from Hardy-Weinberg equilibrium (HWE) and GenBank accession reference number (Accession No). *Significantly deviated from Hardy-Weinberg equilibrium at \(P < 0.05\). Monomorphic loci could not be evaluated for HWE, and are designated NA.

<table>
<thead>
<tr>
<th>Locus ID</th>
<th>Forward Primer</th>
<th>Reverse Primer</th>
<th>(T_a)</th>
<th>Repeat Motif</th>
<th>N</th>
<th>Na</th>
<th>Allele Size Range</th>
<th>(H_o)</th>
<th>(H_e)</th>
<th>HWE</th>
<th>Accession No</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP34</td>
<td>ACCGCTGGTTCTAAACCTGTAATTA</td>
<td>AATTTTGTAGGGTCCACAATGAGG</td>
<td>57</td>
<td>TAT</td>
<td>21</td>
<td>4</td>
<td>156-172</td>
<td>0.524</td>
<td>0.557</td>
<td>NS</td>
<td>KC122214</td>
</tr>
<tr>
<td>IP36</td>
<td>CAATATTGAAATGTGAGGTAGTGG</td>
<td>TCCACATATTGCTGCAGCATACACC</td>
<td>57</td>
<td>TGTGAA</td>
<td>21</td>
<td>3</td>
<td>90-102</td>
<td>0.619</td>
<td>0.591</td>
<td>NS</td>
<td>KC122215</td>
</tr>
<tr>
<td>IP41</td>
<td>GTCAACCTCTCCTCTCGAGGTATT</td>
<td>CTCAACACCTCATTTTGTAGAATCTCA</td>
<td>57</td>
<td>TTTGTT</td>
<td>21</td>
<td>2</td>
<td>152-158</td>
<td>0.143</td>
<td>0.495</td>
<td>*</td>
<td>KC122216</td>
</tr>
<tr>
<td>IP24</td>
<td>GAACACATCTTTATCGACCTTGA</td>
<td>TCTCTTTGTAATTCAGTAGTGGAGAAGAA</td>
<td>57</td>
<td>ATT</td>
<td>21</td>
<td>3</td>
<td>170-176</td>
<td>0.143</td>
<td>0.316</td>
<td>*</td>
<td>KC122210</td>
</tr>
<tr>
<td>IP44</td>
<td>CCAGTCATCTATGTTCTTGTGATC</td>
<td>AAAAATCAACAGAACCCTATAGAATAGC</td>
<td>57</td>
<td>TTTTTG</td>
<td>21</td>
<td>1</td>
<td>168</td>
<td>0.000</td>
<td>0.000</td>
<td>NA</td>
<td>KC122218</td>
</tr>
<tr>
<td>IP8</td>
<td>ATGCTTTCATTCACACCTCATT</td>
<td>AGGAATGCAAAAAGCTTCATTTTG</td>
<td>57</td>
<td>TAT</td>
<td>21</td>
<td>3</td>
<td>171-180</td>
<td>0.190</td>
<td>0.359</td>
<td>*</td>
<td>KC122225</td>
</tr>
<tr>
<td>IP13</td>
<td>TACATGTCTGTAATCTGCAACCA</td>
<td>TGAATGCTCAAGATTTATGATGACTT</td>
<td>57</td>
<td>AAT</td>
<td>21</td>
<td>3</td>
<td>171-180</td>
<td>0.190</td>
<td>0.251</td>
<td>NS</td>
<td>KC122222</td>
</tr>
<tr>
<td>IP18</td>
<td>TAGGGGCTAGGAAACACAGATTGA</td>
<td>GGACACATAGGTTCCACTACATGC</td>
<td>57</td>
<td>AGG(1)AT(1)</td>
<td>21</td>
<td>2</td>
<td>217-220</td>
<td>0.095</td>
<td>0.308</td>
<td>*</td>
<td>KC122223</td>
</tr>
<tr>
<td>IP1</td>
<td>CAATTCTTTGCAATCCCTCTACGG</td>
<td>TGTGGTTTGGAAGACAGATGA</td>
<td>57</td>
<td>ATA</td>
<td>21</td>
<td>1</td>
<td>276</td>
<td>0.000</td>
<td>0.000</td>
<td>NA</td>
<td>KC122221</td>
</tr>
<tr>
<td>IP45</td>
<td>GTCAATCATGTCTGGTGAGTTGACAA</td>
<td>TTTTGAATTTGGAATCTGTTGTTAT</td>
<td>57</td>
<td>AAAAT</td>
<td>21</td>
<td>3</td>
<td>164-174</td>
<td>0.381</td>
<td>0.387</td>
<td>NS</td>
<td>KC122219</td>
</tr>
<tr>
<td>IP42</td>
<td>TACATGCTTAATCAGCTCCATTGA</td>
<td>TGGTTCTATGATTTATATGTCAGG</td>
<td>57</td>
<td>TTTTCT</td>
<td>21</td>
<td>2</td>
<td>132-138</td>
<td>0.048</td>
<td>0.046</td>
<td>NC</td>
<td>KC122217</td>
</tr>
<tr>
<td>IP12</td>
<td>GTTTTTAACCCACCACAAAGATT</td>
<td>AAGCTGTTTTGCGCTTTGTTTACT</td>
<td>57</td>
<td>AAT</td>
<td>21</td>
<td>3</td>
<td>166-172</td>
<td>0.429</td>
<td>0.359</td>
<td>NS</td>
<td>KC122208</td>
</tr>
<tr>
<td>IP21</td>
<td>CATAACATATGCTTCAACAACAGGG</td>
<td>AAGACATAATGCGCTCTTGATCAA</td>
<td>57</td>
<td>ATA(1)ATG(1)</td>
<td>21</td>
<td>3</td>
<td>180-189</td>
<td>0.048</td>
<td>0.135</td>
<td>NS</td>
<td>KC122209</td>
</tr>
<tr>
<td>IP2</td>
<td>GCAAACAGAGAGCTAGCAACAT</td>
<td>GCTGCTAGAGAGCTTAAATGCG</td>
<td>57</td>
<td>ATA</td>
<td>21</td>
<td>3</td>
<td>106-139</td>
<td>0.429</td>
<td>0.635</td>
<td>NS</td>
<td>KC122206</td>
</tr>
<tr>
<td>IP9</td>
<td>TAAATCTTGCAGACCAGACCACT</td>
<td>GTTGGATGAGCTCAAGATTTC</td>
<td>57</td>
<td>TCT</td>
<td>21</td>
<td>2</td>
<td>164-173</td>
<td>0.471</td>
<td>0.472</td>
<td>NS</td>
<td>KC122207</td>
</tr>
<tr>
<td>IP51</td>
<td>GGAACAGCAAGCTTTAACATTT</td>
<td>TGTGCTACAAAGAATTGCTTTT</td>
<td>57</td>
<td>GA</td>
<td>21</td>
<td>3</td>
<td>92-102</td>
<td>0.095</td>
<td>0.254</td>
<td>NS</td>
<td>KC122213</td>
</tr>
<tr>
<td>IP27</td>
<td>CATCTTACAGGCTCAATACATGGC</td>
<td>CTACGCGGAAACAGTCTATCATG</td>
<td>57</td>
<td>CAG</td>
<td>21</td>
<td>3</td>
<td>164-191</td>
<td>0.429</td>
<td>0.475</td>
<td>NS</td>
<td>KC122212</td>
</tr>
<tr>
<td>IP47</td>
<td>TGAATTCTTCTCTGAGCCATGATT</td>
<td>TCTGATGTATCTACGCAAGTTTTT</td>
<td>57</td>
<td>AAAAT</td>
<td>21</td>
<td>2</td>
<td>123-128</td>
<td>0.095</td>
<td>0.172</td>
<td>NS</td>
<td>KC122220</td>
</tr>
<tr>
<td>IP6</td>
<td>GCTTATGATGAGAAGTGTGTTT</td>
<td>TCAAATCAAGAAATCAGATCTCAGAC</td>
<td>57</td>
<td>ATT</td>
<td>21</td>
<td>3</td>
<td>173-185</td>
<td>0.238</td>
<td>0.214</td>
<td>NS</td>
<td>KC122224</td>
</tr>
<tr>
<td>IP26</td>
<td>TTTATCCACTTTTGGCTTTTGTG</td>
<td>TAAGGGGAGACACCAAATCAGT</td>
<td>57</td>
<td>ATT</td>
<td>21</td>
<td>2</td>
<td>242-245</td>
<td>0.095</td>
<td>0.308</td>
<td>*</td>
<td>KC122211</td>
</tr>
</tbody>
</table>