Genic microsatellite markers derived from EST sequences of switchgrass (Panicum virgatum L.)

Christian M. Tobias  
USDA, ARS, Western Regional Research Center, Genomics and Gene Discovery Unit, 800 Buchanan Street, Albany, CA

Daniel M. Hayden  
USDA, ARS, Western Regional Research Center, Genomics and Gene Discovery Unit, 800 Buchanan Street, Albany, CA

Paul Twigg  
University of Nebraska at Kearney, Kearney, Nebraska

Gautam Sarath  
University of Nebraska - Lincoln, gsarath1@unl.edu

Follow this and additional works at: http://digitalcommons.unl.edu/usdaarsfacpub

Part of the Agricultural Science Commons

http://digitalcommons.unl.edu/usdaarsfacpub/43

This Article is brought to you for free and open access by the USDA Agricultural Research Service --Lincoln, Nebraska at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Publications from USDA-ARS / UNL Faculty by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
Genic microsatellite markers derived from EST sequences of switchgrass (*Panicum virgatum* L.)

CHRISTIAN M. TOBIAS,* DANIEL M. HAYDEN,* PAUL TWIGG† and GAUTAM SARATH‡

*USDA, ARS, Western Regional Research Center, Genomics and Gene Discovery Unit, 800 Buchanan Street, Albany, CA 94710, USA, †Biology Department, Bruner Hall of Science, University of Nebraska at Kearney, Kearney, Nebraska 68849, USA, ‡USDA, ARS, Wheat, Sorghum and Forage Research Unit, Keim Hall, E.C. University of Nebraska, Lincoln, Nebraska 68583, USA

**Abstract**

Switchgrass is a large, North American, perennial grass that is being evaluated as a potential energy crop. There is a need to assess genetic diversity in stored accessions and in remaining native stands to assist breeding and conservation efforts. Marker development will also be necessary for genetic linkage mapping. Toward this end, 32 switchgrass genic di-, tri- and tetranucleotide repeat microsatellites were identified from expressed sequence tags (ESTs). These microsatellites were used to screen individuals from two different named cultivars. The markers displayed a high level of polymorphism consistent with the tetraploid, allogamous behaviour of the cultivars tested.

**Keywords:** energy crop, genetic diversity, polyploid, simple sequence repeat

Received 25 August 2005; revision accepted 12 September 2005

Switchgrass is a member of the subfamily Pooidae closely related to proso millet, maize, sorghum and sugarcane. It is utilized as a summer forage, and also has promise as an herbaceous energy crop due to its large size and perennial nature. Improved genetics through breeding is the most economical route to improved value in the low input systems of cultivation employed for forage grasses. Inexpensive genetic marker systems for switchgrass are therefore required for marker-assisted selection to be successful. Such markers will also be useful for guiding conservation strategies to preserve remaining North American native prairie sites. Based on nuclear DNA content and analysis of segregation ratios, switchgrass has been shown to include both tetraploid and octoploid ecotypes. Restriction fragment length polymorphism (RFLP) and random amplified polymorphic DNA (RAPD) markers have been used for linkage studies and to discriminate ecotypes (Missaoui et al. 2005), but no microsatellite markers have been developed for this species. Expressed sequence tag (EST)-derived microsatellites have the added advantages of assaying functional diversity in transcripts, allowing greater transferability across species, and in some cases having additional utility as anchor markers for comparative mapping (Varshney et al. 2005). In this study, ESTs from cv. Kanlow (Tobias et al. 2005) were used as a source of sequence information for the development of microsatellites. Available sequence data from switchgrass were screened using the program ssrit to identify microsatellite sequences larger than 20 bp (Temnykh et al. 2001). Primers were then designed using primer 3.0 (Rozen & Skaletsky 2000) with an optimal primer $T_m$ across the gene-specific regions of 60 °C and a target product size of 200 bases.

We isolated genomic DNA using a modification of a urea prep method (Shureet al. 1983). Forward primers were tailed with M13 (5′-CACGACGTTGTAACACGAC-3′) to allow fluorescent labelling via incorporation of either a FAM- or a VIC-labelled M13 primer into the amplification mixture (Boutin-Ganache et al. 2001). The polymerase chain reactions (PCRs) were carried out in a final volume of 15 µL on an MJ Research PTC-225 thermocycler with a thermal profile consisting of a 2-min initial denaturation step at 94 °C followed by 35 cycles of 20 s at 94 °C, 20 s at the primer-specific annealing temperature (Table 1) and 1 min at 72 °C. A final 72 °C extension step of 30 min was included to promote nontemplated nucleotide addition at the 3′ end of the PCR product (Clark 1988). Reactions were carried out in 10 mM Tris-HCl (pH 9.0 at 25 °C), 50 mM KCl, 0.1% (v/v) Triton X-100, 1% (v/v) DMSO, 0.01% (w/v) gelatin, 0.75–2.5 mM

Correspondence: Christian M. Tobias, Fax: (510) 559–5818; E-mail: ctobias@pw.usda.gov

© 2006 Blackwell Publishing Ltd
**Table 1** Characterization of 32 microsatellites isolated from switchgrass cv. ‘Kanlow’ ESTs and tested on 10 individuals

<table>
<thead>
<tr>
<th>Marker name</th>
<th>Repeat motif in library</th>
<th>Primer sequences (5′–3′)</th>
<th>Size range (bp)</th>
<th>Tm (°C) [MgCl2 (mM)]</th>
<th>Total no. of bands</th>
<th>No. of bands per individual</th>
<th>Diversity index D</th>
<th>GenBank Accession no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>4821_A04</td>
<td>(CAG)$_7$</td>
<td>F: M13-CTCCGCTGCTGATCAGATCAAACA</td>
<td>203–212</td>
<td>56 [1.5]</td>
<td>3</td>
<td>1–2</td>
<td>0.76</td>
<td>DN144448</td>
</tr>
<tr>
<td>4806_H12</td>
<td>(GCC)$_7$</td>
<td>F: M13-CCCTGAAATGACAGCGCTGA</td>
<td>503–519</td>
<td>58 [1.5]</td>
<td>5</td>
<td>1–2</td>
<td>0.58</td>
<td>DN143384</td>
</tr>
<tr>
<td>4808_F04</td>
<td>(GCAT)$_3$</td>
<td>F: M13-GTTCACCTCGTCAGTATCC</td>
<td>221–233</td>
<td>55 [1.5]</td>
<td>4</td>
<td>1–2</td>
<td>0.74</td>
<td>DN143517</td>
</tr>
<tr>
<td>4816_C03</td>
<td>(AG)$_16$</td>
<td>F: M13-AGTGGAAAGTGGCTGCTGAGA</td>
<td>215–279</td>
<td>51 [1.5]</td>
<td>8</td>
<td>1–2</td>
<td>0.88</td>
<td>DN144060</td>
</tr>
<tr>
<td>4827_D09</td>
<td>(AG)$_4$CC</td>
<td>F: M13-CTTCGCTGATTTGGAACCTG</td>
<td>230–261</td>
<td>56 [1.5]</td>
<td>11</td>
<td>1–2</td>
<td>0.9</td>
<td>DN144982</td>
</tr>
<tr>
<td>4828_E03</td>
<td>(TA)$_18$</td>
<td>F: M13-CTTATATCGCATGTTCTTCTT</td>
<td>198–215</td>
<td>54 [2.5]</td>
<td>3</td>
<td>1–2</td>
<td>0.68</td>
<td>DN145072</td>
</tr>
<tr>
<td>4842_D11</td>
<td>(GG)$_9$</td>
<td>F: M13-GATGCGATGAGAGAGAGAGAGA</td>
<td>181–193</td>
<td>54 [1.25]</td>
<td>5</td>
<td>1–2</td>
<td>0.74</td>
<td>DN146202</td>
</tr>
<tr>
<td>4848_B04</td>
<td>(AT)$_13$</td>
<td>F: M13-TCCAGATGATGCCCCAGGAC</td>
<td>214–281</td>
<td>54 [1.25]</td>
<td>14</td>
<td>1–3</td>
<td>0.88</td>
<td>DN146641</td>
</tr>
<tr>
<td>4853_D04</td>
<td>(TA)$_11$</td>
<td>F: M13-CGGAATGCTGCTCAGTACA</td>
<td>232–238</td>
<td>55 [1.5]</td>
<td>3</td>
<td>1–2</td>
<td>0.48</td>
<td>DN147088</td>
</tr>
<tr>
<td>4926_E12</td>
<td>(TC)$_16$</td>
<td>F: M13-GCTGAAACAGGAAAACGAAA</td>
<td>249–273</td>
<td>54 [2.5]</td>
<td>6</td>
<td>1–2</td>
<td>0.82</td>
<td>DN148200</td>
</tr>
<tr>
<td>4933_G06</td>
<td>(GCC)$_7$</td>
<td>F: M13-CACACATGTCATCTCATGCGT</td>
<td>210–219</td>
<td>53 [2.0]</td>
<td>2</td>
<td>2</td>
<td>0.18</td>
<td>DN148479</td>
</tr>
<tr>
<td>4934_H04</td>
<td>(CG)$_7$</td>
<td>F: M13-CAGTGGAGGAGAGAGAGAGAGA</td>
<td>167–204</td>
<td>54 [2.5]</td>
<td>4</td>
<td>2–3</td>
<td>0.74</td>
<td>DN148576</td>
</tr>
<tr>
<td>4937_G11</td>
<td>(ACA)$_7$</td>
<td>F: M13-TGCTGACAGAGAAAAAGGTTA</td>
<td>194–207</td>
<td>56 [1.5]</td>
<td>1</td>
<td>2–3</td>
<td>0.88</td>
<td>DN148799</td>
</tr>
<tr>
<td>4941_F01</td>
<td>(GA)$_17$</td>
<td>F: M13-CTGAGACGACGACAGACAGAC</td>
<td>229–231</td>
<td>54 [1.25]</td>
<td>3</td>
<td>1–3</td>
<td>0.46</td>
<td>DN149096</td>
</tr>
<tr>
<td>4942_H03</td>
<td>(ATCC)$_7$</td>
<td>F: M13-GTTGATGCTGCTGATGAGAC</td>
<td>202–210</td>
<td>54 [2.5]</td>
<td>3</td>
<td>1–2</td>
<td>0.34</td>
<td>DN149194</td>
</tr>
<tr>
<td>4943_D09</td>
<td>(CCG)$_8$</td>
<td>F: M13-ACAGAAGCCGACGACGATG</td>
<td>217–221</td>
<td>52 [1.5]</td>
<td>4</td>
<td>1–2</td>
<td>0.78</td>
<td>DN149247</td>
</tr>
<tr>
<td>5048_B06</td>
<td>(GCC)$_9$</td>
<td>F: M13-CCCTCTCTGCTCTGCTGCT</td>
<td>328–354</td>
<td>54 [1.25]</td>
<td>5</td>
<td>1–3</td>
<td>0.78</td>
<td>DN142309</td>
</tr>
<tr>
<td>5005_B08</td>
<td>(AT)$_29$</td>
<td>F: M13-AGGACATCTGACTCAGTATCT</td>
<td>217–225</td>
<td>54 [2.5]</td>
<td>4</td>
<td>1–2</td>
<td>0.76</td>
<td>DN140825</td>
</tr>
<tr>
<td>5008_B05</td>
<td>(GCC)$_9$</td>
<td>F: M13-CCTCCTCTCTGTCCACGCAC</td>
<td>206–218</td>
<td>52 [1.5]</td>
<td>7</td>
<td>1–3</td>
<td>0.88</td>
<td>DN140953</td>
</tr>
<tr>
<td>5013_F08</td>
<td>(CCG)$_9$</td>
<td>F: M13-CCTCCTGTCTGCTGACCCAC</td>
<td>204–255</td>
<td>54 [2.0]</td>
<td>9</td>
<td>1–2</td>
<td>0.86</td>
<td>DN141358</td>
</tr>
<tr>
<td>5023_A02</td>
<td>(CAG)$_9$</td>
<td>F: M13-CTCCAGATCTCAACACCAACA</td>
<td>294–315</td>
<td>56 [1.25]</td>
<td>8</td>
<td>1–3</td>
<td>0.88</td>
<td>DN141376</td>
</tr>
<tr>
<td>5028_C11</td>
<td>(AG)$_3$</td>
<td>F: M13-CCTCCTCTCTCTCTCTCTCT</td>
<td>136–191</td>
<td>56 [1.25]</td>
<td>10</td>
<td>2–3</td>
<td>0.90</td>
<td>DN141676</td>
</tr>
<tr>
<td>5036_D10</td>
<td>(AG)$_3$</td>
<td>F: M13-CCTCCCTCTCTGACTCACTCA</td>
<td>140–173</td>
<td>54 [2.5]</td>
<td>8</td>
<td>1–2</td>
<td>0.88</td>
<td>DN141907</td>
</tr>
<tr>
<td>5052_C08</td>
<td>(CT)$_12$</td>
<td>F: M13-AGGGAAATAGGCCGTCGAGA</td>
<td>248–270</td>
<td>47 [1.25]</td>
<td>13</td>
<td>1–4</td>
<td>0.88</td>
<td>DN142646</td>
</tr>
<tr>
<td>5054_F06</td>
<td>(TC)$_12$</td>
<td>F: M13-CGACGCAAACACGACGACGAC</td>
<td>157–191</td>
<td>52 [1.5]</td>
<td>8</td>
<td>1–2</td>
<td>0.86</td>
<td>DN143022</td>
</tr>
<tr>
<td>5057_A04</td>
<td>(AG)$_15$</td>
<td>F: M13-CCACCGCCTCCTTCTCTGAGA</td>
<td>285–339</td>
<td>52 [0.75]</td>
<td>8</td>
<td>1–2</td>
<td>0.86</td>
<td>DN150540</td>
</tr>
<tr>
<td>5205_E06</td>
<td>(CG)$_8$</td>
<td>F: M13-ACAGGTCCTGAACTCAGGGAGA</td>
<td>175–186</td>
<td>56 [1.5]</td>
<td>5</td>
<td>1–2</td>
<td>0.76</td>
<td>DN150540</td>
</tr>
<tr>
<td>5211_B07</td>
<td>(AG)$_8$</td>
<td>F: M13-CCTCCATTGCTCCAAAGAGA</td>
<td>244–255</td>
<td>55 [1.5]</td>
<td>4</td>
<td>1–2</td>
<td>0.78</td>
<td>DN150631</td>
</tr>
</tbody>
</table>

**Notes**: no. of bands per individual is in the range 1–2 except for three repeats (2–3).
Repeat motif is listed 5′–3′ with respect to the forward primer (F). D is Nei’s diversity index value for each marker. $T_a$ is the annealing temperature. Numbers in parentheses indicate the MgCl₂ concentration used for each marker.

MgCl₂ (Table 1), 200 μM dNTPs in the presence of 0.4 U Taq polymerase (Promega), 15 ng template genomic DNA, 0.5 μM marker-specific reverse primer, 0.033 μM marker-specific M13-tailed forward primer and 0.5 μM VIC- or FAM-labelled M13 primer. Reaction products were polyethylene glycol (PEG)-precipitated, (Lis 1980) washed in 80% (v/v) ethanol and resuspended in 1 mm Tris pH 8.0 and 0.1 mm EDTA. Dilutions were subjected to capillary electrophoresis on an Applied Biosystems ABI 3730x DNA analyser along with prepared PET-labelled size standards of 133–433 bp.

A total of 96 primer pairs were designed and tested; of these 52 primer pairs gave no or very little amplification. Twelve primer pairs were eliminated due to excessive stutter artefacts, irreproducibility or an excessive number of unsuccessful amplifications. Of the 32 usable markers, no single marker discriminated between populations. We sequenced PCR products amplified with primers for one marker (5232_A03) and found that the sequence agreed with its corresponding GenBank Accession and that observed length variations were exclusively due to gain or loss of AG repeats. Results from five individuals each of the tetraploid switchgrass cultivars ‘Alamo’ and ‘Kanlow’ were used to calculate the total number of polymorphic bands, the number of bands per individual and Nei’s diversity index D (calculated as $1 - \sum p_i^2$) for each marker (Weir 1990). In most cases, the samples were highly polymorphic with markers containing dinucleotide repeats being on average slightly more variable. Tetraploid switchgrass cultivars behave largely as autopolyploids with a high degree of preferential pairing between chromosomes (Missaoui et al. 2005). However, in this study, partial heterozygotes could not be easily discriminated based on peak height, and thus determination of deviation from Hardy–Weinberg equilibrium, and linkage equilibrium was prevented. These markers will facilitate current mapping efforts in this species and population discrimination.

## Acknowledgements

Supported by the United States Department of Agriculture, Agricultural Research Service CRIS 5325-21000-013-00, NP307 Biofuel and Bioenergy Alternatives. This work was also supported in part by NIH Grant P20 RR16569 from the BRIN Program of the National Center for Research Resources and by a University of Nebraska at Kearney Research Services Council grant.

## References