PRIMER NOTE

DEVELOPMENT AND CHARACTERIZATION OF MICROSATELITE MARKERS FOR CENTRAL AMERICAN Begonia sect. Gireoudia (Begoniaceae)1

ALEX D. TWYFORD2,3,5, RICHARD A. ENNOS4, AND CATHERINE A. KIDNER2,3

1Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh EH3 5LR, United Kingdom; 2Institute of Molecular Plant Sciences, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JH, United Kingdom; and 3Institute of Evolutionary Biology, School of Biological Sciences, Ashworth Laboratories, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom

• Premise of the study: Transcriptome sequence data were used to design microsatellite primers for two widespread Central American Begonia species, B. heracleifolia and B. nelumbiifolia, to investigate population structure and hybridization.

• Methods and Results: The transcriptome from vegetative meristem tissue from the related B. plebeja was mined for microsatellite loci, and 31 primer pairs amplified in the target species. Fifteen primer pairs were combined in two multiplex PCR reactions, which amplified an average of four alleles per locus.

• Conclusions: The markers developed will be a valuable genetic resource for medium-throughput genotyping of Central American Begonia sect. Gireoudia. A subset of these markers have perfect sequence matches to Asian B. venusta, and are promising for studies in other Begonia sections.

Key words: Begonia heracleifolia; Begonia nelumbiifolia; Begoniaceae; hybridization; microsatellite primers; transcriptome sequences.

Methods and Results

Microsatellite markers were designed from the transcriptome sequence of vegetative meristem tissue from B. plebeja Liebm., a related species from Begonia sect. Gireoudia (European Nucleotide Archive Sequence Read Archive accession number: ERP001195; Brennan et al., 2012). The QDD bioinformatic pipeline (Meglécz et al., 2010), which integrates microsatellite detection, a redundancy check to avoid amplifying multiple PCR products, and designs primers, was used according to Lepais and Bacles (2011). A FASTA file of the B. plebeja transcriptome sequence assembly was analyzed in QDD version 1.3 using default parameters: selecting only primers that amplify a PCR product between 90 and 320 bp in length, with a repeat motif of 2–6 bp repeats, and a minimum length of four repeat units. To make microsatellite amplification in other species more likely, primers were excluded if they did not have a perfect BLAST match to the transcriptome of B. conchifolia A. Dietr. (sect. Gireoudia; Brennan et al., 2012). Reads from which the primers were designed were BLAST searched against the Arabidopsis Information Resource (TAIR) database (http://www.arabidopsis.org) to investigate the putative function of each locus.

Thirty-one primer pairs detected in QDD were tested for amplification in B. heracleifolia Cham. & Schltdl and B. nelumbiifolia Cham. & Schltdl. These species were chosen because they are two of the most widespread Begonia species in a genus of mostly rare endemics (Hughes and Hollingsworth, 2008). The species are known to hybridize (Burt-Utley, 1985), facilitating studies of species boundaries. Primer amplification was tested in seven individuals of the two species (Appendix 1). A subset of polymorphic markers that amplified reliably in both species was then tested for multiplex compatibility by mixing equimolar ratios of each primer. The PCR multiplexes were then tested on a population of each species (20 individuals) to estimate the genetic diversity of the markers. The primer sequences were BLAST searched against the transcriptome sequence of the divergent Asian species B. venusta King (sect. Platycentrum) to test for likely cross-amplification of primers in other Begonia species.

Approximately 15 mg of silica-dried leaf material was extracted using DNeasy 96-sample kit (QIAGEN, Germantown, Maryland, USA). To overcome an

Begonia L. is a diverse tropical genus with over 1500 species. Evolutionary research has focused on the early-diverging African species (e.g., Hughes and Hollingsworth, 2008) and the more derived Asian species (e.g., Thomas et al., 2011), with the American species largely overlooked. The most recent common ancestor of Central American Begonia is likely to be relatively recent (Miocene; Dewitte et al., 2011), and subsequent speciation has resulted in high species richness (total c. 690 species; Goodall-Copestake et al., 2010). Population studies of Central American Begonia species will shed light on the evolution of species richness in a morphologically diverse group of neotropical herbs; but to date, studies have been limited by the availability of suitable nuclear markers to complement plastid microsatellite markers (Twyford et al., 2013).

In this study, we describe the development of nuclear microsatellite markers to study gene flow within and between Central American Begonia species. This requires markers that amplify over a broad phylogenetic scope, which can then be cross-amplified in divergent species.
| Locus | Primer sequences (5'–3') | Multiplex | Fluorescent dye | T<sub>m</sub> (°C) | Repeat motif<sup>c</sup> | Allele sizes (bp)<sup>d</sup> | Putative function<sup>e</sup> | E-value |
|-------|--------------------------|-----------|-----------------|-----------------|-----------------|-----------------|-----------------|---------|
| BI4329 | F: M13-CACCACCAACAAATGCAAGCTTT | 1 | FAM | 59 | (GGA)<sub>6</sub> | 4 2 89–104 | immunoglobulin E-set superfamily protein | 2E-13 |
| BI3043 | F: M13-CGACATTCAACCAATCTGG | 1 | FAM | 60 | (TC)<sub>5</sub> | 1 2 173–179 | — | — |
| BC432 | F: M13-AAAACCGAATGCAAGCTTT | 1 | FAM | 60 | (TG)<sub>4</sub> | 1 1 261–263 | endotransglucosylase/hydrolase | 5E-18 |
| BC344 | F: M13-GGGGACTCTTGCTTTGAG | 1 | VIC | 60 | (GCA)<sub>3</sub> | 1 1 105–108 | chitinase-like protein | 3E-07 |
| BI6278 | F: M13-TGTAGTGTGTTGATGACGAATCTTTG | 1 | VIC | 59 | (TCC)<sub>3</sub> | 1 3 238–253 | DOF zinc finger protein | 3E-25 |
| BI5347 | F: M13-TCTGCAATTTCTTATCAAGACC | 1 | VIC | 59 | (CTT)<sub>6</sub> | 2 1 171–183 | unknown gene | 0.000002 |
| BC552 | F: M13-TGCTGAATGGAATCTCGCC | 1 | NED | 60 | (GT)<sub>2</sub> | 2 2 271–273 | — | — |
| BI3348 | F: M13-M13-TCAGCTGTTTTCCTGTGAGGA | 1 | PET | 60 | (CT)<sub>5</sub> | 3 3 279–283 | — | — |
| BI1112 | F: M13-M13-ATCCAAATGTCNACCTCTCG | 2 | FAM | 60 | (TCC)<sub>3</sub> | 2 2 109–115 | — | — |
| BI3820 | F: M13-AGGACCAATTTTGCACGCTTA | 2 | FAM | 59 | (CTT)<sub>7</sub> | 5 2 158–176 | LOB domain-containing protein | 2E-39 |
| BI1.34 | F: M13-M13-ATCAAGCTCCTATCTCTCCT | 2 | VIC | 60 | (CT)<sub>5</sub> | 4 2 306–314 | — | — |
| BI4004 | F: M13-TGCTGGAATATCCTGTTGACG | 2 | VIC | 59 | (AT)<sub>3</sub> | 2 3 155–169 | O-fucosyltransferase family protein | 1E-32 |
| BI362 | F: M13-M13-ATCCACCTGCTGTGCAACAC | 2 | NED | 60 | (ATG)<sub>4</sub> | 4 4 147–159 | Acyl-CoA N-acyltransferases (NAT) superfamily protein | 1.0E-45 |
| BC332 | F: M13-M13-AGACCAAGATGCAAGGTTCA | 2 | PET | 59 | (TCA)<sub>5</sub> | 4 2 188–200 | ATPase | 1.0E-122 |

**Additional loci tested**

| Locus | Primer sequences (5'–3') | Multiplex | Fluorescent dye | T<sub>m</sub> (°C) | Repeat motif<sup>c</sup> | Allele sizes (bp)<sup>d</sup> | Putative function<sup>e</sup> | E-value |
|-------|--------------------------|-----------|-----------------|-----------------|-----------------|-----------------|-----------------|---------|
| BC672 | F: M13-CCTTGGCAAGAAGAAGCC | 60 | (CTT)<sub>8</sub> | 3 1 | 152–158 | cellulose-synthase-like C12 | 2E-57 |
| BI4477 | F: M13-GGTCCTCTCCTGCTTCTGTTG | 60 | (CT)<sub>5</sub> | 4 2 | 111–119 | — | — |
| BI06604 | F: M13-ATTTTTTCCAAGAGAGGCC | 59 | (AT)<sub>6</sub> | 6 1 | 111–127 | catalytic LigB subunit of aromatic ring-opening dioxygenase family | 3E-13 |
| BI6294 | F: M13-TGCTGTTGCTGATATTTAACTCA | 59 | (AT)<sub>10</sub> | 1<sup>M</sup> 1<sup>M</sup> | 148 | — | — |
| BI6701 | F: M13-AGGACGTTTCTCAGCTGCAC | 60 | (GA)<sub>6</sub> | 1<sup>M</sup> 1<sup>M</sup> | 195 | — | — |
| BI05710 | F: M13-AGAACTTTGAAATGGGTTTACCA | 60 | (GA)<sub>3</sub> | 3 1 | 178–184 | — | — |
| BI4848 | F: M13-CACCACCCCTCTCAAGAAGCAA | 59 | (AG)<sub>5</sub> | 4 2 | 71–74 | arabinogalactan protein | 6E-07 |
| BC402 | F: M13-TTACTGAGCTGTAAGGAGGC | 60 | (AT)<sub>5</sub> | 1<sup>M</sup> 1<sup>M</sup> | 92 | bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein | 3E-09 |
| BC932 | F: M13-AGATGCTTCTCACTGACTCCAT | 60 | (GA)<sub>6</sub> | 2 1 | 660–662† | cysteine proteinase superfamily protein | 0.000001 |

<sup>a</sup> Multiplexed loci

<sup>b</sup> Fluorescent dye

<sup>c</sup> Repeat motif

<sup>d</sup> Allele sizes (bp)

<sup>e</sup> Putative function

<sup>f</sup> Putative function

† Cysteine proteinase superfamily protein.

TABLE 1. Characterization of nuclear microsatellites for Central American *Begonia* species.
TABLE 1. Continued.

| Locus | Primer sequences (5′–3′) a | Multiplex b | Fluorescent dye | T° C) Repeat motif c | Sizes (bp) d | Putative function e |
|-------|--------------------------|-------------|-----------------|---------------------|-------------|---------------------|
| BI3069 | F: M13-AACCACAGTAATCATCCGGC | 60 (CA) 5 1 1 | 184–192 | — | — | — |
|       | R: TGTCCGGTAACTGTGGTGAA | 60 (AGG) 5 MP MP | — | — | — | — |
| BI5174 | F: M13-GTCGCAGGGTTTGTCTAGGA | 60 (CTT) 5 1 1 | 118–121 | stromal cell-derived factor 2-like protein precursor | — | — |
|       | R: GGAAATCAGAGTGCTGGCTC | 59 (TC) 6 3 2 | 148–164 | — | — | — |
|       | R: GAAGGAGATGATTATGACGAA | 59 (TG) 6 3 2 | 147–173 | — | — | — |
| BI7247 | F: M13-CTCTTATTCCGCGTCAAAGC | 60 (GA) 5 2 1 | 178–180 | — | — | — |
|       | R: AGCGGAGAAGTCGAAAACAG | — | — | — | — | — |
| BC312  | F: M13-ATTTCCTTCTGCGAACGATG | 60 (GA) 5 2 1 | 178–180 | — | — | — |
|       | R: ATCGGAACTCTGAGCCTGAA | — | — | — | — | — |

Note A

A B. heracleifolia; MP = multiple PCR products amplified; nel = B. nelumbiifolia; Tm = primer melting temperature when amplified individually.

a M13 sequence is: CACGACGTTGTAAAACGAC.
b Multiplex to which the primer was assigned.
c Repeat motif in

d The observed range of PCR product sizes excluding the M13 motif.
e Putative function in Arabidopsis.

CONCLUSIONS

We have described the development of nuclear microsatellite primers that amplify in two divergent Central American Begonia species. Some of the primers have exact BLAST matches in the transcriptome of the Southeast Asian species B. venusta, including both the forward and reverse primers for loci BI3348, BC932, and BC552.

Table 2. Genetic diversity in population samples of Begonia heracleifolia and B. nelumbiifolia.

| Locus | A | Hs | He |
|-------|-------|-------|-------|
| BI4329 | 3 | 0.400 | 0.524 |
| BEI0340 | 4 | 0.000 | 0.444 |
| BEC432 | 2 | 0.100 | 0.097 |
| BEC344 | 1 | — | — |
| BEI6278 | 1 | — | — |
| BEI5347 | 3 | 0.300 | 0.449 |
| BEC552 | 1 | — | — |
| BEI3348 | 4 | 0.579 | 0.604 |
| BEI06534 | 5 | 0.500 | 0.750 |
| BEI7112 | 2 | 0.400 | 0.467 |
| BEI3820 | 5 | 0.600 | 0.623 |
| BEC134 | 4 | 0.611 | 0.732 |
| BEI04004 | 2 | 0.059 | 0.059 |
| BIC362 | 2 | 0.050 | 0.050 |
| BEC332 | 4 | 0.250 | 0.483 |
| Mean | 3.333 | 0.321 | 0.440 |
| SD | 1.155 | 0.228 | 0.246 |

Note: A = number of alleles per locus; Hs = total alleles observed in the two species; He = expected heterozygosity; Ht = observed heterozygosity.
multiplexed assay of 15 loci should enable accurate assignment to hybrid classes (e.g., F1, backcross). Future studies will use these loci to estimate the genetic structure of populations, the frequency of hybrids, and the extent of introgression in hybrid swarms.

LITERATURE CITED

Brennan, A. C., S. Bridgett, M. S. Ali, N. Harrison, A. Matthews, J. Pellicer, A. D. Twyford, and C. A. Kinder. 2012. Genomic resources for evolutionary studies in the large, tropical genus Begonia. *Tropical Plant Biology* 5: 261–267.

Burt-Uhtley, K. 1985. A revision of the Central American species of Begonia section Gireoudia (Begoniaceae). *Tulane Studies in Zoology and Botany* 25: 1–131.

Dewitte, A., A. D. Twyford, D. C. Thomas, C. A. Kinder, and J. van Huylenbroeck. 2011. The origin of diversity in *Begonia*: Genome dynamism, population processes and phylogenetic patterns. In O. Grillo and G. Venora [eds.], *The dynamical processes of biodiversity: Case studies of evolution and spatial distribution*. InTech Press, New York, New York, USA.

Goodall-Copestake, W., S. Perez-Espona, D. J. Harris, and P. M. Hollingsworth. 2010. The early evolution of the mega-diverse genus *Begonia* (Begoniaceae) inferred from organelle DNA phylogenies. *Biological Journal of the Linnean Society* 101: 243–250.

Hughes, M., and P. M. Hollingsworth. 2008. Population genetic divergence corresponds with species-level biodiversity patterns in the large genus *Begonia*. *Molecular Ecology* 17: 2643–2651.

Lepais, O., and C. F. E. Bacles. 2011. Comparison of random and SSR-enriched shotgun pyrosequencing for microsatellite discovery and single multiplex PCR optimization in *Acacia harpophylla*. F. Muell. ex Benth. *Molecular Ecology Resources* 11: 711–724.

Mieglecz, E., C. Coste-Doat, V. Dubut, A. Gilles, T. Malaua, N. Pich, and J.-F. Martin. 2010. QDD: A user-friendly program to select microsatellite markers and design primers from large sequencing projects. *Bioinformatics (Oxford, England)* 26: 403–404.

Schuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18: 233–234.

Thomas, D. C., M. Hughes, T. Phuthai, W. H. Ardi, S. Rajhandary, R. Rubite, A. D. Twyford, J. E. Richardson. 2011. West to east dispersal and subsequent rapid diversification of the mega-diverse genus *Begonia* (Begoniaceae) in the Malesian archipelago. *Journal of Biogeography* 39: 1365–1399.

Twyford, A. D., C. A. Kinder, N. Harrison, and R. A. Ennos. 2013. Population history and seed dispersal in widespread Central American *Begonia* species (Begoniaceae) inferred from plastome-derived microsatellite markers. *Botanical Journal of the Linnean Society* 171: 260–276.
**Appendix 2. Continued.**

| Locus | Forward primer sequence (5’–3’) | Reverse primer sequence (5’–3’) | Repeat motif |
|-------|----------------------------------|---------------------------------|-------------|
| BI1733 | GTTCCAACCTCCAATGCTGTTT | CGAGTTTGGCTTTCGAAATCTTG | (GCCACA)_5 |
| BI1816 | GTTGTGCGGTTGAGGGTTGTT | CAAATGCTATCTCCAGCTATGG | (GAT)_1 |
| BI1937 | TCATTCTCCGAGCGAGGAA | GGAAGGTGAGGGAGG | (GGA)_5 |
| BI1948 | CAAGACCGTGGAGGAGGGA | CTGAGTGGAGGAGG | (TA)_5 |
| BI2413 | GAAGGGAGGACGGCTCAG | CGAGAGTGGCTTACATCC | (AGA)_5 |
| BI2675 | TTGCATATACCTCAGCCG | GGTTCTTCTCAGGGTAC | (GA)_5 |
| BI2875 | CCAATCTCCGATGTGTTGC | AGTGGAGACGGCTCAG | (TC)_5 |
| BI2935 | TGAGAAAGGTGGTCTCATATAAGTCA | CATGTGTCGTTGCCATTTT | (CA)_5 |
| BI2946 | ATTTGAGCAACCTAGGTTGCT | AAGGCTGAGGGGAGG | (GAA)_5 |
| BI2961 | TGCCGGAAGGAGGAGGAG | GAGGAGGAGGAGGAG | (AG)_5 |
| BI2967 | TGCGGCGGTTGAGGTAGGAT | TGGAGGAGGAGGAGGAG | (AGA)_5 |
| BI2994 | GATCCATCGGGACGAGAAGA | AAAACATCGGAGGACGAGA | (CT)_5 |
| BI3043* | CGACCATCGGACAACAACTG | CGGATTAGGAGGAGG | (TC)_5 |
| BI3069* | AACAGCATATCAGTCCGACG | TGTCGGAATGCTTGTTGAA | (CA)_5 |
| BI3131 | ACAAATGCCTGCAACGGG | GAAGCTGAGGGGGGAGG | (AAAG)_5 |
| BI3153 | ATGAAGAGGTAGGAGGAGG | GGTGGCTTGTACGGTGAGAT | (TC)_5 |
| BI3286 | TCTCATGAGCTATCCGAGCA | AACGCTGAGGGGAGG | (GAA)_5 |
| BI3301 | GCATTGAGATTGCGCAAGAT | TGTAGTCTGGAGGAGG | (TA)_5 |
| BI3348* | ACTGTGCTTCTGTTGGAGGC | CTGACGGCCGATAGCTTAAA | (CT)_5 |
| BI3377* | AAACACATCATACGCGGAG | AGAACGAGTTAAGTACG | (AGA)_5 |
| BI3384 | ATAACTGTCTGATGTGAGGG | GCTTGTGCTGCTGAGG | (TC)_5 |
| BI3403 | TGTTAGAGGAAGGAGGAGG | GCTGAGTGGCTTACATCC | (AGA)_5 |
| BI3519 | TCCAGAGGGCTTTGTTGTT | AGCAGTATCGGAGACGAT | (CT)_5 |
| BI3865 | ACCCTCACTACATGGCTGTA | TACAGTACGTTTGCGA | (CT)_5 |
| BI3970 | GTGTTGCTGCTCTGTTCT | TCCCTACGCTGAGGCA | (GAA)_5 |
| BI4004* | CGGAGAATTCCTTGAGGGG | TCCCTATCGTGAAGCTAC | (AT)_5 |
| BI4013 | AGGCCAGATACACCCAAGGG | CGGTTGCTCTCTTCTTCTT | (TA)_5 |
| BI4021 | TGGTGGCTGGCTGCGAGTGA | GGAACCTTCTGAGCCTCA | (AT)_5 |
| BI4028 | GTCTTCTCCCATGCTTGAAG | GGCTTGGAGGCCAATCCT | (CT)_5 |
| BI4031 | TCTTCTGCTGTCAGGTTGCT | AATAAGTGCACAACTGAGG | (TC)_5 |
| BI4088 | GTTTGCTGACATGAGGCTC | TTGAGAAATATCCCTCCC | (GGC)_5 |
| BI4128 | AGAACAGAGTTGCCAAGGG | AGGAGCACTGCAAGGAGG | (AGA)_5 |
| BI4166 | CGGGCAAAATGTGAGGG | AAAATGAGGAGGAGG | (TA)_5 |
| BI4175 | GCCAATCAAAGGGTGATTTA | CAGATTGATCTTGAGC | (TA)_5 |
| BI4233 | ATGCAGCTCAAGTCAGG | CAAGTCTGGTGCTAGG | (TG)_5 |
| BI4279 | GCAGAGACGAGGAGGAG | CAGATGACGAGGAGG | (AG)_5 |
| BI4329* | CAACAACATCAAGGGAGCT | CATGAGGATCAAGGGAG | (AGA)_5 |
| BI4360 | CCAGGACCTCCTCCATGAGA | TATTCGCTCAATCCT | (TG)_5 |
| BI4477* | GGATCTGCTGTCCTGTCG | GGGCCAGACGAGAGAAATTTG | (CT)_5 |
| BI4594 | CGGAGATGACGAGGAGG | CGTTAGTCACTTGAG | (GT)_5 |
| BI4600 | GCTATGGGAAATGGTGAGA | AGCCTCTCCTCTCCTTG | (AGA)_5 |
| BI4641 | GCCAGACATTGTCCTGCTGAT | GTGCAAACCCGGAGCATGA | (CT)_5 |
| BI4721 | ACTACCTCTCCAAAGGCTGTT | GCCGAAATGCACTTACG | (TC)_5 |
| BI4740 | AGGCCACCTCCAAGAATGAAT | GCCCTGATCTGCAATGGGA | (GA)_5 |
| BI4976 | GTCGAGACGAGGAGG | TGAGCTTAATGCGAGG | (AGA)_5 |
| BI4779 | CGAAGGGAGGAGGAGGAT | TGCCACATATTACCACT | (AAGC)_5 |
| BI4793 | CAGTCCGCTGCTACCTTC | GGAAGGCACTGCTGAGG | (GAA)_5 |
| BI4804 | TGCGCTGATGTTGTTGGA | AGGAGACGAGAGAATGAG | (TA)_5 |
| BI4848* | GACAGCTCCTCAGAAGGAA | GACGTTGAGGTCTCTCAG | (AT)_5 |
| BI4899 | CCAATGGTTTCTCCTAAAACT | GAGTGGAGGACGACGACT | (GA)_5 |
| BI4987 | AGTGAACACTGTTGCAAGCC | ACCCTTTTCTTACCTCAGG | (GGA)_5 |
| BI5091 | TGCGCTTGAGGGTTGGAAT | GCCAGGCTGGAATTTTGTG | (AGA)_5 |
| BI5107 | CGGGTTTATACAGGCTAGGA | GATGTGGAAGATTGATGTA | (AT)_5 |
| BI5115 | AGACGGACGAGGAGGAAA | TCCCTGCTTCTACAGGTTG | (TA)_5 |
| BI5162 | CTGAGACGAGGAGGAGG | GCTGACGAGGAGGAGG | (AGA)_5 |
| BI5174* | GTCGAGGTTTTCTTCTAGGA | GGGTACAGGGTCTGAGC | (CTT)_5 |
| BI5285 | GGTCAATAGGTTACCTGAA | CTTGCTATCTCGTCGTCAT | (GGT)_5 |
| BI5317 | GCCCTCAATGCTTCTCATCAT | GGGACCACGATATTCCCAT | (AT)_5 |
| BI5325 | TGGGCTGAGGAAAATAAGGT | GGTAGTAGGTGTTGAAGT | (TC)_5 |
| BI5347* | TCACTGTGTTTCTATGCAAGCC | CTGCTATATGCGAAGATGAC | (CTT)_5 |
| BI5377* | ATGGGAGGAGGAGGAGG | GGTGGAGGAGGAGGAG | (AGA)_5 |
| BI5414 | GCGAAGGGAGGAGGAGGAA | GCCAGCAGTCTGAGCTTAA | (AT)_5 |
| BI5423 | GCTTCCATGATGCAAAACT | GAGAAGGCGGAGGAGG | (AGA)_5 |

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

5 of 6
## APPENDIX 2. Continued.

| Locus     | Forward primer sequence (5′–3′)                  | Reverse primer sequence (5′–3′)                  | Repeat motif |
|-----------|-------------------------------------------------|-------------------------------------------------|--------------|
| BI5561    | GTGAGCTCTGCTCCCTCCTCAGTGGG                        | GTTCGTTTCTCAGGTAATCTTC                           | (CTT)$_3$    |
| BI5588    | CAGCTGTTGAGAAACCGTGA                             | AATCTATCGCTCAAGGATCAA                            | (TC)$_5$     |
| BI5593    | ACTCCAATTGGTCGTCGTCGTC                           | AGATAGGAAGGAAAGGAGA                              | (AG)$_6$     |
| BI5638    | GCTTCTTCCCTCCCTCCTTCCC                          | TTACGSCCTCCAGATCTCTGCT                           | (TCT)$_7$    |
| BI5668    | ACTCTGTTGAGAAACCCGTCAGATCTTC                     | AGGAGGAGGAGGAGGAG                              | (AGC)$_8$    |
| BI5710    | GAAATTTTTGGAGGAAGCCC                            | TGGAGATGATGACAGAAGGTTATGTC                      | (GA)$_7$     |
| BI5800    | CGGCTCCCATATCTCCTGAAA                           | GGAAGCTGAGTGGTTGTTGCT                          | (TCT)$_3$    |
| BI5813    | CGGTTAGTGATGAAAGGAGA                             | AGCATGCGCTCAAGGATGGCT                           | (AG)$_6$     |
| BI6067    | CATCTGGAGGAATCTCAGAGCC                          | AGGCCGGAATAGCAAGGATGAC                         | (TA)$_8$     |
| BI6141    | GTGCCATGACGATAAGGTTT                            | TCTGACCATGAGAGTGAGC                            | (AG)$_8$     |
| BI6227    | GAGGCCGGGGAATGAGAAGGAGA                         | ATAGCATGGAGGAGGAGAA                              | (TCT)$_6$    |
| BI6278    | TGAGATGATGAGTGAGGAGAAC                          | CATGATGAGGAGGAGGAGAA                            | (TCC)$_7$    |
| BI6294    | TGCTGCTCTGCTAGCTTTTAATC                         | TGGGACCTGACTCCTTTACC                           | (AT)$_6$     |
| BI6299    | CATGCCTCTTATGAGCTCTACT                          | CTTGAGACCCTGACTTACCA                            | (AT)$_3$     |
| BI6399    | CGTCATCAGCTCAGCTCAACATCA                        | CGAGAGAAAGGAGGAGGAGTAC                        | (TC)$_3$     |
| BI6422    | TTTAGTTGGAGAGTATTAGTGAAGA                       | AGGCGGAATACGTTCTGCT                           | (TTC)$_3$    |
| BI6423    | ATATGGAGAGATCCGAGGAGGAGA                       | CTGAGGAGGAGGAGGAGGAGA                          | (AG)$_6$     |
| BI6469    | TCTAGTGGCCTCCAGAAGGAGA                          | CTTCCCTACCTATGCGGAT                           | (GA)$_7$     |
| BI6534    | CGTCTGTCTGCTCTCTACCT                            | AGATACAGCCACCGGATATC                           | (TC)$_5$     |
| BI6535    | AAAAGGGGAAAGCAGGAAAGA                           | GGGGATGAGGAGGAGGAGGAGA                         | (GAA)$_7$    |
| BI6561    | CCTCTGAGACTGCTGACCGCC                          | TAGCTGCGCCTCAAAACCC                            | (GTA)$_3$    |
| BI6581    | TTGCTTTTCTCTTCAGTCTC                           | CGGATCTCCAGCTATCAC                             | (CT)$_6$     |
| BI6604    | ATTTTCCACAGAAGAGGCC                             | GGCAGAAGCCGAGATATAC                           | (TA)$_8$     |
| BI6605    | TCAAAAGGCTGTCCCTCCATT                           | GGAAGGAGCAAGATGCTGAGG                          | (TTC)$_3$    |
| BI6701    | AGATACCCCAGACTCAGGCC                           | GAGATGATGAGGAGGAGCC                            | (GA)$_3$     |
| BI6717    | GATCTGCGTGGATGATTTTTAT                           | AGTCCGACAGCAGGATGTAAC                          | (CT)$_8$     |
| BI6761    | GTGTTTCTTCTGCTCCATC                             | ACATGCTCTCTGCTGCTT                            | (TC)$_6$     |
| BI6776    | CCAACAGCAGAACACTGCACTGCA                       | GCTGCCAGACCTGCTGAAGTAC                        | (AG)$_5$     |
| BI6828    | TGGCTCTTCTCTTGTCTTCTC                          | GGTGCTGCTGCTGCTTC                              | (CT)$_3$     |
| BI6849    | CCTGAGACTGAGAGGAGGAGA                          | GCCCTTTCTCTTTAAGTTAC                          | (TA)$_5$     |
| BI6886    | TTCCTCAGCCTGCCTCATT                             | TGGAAATCAGGAAAGAAGC                            | (CT)$_5$     |
| BI6901    | CGAAGTGGAGAAGAAGACTCATAAC                      | GTCGAGGGACGGGAGATGATG                         | (AG)$_6$     |
| BI6984    | GTATGGCAAGGAGAGGAGGCC                          | TTGTCAATCTCCATGACAGA                           | (TC)$_6$     |
| BI7015    | TGGCAGCATATGATCAGCAGC                          | TCTCCTGGGCTCACGTCGAC                           | (GAA)$_5$    |
| BI7023    | TTAAGGCGGTTGAGGAGA                              | CTTTCTGCTGCAADGATG                             | (GAA)$_5$    |
| BI7036    | TTGAGGAGGAGGAGGAGGAGGAGA                        | ATGCTGAGGAGGAGGAGGAGA                          | (CT)$_3$     |
| BI7059    | CTCTCCGCGCCTCCGCTCATAAC                        | TGGCTCTTCTGCGGAGTTT                           | (CT)$_3$     |
| BI7085    | ACTGCGAGAATACGCTCCGAA                          | CACCTCTCTAGCTGCTCCTC                          | (GA)$_3$     |
| BI7112    | ATCCAAATGTCACCTCCCTC                            | GTGATAGTATGTCCTGCTT                            | (TTC)$_6$    |
| BI7149    | CCGAGAATGTCACCTCCGAGT                          | CCTGGAAGGAGGAGGAGGAGA                          | (CT)$_3$     |
| BI7165    | AATGCGACGACGACCTGCTTT                          | GAGAATTTTTTTGGACGCTGCTGA                        | (AG)$_3$     |
| BI7247    | CAGGAGGCAAGGAGGAGGAGGAGA                       | AGGAGGAGGAGGAGGAGGAGA                          | (AG)$_3$     |
| BI7287    | TGGGAGCAGAAACAGGAGA                            | CAGTGCTTTCTCTACACACACACACAC                   | (TGA)$_5$    |

*Indicates markers tested for amplification and polymorphism.