Applications in Plant Sciences

Applications in Plant Sciences 2013 1(5): 1200499

Primer Note

Development and characterization of microsatellite markers for Central American *Begonia* sect. *Gireoudia* (Begoniaceae)[1]

Alex D. Twyford[2,3,5], Richard A. Ennos[4], and Catherine A. Kidner[2,3]

[2]Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh EH3 5LR, United Kingdom; [3]Institute of Molecular Plant Sciences, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JH, United Kingdom; and [4]Institute 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 species of *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.

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.

[1]Manuscript received 21 September 2012; revision accepted 10 November 2012.

The authors thank H. Domínguez, K. Gardner, M. Hughes, and A. Sanchez for help with fieldwork, and A. Brennan, R. Crichton, M. Hollingsworth, R. McGregor, and M. Ruhsam for help and advice in the laboratory. This work was funded by a BBSRC studentship to A.D.T.

[5]Author for correspondence: adtwyfor@syr.edu

doi:10.3732/apps.1200499

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

Applications in Plant Sciences 2013 1(5): 1200499
doi:10.3732/apps.1200499

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

Locus	Primer sequences (5′–3′)[a]	Multiplex[b]	Fluorescent dye	T_m (°C)	Repeat motif[c]	*A* her	*A* nel	Allele sizes (bp)[d]	Putative function[e]	*E*-value
Multiplexed loci										
BI4329	F: M13-CAACCAACAATGGCAGCTT R: CATGGAGATAATGGAGCTGG	1	FAM	59	$(GGA)_6$	4	2	89–104	immunoglobulin E-set superfamily protein	2E-13
BI3043	F: M13-CGACATCCAACCAAACCTG R: TTGATAGATGGAAGGGTCGC	1	FAM	60	$(TC)_5$	1	2	173–179	—	—
BC432	F: M13-AAACTCCGATGGATTCAGCA R: TGAAATAAACACACAAACAAAGACA	1	FAM	60	$(TG)_5$	1	1	261–263	endotrans glucosylase/hydrolase	5E-18
BC344	F: M13-GAGGGAGGGTCCCTTGTTAG R: CCGTCTTACGTTGCATCATC	1	VIC	60	$(GCA)_5$	1	1	105–108	chitinase-like protein	3E-07
BI6278	F: M13-TGTAGTTGTTGTAGTAGCAGAACTTTG R: CAGATGGGTCGGAGATTTTG	1	VIC	59	$(TCC)_7$	1	3	238–253	DOF zinc finger protein	3E-25
BI5347	F: M13-TCAGTCCATTTTCTTAATCAGACC R: CTCTATCATTTCCAAGCGATTTC	1	VIC	59	$(CTT)_6$	2	1	171–183	unknown gene	0.000002
BC552	F: M13-TGTCTGAGATGGAAACTGCG R: TAGTCGAAGGGATCCGAATG	1	NED	60	$(GT)_5$	2	2	271–273	—	—
BI3348	F: M13-ACTTGTTTCTCGTTGGGAGC R: CTGCAGCCCAGTGGATTTAC	1	PET	60	$(CT)_6$	3	3	279–283	—	—
BI06534	F: M13-CGTTGCTCTGCTCTAACCCT R: AGATACAGCCAACCGGATTC	1	PET	59	$(TC)_6$	6	2	97–107	sterol 4-alpha-methyl-oxidase 2-1	7E-57
BI7112	F: M13-ATCCAATGTCAACCTCTCGG R: GTGCATTAGAGTCCCGTGGT	2	FAM	60	$(TCC)_6$	2	2	109–115	—	—
BI3820	F: M13-AGGACCAGTTTTGACGGCTA R: GAAGCTTTTGCTCTTCTGTTGA	2	FAM	59	$(CTT)_7$	5	2	158–176	LOB domain-containing protein	2E-39
BI134	F: M13-ATCAGCTCACTCCCTATCCTCT R: TGCAATCTCCTTCGGTTCTT	2	VIC	60	$(CT)_6$	4	2	306–314	—	—
BI4004	F: M13-TCAGGAAATATTCGATTGGGA R: GCATTCCTCTGTGTACAATGC	2	VIC	59	$(AT)_5$	2	3	155–169	O-fucosyltransferase family protein	1E-32
BI362	F: M13-CTTCACCTCGCCTGAACAAC R: GAGGCGAAATATTATGCGGA	2	NED	60	$(ATG)_6$	4	4	147–159	Acyl-CoA N-acyltransferases (NAT) superfamily protein	1.00E-45
BC332	F: M13-GAACCAGAAGTCAAGGGTTCA R: AAACATGATTTTCCTCATCCAA	2	PET	59	$(TCA)_5$	4	2	188–200	ATPase	1.00E-122
Additional loci tested										
BC672	F: M13-CCTTGATCGAGAAAGAACCG R: AAAGCCAGCTCCTTCCTGTA			60	$(CTT)_8$	3	1	152–158	cellulose-synthase-like C12	2E-57
BI4477	F: M13-GGATCTCCTCTGCTTTGCTG R: GGCGAGACCAGAAGAAAAGTT			60	$(CT)_9$	4	2	111–119	—	—
BI06604	F: M13-ATTTTTCCACAGAAGAGCCC R: GGCAGAACCCGCAGTATATC			59	$(AT)_8$	6	1	111–127	—	—
BI6294	F: M13-TGCTGGTCTGAATCTTTAATCA R: TGGGGTCTTGGTACTCTTTCC			59	$(AT)_{10}$	1[M]	1[M]	148	catalytic LigB subunit of aromatic ring-opening dioxygenase family	3E-13
BI6701	F: M13-AGAATCCCCACTCACTGCAC R: GAGATGATGAGGGTTCAGGC			60	$(GA)_6$	1[M]	1[M]	195	—	—
BI05710	F: M13-GAAAGTTTTGGAGGAAGCCC R: TGGAAGAGATCAGAAGGTACA			60	$(GAA)_7$	3	1	178–184	—	—
BI4848	F: M13-CGACGCCTCTCAAAGAAGAA R: GAGCTTTGAATTTCGCTACG			59	$(AG)_6$	4	2	71–74	arabinogalactan protein	6E-07
BC402	F: M13-TTACTCGAGCTAGAAGCCGC R: AGGGCTTGGAGAGCTAGAGG			60	$(AT)_5$	1[M]	1[M]	92	bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	3E-09
BC932	F: M13-GTAGTCCATCAGTCCGCCAT R: GAGTGATGAAGGCGAAGAGG			60	$(GA)_5$	2	1	660–662†	cysteine proteinase superfamily protein	0.000001

TABLE 1. Continued.

Locus	Primer sequences (5′–3′)[a]	Multiplex[b]	Fluorescent dye	T_m (°C)	Repeat motif[c]	*A* her	*A* nel	Allele sizes (bp)[d]	Putative function[e]	*E*-value
BI3069	F: M13-AACCACAGTAATCATCCGGC			60	$(CA)_5$	1	1	184–192	—	—
	R: TGTCCGGTAACTGTGGTGAA									
BI3377	F: M13-AACACAATCATCAGCCGGAC			60	$(AGG)_5$	MP	MP		—	—
	R: GAAGGAGATGATTATGACGAA									
BI5174	F: M13-GTCGCAGGGTTTGTCTAGGA			60	$(CTT)_5$	1	1	118–121	stromal cell-derived factor 2-like protein precursor	8E-07
	R: GGAAATCAGAGTGCTGGCTC									
BC42	F: M13-GCTATGCAGGTTCTGGTGGT			59	$(TGG)_6$	3	2	147–173	—	—
	R: ACTGGTTGTCACTACTGCCG									
BI6984	F: M13-GAAGGGGTTTCTTGGTCTCA			59	$(TC)_6$	3	2	148–164	—	—
	R: TTGTCAATTCTCACCAGACACA									
BI7247	F: M13-CTCTTATTCCGCGTCAAAGC			60	$(AG)_6$	1[M]	1[M]	135	—	—
	R: AGCGGAGAAGTCGAAAACAG									
BC312	F: M13-ATTTCCTTCTGCGAACGATG			60	$(GA)_5$	2	1	178–180	—	—
	R: ATCGGAACTCTGAGCCTGAA									

Note: *A* = number of alleles per locus; her = *B. heracleifolia*; MP = multiple PCR products amplified; nel = *B. nelumbiifolia*; T_m = primer melting temperature when amplified individually.
[a] M13 sequence is: CACGACGTTGTAAAACGAC.
[b] Multiplex to which the primer was assigned.
[c] Repeat motif in *B. plebeja*.
[d] The observed range of PCR product sizes excluding the M13 motif.
[e] Putative function in *Arabidopsis*.
[M] Monomorphic in all individuals tested.
† Large product size assumed to be caused by an intron.

unknown PCR inhibitor that coelutes with DNA extractions in *Begonia*, extractions were diluted 100-fold with Millipore dH_2O to a final DNA concentration of ~0.1–1.5 µg/mL. PCR reactions were performed using the M13-tailed primer method (Schuelke, 2000) in a final reaction volume of 10 µL containing: 0.5 µL of 1 mM M13-tailed forward primer (Invitrogen, Grand Island, New York, USA), 1 µL reverse primer (1 mM), 1 µL of 1 mM M13 fluorescently modified primer (6-FAM,VIC, NED, PET), 0.25 µL bovine serum albumin (BSA, 0.4%), 1 µL of 10× reaction buffer, 1 µL of 2 mM dNTPs, 0.6 µL of 25 mM $MgCl_2$, 0.05 µL BIOTAQ polymerase (Bioline, London, United Kingdom), 1 µL dilute DNA template, and made up to the final volume using dH_2O. PCR cycles consisted of an initial denaturation of 1 min at 95°C, followed by 40 cycles of denaturation for 1 min at 95°C, annealing for 1 min at 57°C, and extension for 1 min at 72°C. Five microliters of each PCR product labeled with the four fluorescent dye colors was pooled and diluted 2× in Millipore dH_2O, and the GeneScan 500 LIZ internal size standard (Applied Biosystems, Foster City, California, USA) was added prior to fragment analysis on the ABI 3730xl analyzer (Applied Biosystems; analysis was performed at GenePool, University of Edinburgh, Edinburgh, United Kingdom). Fluorescent traces were analyzed automatically with manual editing using GeneMapper version 4.0 (Applied Biosystems).

A total of 136 primer pairs were located in the *B. plebeja* transcriptome using the QDD bioinformatic pipeline (Appendix 2). All 31 of the subset of primers tested for amplification yielded a PCR product (Table 1). Sixteen loci had a significant (<E-5) BLAST match in the TAIR database (Table 1). Of these loci, four loci were monomorphic (BI6701, BC402, BI6294, and BI7247) and one amplified multiple PCR products (BI3377). Two PCR multiplex reactions were designed to amplify a total of 15 polymorphic loci (Table 1). All loci were polymorphic in at least one of the populations tested, and showed moderate genetic diversity, with the number of alleles per species ranging from one to five and the expected within-population heterozygosity between 0 and 0.75 (Table 2). Twenty-one of the 62 primers (34%) had perfect BLAST matches in the transcriptome of the divergent *B. venusta*, including both the forward and reverse primers for loci BI3348, BC932, and BC552.

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* and, therefore, may be transferable more widely across the genus. The transferability of markers is important for the study of natural hybrids, and the development of a

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

	B. heracleifolia			*B. nelumbiifolia*			
Locus	*A*	H_o	H_e	*A*	H_o	H_e	A_t
BEI4329	3	0.400	0.524	3	0.500	0.537	5
BEI03043	4	0.000	0.444	3	0.500	0.630	4
BEC432	2	0.100	0.097	2	0.000	0.097	3
BEC344	1	—	—	2	0.000	0.097	2
BEI6278	1	—	—	3	0.353	0.668	4
BEI5347	3	0.300	0.449	1	—	—	4
BEC552	1	—	—	3	0.050	0.229	3
BEI3348	4	0.579	0.604	4	0.500	0.665	5
BEI06534	5	0.500	0.750	4	0.105	0.201	7
BEI7112	2	0.400	0.467	3	0.278	0.522	4
BEI3820	5	0.600	0.623	2	0.000	0.108	6
BEC134	4	0.611	0.732	3	0.050	0.145	5
BEI04004	2	0.059	0.059	3	0.188	0.623	4
BIC362	2	0.050	0.050	2	0.000	0.097	2
BEC332	4	0.250	0.483	3	0.154	0.495	5
Mean	3.333	0.321	0.440	2.857	0.191	0.365	4
SD	1.155	0.228	0.246	0.663	0.199	0.243	1.327

Note: *A* = number of alleles per locus; A_t = total alleles observed in the two species; H_e = expected heterozygosity; H_o = 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. Kidner. 2012. Genomic resources for evolutionary studies in the large, diverse, tropical genus, *Begonia*. *Tropical Plant Biology* 5: 261–276.

Burt-Utley, K. 1985. A revision of 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. Kidner, 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. Pérez-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.

Meglécz, E., C. Costedoat, V. Dubut, A. Gilles, T. Malausa, N. Pech, 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. Phutthai, W. H. Ardi, S. Rajbhandary, 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: 1365–2699.

Twyford, A. D., C. A. Kidner, 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 1. Information on Mexican *Begonia* voucher specimens deposited in the herbarium at the Royal Botanic Garden Edinburgh (E). Information presented: taxon, collection number, collection locality, GPS coordinates.

Begonia heracleifolia Cham. & Schltdl.: AT48, San Andrés Tuxtla, 18.47850, –95.17802; AT244, Agua Azul, 17.22117, –92.11073; AT375, Ocozocoautla, 16.90533, –93.45153; AT505, Berriozábal, 16.86693, –93.32781; AT819, Santa María Jacatepec, 17.85819, –96.21853; AT922, Motzorongo, 18.66953, –96.78714; AT1080, Santa María Xanabi, 15.98808, –96.11061.

Begonia nelumbiifolia Cham. & Schltdl.: AT28, Los Tuxtlas, 18.59026, –95.07876; AT125, San Andrés Tuxtla, 18.50660, –95.16607; AT619, Huatusco, 19.20367, –96.74256; AT683, Josaa, 16.01419, –96.11289; AT771, Los Cantiles, 17.74356, –96.32803; AT958, Motzorongo, 18.66953, –96.78714; AT1029, San Jeronimo Zoochina, 17.22117, –95.23547.

Appendix 2. Microsatellite loci in the transcriptome of *Begonia plebeja*.

Locus	Forward primer sequence (5′–3′)	Reverse primer sequence (5′–3′)	Repeat motif
BC134*	ATCAGCTCACTCCCTATCCTCT	TGCAATCTCCTTCGGTTCTT	$(CT)_6$
BC192	AAGTCAAACCTGTTGACCCG	ATCCTCATCGGATTCGTCAT	$(GAT)_9$
BC232	TGGAAATGCTGTCGTTGAAT	ATTGGAGAAAAGGCAAAGCA	$(TCT)_8$
BC312*	ATTTCCTTCTGCGAACGATG	ATCGGAACTCTGAGCCTGAA	$(GA)_5$
BC332*	GAACCAGAAGTCAAGGGTTCA	AAACATGATTTTCCTCATCCAA	$(TCA)_5$
BC344*	GAGGGAGGGTCCCTTGTTAG	CCGTCTTACGTTGCATCATC	$(GCA)_5$
BC362*	CTTCACCTCGCCTGAACAAC	GAGGCGAAATATTATGCGGA	$(ATG)_6$
BC42*	GAAGGGGTTTCTTGGTCTCA	TTGTCAATTCTCACCAGACACA	$(TGG)_6$
BC402*	TTACTCGAGCTAGAAGCCGC	AGGGCTTGGAGAGCTAGAGG	$(AT)_5$
BC432*	AAACTCCGATGGATTCAGCA	TTGAAATAAACACACAAACAAAGACA	$(TG)_5$
BC532	TCATTCCGCTTCTATGCTCC	CGTCATCGTCAATATCATCCTC	$(TGA)_6$
BC552*	TGTCTGAGATGGAAACTGCG	TAGTCGAAGGGATCCGAATG	(TG)5
BC602	GCAAAGCAGGTAACTTTTAGCC	ACTCACCGAACTTTGGCAAC	$(CAG)_5$
BC632	CATAGCGCTCAGCTTGCTC	GAGATCTTATACGAGCTACTGGATAGT	$(TC)_9$
BC643	GGAGGAGCTCGGTCATTAGA	AACCACCGGTACCCTCATTT	$(CT)_6$
BC652	TTTCGTCCATGAAGAAAGGC	TCCAGGGAACTCCATCACTC	$(GAA)_5$
BC672*	CCTTGATCGAGAAAGAACCG	AAAGCCAGCTCCTTCCTGTA	$(GAA)_8$
BC692	AACATGGCCGTCACTAGTCC	CAGGCAGACAAAGAAGATTCC	$(AG)_{11}$
BC752	GGCAGATTTTACTGGGACGA	CGCCCATCTATCTGTATCCAA	$(TTC)_5$
BC762	CAACTCTGCAAATGCAAGGA	ACCCATGACAGCATGAACAA	$(CT)_5$
BC932*	GTAGTCCATCAGTCCGCCAT	GAGTGATGAAGGCGAAGAGG	$(GA)_5$
BI0537	CAGATCAACCCTCTTCCTGC	ATCGAAAACCCATTGACTGC	$(CCT)_6$
BI1195	TGCTGCAGAAACTTTAGCCA	CGGTGATTAAAGAAGAGCAAGAA	$(GA)_{11}$
BI1430	CACAATTCGTGAAAACACGG	TTCTGCATGATGTTGGCTTT	$(GA)_5$

Appendix 2. Continued.

Locus	Forward primer sequence (5′–3′)	Reverse primer sequence (5′–3′)	Repeat motif
BI1733	GTTCACCACTCCAATGGCTT	CGAGTTTGCCTTCGAATCTC	$(GCCACA)_5$
BI1816	GTTTTGCGGTTGAGTTTGGT	CAAATGAATCTTCTTCATCCAGTG	$(GAT)_7$
BI1937	TCATCATCGCAGCAGAAGAC	CGAAGCTGGGAGTGAGTTTC	$(GGA)_6$
BI1948	CAAAACTGGCTTTGCAGACA	CACGGGCACTTTCAATTTCT	$(TA)_5$
BI2413	GAATGAAGAGCGAATCGACG	CAGAGCTCCGGAATCTCATC	$(AGA)_5$
BI2675	TTCCATTTACTCTCAGCCGC	CGTTCTCCTTCGAGGACTTG	$(GA)_7$
BI2875	CCCAATCTCCCTGTCTATCG	AAGCTGACGAAGCTCTTCCA	$(TC)_5$
BI2935	TGGAAGAAGGTCTCCATATAAGTCA	CATGTTTTCTTCGCCCATTC	$(CAC)_9$
BI2946	ATTTGAAGCCATTGGGTCTG	AAGACGGGAAAGGGTGAGAG	$(TC)_6$
BI2961	TCGCAAAAGAAGAAATCACAAA	TCCTCCGGCACAATAATCTC	$(GAA)_6$
BI2967	GGTGGCTTGTACGGTGAGAT	TCGATTCTCAAATGCCTTCA	$(GAA)_5$
BI2994	GATTTCCGTGGAGGAAACAA	AAACATCACCAGAGCACAACA	$(CT)_5$
BI3043*	CGACATCCAACCAAACCTG	TTGATAGATGGAAGGGTCGC	$(TC)_5$
BI3069*	AACCACAGTAATCATCCGGC	TGTCCGGTAACTGTGGTGAA	$(CA)_5$
BI3131	ACATTGTGTTCAATGGCGAA	GAGCTCATGCAATGCTTCAA	$(GAA)_6$
BI3233	TATGAAGGACGTGGGAGGAG	GGGAATCAGAAGCCAATCAA	$(GA)_5$
BI3234	AAACAAGGAACGCTCAATCC	GCTCGAGTTGGCTTCATTTC	$(AG)_5$
BI3286	CCTATGATGATAGCGTCCGA	AGGCCGACATTCTTTTCCTT	$(CT)_{10}$
BI3301	GCATGGAGATTGCCAGATTT	CTATTGCTCAGCGGAGAAGG	$(GAA)_5$
BI3348*	ACTTGTTTCTCGTTGGGAGC	CTGCAGCCCAGTGGATTTAC	$(CT)_6$
BI3377*	AACACAATCATCAGCCGGAC	ACGAAGGAGATGATTATGACGAA	$(AGG)_5$
BI3384	ATAATTGGGCTAGGGTTCGG	GCTTTTGGTTGCTTCAGAGG	$(TC)_5$
BI3403	TGTAGGAACAACGGTTAGCG	CGTAGAGACGATTTCCTTAGCC	$(GAA)_6$
BI3519	TTCAGAGCGCTTTTGGTTTT	ACGCACTATGCCGTTCTTCT	$(TA)_6$
BI3553	TCTGAAATAGCACCGCTTCC	TTTCTTCGATGAACGCACTG	$(AAG)_5$
BI3600	CATTATTTCCTGTCGGGACG	TGCTGAAAAGTTGCAGGAAA	$(TGTT)_5$
BI3727	CCTCCACCAGATTTGCTTAAA	AACAGAAACATTTGCCGGTG	$(TC)_{12}$
BI3741	GCAACACAGCTCCTCTTCGT	GGTCGGAATCGTCGAGTAAA	$(CT)_7$
BI3820*	AGGACCAGTTTTGACGGCTA	GAAGCTTTTGCTCTTCTGTTGA	$(CTT)_7$
BI3865	ACCTCACTCAACCGCCATAG	TTCAGCATCTGTTGCAGGAC	$(CT)_5$
BI3970	TGTGTTCACTCAATTCTGCCA	TCCTTCACCTGAGACGACAA	$(TC)_5$
BI4004*	TCAGGAAATATTCGATTGGGA	GCATTCCTCTGTGTACAATGC	$(AT)_5$
BI4013	AAGCCAAGATACCCCAAAGG	CCGCTTGTCCTTTCTTCTTG	$(AG)_5$
BI4021	TGTGTTGCCCTGCAAGTAGA	GGAAACCTTTCAGAGCTCCA	$(AG)_5$
BI4028	GTCTTCTCCCCATCGTTGAA	GGGCTTTGGAAACATCTCCT	$(CT)_{15}$
BI4031	TCTTCGCTCTAAAGGCTTGC	AAATTTCGCCAAACATGGAG	$(TC)_5$
BI4088	GGTTTCGAGATATGGCCTCA	TTGGCAATTTATCCCTCCTC	$(GGC)_5$
BI4128	AAGACAACGCCATTCCAAAC	AGGGACGACCGGAAGTAGAG	$(CT)_5$
BI4166	CGGGACAAATGTTAAGCGAT	CAATAAAGAACTTCCGGCGA	$(TG)_5$
BI4175	GGCGATCAAAGGGTGATTTA	CGATTAGCCTCTTCTCGACG	$(AG)_5$
BI4233	ATGCAGACGTAATCGAAGGC	CAAGTTGGTTGGCAAAGACA	$(AG)_{12}$
BI4279	GGGAGGAAGAGGAAGAAGCA	TCAGATTCAGCGTCATCAGAA	$(AGG)_6$
BI4329*	CAACCAACAATGGCAGCTT	TCATGGAGATAATGGAGCTGG	$(GGA)_6$
BI4360	CCGCAGATCCTCCATTAGAA	TTATGTCCCAAACTCCGCTC	$(TGT)_5$
BI4477*	GGATCTCCTCTGCTTTGCTG	GGCGAGACCAGAAGAAAAGTT	$(CT)_9$
BI4594	CCAGAATCGTGGTCACTTCC	CGTGAATCGAAACTTCTCCC	$(TC)_9$
BI4600	GCTATGGGAAGTTGCTTGGA	AGCTCTTCCTCCCTTTCTGG	$(AGA)_7$
BI4641	GCCACAGTTTTAGCTGTGCTAT	CTGCAACCACGAGGAGTTTA	$(CT)_5$
BI4721	ACTACCCTCCCAAGGCTGTT	GGCCAGAAGTCAAACCTCAA	$(TC)_8$
BI4740	AGGCACCCTCCCAAAGTAAT	GCCTGTATCTGAAATTGGCA	$(GA)_7$
BI4746	GTCGGAGTCAGCGAGGGA	TGATCCTATGCACTCGTGGT	$(AG)_5$
BI4779	CGAAGGAGGAAGAGACGATG	TGGCACTATAATTCCAAGCTCC	$(AACG)_5$
BI4793	CAGTCCCCGCACTAATCTTC	GAAAGACCAGCTTCGTTTGC	$(GA)_5$
BI4804	TCGCTGATGATTTGTTTGGA	AGAATGCCGACGAAATTGAG	$(TCT)_{10}$
BI4848*	CGACGCCTCTCAAAGAAGAA	GAGCTTTGAATTTCGCTACG	$(AG)_6$
BI4899	CCCATTTGCTTCCAAAACAT	GAGTCGAGGAGCAGCACTCT	$(GAA)_5$
BI4987	AGTGAAAACCTTGGCACCAC	ACCCTTTTCCTATTCCACGG	$(GAG)_5$
BI5091	TGCTTTCCAGGTTCATAGGG	GGCAAGCTTGGAACTTTTGT	$(AGA)_7$
BI5107	CGCGTTTTACATGGCTGAAT	CGATTGAAAACCTTGAAGATGA	$(AT)_5$
BI5115	AGACCGATGACCGAACAATC	TCCGTCGTTTCTAACCGTTC	$(TC)_5$
BI5162	CTCTGAAACTCGCTCATCCC	GCTCTTTCCGTCTCATTTGC	$(AGG)_5$
BI5174*	GTCGCAGGGTTTGTCTAGGA	GGAAATCAGAGTGCTGGCTC	$(CTT)_5$
BI5285	GGTCAAATGGGTAACATGCC	CTGGTTCATCATCGCTGCTA	$(GGT)_5$
BI5317	GCCCTCAAGTTCCTCCATCT	GGGACCGTCGATTATTCTCA	$(AT)_5$
BI5325	TTCCGGACTGAAAGAAATGG	CGTGAGTGGAGTGGTGATTG	$(TC)_5$
BI5347*	TCAGTCCATTTTCTTAATCAGACC	CTCTATCATTTCCAAGCGATTTC	$(CTT)_6$
BI5377	ATCCTCTTCCTATCCACCGC	GGGAGACGGTGAAACTCTGA	$(TC)_5$
BI5414	GCAAAGCAAAGCTGAAAACC	GGCCCAGTCTACCTGCAATA	$(AT)_5$
BI5423	GCTTCCAATGATGCAAACCT	GAGAAGCGCAGGAGAGCTTA	$(AG)_5$

Appendix 2. Continued.

Locus	Forward primer sequence (5′–3′)	Reverse primer sequence (5′–3′)	Repeat motif
BI5561	GTTGACTCGTCCTCGTCTCC	GTCGTTTCTGCCGATTCTTC	$(CTT)_5$
BI5588	CAGCTGGTTGAGAAACGTGA	AATCATATCGCCGATCAAGG	$(TC)_5$
BI5593	ACTCCAAATTAGGTGCGTGG	AGATAACGAAGCAAAGCGGA	$(AG)_9$
BI5638	GCTTCTTCGTCCTCTTCTTCC	TTACGGCTCCAGATTCTGCT	$(TCT)_7$
BI5668	TATGGGTCCGGATATGGAAA	AGGAAGAGCTCGAAGAAGCC	$(GCG)_5$
BI5710*	GAAAGTTTTGGAGGAAGCCC	TGGAAGAGATCAGAAGGTACA	$(GAA)_7$
BI5800	CGCCTCCCATATCTCGTAAA	GGAAGGTGATGGTTGTTGCT	$(TCT)_5$
BI5813	CGGTAGATTGAATGGGGAGA	AGCATCGCCTCAAGTTGTCT	$(AG)_5$
BI6067	CAGCTTGGAAAATCAGACCC	AGGGGCGTAAGCATAAAGGT	$(TA)_5$
BI6141	GTCGCCATGACGATAAGGTT	TCTGACCCTGAAGATGGACC	$(AG)_{10}$
BI6227	GACGCGACGAAGATAAGGTA	ATACATCGGAGGGAAGCAAA	$(TCT)_5$
BI6278*	TGTAGTTGTTGTAGTAGCAGAACTTTG	CAGATGGGTCGGAGATTTTG	$(TCC)_7$
BI6294*	TGCTGGTCTGAATCTTTAATCA	TGGGGTCTTGGTACTCTTTCC	$(AT)_{10}$
BI6299	CATCGCTCTATGAAGCTGCTACT	CCTGAGACCCTGCTATTCCA	$(AT)_5$
BI6399	CTGTCATCATCCCCATCACA	CAGTGAGAAATGCAGGGTCA	$(TC)_5$
BI6422	TTTGATGGAGAAGATTAGTGAGAAGA	AGGCGGAATACCTTGTCCTT	$(TTC)_5$
BI6423	ATATTGGACATGCCAGCACA	CATGAAACAAGAACTCTGGAGAA	$(AG)_5$
BI6469	TCTAGGCGCCAAAAGAAAGA	CTCCCTCATCACTTGCGAAT	$(GA)_{13}$
BI6534*	CGTTGCTCTGCTCTAACCCT	AGATACAGCCAACCGGATTC	$(TC)_6$
BI6535	AAAGGGGAAAGCAAGGAAAA	GGGATGGATGGCTGATTAAA	$(GAA)_7$
BI6561	CTTCTGAGACTCGTACCGGC	TAGCTCGGTTCAAAACACCC	$(GTG)_5$
BI6581	TTGCTTTTCCTTTCTCATCCA	CCGATTCCAGCTCTATCAGC	$(TTC)_6$
BI6604*	ATTTTTCCACAGAAGAGCCC	GGCAGAACCCGCAGTATATC	$(TA)_8$
BI6605	TCAAAGCTTCGTTCCCATTC	GGAAAGCGTCAGAGTTGAGG	$(TTC)_5$
BI6701*	AGAATCCCCACTCACTGCAC	GAGATGATGAGGGTTCAGGC	$(GA)_6$
BI6717	GATCTCGGGGATTTGGATTT	ACTGCCATAGCCTCCATCAC	$(GTG)_5$
BI6761	TGTTCTTCCGCTCTCCACTT	ACATGCTCTTCCTGGCTTGT	$(TC)_5$
BI6776	CCAAACAGCAAAACTCTTCG	GTTTTGTGGAAGGGTGGCTA	$(AG)_5$
BI6828	TCGTCTCCTTCTTCGTCTCC	GGTCGTCGCTCTGATTCTTC	$(CTC)_5$
BI6849	CCTCAGATCCAGAGGAAGGG	GCGCCTTTTCCTTTAAGTCC	$(TA)_6$
BI6886	TCTTCTCACGGCTCTCCATT	TGGAAATCAAGGAAAGCACC	$(CTT)_5$
BI6901	CGAACTGGAAGAAGACTACAATCA	GCTGCAGCACGGAGTTTTAG	$(AG)_8$
BI6984*	GTATGCAAAGGAGAGCCGAG	TTGTCAATTCTCACCAGACACA	$(TC)_6$
BI7015	TGGTCCAGATTATGATCAGCC	TCTTCTCCGATTCCGATCAC	$(GAA)_5$
BI7023	TTAAGGCGGTGACACAGAGA	CCTTTCGTCTGCAAATGGAT	$(GAA)_5$
BI7036	TTGAGCAGGCTTCCAAACTT	ATTCGAAGGAAGAAGACGGC	$(CTT)_5$
BI7059	CTCCCTCCGACCTCCATAAC	TAGCCTTCTGCGGAGTGTTT	$(CT)_5$
BI7085	ACTCGCGAATATCTCCGAAA	CACCTCTTCAGCTCGTCTCC	$(GA)_5$
BI7112*	ATCCAATGTCAACCTCTCGG	GTGCATTAGAGTCCCGTGGT	$(TTC)_6$
BI7149	CGGAGAATCGAACCTCTGAT	CCCTGAACGATGGAACTCAT	$(CT)_5$
BI7165	AATGAGCACGAACCTGCTTT	GAGGAATTTGGACCGTCTGA	$(AG)_5$
BI7247*	CTCTTATTCCGCGTCAAAGC	AGCGGAGAAGTCGAAAACAG	$(AG)_6$
BI7287	TTGGGGACAACAAATGATGA	CAGTGCTTTCTTTAACAACGCTT	$(TGA)_5$

* Indicates markers tested for amplification and polymorphism.