Cost-Effective Microsatellite Markers for Banksia integrifolia (Proteaceae)

Authors: Fatemi, Mohammad, Houliston, Gary J., Haddadchi, Azadeh, and Gross, Caroline L.

Source: Applications in Plant Sciences, 1(2)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1200130

Applications in Plant Sciences 2013 1(2): 1200130

PRIMER NOTE

COST-EFFECTIVE MICROSATELLITE MARKERS FOR *BANKSIA INTEGRIFOLIA* (PROTEACEAE)[1]

MOHAMMAD FATEMI[2,4], GARY J. HOULISTON[3], AZADEH HADDADCHI[2], AND CAROLINE L. GROSS[2]

[2]Ecosystem Management, University of New England, Armidale, New South Wales 2351, Australia; and [3]Ecological Genetics, Landcare Research, Gerald Street, Lincoln 7640, New Zealand

- *Premise of the study:* Microsatellite markers were developed to assess the level of genetic variation and population structure in *Banksia integrifolia*, a widespread species endemic to eastern Australia.
- *Methods and Results:* We used next-generation sequencing approaches to identify and develop 11 polymorphic microsatellite markers with perfect tri- and tetranucleotide repeats. We tested these markers with 71 specimens from three populations. Observed and expected heterozygosities ranged from 0.0 to 0.875 and 0.0 to 0.763, respectively.
- *Conclusions:* The developed markers will be valuable for studies of the population structure, mating system, and selection of provenances for restoration projects involving *B. integrifolia*.

Key words: *Banksia integrifolia*; coast banksia; genetic variation; next-generation sequencing; Proteaceae; simple sequence repeat markers.

Banksia integrifolia L.f. (Proteaceae), the coast banksia, is one of the most widely distributed tree species along the eastern seaboard of Australia. It occurs over a remarkable climatic range from the tropics to the cool temperature latitudes of Bass Strait. It grows in a wide variety of woodland and scrub habitats, including dune systems, granite outcrops, and on soils derived from basalt (Brewer and Whelan, 2003). *Banksia integrifolia* is considered a keystone species in many communities (Peel, 2010) including endangered ecological communities (e.g., Warkworth Sands Woodland in New South Wales, Australia). The barrel-shaped inflorescences provide a nectar resource for a variety of animals including blossom bats, birds, and nonflying mammals throughout the year, during autumn and winter in cold climates and at times when other species are not in flower (Hackett and Goldingay, 2001). Its proteoid root system forms a mat in the soil, which enhances nutrient solubilization and uptake (Grierson and Attiwill, 1989), making it an important species for restoration projects. It has been introduced to New Zealand and Western Australia where it is considered an emerging weed.

Microsatellites are useful genetic markers for the study of population genetics and mating systems, for the conservation and management of natural populations, and for addressing evolutionary biology questions. This utility is because microsatellites are codominant, highly polymorphic, and able to be scored relatively easily. However, the major drawback of the traditional approaches used to develop microsatellites is that the de novo generation of markers is highly time-consuming and/or expensive due to the requirement to enrich genomic DNA for repeated motifs, cloning, and sequencing. The identification of microsatellite markers using next-generation sequencing (NGS) is a fast and cost-effective approach that is becoming economically feasible (Zalapa et al., 2012). We describe the development of species-specific microsatellite loci for *B. integrifolia* using 454 sequencing. Our aim was to develop and test microsatellite markers by combining the 454 sequencing with the screening of markers using the QIAxcel gel electrophoresis system (QIAGEN, Melbourne, Victoria, Australia; Gross et al., 2012) before labeling with the cost-effective M13 universal labeling technique.

METHODS AND RESULTS

Total genomic DNA was extracted from leaf samples (dried in silica gel) using the DNeasy Plant Mini Kit (QIAGEN) according to the manufacturer's protocol. The quality and quantity of DNA in the samples was tested on a 1% agarose gel along with standard concentrations of λ DNA (5, 100, 250, and 500 ng), followed by visual assessment of each sample. DNA was quantified using a spectrophotometer (NanoDrop ND-1000; NanoDrop Technologies, Wilmington, Delaware, USA) and diluted to a concentration of 20 ng/μL in PCR-grade water. A DNA sample from South Auckland, New Zealand (Appendix 1), was used to create a shotgun library using a Roche 454 GS Junior Titanium emPCR Kit (Lib-L) and was sequenced on a Roche 454 GS Junior Titanium Genome Sequencer (Roche Diagnostics, Pleasanton, California, USA) using Roche Titanium sequencing chemistry at Landcare Research, Auckland, New Zealand. We obtained 126 393 reads in FASTA format with a mean read length of 358 bp and a total yield of 46 Mbp. Sequences were directly screened for perfect di-, tri-, and tetranucleotide repeats, using MSATCOMMANDER version 0.8.2 (Faircloth, 2008). A total of 4615 reads contained microsatellite repeats, of which 2402 reads had enough flanking regions (at least 30 bp) to design primers. The 'Design Primer' option was selected in MSATCOMMANDER so that the program was searching and identifying primer annealing sites in one single step. MSATCOMMANDER uses Primer3 (Rozen and Skaletsky, 2000) for

[1]Manuscript received 14 March 2012; revision accepted 23 July 2012.

The authors thank C. Winks and I. Simpson for collecting the samples. D. Park carried out the library preparation and sequencing. Funding was provided for this work to G.J.H. via Landcare Research contract number CF1112-94-03 and to C.L.G. by Rio Tinto Coal & Allied.

[4]Author for correspondence: mfatemi@une.edu.au

doi:10.3732/apps.1200130

TABLE 1. Characteristics of 11 polymorphic microsatellite loci in *Banksia integrifolia*.[a]

Locus	GenBank accession no.	Primer sequences (5′–3′)	Fluorescent label	Repeat motif	*A*	Allele size (bp)
Bint02	JQ717058	F: GTACCTACCAAGGGCGGTG R: TCCTGGATTGCAAAATCACAAAC	NED	$(AAAG)_5$	5	202
Bint03	JQ717059	F: GCCAGCCCTACAGATCAGG R: AGAGGCAAAACAAACTCGGAC	PET	$(AAAT)_7$	3	160
Bint05	JQ717060	F: TTCACCGGCCTCTTATCCC R: TCGTTGCGGATTGCAGTTG	VIC	$(AAC)_8$	4	226
Bint07	JQ717061	F: TCCAAGCCGGGAAGAATTG R: AGAAACTGCGAGCAGAACTC	FAM	$(AAG)_5$	4	156
Bint12	JQ717062	F: AAGGGCCATACCTGGTGC R: AGAGCGTGGATAGGACACG	NED	$(AAT)_{13}$	7	242
Bint14	JQ717063	F: TCCATATAGCTGATCACATTAAGTTGG R: TCAGATGAACACGTAGCAATCG	PET	$(AAT)_9$	7	156
Bint16	JQ717064	F: AGAGGAGCGAGTCGAAGTG R: GAAACAGGTAGCGTGTCCG	VIC	$(ACGG)_6$	5	241
Bint17	JQ717065	F: GAAGGCCTGAAATGTCGCC R: GAACCTGGCTGGAATGCTG	FAM	$(AGG)_{10}$	7	156
Bint19	JQ717066	F: GGTAGCAACAAGACGTTGGG R: ACCACCATCACCCACTACC	NED	$(GGT)_{10}$	4	149
Bint24	JQ717067	F: GTGCCTGCACAGAAAGGTG R: AATGGACTGGACCTCAGCC	PET	$(AGT)_{11}$	4	220
Bint25	JQ717068	F: ATCTCCCCAGCAACTCGTG R: AGATCGGCCATTCTGGGTC	VIC	$(AGG)_7$	3	160

Note: *A* = number of alleles.

[a] Sequences are deposited in the GenBank Sequence Read Archive (SRA053681).

primer design, and the following criteria were met in primer design: optimal melting temperature of 60°C, optimal GC content of 50%, possession of 1 bp GC clamp, and low levels of self- or pair-complementarity and maximum end stability (Faircloth, 2008). From a total of 126 393 reads we detected 1434 dinucleotide, 860 trinucleotide, and 108 tetranucleotide repeats. We designed 25 markers (20 trinucleotides and five tetranucleotides).

Initial primer testing included 12 samples of *B. integrifolia* from three populations in the Hunter Valley region, Australia (Appendix 1). The PCR volume of 15 μL consisted of approximately 10 ng of genomic DNA, 1× Master Mix (Promega Corporation, Alexandria, New South Wales, Australia), and 0.01 μM of each forward and reverse primer. Amplification was carried out on a GenePro Thermal Cycler (Bioer Technology, Hanghzhou, Zhejiang, China) with initial denaturing at 94°C for 10 min, followed by two cycles of 94°C for 30 s, 64°C for 30 s, and 72°C for 90 s, and then another two cycles in which the annealing temperature was reduced by 3°C to 61°C, followed by 40 cycles of 94°C for 30 s, 58°C for 30 s, and 72°C for 90 s, with a final extension at 72°C for 30 min. PCR products were run on a QIAxcel gel electrophoresis system with a high-resolution gel cartridge, and data were analyzed using the QIAxcel BioCalculator version 3.2 (QIAGEN).

Primers that produced clear and polymorphic bands were selected for further testing by labeling them with M13 fluorophore-conjugated primers (6-FAM, NED, PET, and VIC). The PCR volume of 10 μL consisted of approximately 10 ng of genomic DNA, 1× Master Mix (Promega Corporation), 2.5 nM forward primer, and 0.01 μM of reverse and M13 primer. PCR amplification was carried out with an initial denaturing for 5 min at 94°C, followed by eight cycles of 30 s at 94°C, 30 s at 58°C, and 90 s of elongation, then 30 cycles of 30 s at 94°C, 30 s at 53°C, and 90 s of elongation to incorporate M13 fluorophore-conjugated primers, and a final extension at 72°C for 30 min. Fragments were

TABLE 2. Results of primer screening in three Australian populations[a] of *Banksia integrifolia*.

	WM (*N* = 24)					EW (*N* = 24)					AF (*N* = 23)				
Locus[b]	*A*	H_o	H_e	F_{IS}	Null	*A*	H_o	H_e	F_{IS}	Null	*A*	H_o	H_e	F_{IS}	Null
Bint02**	3	0.125	0.318	0.599	0.17	4	0.167	0.365	0.534	0.19	4	0.174	0.617	0.712	0.27
Bint03	3	0.261	0.309	0.138	0.16	2	0.042	0.042	–0.021	0.00	1	0	0	—	—
Bint05**	4	0.792	0.575	–0.405	0.49	4	0.875	0.689	–0.297	0.41	4	0.696	0.669	–0.064	0.48
Bint07	2	0.375	0.361	–0.061	0.00	3	0.292	0.263	0.131	0.00	3	0.348	0.305	–0.165	0.00
Bint12**	6	0.154	0.702	0.772	0.35	4	0.048	0.645	0.924	0.36	3	0	0.557	1.000	0.35
Bint14**	6	0.208	0.563	0.622	0.25	4	0.292	0.610	0.512	0.21	4	0.087	0.312	0.715	0.20
Bint16*	5	0.500	0.556	0.081	0.01	2	0.417	0.383	–0.111	0.20	3	0.130	0.335	0.602	0.21
Bint17**	5	0.167	0.270	0.368	0.09	5	0.417	0.763	0.443	0.31	4	0.478	0.747	0.345	0.13
Bint19*	0	0	0	—	—	3	0.150	0.304	0.494	0.16	3	0.696	0.590	–0.205	0.00
Bint24**	3	0.417	0.643	0.338	0.56	3	0.333	0.507	0.329	0.44	3	0.435	0.527	0.156	0.34
Bint25	2	0.182	0.169	–0.100	0.00	2	0.042	0.042	–0.021	0.00	1	0	0	—	—
Mean	3.54	0.289	0.406	0.223	—	3.27	0.279	0.419	0.267	—	3	0.277	0.423	0.281	—

Note: *A* = number of alleles; F_{IS} = inbreeding coefficient; H_e = expected heterozygosity; H_o = observed heterozygosity; *N* = number of sampled individuals; Null = null allele frequency estimate.

[a] Populations are abbreviated as: WM = Wards Mistake; EW = Ex Wambo; AF = Archerfield.

[b] Asterisks indicate significant departure from Hardy–Weinberg equilibrium after Bonferroni correction: $*P \le 0.05$, $**P \le 0.01$.

separated on an ABI 3730 genetic analyzer with GeneScan 500 LIZ (Applied Biosystems, Carlsbad, California, USA) as an internal size standard. In total, 11 primers were selected and tested on 71 samples from three populations (Table 1). For these 71 samples, we used GenAlEx 6.41 (Peakall and Smouse, 2006) to estimate the number of alleles, expected and observed heterozygosities, and inbreeding coefficient (F_{IS}). The number of alleles ranged from three to seven with an average of 4.8 alleles per locus. Observed and expected heterozygosities were 0.0–0.875 and 0.0–0.763, respectively (Table 2). Null allele frequencies calculated using CERVUS ranged from 0.0 to 0.49. The presence of null alleles was further tested using a pedigree analysis comprising known plants and their offspring. Pedigree analysis results indicated that null alleles are present only in three loci (Bint05, Bint07, and Bint19). These loci are still of use in parentage assignment, particularly when only heterozygous offspring genotypes are compared to the candidate parent genotypes. Eight loci deviated from Hardy–Weinberg equilibrium (HWE) when tested using a Markov chain method (after Bonferroni corrections; Table 2) both at population and global levels; this could be attributed to a nonrandom pollinator visitation. Mean F_{IS} values for studied populations ranged from 0.223 to 0.281. Although many *Banksia* species are typically considered as outcrossing plants (Broadhurst and Coates, 2004), our results indicate substantial amounts of self-pollination and inbreeding within the same tree (i.e., correlated mating [range of mean F_{IS} values without considering loci with null alleles: 0.352–0.588]) in the populations of *B. integrifolia* that we studied. Further studies that assess pollinator efficacy and mating systems are now required.

CONCLUSIONS

We developed 11 species-specific microsatellite markers based on next-generation sequencing for *B. integrifolia*. The developed markers will be valuable in studies of population structure, mating system, and selection of provenances for restoration projects. These markers will also be useful for assessing genetic variation in the introduced range.

LITERATURE CITED

Brewer, I. M., and R. J. Whelan. 2003. Changes in dune vegetation over 60 years in a sand-mined area of the NSW lower North Coast. *Cunninghamia* 8: 85–92.

Broadhurst, L. M., and D. J. Coates. 2004. Genetic divergence among and diversity within two rare *Banksia* species and their common close relative in the subgenus *Isostylis* R.Br. (Proteaceae). *Conservation Genetics* 5: 837–846.

Faircloth, B. C. 2008. MSATCOMMANDER: Detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources* 8: 92–94.

Grierson, P. F., and P. M. Attiwill. 1989. Chemical characteristics of the proteoid root mat of *Banksia integrifolia* L. *Australian Journal of Botany* 37: 137–143.

Gross, C. L., P. A. Nelson, A. Haddadchi, and M. Fatemi. 2012. Somatic mutations contribute to genotypic diversity in sterile and fertile populations of the threatened shrub, *Grevillea rhizomatosa* (Proteaceae). *Annals of Botany* 109: 331–342.

Hackett, D. J., and R. L. Goldingay. 2001. Pollination of *Banksia* spp. by non-flying mammals in north-eastern New South Wales. *Australian Journal of Botany* 49: 637–644.

Peakall, R., and P. E. Smouse. 2006. GenAlEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288–295.

Peel, B. 2010. Rainforest restoration manual for south-eastern Australia. CSIRO Publishing, Collingwood, Victoria, Australia.

Rozen, S., and H. J. Skaletsky. 2000. Primer3 on the WWW for general users and for biologist programmers. *In* S. Misener and S. A. Krawetz [eds.], Methods in molecular biology, vol. 132: Bioinformatics methods and protocols, 365–386. Humana Press, Totowa, New Jersey, USA.

Zalapa, J. E., H. Cuevas, H. Zhu, S. Steffan, D. Senalik, E. Zeldin, B. McCown, et al. 2012. Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. *American Journal of Botany* 99: 193–208.

Appendix 1. Locations and herbarium voucher information for *Banksia integrifolia* populations used in this study.

Population	Voucher	Geographic coordinates	Herbarium
South Auckland	C. Winks	36°55′11″S, 174°46′03″E	CHR618143
Wards Mistake	I. Simpson WSW-DNA-11	30°08′18″S, 151°51′36″E	NE098953A
Ex Wambo	I. Simpson WSW-DNA-12	32°35′51″S, 151°1′49″E	NE098954A
Archerfield	I. Simpson WSW-DNA-7	32°30′17″S, 151°02′55″E	NE098950A

Note: CHR = Allan Herbarium; NE = N. C. W. Beadle Herbarium.