Development of Microsatellite Loci for the Riparian Tree Species Melaleuca argentea (Myrtaceae) Using 454 Sequencing

Authors: Nevill, Paul G., Williams, Anna, Krauss, Siegfried, Bradbury, Donna, Samaraweera, Sunil, et al.

DEVELOPMENT OF MICROSATELLITE LOCI FOR THE RIPARIAN TREE SPECIES *MELALEUCA ARGENTEA* (MYRTACEAE) USING 454 SEQUENCING[1]

PAUL G. NEVILL[2,3,7], ANNA WILLIAMS[2,3], SIEGFRIED KRAUSS[2,3], DONNA BRADBURY[2,3], SUNIL SAMARAWEERA[4], AND MICHAEL G. GARDNER[5,6]

[2]Botanic Gardens and Parks Authority, Kings Park and Botanic Garden, West Perth 6005 Western Australia, Australia; [3]School of Plant Biology, University of Western Australia, Nedlands 6009 Western Australia, Australia; [4]Rio Tinto Iron Ore, Level 22, 152–158 St. Georges Terrace, Perth 6000 Western Australia, Australia; [5]School of Biological Sciences, Flinders University, GPO Box 2100, Adelaide 5001 South Australia, Australia; and [6]Australian Centre for Evolutionary Biology and Biodiversity, University of Adelaide, Adelaide 5005 South Australia, Australia

- *Premise of the study:* Microsatellite primers were developed for *Melaleuca argentea* (Myrtaceae) to evaluate genetic diversity and population genetic structure of this broadly distributed northern Australian riparian tree species.
- *Methods and Results:* 454 GS-FLX shotgun sequencing was used to obtain 5860 sequences containing putative microsatellite motifs. Two multiplex PCRs were optimized to genotype 11 polymorphic microsatellite loci. These loci were screened for variation in individuals from two populations in the Pilbara region, northwestern Western Australia. Overall, observed heterozygosities ranged from 0.27 to 0.86 (mean: 0.52) and the number of alleles per locus ranged from two to 13 (average: 4.3).
- *Conclusions:* These microsatellite loci will be useful in future studies of the evolutionary history and population and spatial genetic structure in *M. argentea*, and inform the development of seed sourcing strategies for the species.

Key words: 454 GS-FLX; *Melaleuca argentea*; northern Australia; nuclear SSR markers; riparian tree; shotgun sequencing.

Melaleuca argentea W. Fitzg. (Myrtaceae) is a small to medium (3–25 m high) tree species with a broad distribution throughout northern Australia. *Melaleuca argentea* is shallow rooted and an obligate phreatophore and therefore reliant on shallow water tables or surface water (Graham, 2001). Surface and near-surface water are scarce in the semiarid Pilbara region, and larger tree species like *M. argentea* are only found in or along riparian communities. Changes to the spatial and temporal aspects of water availability, associated with mining activities in the region, may have the potential to impact some areas of the riparian communities, requiring rehabilitation of areas where loss of vegetation may occur. Therefore, information on the spatial genetic structure of *M. argentea* is needed to inform the development of seed sourcing strategies. There are no species-specific molecular markers available for *M. argentea* and there have been no previous studies on genetic variation in the species. Here, we report the isolation and characterization of 11 nuclear microsatellite loci using next-generation sequencing that will be used for population genetic studies in this species.

METHODS AND RESULTS

Genomic DNA (5 μg) was isolated from the leaf tissue of one individual of *M. argentea* following the protocol of Glaubitz et al. (2001), modified with the addition of a wash buffer (Wagner et al., 1987). The DNA was then sent to the Ramaciotti Centre for Gene Function Analysis (University of New South Wales, Sydney, Australia) for shotgun sequencing on a Roche 454 GS-FLX sequencer with titanium chemistry (Roche Applied Science, Indianapolis, Indiana, USA) following Gardner et al. (2011). The sample occupied 12.5% of a plate and produced 274,938 individual sequences, with an average read length of 362 bp, of which 5860 contained microsatellites. We used the program QDD version 1 (Meglécz et al., 2010) to screen the raw sequences for those that contained eight or more di-, tri-, tetra-, or pentabase repeats, remove redundant sequences, and design primers (automated in QDD using Primer3 [Rozen and Skaletsky, 2000]). Software running parameters were set to default values with the exception of PCR product lengths, which was set to 90–450 bp. Primer pairs were designed for 548 different loci and from these we excluded immediately all loci that contained imperfect repeats, had a greater than 2°C difference between the forward and reverse primer annealing temperatures, a GC content less than 40% or greater than 60%, polynucleotide runs of four or more runs in the flanking regions, and short repeat motifs within the flanking region or primer sequence. From the remaining 355 loci, we randomly chose for further development 32 loci containing either dinucleotide or trinucleotide repeats (GenBank accession no.: JX424003–JX424034). These 32 loci were trialed for amplification with the cost-effective approach of Schuelke (2000) using a QIAGEN Multiplex PCR Kit (QIAGEN, Hilden, Germany). Loci were amplified in individual, 20-μL reactions containing 10 μL QIAGEN Multiplex PCR Master Mix; 2.5 μL Q-solution; forward primer (with sequence tag at 5′ end, unlabeled) 0.05 μM, reverse primer (unlabeled) 0.2 μM, and 5′ sequence tag (labeled; unique to primer) 0.2 μM; 10–50 ng DNA; plus sterile H_2O to 20 μL. PCR cycling was performed in a Corbett Gradient Palm-Cycler (Corbett Life Science, Sydney, Australia) according to the manufacturer's protocol as follows: *Taq* activation at 95°C for 15 min; followed by 35 cycles of denaturation at 94°C for 30 s, annealing at 60°C for 90 s, and extension at 72°C for 90 s; followed by a final extension at 60°C for 30 min. PCR products were visualized

TABLE 1. Characteristics of 11 nuclear microsatellite primers developed in *Melaleuca argentea*.[a,b]

Locus	Primer sequences (5′–3′)[c]	Repeat motif	Size range (bp)[d]	Label (Multiplex)[e]	GenBank accession no.
KPMA07	F: CAGTGCGACTGATCCTTTGA R: AAAATTAGATGACGCCGAGC	$(AG)_{11}$	119–127	FAM (1)	JX424009
KPMA11	F: AACATGGCTACCAACCAACC R: ACGATCAGGGACAGCAAGAT	$(GCC)_8$	126–144	PET (2)	JX424013
KPMA14	F: CGACAACAGAAGCCGGTAAT R: CGAAGAGAAACGAAGGAACG	$(GAA)_9$	202–208	VIC (1)	JX424016
KPMA16	F: CGTAGAACTTGGCGAGGAAG R: CCCTACTGTCCGACGAAGAA	$(CTT)_9$	209–221	FAM (2)	JX424018
KPMA17	F: CCTGGTGGGTAAAGGGTCTT R: GGTCTAGAGTTCAGCGCTACAAA	$(CT)_8$	132–140	NED (1)	JX424019
KPMA18	F: TTGAAGGTTTGTTTGCACCA R: CTGCCTTGCTTGATGTTTCA	$(AG)_{10}$	208–216	NED (2)	JX424020
KPMA20	F: CTCCGCATCTTGATAGCGTT R: TACGTAAATTCGGGCTCACC	$(GCA)_8$	221–230	PET (1)	JX424022
KPMA23	F: ATGAATGGAGGCCAGACAAC R: CCCCTCTTTCACCTCTTCCT	$(GCG)_8$	126–138	FAM (2)	JX424025
KPMA25	F: GCTCATTCGTTTGCTTCCAT R: CGCTCCTCCTAAAGTTTCCC	$(GA)_{13}$	142–188	NED (2)	JX424027
KPMA26	F: GCATCCTAGCGGACGAGTAG R: TAGCTGCTTCAAAGATGGCA	$(CTT)_{11}$	238–258	NED (2)	JX424028
KPMA28	F: ACCAAAACATTTCCCCATGA R: TATACGCCCTCCTTCAATGC	$(GA)_{13}$	152–172	PET (2)	JX424030

[a] All values are based on samples from two populations in the Pilbara region of Western Australia (WW-BO: UTM coordinates 720216E 7448488N, n = 25; WW-LP: UTM coordinates 725022E 7462878N, n = 21).

[b] An annealing temperature of 56°C was used for all primers.

[c] Forward primer sequences do not include the 5′-forward tail.

[d] Size range of the PCR products in base pairs includes the 5′-forward tail.

[e] 5′ tags: FAM 5′-AATACGACTCACTATAG; NED 5′-ATTTAGGTGACACTATAG; PET 5′-GCGGATAACAATTTCACACAGG; VIC 5′-ATTA-ACCCTCACTAAAG.

on a 2% agarose gel stained with SYBR Safe (Invitrogen Corporation, Carlsbad, California, USA). Twenty-four loci amplified an unambiguous product of the expected size. These amplifiable loci were then screened on eight individuals to test for polymorphism and alleles sized using an ABI 3730 genetic analyzer (Applied Biosystems, Foster City, California, USA) and GENEMAPPER version 4.0 software (Applied Biosystems).

Of the 24 loci, 11 were polymorphic and scorable, seven did not amplify consistently, and alleles at six loci were unclear. We screened the 11 polymorphic and scorable loci for variation in 20 or more individuals from each of two populations of *M. argentea* from the Pilbara region (Table 1) (WW-BO: Universal Tranverse Mercator [UTM] coordinates 720216E 7448488N, collection no. PN103; WW-LP: UTM coordinates 725022E 7462878N, collection no. PN104; herbarium material is deposited at the Kings Park and Botanic Garden Herbarium, Perth, Western Australia [KPBG]). The loci were amplified in two multiplex PCRs, and conditions were as per the initial screening with the exception of a reduction of the annealing temperatures to 56°C to improve amplification. Genetic diversity parameters were calculated using GenAlEx version 6.4 (Peakall and Smouse, 2006), and deviation from Hardy–Weinberg equilibrium (HWE) was determined using GENEPOP 3.4 (Raymond and Rousset, 1995) (Tables 1 and 2). P values from HWE tests were adjusted for multiple tests of significance using the sequential Bonferroni method. Overall, the number of alleles observed for the 11 loci ranged from two to 13 with an average of 4.3 alleles per locus (Table 2). The observed and expected heterozygosities ranged from 0.23 to 0.86 and 0.29 to 0.78, respectively (Table 2). Significant departures from HWE were detected in the KPMA07 and KPMA25 loci in WW-BO and WW-LP (Table 2). This may be a result of heterozygote deficiency, potentially resulting from the presence of null alleles and/or population structure. We checked all pairs of loci for linkage disequilibrium in GENEPOP, and none were significant after sequential Bonferroni adjustment.

a species that may be impacted by hydrological changes associated with mining in the Pilbara region of northwestern Western Australia. Specifically, research is planned that will examine genetic connectivity between geographically proximate riparian systems, which will inform the development of seed sourcing guidelines for the species.

TABLE 2. Results of primer screening in two populations of *Melaleuca argentea*.[a]

	WW-BO			WW-LP		
Locus	A	H_o[b]	H_e	A	H_o[b]	H_e
KPMA07	5	0.61*	0.55	5	0.23*	0.48
KPMA11	3	0.44	0.50	3	0.33	0.29
KPMA14	2	0.29	0.43	2	0.44	0.42
KPMA16	3	0.58	0.58	3	0.55	0.52
KPMA17	3	0.67	0.57	5	0.44	0.62
KPMA18	3	0.27	0.49	4	0.38	0.70
KPMA20	3	0.44	0.45	3	0.65	0.51
KPMA23	2	0.56	0.50	5	0.57	0.56
KPMA25	8	0.50*	0.66	10	0.52*	0.67
KPMA26	4	0.72	0.64	6	0.86	0.78
KPMA28	6	0.64	0.64	6	0.76	0.76

Note: A = number of alleles; H_e = expected heterozygosity; H_o = observed heterozygosity.

[a] All values are based on samples from two populations in the Pilbara region of Western Australia (WW-BO: UTM coordinates 720216E 7448488N, n = 25; WW-LP: UTM coordinates 725022E 7462878N, n = 21).

[b] Significant deviations from Hardy–Weinberg equilibrium (*P < 0.01) after correction for multiple tests (sequential Bonferroni procedure) are reported.

CONCLUSIONS

The microsatellite loci presented in this study will be useful in the examination of the population genetic structure of *M. argentea*,

LITERATURE CITED

GARDNER, M. G., A. FITCH, T. BERTOZZI, AND A. J. LOWE. 2011. Rise of the machines: Recommendations for ecologists when using second generation sequencing for microsatellite development. *Molecular Ecology Resources* 11: 1093–1101.

GLAUBITZ, J. C., L. EMEBIRI, AND G. MORAN. 2001. Dinucleotide microsatellites from *Eucalyptus sieberi*: Inheritance, diversity and improved scoring of single-base differences. *Genome* 44: 1041–1045.

GRAHAM, J. 2001. The root hydraulic architecture of *Melaleuca argentea*. Honors Thesis, Department of Botany, University of Western Australia, Perth, Australia.

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.

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.

RAYMOND, M., AND F. ROUSSET. 1995. GENEPOP (version 1.2): Population genetics software for exact tests and ecumenism. *Journal of Heredity* 86: 248–249.

ROZEN, S., AND H. SKALETSKY. 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology (Clifton, N.J.)* 132: 365–386.

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

WAGNER, B. D., G. R. FURNIER, S. M. WILLIAMS, M. A. SAGHAI-MAROOF, B. P. DANIK, AND R. W. ALLARD. 1987. Chloroplast DNA polymorphisms in lodgepole and jack pines and their hybrids. *Proceedings of the National Academy of the United States of America* 84: 2097–2100.