Microsatellite Letter for Conservation Genetic Resources

Word Count 830

Characterisation of microsatellite DNA markers for *Grevillea globosa* C. A. Gardner.

Melissa A Millar[1, 2], Margaret Byrne[1, 3], David J Coates[1,3], J Dale Roberts[4]

[1]Science and Conservation Division, Department of Parks and Wildlife, Locked Bag 104, Bentley Delivery Centre, Bentley, WA 6983, Australia

[2]School of Animal Biology, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia

[3]School of Plant Biology, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia

[4]School of Animal Biology, Centre for Evolutionary Biology and Centre of Excellence in Natural Resource Management, The University of Western Australia, PO Box 5771, Albany , WA 6332, Australia

Keywords: microsatellites, M13 tagged primer, Western Australia,

*Corresponding Author. Melissa A. Millar, Science and Conservation Division, Department of Parks and Wildlife, Locked Bag 104, Bentley Delivery Centre, Bentley W.A. 6983, Australia

Phone +618 92919 9083, Fax +618 9334 0327, email: melissa.millar@dec.wa.gov.au

Running Title: Microsatellite DNA markers for *Grevillea globosa*

Abstract

A genomic library was constructed and 14 novel polymorphic di- and tri-nucleotide nuclear microsatellite markers were developed for *Grevillea globosa*, an endemic shrub of southwest Western Australia. Populations are patchily distributed and population genetic structure is being investigated to inform appropriate seed collection and restoration strategies. Diversity in a selected population was high, with the number of alleles per locus ranging from three to 13 and expected and observed heterozygosities averaging 0.693 and 0.799 respectively. All loci showed independent inheritance and there was no evidence of possible null alleles.

Microsatellite Letter

Microsatellite markers were developed for *Grevillea globosa*, a shrub with a limited distribution in southwest Western Australia's semiarid rainfall zone. Plants occur on sand and loam in Acacia shrubland, mallee woodland and shrubland, and Eucalypt woodland. Populations are patchily distributed and poorly known hence the taxon is of some conservation significance. DNA was extracted from homogenised, freeze dried material of an individual originating from the centre of the species range using a NucleoSpin Plant II Maxi protocol following the manufacturers' instructions (Macherey-Nagel GmbH & Co, Düren, Germany). Genomic DNA was 454 shotgun sequenced using 1\8 of a picotiter plate on a GS-FLX Titanium machine (Roche Diagnostics Corporation, 454 Life Sciences, Branford, USA) by the Australian Genome Research Facility (Adelaide, Australia). Sequencing, microsatellite identification and primer design followed Gardner et al. (2011).

Amplification and polymorphism were evaluated in eight individuals from four populations using an M13 tailed three primer PCR system. Loci were amplified in a total volume of 7.5 µl per reaction containing 2 ng template DNA, 50 mM KCl, 20mM Tris HCl (pH 8.4), 0.2 mM each dNTP, 0.16 µM M13 tailed fluorescently labelled forward and reverse primer, 0.032 µM M13 tailed forward primer, 0.175 µl of *Taq* DNA polymerase (Invitrogen) and 3.5mM $MgCl_2$. Amplification reactions for all primer pairs were optimised using a program of 95°C for 4 min, 25 cycles of 30 s at 94°C, 30 s step down from 65°C to 50°C, 80 s at 72°C followed by 30 cycles of 15 s at 94°C, 15 s at 50°C, and 45 s at 72°C, with a final step of 8 min at 72°C. Reactions were completed in an Eppendorf™ Thermal Cycler (Eppendorf, New York, USA) and PCR products visualised on 1% agarose gels stained with GelRed™ nucleic acid stain (Biotium, San Fransisco, USA).

Six primers failed to amplify or produced numerous bands that could not be resolved under high stringency conditions. Polymorphism was tested for the final 14 loci using in 20 individuals from a single population located within the centre of the species range (Table 1). One µl of diluted PCR product was added to 12 µl of GeneScan™ LIZ®500(-250) size standard/formamide, and fragments visualised using an 3730 DNA Analyser. Genotypes were scored using Genemapper™ v3.7 (Applied Biosystems). Tests for heterozygotic deficit and excess were conducted using the Markov chain method and tests for Linkage Disequilibrium (LD) were conducted using Fisher's exact tests in GENEPOP v3.4 (Raymond & Rousset 1995). Bonferroni corrections were applied for multiple comparisons. Evidence for large allele drop-out, scoring error due to stutter and null

alleles were assessed with MICRO-CHECKER (Oosterhout et al. 2004). Genetic diversity parameters were estimated using GenAlEx v6.2 (Peakall & Smouse 2006) (Table 1).

Overall, polymorphism was high (Table 1). 100% of loci were polymorphic. Number of alleles per locus averaged 7.357 and the effective number of alleles averaged 3.581. There was no evidence of significant LD ($p < 0.001$) or of possible null alleles at any loci. All loci were in Hardy Weinberg equilibrium and there was no evidence of large allele drop-out or scoring error due to peak stutter. Expected heterozygosity averaged 0.712 and observed heterozygosity averaged 0.749. These markers will be used to study genetic structure in *G. globosa* as part of a broader investigation to inform appropriate seed collection and restoration strategies for a number of keystone species in this landscape.

Acknowledgements

Funding was provided by the Australian Research Council and Karara Mining Limited. We thank Jenny Borger, David Beven and Roger Pitman for assistance with field collections.

References

Doyle J, Doyle J (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin 19:11-15

Gardner M, Fitch A, Bertozzi T, Lowe A (2011) Rise of the machines - recommendations for ecologist when using next generation sequencing for microsatellite development. Mol Ecol Resour 11:1093-1101

Oosterhout C, Hutchinson W, Willds D, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535-538

Peakall R, Smouse P (2006) GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. Mol Ecol Notes 6:288-295

Raymond M, Rousset F (1995) GENEPOP (version 3.4): population genetics software for exact test and ecumenicism. Journal of Heredity 86:248-249

Table 1. Characterisation of 14 microsatellite loci in *Grevillea globosa*. Details are given for locus name, primer sequence, GenBank accession number, repeat motif, the size range of observed alleles in base pairs (including the M13 tail), number of alleles (N), observed (H_o) and expected (H_e) heterozygosities, estimate of the fixation index (F).

Locus	Primer sequence (5′-3′)	GenBank accession no.	Repeat motif	Allele size range	N	N_e	H_o	H_e	F
Gg001	F:CACGCACTGCATCTAGAACG R: CCCTCACCGTAACCCTAACC	KJ470095	$(AG)_8$	142-192	13	4.484	0.842	0.777	-0.084
Gg002	F:CTGCGTCCAGCGTCTTAACT R:TTTCCGATTCCACTTCTCCA	KJ470096	$(AG)_8$	164-174	5	2.837	0.850	0.648	-0.313
Gg004	F: AAGGTTGGCCCTGGATACAT R:TGCCCTTCTGCTTCTCTTTG	KJ470097	$(AaG)_8$	247-264	9	5.405	1.000	0.815	-0.227
Gg005	F: TATCGCTGTTCTGGCTGTGA R:TCTGCAGCTTGCTCTGTCTG	KJ470098	$(ACC)_7$	275-296	7	2.247	0.550	0.555	0.009
Gg006	F:TTGGCAAATATGGAAATGGC R:CCTTCTACGCAAACAATGGC	KJ470099	$(AG)_{10}$	107-121	7	2.424	0.550	0.588	0.064
Gg007	F:ATGCCTGGAATAGACGGGAT R:TCCTCCTTCACCTTCACCAA	KJ470100	(AGT)7	209-231	12	5.674	1.000	0.824	-0.214
Gg009	F:GCTTCATCGGTCCTTGTCAC R:GCTGTCAGGGATTCAGGTTG	KJ470101	$(AG)_7$	110-119	3	2.831	0.895	0.647	-0.383
Gg010	F:TGGACGACGATCTTCCAAGT R:GTGGCTAATCGTGGTTGCTC	KJ470102	$(AC)_9$	188-204	7	3.065	0.850	0.674	-0.262
Gg011	F:TATCAGTTCCTCCGTTCCGA RAGTTAGCGAAACCGAAACCC	KJ470103	$(AAG)_7$	141-161	8	3.187	0.850	0.686	-0.239
Gg012	F:AGCGTGGGTGATGATCTTGT R:AGCAGCATCTGCATCTCCTC	KJ470104	$(AG)_7$	199-207	5	2.156	0.450	0.536	0.161
Gg013	F:AAGAGCTGGATAACCTCGGC R:TGAGCCCTCCCAAATCTACA	KJ470105	$(AG)_8$	214-231	6	2.837	0.900	0.648	-0.390
Gg014	F:CCTGAGAAGCCCGGATAGTT R:GGAAGACGGAAACCTCCCTA	KJ540106	$(AG)_7$	216-237	6	4.324	0.900	0.769	-0.171
Gg018	F:CTATTTCTGCACCGACCACC R:AGTCAAGACAGCCGTTGGAG\	KJ470107	$(AAG)_7$	225-242	7	4.678	0.800	0.786	-0.017
Gg020	F:AGACCGTAGTTGTCGTCCCA R:TTGAAATCGTCCGTAATGGC	KJ470108	$(AG)_7$	113-122	8	3.980	0.750	0.749	-0.002
Mean (standard error)					7.357 (0.708)	3.581 (0.312)	0.799 (0.045)	0.693 (0.025)	-0.148 (0.046)