BioOne COMPLETE

Development and Testing of New Gene-Homologous EST-SSRs for Eucalyptus gomphocephala (Myrtaceae)

Authors: Bradbury, Donna, Smithson, Ann, and Krauss, Siegfried L.

Source: Applications in Plant Sciences, 1(8)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1300004

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

Downloaded From: https://bioone.org/journals/Applications-in-Plant-Sciences on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Applications in Plant Sciences 2013 1(8): 1300004

PRIMER NOTE

DEVELOPMENT AND TESTING OF NEW GENE-HOMOLOGOUS EST-SSRS FOR *EUCALYPTUS GOMPHOCEPHALA* (MYRTACEAE)[1]

DONNA BRADBURY[2,3,4], ANN SMITHSON[2,3], AND SIEGFRIED L. KRAUSS[2,3]

[2]School of Plant Biology, The University of Western Australia, 35 Stirling Highway, Crawley, Western Australia 6005, Australia; and
[3]Botanic Gardens and Parks Authority, Kings Park and Botanic Garden, Fraser Avenue, West Perth, Western Australia 6009, Australia

- *Premise of the study:* New microsatellite (simple sequence repeat [SSR]) primers were developed from *Eucalyptus* expressed sequence tags (ESTs) and optimized for genetic studies of the southwestern Australian tree *E. gomphocephala*, which is severely impacted by tree health decline and habitat fragmentation.
- *Methods and Results:* A total of 133 gene-homologous EST-SSR primer pairs were designed for *Eucalyptus*, and 44 were screened in *E. gomphocephala*. Of these, 17 produced reliable amplification products and 11 were polymorphic. Between two and 13 alleles were observed per locus, and observed heterozygosities ranged from 0.172 to 0.867. All 17 EST-SSRs that amplified *E. gomphocephala* cross-amplified to at least one of *E. marginata*, *E. camaldulensis*, and *E. victrix*.
- *Conclusions:* This set of EST-SSR primer pairs will be valuable tools for future population genetic studies of *E. gomphocephala* and other eucalypts, particularly for studying gene-linked variation and informing seed-sourcing strategies for ecological restoration.

Key words: ecologically important genetic variation; EST-microsatellite; *Eucalyptus gomphocephala*; Myrtaceae; SSR mining; tuart.

The Australasian tree genus *Eucalyptus* L'Hér. (Myrtaceae) comprises more than 700 species (Brooker, 2000) and is economically significant to the forestry industry. Vast genomic resources for model species have allowed molecular markers to be developed in related nonmodel taxa of conservation significance. *Eucalyptus gomphocephala* DC. (subg. *Symphyomyrtus*, common name tuart) is endemic to the Swan Coastal Plain of southwestern Australia, and is a culturally iconic species in the region. Severe pathogen-mediated tree decline (Cai et al., 2010), together with extensive habitat fragmentation in urbanized areas, has required rapid conservation actions for the species. To inform seed sourcing strategies for ecological restoration and forest management in the context of climate change, measures of adaptive population genetic diversity and structure are critical, but no microsatellite markers (simple sequence repeats [SSRs]) have yet been tested or optimized for *E. gomphocephala*. Specifically, SSRs within gene coding regions such as expressed sequence tags (ESTs) will be valuable for investigating variation in functional and potentially adaptive regions of the genome.

[1]Manuscript received 7 January 2013; revision accepted 28 January 2013.

The authors thank E. Veneklaas, M. Wallace, P. Nevill, J. Anthony, M. Foley, and A. Bradbury for field, laboratory, and conceptual assistance and for advice. We also thank E. O'Brien and P. Nevill for providing *E. marginata*, *E. victrix*, and *E. camaldulensis* DNA. This project was funded by an Australian Research Council Linkage Grant (LP0669757) with additional support from Alcoa of Australia and BHP Billiton Worsley Alumina Pty Ltd.

[4]Author for correspondence: donna.bradbury@bgpa.wa.gov.au

doi:10.3732/apps.1300004

Here, we design novel PCR primer pairs for SSRs within *Eucalyptus* ESTs, targeting those that are homologous to annotated genes and therefore those with putative ecological relevance. For the first time, we optimize a subset for utility in population genetic analysis of *E. gomphocephala* to inform future conservation programs. To demonstrate the markers' broader utility in the genus, we report cross-species amplification in three additional ecologically important species (*E. camaldulensis* Dehnh., *E. marginata* Sm., *E. victrix* L. A. S. Johnson & K. D. Hill). This is the first set of gene-homologous EST-SSRs to our knowledge that have been reported and tested in natural populations of nonmodel eucalypt species of ecological and conservation significance. These EST-SSR primers add value to a recently expanding set of EST-SSRs being made available for *Eucalyptus*, increasing the variety of genes that can be investigated in novel population and conservation genetic studies of nonmodel species.

METHODS AND RESULTS

A total of 36,001 *Eucalyptus* EST sequences were downloaded from GenBank in 2007 and screened for microsatellite repeats with Tandem Repeats Finder (Benson, 1999) using default parameters, and organized using Tandem Repeats Database version 2.30 (Gelfand et al., 2007). A total of 1098 repeats were detected in 1073 (3%) of 36,001 ESTs. Sequences were clustered using the sequence assembly program CAP3 (Huang and Madan, 1999) with default parameters, into a set of 128 nonredundant contigs and 401 singletons (Appendix S1 and S2). A total of 154 of 529 (29%) ESTs were homologous to various annotated genes, following BLASTN searches of the National Center for Biotechnology Information's nucleotide (nr/nt) collection, using an *E*-value threshold of 10^{-10}. PCR primers were designed for all 154 sequences using Primer3 (Rozen and Skaletsky, 2000). Annotated gene homologs were specifically targeted

 © 2013 Bradbury et al. Published by the Botanical Society of America. This work is licensed under a Creative Commons Attribution License (CC-BY-NC-SA).

Downloaded From: https://bioone.org/journals/Applications-in-Plant-Sciences on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

TABLE 1. Characteristics of 11 new EST-SSR loci developed and optimized for *Eucalyptus gomphocephala*.

Locus	Primer sequences (5′–3′)	Repeat motif	GenBank accession no.[a]	Source species[b]	Putative function and *E*-value[c]	Dye	A_T	Allele size range (bp)	M[d]	[Final]	T_a (°C)
EGM09	F: ATTTGCTGAAGTGGGTCTCG R: ACAGGTCCAGAAGCATGAGC	$(AG)_{17}$	ES594818	*E. globulus*	*Ricinus communis* glucan endo-1, 3-beta-glucosidase, putative [2.0e-68]	D4	4	160–166	—	0.2	56
EGM12	F: GCGCCGAGAATCAATACG R: GTAGCTGTTGGCAGCTTTGG	$(CAG)_{10}$	CD668471	*E. tereticonis*	*Populus deltoides* CONSTANS-like protein CO1 [3.0e-22]	D3	4	188–197	A	0.2	56
EGM14	F: CACTGCCACTTACCAGAGTCG R: CCTCCACCATCTCGAACG	$(CT)_{18}$	CB968019	*E. grandis*	*Glycine max* heat shock transcription factor 21 [1.0e-17]	D4	8	350–372	B	0.1	54
EGM24	F: CCTGCAACGCTTCTCGTC R: TCTGTATTGAGGCTCGCGTA	$(CT)_{17}$	ES589764	*E. globulus*	*Prunus dulcis* putative S-adenosylmethionine decarboxylase (SAMDC) [1.0e-26]	D2	3	204–212	—	0.2	56
EGM25	F: CCAGAAGCAACCTCAATTTCC R: AGCCACAGCAGGGAGTAGC	$(TC)_{15}$	ES589925	*E. globulus*	*Arabidopsis thaliana* metal transporter Nramp3 [1.0e-109]	D2	2	350–356	C	0.1	56
EGM30	F: AGTGCAGCACCTTTCAGACC R: AAGATTGATTGCTAGATCAGTCACC	$(AG)_{18}$	CU398186	*E. gunnii*	*Ricinus communis* chlorophyll A/B binding protein, putative [5.0e-45]	D3	13	225–255	C	0.1	56
EGM35	F: ATACGCGTCCCAGTGATTTC R: AGGAGCAGACGAACTTGCAT	$(AG)_{18}$	Contig 92*	*E. globulus*	*Ricinus communis* fructose-1,6-bisphosphatase, cytosolic [1.0e-143]	D2	9	196–212	C	0.1	56
EGM37	F: TGAGGTCACTTCAAGCACCAAGA R: GGAAGCGGCAACAACCTTAACA	$(GCTTA)_5$	Contig85*	*E. globulus*	*Gossypium hirsutum* quinone oxidoreductase [1.0e-106]	D4	2	256–261	B	0.025	54
EGM46	F: ATATTCGGCCTCTTCGCATT R: ACCTTGGCGTTGTACTCGAC	$(CAG)_4(AG)_{12}$	Contig18*	*E. gunnii*	*Glycine max* desiccation protectant protein Lea14 homolog [4.0e-79]	D2	13	233–257	A	0.2	56
EGM47	F: TCGTTCGGTTTCTGTTCTGA R: ACATCCTTCGATCCAACCAG	$(AATCG)_6$	Contig79*	*E. grandis* + *E. gunnii*	*Nicotiana tabacum* small GTPase Rab2 [3.0e-65]	D2	3	94–114	C	0.05	56
EGM48	F: TCACACTCCAATCTCCAACG R: CCTGGTAGTCCTTCGACTGG	$(CT)_{12}$	CU396026	*E. gunnii*	*Quercus macrocarpa* aquaporin PIP2;1 [1.0e-125]	D2	6	141–155	A	0.2	56

Note: A_T = total number of alleles observed, based on 60 individuals from two natural populations of *E. gomphocephala*; Dye = WellRED dye label; [Final] = final concentration of primer pairs in PCR reaction (μM); M = PCR multiplex group; T_a = annealing temperature.

[a] GenBank accession number of source EST, or nonredundant contig number. Contigs are marked with an asterisk (*) and were derived from multiple redundant singleton ESTs. Contig sequence assembly is reported in Appendix S1; EST contig FASTA sequences are reported in Appendix S2.

[b] Species from which EST(s) were derived.

[c] Putative EST function based on a BLASTX search of the UniRef100 database; *E*-value of the match is given in brackets.

[d] Loci allocated the same letter (A, B, or C) were amplified together in the same multiplex PCR reaction; a dash (—) indicates the locus was amplified in singleplex.

Downloaded From: https://bioone.org/journals/Applications-in-Plant-Sciences on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

so that loci could inform hypotheses relating to adaptive genetic variation in future studies of *E. gomphocephala.* All primer pairs and putative gene functions are provided in Appendix S3.

We prioritized a random subset of 44 EST-SSRs for further PCR optimization in *E. gomphocephala*. Primer pairs were synthesized by GeneWorks Pty Ltd (Hindmarsh, Australia) and initially tested for amplification products on a screening panel of seven *E. gomphocephala* individuals. DNA was extracted from freeze-dried leaf material using a NucleoSpin 8/96 Plant II Core Kit, with buffer set PL2/3 (Macherey-Nagel GmbH & Co., Düren, Germany), as per the manufacturer's instructions. PCRs were carried out in a total volume of 10 µL, containing 10 ng genomic DNA template, 1× PCR Polymerization Buffer containing dNTPs, 0.2 µM of each primer, 0.44 units *Taq* DNA polymerase, and 2 mM of $MgCl_2$ (reagents from Fisher Biotech, Perth, Western Australia, Australia). Thermocycling was carried out as follows: denaturation at 94°C for 5 min; followed by 30 cycles of 94°C for 1 min, annealing at 55°C for 30 s, and extension at 72°C for 30 s; followed by a final extension at 72°C for 15 min. Annealing temperatures (T_a) were optimized for each locus and products were visualized on 2% agarose gels stained with SYBR Safe (Invitrogen, Carlsbad, California, USA).

A set of 17 (39%) primer pairs reliably amplified a product of expected size in *E. gomphocephala*. These loci were screened for allele size polymorphism and genotyped in multiplex PCR of between two and four loci per reaction, using a commercial kit (PCR Multiplex Kit using Q Solution; QIAGEN, Hilden, Germany). Forward primers were 5′ end-labeled with WellRED fluorescent dyes (D2, D3, D4; Sigma-Aldrich, St. Louis, Missouri, USA). Reactions were carried out in a total volume of 12.5 µL, containing 5–30 ng DNA template, 1× QIAGEN Multiplex PCR Master Mix, 0.5× Q-solution, and 0.2 µM of each primer (with some exceptions, Table 1). PCR was conducted with an initial activation step at 95°C for 15 min; followed by 30 cycles of 94°C for 30 s, annealing at T_a (Table 1) for 90 s, and extension at 72°C for 60 s; followed by final extension at 72°C for 10 min. Products were genotyped using a CEQ 8800 Genetic Analysis System and analyzed using CEQ Fragment Analysis Software (Beckman Coulter, Brea, California, USA). Two loci (EGM09 and EGM24) exhibited amplification failures or dubious peaks in multiplex and were analyzed in singleplex for further reactions. Following genotyping tests, 11 EST-SSRs were polymorphic and produced reliable electrophoretic profiles in *E. gomphocephala* (Tables 1 and 2). For these loci, we further investigated the most likely gene annotations by conducting a BLASTX search against proteins in the UniRef100 database (European Molecular Biology Laboratory–European Bioinformatics Institute [EMBL-EBI], http://www.ebi.ac.uk/Tools/blastall/) (Table 1).

To screen genetic diversity of the 11 EST-SSRs, we genotyped 60 *E. gomphocephala* individuals from two natural populations: 'Yalgorup' (voucher: *R. Davis 1390*, PERTH *04473167*) and 'Ludlow' (voucher: *C. A. Gardner*, PERTH *01350501*) (Table 2). Genetic diversity parameters were calculated with GenAlEx version 6.4.1 (Peakall and Smouse, 2006). Deviation from Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium among loci were calculated with GENEPOP version 4.0.10 (Raymond and Rousset, 1995; Rousset, 2008), using the Bonferroni correction for multiple testing. The total number of alleles per locus ranged from two to 13 (mean = 6) (Table 1). Observed and expected heterozygosities ranged from 0.172 to 0.867 and 0.238 to 0.861, respectively (Table 2). Significant deviation from HWE was detected for EGM09 at Ludlow. We did not detect evidence of linkage disequilibrium between any pair of loci in more than one population. The potential presence of null alleles and their frequency (r) in each population were estimated using MICRO-CHECKER version 2.2.3 (van Oosterhout et al., 2004), and were predicted for EGM09 and EGM24 in both populations (r range = 0.17–0.29) and for EGM35 at Yalgorup (r = 0.09). The potential for stuttering was predicted for EGM09 at Ludlow.

All 17 primer pairs that initially amplified *E. gomphocephala* DNA were screened for cross-amplification in three additional species, belonging to diverse sections within two subgenera: *E. camaldulensis* (sect. *Exsertaria*) and *E. victrix* (sect. *Adnataria*) (both subg. *Symphyomyrtus*), and *E. marginata* (sect. *Longistylus*) (subg. *Eucalyptus*) according to the methods described above. Samples were collected from natural populations (Appendix 1). Sixteen out of 17 (94%) loci cross-amplified within subg. *Symphyomyrtus*, and 14 out of 17 (82%) cross-amplified across the subgenus boundary to subg. *Eucalyptus*, demonstrating high transferability rates within the genus.

Table 2. Population genetic properties of the 11 newly developed EST-SSRs in two natural populations of *Eucalyptus gomphocephala*.[a]

	Yalgorup National Park (n = 30)			Ludlow Tuart Forest (n = 30)		
Locus	A_T	H_o	H_e	A_T	H_o	H_e
EGM09	4	0.400	0.639	4	0.233	0.558
EGM12	4	0.600	0.629	4	0.621	0.640
EGM14	7	0.733	0.779	8	0.733	0.776
EGM24	3	0.345	0.580	3	0.172	0.506
EGM25	2	0.233	0.255	2	0.207	0.238
EGM30	9	0.667	0.728	12	0.833	0.839
EGM35	9	0.667	0.832	6	0.767	0.748
EGM37	2	0.433	0.495	2	0.300	0.406
EGM46	12	0.867	0.816	9	0.862	0.861
EGM47	3	0.533	0.515	3	0.500	0.452
EGM48	6	0.833	0.776	5	0.862	0.794

Note: A_T = total number of alleles observed; H_e = expected heterozygosity; H_o = observed heterozygosity; n = number of individuals sampled.

[a] Geographic coordinates for each population are: Yalgorup National Park = 32°51′17″S, 115°39′54″E; Ludlow Tuart Forest = 33°34′42″S, 115°29′42″E.

CONCLUSIONS

This new set of 11 polymorphic EST-SSR loci will enable the characterization of population genetic diversity and structure throughout the species' range of *E. gomphocephala* in conjunction with putatively neutral, nuclear genomic SSRs (gSSRs). Given their gene-linked nature, the EST-SSRs can be used to test for signatures of selection in natural populations. The high transferability of the loci demonstrates their suitability for application to other species of ecological, conservation, or economic importance in the genus. The broader set of 133 EST-SSRs represents a resource that could be exploited to optimize additional gene-linked EST-SSRs for *Eucalyptus* and further expand the types of genes available for investigation. Data obtained using these markers will be particularly valuable to inform seed sourcing and conservation strategies in natural *E. gomphocephala* populations.

LITERATURE CITED

Benson, G. 1999. Tandem repeats finder: A program to analyze DNA sequences. *Nucleic Acids Research* 27: 573–580.

Brooker, M. I. H. 2000. A new classification of the genus *Eucalyptus* L'Hér. (Myrtaceae). *Australian Systematic Botany* 13: 79–148.

Cai, Y. F., P. Barber, B. Dell, P. O'Brien, N. Williams, B. Bowen, and G. Hardy. 2010. Soil bacterial functional diversity is associated with the decline of *Eucalyptus gomphocephala. Forest Ecology and Management* 260: 1047–1057.

Gelfand, Y., A. Rodriguez, and G. Benson. 2007. TRDB: The Tandem Repeats Database. *Nucleic Acids Research* 35: D80–D87.

Huang, X., and A. Madan. 1999. CAP3: A DNA sequence assembly program. *Genome Research* 9: 868–877.

Peakall, R., and P. 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 ecumenicism. *Journal of Heredity* 86: 248–249.

Rousset, F. 2008. GENEPOP'007: A complete reimplementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* 8: 103–106.

Rozen, S., and H. 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.

Downloaded From: https://bioone.org/journals/Applications-in-Plant-Sciences on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Steane, D. A., D. Nicolle, and B. M. Potts. 2007. Phylogenetic positioning of anomalous eucalypts by using ITS sequence data. *Australian Systematic Botany* 20: 402–408.

Steane, D. A., D. Nicolle, C. P. Sansaloni, C. D. Petroli, J. Carling, A. Kilian, A. A. Myburg, et al. 2011. Population genetic analysis and phylogeny reconstruction in *Eucalyptus* (Myrtaceae) using high-throughput, genome-wide genotyping. *Molecular Phylogenetics and Evolution* 59: 206–224.

van Oosterhout, C., W. F. Hutchinson, D. P. M. Willa, and P. Shipley. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535–538.

Appendix 1. Cross-species amplification of 17 EST-SSR loci from *Eucalyptus gomphocephala* to three additional eucalypt species from two subgenera.[a,b,c]

Locus	Annotation[d]	Source[e,f]	*E. gomphocephala*	*E. camaldulensis*	*E. victrix*	*E. marginata*
EGM09	Glucan endo-1,3-beta-glucosidase	*E. globulus*	+	+	+	+
EGM12	CONSTANS-like protein CO1	*E. tereticornis*	+	+	+	+
EGM14	Heat shock transcription factor 21	*E. grandis*	+	+	+	+
EGM24	Putative S-adenosylmethionine decarboxylase (SAMDC)	*E. globulus*	+	+	+	+
EGM25	Metal transporter Nramp3	*E. globulus*	+	NA	+	NA
EGM30	Chlorophyll A/B binding protein, putative	*E. gunnii*	+	+	+	+
EGM35	Fructose-1,6-bisphosphatase, cytosolic	*E. globulus*	+	+	+	+
EGM37	Quinone oxidoreductase	*E. globulus*	+	+	+	+
EGM46	Desiccation protectant protein Lea14 homolog	*E. gunnii*	+	+	+	+
EGM47	Small GTPase Rab2	*E. grandis* + *E. gunnii*	+	+	+	+
EGM48	Aquaporin PIP2;1	*E. gunnii*	+	+	+	+
EGM28	Chloroplastic phosphoglycerate kinase	*E. globulus* + *E. grandis*	+	+	+	NA
EGM19	Aquaporin (PIP1)	*E. globulus* + *E. gunnii*	+	+	+	+
EGM33	Synaptobrevin-related protein 1 (SAR1)	*E. tereticornis*	+	+	+	+
EGM34	Rapid alkalinization factor (RALF)	*E. globulus*	+	+	+	+
EGM26	Magnesium transporter CorA-like family protein MRS2-2	*E. globulus*	+	+	NA	NA
EGM42	bZIP (basic leucine zipper) transcription factor	*E. globulus*	+	+	+	+

Note: + = positive amplification; NA = no amplification or inconsistent amplification.

[a] *Eucalyptus gomphocephala* is taxonomically classified within the monotypic sect. *Bolites* (Brooker, 2000) but has been shown to have affinities with sect. *Bisectae I* according to ITS (Steane et al., 2007) and Diversity Arrays Technology (DArT) analyses (Steane et al., 2011).

[b] Taxonomic information for tested species: *E. gomphocephala* DC. = sect. *Bolites/Bisectae I*, subg. *Symphyomyrtus*; *E. camaldulensis* Dehnh. = sect. *Exsertaria*, subg. *Symphyomyrtus*; *E. victrix* L. A. S. Johnson & K. D. Hill = sect. *Adnataria*, subg. *Symphyomyrtus*; *E. marginata* Sm. = sect. *Longistylus*, subg. *Eucalyptus*.

[c] Source locations: *E. camaldulensis* = 23°06′S, 119°34′E (n = 2) and 23°10′S, 119°45′E (n = 1); *E. victrix* = 22°55′S, 119°10′E (n = 3); *E. marginata* = 31°56′S, 115°46′E (n = 1), 32°46′S, 116°26′E (n = 1), and 32°42′S, 116°03′E (n = 1).

[d] Predicted gene annotation based on BLAST searches.

[e] Species from which the original EST was derived.

[f] Taxonomic information for source species: *E. globulus* Labill. = sect. *Maidenaria*, subg. *Symphyomyrtus*; *E. gunnii* Hook. f. = sect. *Maidenaria*, subg. *Symphyomyrtus*; *E. tereticornis* Sm. = sect. *Exsertaria*, subg. *Symphyomyrtus*; *E. grandis* W. Hill = sect. *Latoangulatae*, subg. *Symphyomyrtus*.

Downloaded From: https://bioone.org/journals/Applications-in-Plant-Sciences on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use