Development and Characterisation of Microsatellite Loci in *Eucalyptus globulus* (Myrtaceae)

By D. A. STEANE[1]), R. E. VAILLANCOURT, J. RUSSELL[2]), W. POWELL[2]), D. MARSHALL[2]) and B. M. POTTS

Cooperative Research Centre for Sustainable Production Forestry, University of Tasmania, GPO Box 252-55, Hobart, Tasmania 7001, Australia

(Received 11th January 2001)

Summary

Eucalyptus globulus LABILL. is the premier hardwood plantation species in temperate regions of the world. We developed twelve *E. globulus* microsatellite loci for fingerprinting and future studies in genome mapping, gene flow and genetic diversity. The microsatellites were highly polymorphic in *E. globulus* (average number of alleles per locus, 17.8; average observed heterozygosity, 0.66). The loci were robust, amplifying in other *Eucalyptus* species, *Corymbia* and *Angophora*. The high variability and transferability of these molecular markers make them useful in *E. globulus*, as well as in many of the 700 species of *Eucalyptus*.

Key words: SSR, genetic diversity, molecular markers, primer, Tasmanian blue gum, *Angophora*, *Corymbia*.

Introduction

Eucalyptus globulus is one of the world's premier paper and pulpwood species (ELDRIDGE *et al.*, 1993). With the timber industry moving towards intensive plantation forestry, construction of genetic linkage maps, genetic fingerprinting, and studies of gene flow, population structure and paternity are becoming increasingly important. Such studies may be aimed at maximising outcrossing in seed orchards, monitoring gene flow out of plantations into native forest (genetic pollution), or simply at determining accurately the identity of breeding material (genotype fingerprinting), an important issue in quality control. The high allelic diversity and abundance of microsatellites in the eukaryotic genome (BRONDANI *et al.*, 1998) make these codominant molecular markers popular for such studies. We report the development of 12 microsatellite loci (SSRs) from *E. globulus*. The SSR primer sets developed here were also tested for interspecies amplification on a range of other species of *Eucalyptus*, as well as on the closely related genera, *Angophora* and *Corymbia*.

Materials and Methods

Twelve *Eucalyptus globulus* microsatellites were identified using the protocol of WHITE and POWELL (1997). AC-rich clones were sequenced using an ABI Prism Ready Reaction Dye Terminator Cycle Sequencing Kit. Sequencing products were separated on an ABI 377 automated sequencer. PCR primers were designed using OLIGO software (Molecular Biology Insights Inc., USA). Two sets of primer pairs were developed for locus EMCRC 1 (*Table 1*). PCRs used a total volume of 25 µl containing 10 ng to 20 ng DNA, PCR buffer (67 mM Tris-HCl, pH 8.8, 16.6 mM $(NH_4)_2SO_4$, 0.45% Triton X-100, 0.2 mg/ml gelatine), 0.1 µg/µl BSA, 120 µM dNTPs, 200 nM HEX-, FAM- or TET-labelled forward primer and 200 nM reverse primer, $MgCl_2$ according to *table 1*, and 1.5 U *Taq* polymerase. PCR conditions (using a PTC-100, MJ Research, Inc.) were: denaturation at 94 °C for 5 min; 30 cycles of denaturation at 94 °C for 1 min, annealing at T_a (*Table 1*) for 30 sec, and extension at 72 °C for 30 sec. An ABI 377 automated sequencer was used to separate microsatellite alleles, with TAMARA-500 as a size standard; electrophoretic output was recorded using Genescan 3.1 software (PE Biosystems); alleles were sized using Genotyper 1.1 software (PE Biosystems).

The main study included 88 samples of *E. globulus* (*sensu* BROOKER, 2000) and intergrades (between *E. globulus* and *E. bicostata* or *E. pseudoglobulus*) from 11 of the 13 races identified by DUTKOWSKI and POTTS (1999; 4 to 14 samples per race, excluding Dromedary and Recherche Bay). Mean number of alleles per locus (N_a), mean observed (H_o) and expected (H_e) heterozygosities were calculated using POPGENE (YEH *et al.*, 1997).

The "species-compatibility" study used forty additional samples of *Eucalyptus*, including six other species from subgenus *Symphyomyrtus* (*E. bicostata*, *E. maidenii*, *E. pseudoglobulus* and *E. nitens* [section *Maidenaria*]; *E. urophylla* and *E. grandis* [section *Latoangulatae*]); four species from subgenus *Eucalyptus* (*E. regnans*, *E. obliqua* [section *Eucalyptus*]; *E. risdonii* and *E. amygdalina* [section *Aromatica*]); two species from subgenus *Eudesmia* (*E. gongylocarpa* [section *Limbatae*] and *E. baileyana* [section *Reticulatae*]); 2 species of *Angophora* (*A. floribunda* and *A. melanoxylon*); and 2 species of *Corymbia* (*C. aparrerinja* [section *Blakearia*] and *C. ficifolia* [section *Rufaria*]). Nomenclature follows BROOKER (2000), except that we maintain *Angophora* and *Corymbia* as genera. Five of the *E. globulus* samples from the first study were included as controls. For the species-compatibility survey, PCR products were separated on 2% agarose gels (Agarose 1000, GibcoBRL/Life Technologies), stained with ethidium bromide and visualised with UV light.

Results and Discussion

Out of 105 positive clones sequenced, primer pairs were designed for 13 (12%). However, after mapping the microsatellite loci (BUNDOCK *et al.*, in press), two of these primer pairs were found to amplify the same locus (EMCRC 1a and 1b; *Table 1*). The EMCRC primers worked consistently in control samples of *E. globulus* as well as in the three closely related taxa, *E. maidenii*, *E. bicostata* and *E. pseudoglobulus*. Most primers worked well in *E. nitens* (Series *Globulares*), but success declined as the taxonomic distance from *E. globulus* increased (see STEANE *et al.*, 1999; optimisation of PCRs may increase success rate in some taxa). Overall, the primers are robust, with at least seven out of the twelve (EMCRC 2, 3, 4, 8,

[1]) Correspondence: D. A. STEANE. Fax: +61 (0)3 62262698; Email: Dorothy.Steane@utas.edu.au

[2]) Department of Cell and Molecular Genetics, Scottish Crop Research Institute, Invergowrie, Dundee DD2 5DA, Scotland, UK.

Table 1. – *Eucalyptus globulus* SSR primers used in this study: EMBL accession numbers of the sequenced *E. globulus* clones; forward (fwd) and reverse (rev) primer sequences (5'–3'); PCR parameters (T_a, annealing temperature; Mg^{2+}, $MgCl_2$ concentration); sample size (N, number of individual trees); observed number of alleles per locus (N_a); observed heterozygosity (H_o); expected heterozygosity (H_e).

Locus	EMBL accession no.	Cloned repeat	Primer sequences	Product size range (bp)	T_a (°C)	Mg^{2+} (mM)	N	N_a	Heterozygosity H_o	Heterozygosity H_e
EMCRC 1a	AJ401136	$(CA)_{16}$	fwd GCATGGACACCCTTTTTC	151-203	50	1.5	84	18	0.54	0.88
			rev ATTGAGAATGCTGAACCAAAC							
EMCRC 1b	AJ401136	$(CA)_{16}$	fwd ATGCCGCACTTGGAAGC	295-342	55	1.0	"	"	"	"
			rev GGACTGAAAGCCCATTGAGAA							
EMCRC 2	AJ401137	$(CT)_9$--$(CA)_{10}$	fwd GCGACTGTGTGGCTTTC	157-189	55	2.5	85	17	0.68	0.85
			rev CCCAATCATTTTTCATTTTGA							
EMCRC 3	AJ401138	$(CA)_{10}$	fwd AGATGGGGTTTCTCATGGTTT	109-145	55	2.0	85	14	0.85	0.85
			rev ACCGTACTATGCAGCTGGAAC							
EMCRC 4	AJ401139	$(AC)_{17}$	fwd GTAATCTTTCATTCTCCGACC	178-268	55	2.0	78	20	0.82	0.69
			rev CTCGAGGACATGTTGAGTG							
EMCRC 5	AJ401140	$(CT)_{19}(CA)_{13}$	fwd GTTTCTTCCTCTGCTTGTTGC	202-248	60	1.0	77	19	0.55	0.81
			rev GATGGGTTCGGATTTAGGC							
EMCRC 6	AJ401141	$(CT)_{11}(CA)_{23}$	fwd CTTCAAGGTTCACAGATGG	151-193	50	1.5	84	18	0.67	0.85
			rev TCTTCATAAGTCCCCTAATCA							
EMCRC 7	AJ401142	$(TG)_{15}(AG)_8$	fwd CGAATCAAGTCGACATGTGTG	271-309	60	1.0	81	17	0.68	0.85
			rev CCGTCGACCGCCCTAT							
EMCRC 8	AJ401143	$(CT)_{13}(CA)_{24}$	fwd CCAGATTGTAGCCCTTATGTG	231-265	55	2.0	79	18	0.67	0.92
			rev CATCCCAATCAAACGAAC							
EMCRC 9	AJ401144	$(TG)_{14}$	fwd CTGGGCTGTGCATCTCTGAAA	286-342	55	1.0	74	20	0.31	0.79
			rev GACCCGGTCAACTCCTCAAGA							
EMCRC 10	AJ401145	$(GT)_{19}(GA)_9$	fwd GCTTGGTCGGGTAGGAA	312-344	55	1.0	69	15	0.57	0.88
			rev TCGGGTTGATGTCCTTATTGT							
EMCRC 11	AJ401146	$(TC)_{10}(AC)_{10}$	fwd AACTGACTGTGGATTTGAAGC	221-255	55	1.0	85	17	0.81	0.91
			rev GTGAGTCATTATTTGGCAACC							
EMCRC 12	AJ401147	$(CT)_8(CA)_{14}$	fwd CTCCGACCTCCTCCACT	70-128	50	1.5	82	21	0.79	0.86
			rev AATCGTCTTCATCGAATCAAG							
Mean							80.2	17.8	0.66	0.85
Std. Dev.							5.1	2.0	0.15	0.06

9, 11 and 12) working in the closely related genera, *Angophora* and *Corymbia* (*Table 2*).

All 12 SSR loci were polymorphic in *E. globulus*, with the number of alleles per locus ranging from 14 to 21 (mean 17.8; *Table 1*). This figure is somewhat greater than those reported by BRONDANI *et al.* (1998) in *E. urophylla* and *E. grandis* (mean N_a=11.4, across 15 loci) or by BYRNE *et al.* (1996) in *E. nitens* (mean N_a= 9.5, across 4 loci). Our slightly higher values of N_a can be explained by the relatively large sample size (88 individuals) that spanned the entire geographic range of *E. globulus*, in contrast to the smaller sample sizes in the other two studies (16 individuals each of *E. grandis* and *E. urophylla* [BRONDANI *et al.*, 1998], and 20 individuals of *E. nitens* [BYRNE *et al.*, 1996]). Each locus had a few alleles that occurred with relative-

Table 2. – Compatibility of *E. globulus* microsatellite primers with other taxa. Numbers in parentheses following species names indicate sample sizes greater than 1; 1 = successful PCR; 0 = absence of any PCR product; ? = presence of non-specific PCR products (PCR optimisation required). Taxonomic classification follows BROOKER (2000), except that we maintain *Angophora* and *Corymbia* as genera.

Taxon	EMCRC primers											
	1a	1b	2	3	4	5	6	7	8	9	10	11	12
Eucalyptus													
Subgenus *Symphyomyrtus*													
Section *Maidenaria: E. bicostata* (6), *E. maidenii* (6), *E. pseudoglobulus* (6), *E. nitens* (6)	1	1	1	1	1	1	1	1	1	1	1	1	1
Section *Latoangulatae: E. urophylla* (3), *E. grandis* (3)	?	?	1	1	1	1	1	1	1	1	?	1	1
Subgenus *Eucalyptus*: *E. regnans, E. obliqua, E. risdonii, E. amygdalina*	?	1	1	1	?	0	1	1	1	1	?	1	1
Subgenus *Eudesmia*: *E. gongylocarpa, E. baileyana*	?	0	?	1	1	1	1	1	1	1	?	1	1
Corymbia aparrerinja, C. ficifolia	1	1	1	1	1	0	0	?	1	1	?	1	1
Angophora floribunda, A. melanoxylon	?	1	1	1	?	0	0	0	1	1	?	1	1

ly high frequency across the geographic range, as well as numerous rare alleles. This observation is supported by the fact that the allele frequency distributions were multi-modal (usually bi-modal), exhibiting many rare alleles. There was no correlation between the modality of the allele frequency distribution and whether the SSR was simple or compound. Of the 12 loci, four were perfect simple repeats (EMCRC 1, 3, 4 and 9), seven were perfect compound repeats (EMCRC 5, 6, 7, 8, 10, 11, 12) and one (EMCRC 2) was an imperfect compound repeat (*Table 1*).

As expected, the level of heterozygosity (H_e) for the microsatellites is much higher than that reported in allozyme studies (0.10 to 0.34) of *Eucalyptus* (POTTS and WILTSHIRE, 1997). Our mean values of H_o and H_e (0.66 and 0.85, respectively) are comparable to those based on microsatellite data reported by BRONDANI *et al.* (1998) in *E. urophylla* and *E. grandis* (H_o = 0.57; H_e = 0.89) and by BYRNE *et al.* (1996) in *E. nitens* (H_o = 0.58; H_e = 0.83). In all these studies, the level of observed heterozygosity is less than that expected in a random mating, outbreeding population. Such deviation is observed in the majority of isozyme studies of *Eucalyptus* (POTTS and WILTSHIRE, 1997), and could arise through geographical structure within the species (as has been demonstrated with quantitative genetic data; DUTKOWSKI and POTTS, 1999) or as a result of the mixed mating system of *E. globulus* (HARDNER *et al.*, 1996).

These SSR primers will be invaluable for routine genetic fingerprinting of selections for breeding and deployment, as well as for studies of gene flow in natural and plantation forests of *E. globulus* and related taxa. We are currently using these SSRs to produce a database of genetic diversity within *E. globulus* for fingerprinting and quality control. Such data will be utilised for studies of geographic partitioning of genetic diversity and for quantifying the relationships between, as well as the inbreeding levels within, the various races of *E. globulus*. Such information is required in order to refine quantitative genetic models currently being used for breeding value prediction.

Acknowledgments

This research was funded by Shell Forestry U.K. and the Cooperative Research Centre for Sustainable Production Forestry. We would like to thank Mr. S. SKABO for technical assistance, and Mr. G. DUTKOWSKI and Prof. J. REID for their support and advice.

References

BRONDANI, R. P. V., BRONDANI, C., TARCHINI, R., GRATTAPAGLIA, D.: Development, characterisation and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla*. Theor. Appl. Genet. **97**: 816–827 (1998). — BROOKER, M. I. H.: A new classification of the genus *Eucalyptus* L'Hér. (Myrtaceae). Aust. Syst. Bot. **13**: 79–148 (2000). — BUNDOCK, P. C., HAYDEN, M. and VAILLANCOURT, R. E.: RAPD and microsatellite linkage maps from an intraprovenance cross in *Eucalyptus globulus*. Silvae Genet. **49**: 223–232 (2000). — BYRNE, M., MARQUEZ-GARCIA, M. I., UREN, T., SMITH, D. S. and MORAN, G. F.: Conservation and genetic diversity of microsatellite loci in the genus *Eucalyptus*. Aust. J. Bot. **44**: 331–341 (1996). — BYRNE, M., MORAN, G. F., MURRELL, J. C. and TIBBITS, W. N.: Detection and inheritance of RFLPs in *Eucalyptus nitens*. Theor. Appl. Genet. **89**: 397–402 (1994). — DUTKOWSKI, G. W. and POTTS, B. M.: Geographic patterns of genetic variation in *Eucalyptus globulus* ssp. *globulus* and a revised racial classification. Aust. J. Bot. **47**: 237–263 (1999). — ELDRIDGE, K., DAVIDSON, J., HARWOOD, C. and VAN WYK, G.: Eucalypt domestication and breeding. Oxford University Press, Oxford (1993). — HARDNER, C. M., VAILLANCOURT, R. E and POTTS, B. M.: Stand density influences outcrossing rate and growth of open-pollinated families of *Eucalyptus globulus*. Silvae Genet. **45**: 226–228 (1996). — POTTS, B. M. and WILTSHIRE, R. J. E.: Eucalypt genetics and genecology. In: Eucalypt Ecology: Individuals to Ecosystems. WILLIAMS, J. and WOINARSKI, J. (eds.): 56–91. Cambridge University Press, Cambridge (1997). — STEANE, D. A., MCKINNON, G. E., VAILLANCOURT, R. E. and POTTS, B. M.: ITS sequence data resolve higher level relationships among the eucalypts. Mol. Phylog. Evol. **12**: 215–223 (1999). — WHITE, G. and POWELL, W.: Cross-species amplification of SSR loci in the Meliaceae family. Mol. Ecol. **6**: 1195–1197 (1997). — YEH, F. C., YANG, R.-C., BOYLE, T. B. J., YE, Z.-H. and MAO, J. X.: POPGENE, the user friendly shareware for population genetic analysis. Molecular Biology and Biotechnology Centre, University of Alberta, Canada (1997).