

## Multiplex and characterization of new EST-derived microsatellite and transferability among five *Eucalyptus* species (Myrtaceae)

Multiplex e caracterização de marcadores microssatélites derivados de sequências EST e transferabilidade para cinco espécies de *Eucalyptus* (Myrtaceae)

Julio Cezar Santos Otto<sup>1</sup>, Leonardo Curi Martin<sup>1</sup>, Cesar Augusto Valencise Bonine<sup>2</sup>, Vanusa do Socorro Leite Otto<sup>1</sup>, Bruno Cesar Rossini<sup>1</sup> and Celso Luis Marino<sup>1</sup>

### Abstract

Polymorphic and transferable microsatellite loci among five *Eucalyptus* species (*Eucalyptus grandis*, *Eucalyptus urophylla*, *Eucalyptus tereticornis*, *Eucalyptus saligna* and *Eucalyptus brassiana*) were developed in order to use them in further studies of population structure, conservation, forest management and breeding programs of the genus. Four new microsatellite loci were isolated from the database *Eucalyptus* Sequencing Project Consortium (FORESTs). Allele numbers ranged from two to ten, the observed heterozygosity ( $H_o$ ) ranged from 0 to 1 and the expected heterozygosity ( $H_e$ ) ranged from 0.255 to 0.792. Microsatellite loci described demonstrated a great variation and reliable findings that could support other future studies focusing breeding programs and related areas.

**Keywords:** breeding programs; forest management; genotyping.

### Resumo

Marcadores microssatélites polimórficos e transferíveis para cinco espécies de *Eucalyptus* (*Eucalyptus grandis*, *Eucalyptus urophylla*, *Eucalyptus tereticornis*, *Eucalyptus saligna* e *Eucalyptus brassiana*) foram desenvolvidos para uso em estudos no gênero de estruturação populacional, conservação, programas de manejo e cruzamentos. Quatro novos marcadores foram isolados do banco de dados *Eucalyptus* Sequencing Project Consortium (FORESTs). O número de alelos variou de dois a dez, heterozigosidade observada ( $H_o$ ) de 0 a 1 e a heterozigosidade esperada ( $H_e$ ) de 0.255 a 0.792. Os marcadores microssatélites aqui descritos demonstram uma grande variação e consistência que suportam futuros estudos baseados em programas de melhoramento e áreas correlacionadas.

**Palavras-chave:** genotipagem; manejo florestal; programas de melhoramento.

## INTRODUCTION

The genus *Eucalyptus* comprises a wide diversity of varieties and hybrids, with more than 900 species described (BOLAND et al., 2006). Within the subgenera *Symphomyrtus* are the major and important commercially species, such as *E. grandis*, *E. urophylla*, *E. tereticornis*, *E. saligna*, *E. brassiana*, *E. globulus* and *E. camaldulensis* because of rapid growth, high productivity, adaptability and use for different purposes, from construction to cellulose and papers. These characteristics are the main reason for using this genus in forest plantations in Brazil and others countries (MORA; GARCIA 2000).

Microsatellite markers (or SSR-Simple Sequence Repeats) have provided a powerful tool for population genetics studies and individual identification in *Eucalyptus* and other plants in the neotropics (JONES et al., 2008; XAVIER et al., 2014). Also, SSRs can be found in coding regions and have been characterized for *Eucalyptus* species based on public database of expressed sequences tags (ESTs) (FARIA et al., 2011; HE et al., 2012). The availability of polymorphic and transferable microsatellite loci in the most studied and cultivated species of *Eucalyptus* are important for the

1. Biotechnology Institute, Universidade Estadual Paulista "Júlio de Mesquita Filho", Botucatu, SP, Brazil.

2. Suzano SA, Jacareí, SP, Brazil.

Corresponding author: bu\_rossini@yahoo.com.br

comparative population studies and genetic breeding, especially with the great concern in exploring hybrid combinations. In contrast to SSRs derived from random genomic sequences, those developed from ESTs are generally more robust and obtain a better resolution of alleles and higher interspecific transferability (VARSHNE et al., 2005).

In this study, we report on the development and characterization of four microsatellite markers derived of expressed sequence tags from the non-public database *Eucalyptus* Sequencing Project Consortium (FORESTs). This database has over than 100,000 ESTs generated from different types of tissues of five eucalyptus species (*E. grandis*, *E. urophylla*, *E. globulus*, *E. saligna* and *E. camaldulensis*) (VICENTINI et al., 2005). We also tested these loci in five species widely cultivated in the world (*E. grandis*, *E. urophylla*, *E. tereticornis*, *E. saligna* and *E. brassiana*) and the hybrid *E. grandis* x *E. urophylla*, in order to select the markers with full transferability between species, germplasm characterization and genetic map construction.

## MATERIALS AND METHODS

### Sampling and DNA extraction

Genomic DNA of 89 individuals (*E. grandis* n= 7, *E. urophylla* n=7, *E. tereticornis* n=7, *E. saligna* n=7, *E. brassiana* n=7 and the hybrid *E. grandis* x *E. urophylla* n=54) from breeding programs of forest companies in Brazil were extracted from leaves using modified CTAB method (incubation time in extraction buffer for 60 minutes; inclusion of a precipitation step with CTAB 10% and NaCl 1,4M - 60 µl - and increased centrifugation time for all steps - 20 minutes; Ferreira and Grattapaglia (1998)).

### Markers development and PCR amplification

Microsatellite markers in ESTs were searched with di, tri, tetra and pentanucleotide repeats in the database FORESTs in regions related to abiotic stress and primers designed by Primer 3 software (UNTERGASSER, 2012).

Polymerase chain reactions were carried out in 10 µl at the following concentrations: 1.5 mM MgCl<sub>2</sub>, 0.2 mM of each dNTP, 0.5 mM of each primer, 1 U Taq DNA polymerase (Promega) and 30 ng of genomic DNA. Thermocycling followed: 3 min at 95 °C, 32 cycles of 1 min at 95 °C, 55 °C for 1 min, 72 °C for 1 min and final extension at 72 °C for 10 min. Also, two PCR multiplex systems consisting of three pairs of primers each could be used as follows: (1) VCP3, VCP8 and VCP10 and (2) VCP12, VCP16 and VCP17. Both multiplex had better results when using 54°C of annealing temperature. Successful amplification was determined by 1,5% agarose gel electrophoresis stained with 1:20 ethidium bromide (Life Technologies).

### Data analysis

Genotyping was performed on an automated sequencer ABI 3100 (Applied Biosystems), consisting of 1µl of the PCR product, 0.4µl of GeneScan 500 ROX and 8.6µl of Hi-Di formamide (Applied Biosystems). Primer sequences, annealing temperature and repeat motifs are shown in Table 1. Number of alleles (A), observed heterozygosity (Ho), expected heterozygosity (He), and fixation index (F) were calculated using the software POPGENE v1.32 (YEH et al., 1997). Blastx tool was used for annotation of the sequences, as also the BLAST tool in Phytozome v12.1.5 (GOODSTEIN et al., 2012).

## RESULTS AND DISCUSSION

A total of 520 probable SSR-ESTs were identified by software Genemarker (Softgenetics; data not shown). SSR-ESTs with trinucleotide motifs were the most abundant 362 (69.6%), followed by di- 117 (22.5%) and tetra- 29 (5.57%), while 12 pentanucleotide (2.33%) had a lower representative rate. From these, 19 pairs of primer were selected and tested, but only six primers pairs had successful amplification and polymorphism. Forward sequences were then marked at the 5' end with fluorophores (NED, 6-FAM or HEX). In a search in the NCBI EST database to avoid duplication of markers, two of them had some significant similarity against previous described microsatellite markers (VCP10 similar to EUCeSSR479 from He *et al.*, 2012 and VCP16 similar to EMBRA1732 from Grattapaglia et al., (2015)) and were excluded from description, but not from transferability analysis, being also possible to multiplex them with the novel markers (Table 2). Among all these polymorphic loci

**Table 1.** Characteristics of the four novel transferable microsatellite loci developed for *Eucalyptus* genus.

**Tabela 1.** Características dos quatro novos marcadores microsatélites desenvolvidos para o gênero *Eucalyptus*.

Locus	Primer sequence (5' - 3')	Repeat motif	Annealing temperature (°C)	Putative function	BLASTX E-value	GenBank accession number	Phytozome BLAST ( <i>Eucalyptus grandis</i> )
VCP3	F CAAAAACCCTGCAAAAACCTC	(CAG) <sub>7</sub>	55	uncharacterized protein LOC104456484 [Eucalyptus grandis]	2e-83	XP_010069593	Eucgr.H00722.1 (Chr08)
	R GAGTCGAGGAACGAGACGAG						
VCP8	F GCTCGACTTGATGCCAGTG	(CGG) <sub>8</sub> (AGG) <sub>5</sub>	55	heat stress transcription factor B-2b-like [Eucalyptus grandis]	7e-147	XP_010061278.1	Eucgr.A02804.1 (Chr01)
	R CCTCCCTTCTCTGATTCC						
VCP12	F CGATCCATTCGCCACTTAC	(ACG) <sub>6</sub>	55	PREDICTED: Eucalyptus grandis probable inactive shikimate kinase like 1, chloroplastic (LOC104419115), mRNA	0.0	XM_010030661.2	Eucgr.i01618.1 (Chr09)
	R GATTCCGACGAAGAGATGC						
VCP17	F CTCGATGGTGCCGAAAATC	(CCT) <sub>6</sub> (CCT) <sub>8</sub>	55	PREDICTED: Eucalyptus grandis uncharacterized LOC104443136 (LOC104443136), mRNA	0.0	XM_010056484.2	Eucgr.D01909.1 (Chr04)
	R GTGATCTCGACGGAGACTGG						

**Table 2.** Characteristics of the six microsatellite loci multiplex system for *Eucalyptus*. A = number of alleles sampled; He = expected heterozygosity; Ho = observed heterozygosity; N = number of individuals sampled. \*a = sequence with marker EUCeSSR479 previously described by He et al., (2012); \*b = sequence with marker EMBRA1732 previously described by Grattapaglia et al., 2015.

**Tabela 2.** Características dos seis locos microssatélites utilizados em sistema multiplex para *Eucalyptus*. A = número de alelos; He = heterozigosidade esperada; Ho = heterozigosidade observada; N = número de indivíduos amostrados; \*a = sequência com marcador EUCeSSR479 previamente descrito por He et al., (2012); \*b = sequência com marcador EMBRA1732 previamente descrito por Grattapaglia et al., 2015.

Locus	Allele size range (bp)	<i>E. grandis</i> x <i>E. urophylla</i> (N= 54)			<i>E. grandis</i> (N= 7)			<i>E. urophylla</i> (N= 7)		
		A	Ho	He	A	Ho	He	A	Ho	He
VCP3	166-199	6	0,922	0,747	4	1	0,722	2	1	0,5
VCP8	166-196	7	0,843	0,725	2	0,333	0,278	4	0,667	0,722
VCP10*a	217-244	7	0,745	0,792	5	0,714	0,735	4	0,286	0,694
VCP12	176-212	7	0,902	0,788	3	0,286	0,255	4	0,667	0,625
VCP16*b	188-245	10	0,904	0,756	3	0,571	0,5	4	1	0,708
VCP17	152-200	7	0,765	0,636	3	0,5	0,5	3	0,286	0,571
Locus	Allele size range (bp)	<i>E. saligna</i> (N= 7)			<i>E. tereticornis</i> (N= 7)			<i>E. brassiana</i> (N= 7)		
		A	Ho	He	A	Ho	He	A	Ho	He
VCP3	166-199	3	1	0,594	2	1	0,5	5	1	0,75
VCP8	166-195	3	1	0,602	4	0,5	0,597	6	0,714	0,745
VCP10*a	217-244	5	0,5	0,764	4	0,667	0,681	5	0,429	0,704
VCP12	176-212	3	0,857	0,541	2	0	0,32	3	0,286	0,439
VCP16*b	188-245	3	0,714	0,52	3	0,286	0,357	4	0,5	0,597
VCP17	152-200	3	0	0,625	2	0	0,32	5	0,5	0,764

tested, the number of alleles found ranged from two to ten, the observed heterozygosity (Ho) ranged from 0 to 1, the expected heterozygosity (He) ranged from 0.255 to 0.792 for all species tested. For *E. grandis* x *E. urophylla* the number of alleles ranged from six to ten, with observed heterozygosity ranging from 0.745 to 0.922 and expected heterozygosity from 0.636 to 0.792.

The six microsatellite loci were highly transferable among *Eucalyptus* species tested and also can be used as multiplex system. These markers were classified as transferable because there was amplification and polymorphism in all species and the hybrid tested. Loci with high transferability between various species are useful genetic markers for interspecific research on topics such as speciation and hybridization (THOMSON et al., 2015). Aside from two markers that were already described by other studies, the loci EUCeSSR479 were only tested for *E. urophylla* and *E. tereticornis* and EMBRA1732 only for *E. grandis* and *E. globulus*, demonstrating herein that they can be used at least in other two species (*E. saligna* and *E. brassiana*) and a hybrid (*E. grandis* x *E. urophylla*) (Table 2).

## CONCLUSIONS

These polymorphic microsatellite loci may become useful tools for future studies on mating system and population structure of *Eucalyptus* species cited. The characterization of microsatellite loci transferable also applies to the conservation, management, forest breeding and research on topics such as hybridization and speciation.

## REFERENCES

- BOLAND, D.; BROKER, M. H.; CHIPPENDALE, G. M.; HALL, N.; HYLAND, B. P. M.; JOHNSON, R. D.; KLEINING, D. A.; MCDONALD, M. W.; TURNER, J. D. **Forest Trees of Australia**. 5 ed. Melbourne: CSIRO, 2006. 736 p.
- FARIA, D. A.; MAMANI, E. M. C.; PAPPAS, G. J.; GRATTAPAGLIA, D. Genotyping systems for *Eucalyptus* based on tetra-, penta-, and hexanucleotide repeat EST microsatellites and their use for individual fingerprinting and assignment tests. **Tree Genetics & Genomes**, v. 7, n. 1, p 63-77, 2011.
- FERREIRA, M.E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análises genéticas**. Brasília: Embrapa-Cenagen, 1998. 220 p.

- GOODSTEIN, D. M.; SHU, S.; HOWSON, R.; NEUPANE, R.; HAYES, R. D.; FAZO, J.; MITROS, T.; DIRKS, W.; HELLSTEN, U.; PUTNAM, N.; ROKHSAR, D. S. Phytozome: a comparative platform for green plant genomics. **Nucleic Acids Research**, v.40, n. 1, p .1178 - 1186, 2012.
- GRATTAPAGLIA, D.; MAMANI, E. M. C.; SILVA JUNIOR, O. B.; FARIA, D. A. A novel genome-wide microsatellite resource for species of *Eucalyptus* with linkage-to-physical correspondence on the reference genome sequence. **Molecular Ecology Resources**, v. 15, n. 2, p. 437-448, 2015.
- HE, X.; WANG, Y.; LI, F.; WENG, Q.; LI, M.; XU, L.; SHI, J.; GAN, S. Development of 198 novel EST-Derived microsatellites in *Eucalyptus* (Myrtaceae). **American Journal of Botany**, v. 99, n. 4, p 134-148, 2012.
- JONES, M. E.; SHEPHERD, M.; HENRY, R.; DELVES, A. Pollen flow in *Eucalyptus grandis* determined by paternity analysis using microsatellite markers. **Tree Genetics & Genomes**, v. 4, n.1, p 37-47, 2008.
- MORA, A. L.; GARCIA, C. H. **A cultura do eucalipto no Brasil**. São Paulo: Sociedade Brasileira de Silvicultura, 2000, 112 p.
- THOMSON, A. M.; DICK, C. W.; PASCOINI, A. L.; DAYANANDAN, S. Despite introgressive hybridization, North American birches (*Betula* spp.) maintain strong differentiation at nuclear microsatellite loci. **Tree Genetics & Genomes**, v. 11, n. 5, p. 1 - 12, 2015.
- UNTERGASSER, A.; CUTCUTACHE, I.; KORESSAAR, T.; YE, J.; FAIRCLOTH, B. C.; REMM, M.; ROZEN, S. G. Primer3 - new capabilities and interfaces. **Nucleic Acids Research**, v. 40, n. 15, p e115, 2012.
- VARSHNEY, R. K.; GRANER, A.; SORRELLS, M.E. Genic microsatellite markers in plants: features and applications. **Trends Biotechnology**, v. 23, n. 1, p 48-55, 2005.
- VICENTINI, R.; SASSAKI, F.T.; GIMENES, M.A.; MAIA, I.G.; ENOSSI, M. *In silico* evaluation of the *Eucalyptus* transcriptome. **Genetics and Molecular Biology**, v. 28, n. 3, p. 487-495, 2005.
- XAVIER, M. A.; PINTO, L. R.; FÁVERO, T. M.; PERECIN, D.; CARLINI-GARCIA, L. A.; LANDELL, M. G. Paternity identification in sugarcane polycrosses by using microsatellite markers. **Genetics and Molecular Research**. v. 13, n. 1, p. 2268-2277, 2014.
- YEH, F. C.; YANG, R. C.; BOILEY, T.; YE, Z. H.; MAO, J. X. **Popgene, the user-friendly shareware for population genetic analysis**. Canada: Molecular Biology and Biotechnology Center, University of Alberta, 1997.

Received: 2018/28/03  
Accepted: 2019/29/01