

# Development of SCAR Molecular Markers in *Eucalyptus saligna* and *Eucalyptus tereticornis*

Kettener K.<sup>1,\*</sup>, Fuchs M. C. P.<sup>1</sup>, Madacki A. C. A.<sup>1</sup>, Gonzales E.<sup>2</sup>, Souza I. C. G.<sup>3</sup>, Oda S.<sup>3</sup>, Marino C. L.<sup>1</sup>

<sup>1</sup>Department of Genetics, Institute of Biosciences, Botucatu, São Paulo, Brazil

<sup>2</sup>Futuragene, Itapetininga, São Paulo, Brazil

<sup>3</sup>Suzano Pulp and Paper, Itapetininga, São Paulo, Brazil

**Abstract** The genus *Eucalyptus* includes over 700 species, some of which are the most widely planted hardwoods worldwide. Each species of *Eucalyptus* present different characteristics regarding its wood quality and yield. This fact makes it very important to work with known species to optimize handling and conservation of forest resources. Some of them are morphologically similar, making it difficult to differentiate by simple observation. An alternative approach is to develop molecular methods for the species differentiation. Using a Bulk Segregant Analysis (BSA) with 59 RAPD (Random-Amplified Polymorphic DNA) primers of Operon Technologies Inc. Kits, polymorphic DNA fragments between *Eucalyptus* species were isolated and SCAR (Sequence Characterized Amplified Regions) markers designed for *Eucalyptus saligna* and *Eucalyptus tereticornis*.

**Keywords** Bulk Segregant Analysis, RAPD, Myrtaceae, SCAR

## 1. Introduction

*Eucalyptus* genus is an important forest culture in the world economy, due to certain characters that confer advantages, both in their introduction and maintenance in different regions [1]. The genus presents wide species-diversity, with many varieties and hybrids (more than 900) [2, 4].

*Eucalypts* are native to Australia and north islands, occurring from the tropics to latitude 43° south [5]. The latest taxonomic revision [3] of the *eucalypts* recognizes over 700 species that belong to 13 main evolutionary lineages. Most species belong to the subgenus *Symphymyrtus*, and it is mainly species from three sections of this subgenus that are used in plantation forestry such as *Eucalyptus grandis* and *Eucalyptus urophylla* (section *Transversaria*), *Eucalyptus globulus* (section *Maidenaria*), *Eucalyptus camaldulensis* (section *Exsertaria*), *Eucalyptus saligna* (section *Latoangulatae*) and *Eucalyptus tereticornis*. *E. saligna*, native to the southern coast of New South Wales and south Queensland, Australia, is to be found in a region where frost is frequent (more than 60% of the year), which justifies its wide use in breeding programs aimed at cloning frost-resistant species, with the specific characteristics of increased growth and density in cold regions [1, 7]. The natural occurrence of *E. tereticornis* in Papua New Guinea

and Australia (Victoria, New South Wales and Queensland), regions with dry periods of up to seven months during the year, explains its importance in drought-tolerant-clone development [1]. Besides drought tolerance, other outstanding characteristics are high disease-resistance potential and wood density [7].

Each species of *Eucalyptus* present different characteristics regarding its wood quality and yield. This fact makes it very important to work with known species to optimize handling and conservation of forest resources. There are two moments in which the correct identification between these species is very important and cannot be performed by visual methods: in the nursery, as a seedling before planting it, and in debarked wood.

The precise classification requires tools comprising specific molecular-biology techniques [4], applicable to morphological identification. The development of species-specific molecular markers becomes a feasible alternative in solving this conflict accurately, quickly and at low cost.

Forestry companies, by using species-specific molecular markers to determine matrices and launch authentic hybrids [9-12], manage to avoid taxa introgression, with the possible aftermath of negative consequences, such as variability-loss and genetic-assimilation. ISSR (Inter-Simple Sequence Repeat) molecular markers have already been developed for *E. urophylla*, *E. grandis* and *E. camaldulensis* species [13]. The RAPD (Random-Amplified Polymorphic DNA) analysis has become a method for estimating genetic diversity in plant populations or cultivars [14, 15]; it was also used by Paran and Micheltore to develop a technique

\* Corresponding author:

karinekettener@gmail.com (Kettener K.)

Published online at <http://journal.sapub.org/ijaf>

Copyright © 2014 Scientific & Academic Publishing. All Rights Reserved

known as sequence characterized amplified regions (SCAR) [16].

We have used this technology to develop SCAR markers for *E. saligna* and *E. tereticornis* to quick differentiation of these two species. These specific primers lead to positive or negative amplification in target-containing and non-target - containing samples, respectively; they also can be used to generate amplification products of different sizes in closely related samples.

## 2. Materials and Methods

### 2.1. Plant Material

SCAR marker development was carried out in two groups of four-month-old *Eucalyptus* sp. seedlings (Table 1), supplied by Suzano Pulp and Paper breeding program.

**Table 1.** Origin and number of individuals from each *Eucalyptus* species used in the SCAR marker development process and marker validation

Species	Origin	Number of plants
<i>Eucalyptus brassiana</i>	Embrapa-CSIRO 10972 (North Moreton, QLD, Australia)	111
<i>Eucalyptus saligna</i>	Coffs Harbour (Australia)	114
<i>Eucalyptus grandis</i>	Coffs Harbour (Australia)	125
<i>Eucalyptus urophylla</i>	IPEF-Timor	95
<i>Eucalyptus tereticornis</i>	Embrapa-CSIRO 10975-8140 (Cooktown and Laura, QLD, Australia)	105

### 2.2. DNA extraction

DNA extraction was based on [23] protocol, with a reduction in CTAB concentration of 10% to 5%, and by using twice volume. Extracted DNA was quantified by a spectrophotometer and comparison of band intensities with known standards of GeneRuler 1 KB DNA Ladder (Thermo Fisher Scientific, GA, USA) on 0.8 % agarose gels. Each DNA concentration was adjusted to 20 ng/μl in sterile miliQ water and stored at -20 °C.

### 2.3. BSA (Bulk-Segregant Analysis) and RAPD (Random-Amplified Polymorphic DNA)

Bulk Segregant Analysis (BSA) Technique [17] was used to identify RAPD Markers. Two separate DNA bulks were prepared, each containing an equal amount of DNA (500ng) from ten individuals for each eucalyptus species, in a final concentration of 50ng/μl. Each bulk was identified with the first letter of the species epithet. A total of 59 RAPD primers (Operon Technologies Inc.) were screened between the pools. RAPD reactions were performed in a 96-well thermal cycler (MJ Research - PTC 100), with one step at 96°C for 3 min and 41 cycles at 92°C for 1 min, 35°C for 1 min, and 72°C for 2 min and 30s, followed by one step at 72°C for 10 min. Following 1.5% agarose gel electrophoresis and ethidium bromide staining, amplified patterns were visualized over a

UV transilluminator and photographed by a digital camera. RAPD amplified bands were scored visually according to their presence or absence for the species studied. Only clear, unambiguous and reproducible RAPD molecular markers were taken into account. The reproducibility of each scored marker was checked by two RAPD experiments.

### 2.4. SCAR Marker Development

DNA fragments of selected RAPD markers were isolated from agarose gel with a Gel Band Purification kit (Amersham), as recommended by the manufacturers. Purified fragments were cloned into a pGEM-T Easy Vector System I vector (Promega), and then inserted into competent cells of a DH5α-FT UltraMax strain (Life Technologies, GibcoBRL), as recommended by the manufacturers.

DNA sequencing reactions were obtained using 400 to 500 ng of plasmid DNA with forward and reverse M13 primers (1 mM), according to the protocol of ABI Big Dye Terminator Version 3.1 Cycle Sequencing kit (Applied Biosystems). Amplification conditions in a thermocycler (MJ Research – PTC 100) were an initial 2 minutes at 96°C, followed by 40 cycles, each of 30 seconds at 96°C, 30 seconds at 55°C and a final 4 minutes at 60°C. The reaction was purified by adding 80 μL of 75% isopropanol, incubated for 15 minutes at room temperature, and then centrifuged for 45 minutes at 4000 rpm. The supernatant was discarded, and 1ml of 70% ethanol subsequently added for washing. Centrifugation was repeated for 10 minutes at 4000 rpm, whereupon the supernatant was again discarded and the pellet dried at room temperature. After resuspension in 10μL of formamide, the pellet was sequenced on an automated ABI / Hitachi 3100 Genetic Analyzer sequencer (Applied Biosystems).

Nucleotide sequence data were compared against the GenBank nucleotide sequence database (BLAST search) and the Phytozome database (<http://www.phytozome.net>). They were then analyzed with the Primer3 [18] software in order to design pairs of PCR primers (approximately 20-mers) to obtain SCAR molecular markers characteristic of *E. saligna* and *E. tereticornis*. Oligonucleotides were synthesized by Sigma-Aldrich.

PCR conditions for the amplification of SCAR molecular markers were prepared in a final volume of 25 μl containing 150 ng DNA, 1x PCR buffer, 1.92 mM MgCl<sub>2</sub>, 1μg/μL of Bovine Serum Albumin, 0.8 mM dNTPs, 15 ng / primer (Operon Technologies Inc.) and 1U Taq DNA polymerase (Invitrogen - Life Technologies). The final reaction volume was completed with 13 μL of autoclaved di-ionized water. All amplifications were repeated at least in three independent experiments.

## 3. Results and Discussion

59 RAPD primers were tested on two bulks with ten individuals per bulk of *E. saligna* and *E. tereticornis* to select a set of RAPD primers that produced reliable and

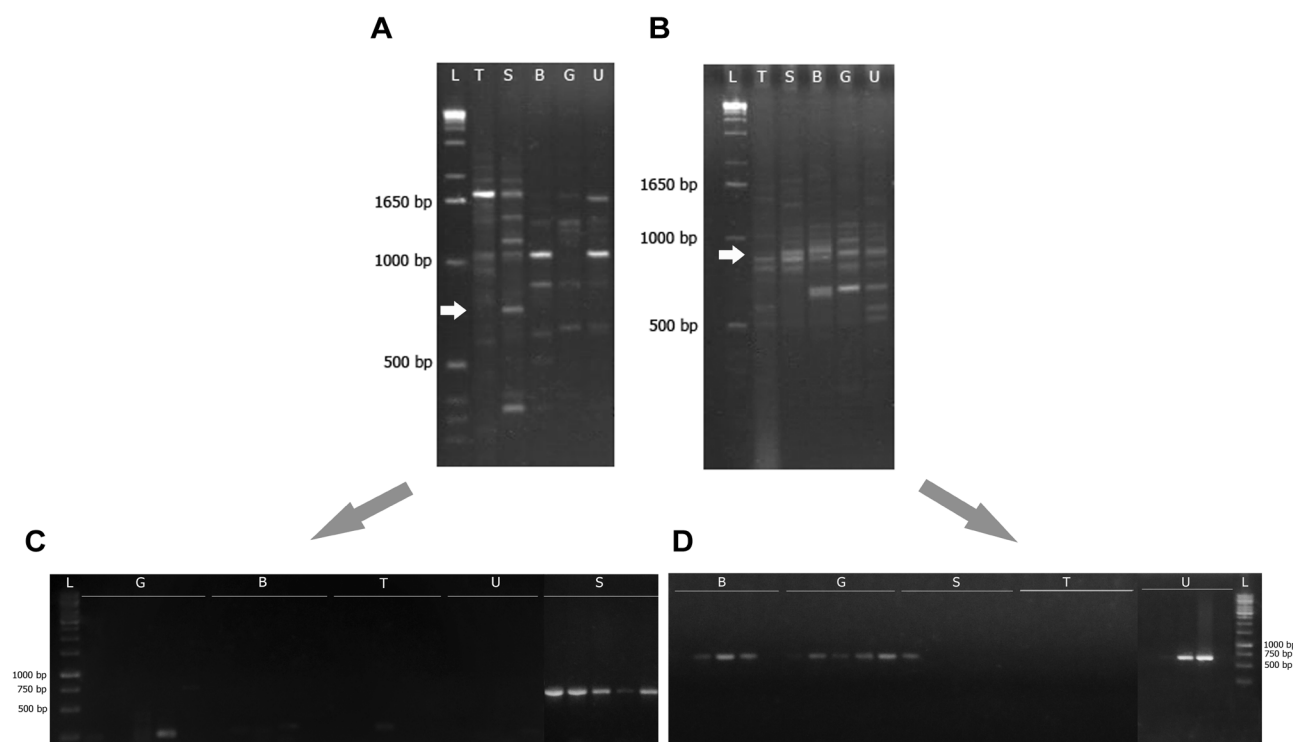
reproducible fingerprints for these two species. Reproducibility of the amplification pattern was checked by repeating each reaction at least twice without alteration in the protocol.

Only two RAPD primers (OPAD-01 and OPH-03; Table 2) among the 59 tested produced RAPD patterns that allowed differentiation of the *E. saligna* and *E. tereticornis* species. Selection criteria was a sufficient DNA length in order to maximize the availability of convenient sites for designing PCR primers and a well-defined DNA bands to increase the chance of cloning the targeted molecular marker.

OPAD-01 proved to have a RAPD product to *Eucalyptus saligna* around 750 bp (Figure 1A). OPH-03 had no RAPD product in *E. tereticornis* at 700 bp region (Figure 1B), implies a negative marker for this species. The bulks were screened for polymorphism confirmation and marker selection, and thus, SCAR marker development.

Two SCAR markers were developed, CAS and CHT. CAS

presented a common band to all individuals of *E. saligna* and absence in all individuals of the other species studied. The blast results as compared to sequences deposited in GenBank revealed that most of the RAPD product sequences had no homology with known sequences at different sequence-similarity levels (data not shown). CHT presented absence bands, which means a negative marker for *E. tereticornis*. The PCR primer pairs were further challenged using 40 individuals of each species-specific marker, plus 20 individuals of each of the remaining species. The percentage of efficiency was calculated by the primer pairs that did not amplified DNA from the other species. The CAS primer demonstrated 90% efficiency in *E. saligna* detection and band-absence on other species (Figure 1C), although CHT revealed band-absence in *E. tereticornis*, band-presence on different percentages at *E. saligna* and *E. urophylla* - 20%, *E. brassiana* - 40% and *E. grandis* - 60% (Figure 1D).



**Figure 1.** RAPD marker profiles: (A) OPAD-01 primer of *E. saligna*, arrows show bands characteristic of species (white arrow), (B) OPH-03 primer with band-absence in *E. tereticornis* (white arrow). SCAR marker profiles of CAS (C) and CHT (D) SCAR markers in five individuals of each species. Molecular size marker 1kbp ladder (L); *E. tereticornis* (T); *E. saligna* (S); *E. brassiana* (B); *E. grandis* (G), and *E. urophylla* (U)

**Table 2.** Characteristics of SCAR primers developed. The random-primer sequence used is underlined

Random primer	SCAR primer	SCAR sequence (5'→3')	Species	Amplification	Annealing temperature
OpAD-01	CAS-F	<u>CAAAGGGCGGAATGGTTC</u>	<i>E. saligna</i>	Positive	59°C
	CAS-R	<u>CAAAGGGCGGCGGTAAAG</u>			
OpH-03	CHT-F	<u>AGACGTCCACATAGGAGA</u>	<i>E. tereticornis</i>	Negative	57°C
	CHT-R	<u>AGACGTCCACCCATGTAAG</u>			

CAS marker achieved 90% efficiency in *E. saligna* identification, thus proving to be a reliable tool in breeding-programs. Analysis of this consensus-sequence marker in the *E. grandis* genome, revealed its localization between the 3' region of the AMP-binding-protein gene and the adjacent intergenic region. As in this region, there is identity with only one primer (CAS-R), amplification, not only in *E. grandis* itself, but also in other species, is impossible. Due to possible mutation in the intergenic region in *E. saligna*, the CAS-F primer probably presents specificity. Although not fixed in the *E. saligna* genome, this mutation is very frequent population wise (90%).

Amplification of the negative species-specific marker CHT for *E. tereticornis* in other commercial species ranged from 20 to 60% (*E. grandis* 60%, *E. brassiana* 40%, and both *E. saligna* and *E. urophylla*, 20%). Even though *in silico* analysis of this marker in the *E. grandis* genome indicated its location in an intergenic region, individually sequenced representatives of the species presented no identity with CHT primers, thereby indicating it to represent 40% of the *E. grandis* population without a tag. This implies that, although marker-detector alleles are absent in *E. tereticornis*, this may not be so in other species, thereby demonstrating the need for developing additional and more efficient strategies for identification.

One problem encountered in silviculture is the precise identification of both pure-species and interspecific-hybrid origin. Combining species-specific molecular marker usage and morphological analysis can do this. With a view to minimizing the economic losses involved, emphasis was given to the potential use of these markers in initial screening in breeding-company *Eucalyptus* populations. Screening would facilitate individual selection according to band presence or absence, and be of assistance in guiding and monitoring breeding and hybridization programs.

## 4. Conclusions

A set of molecular markers (CAS and CHT) specific to *E. saligna* and *E. tereticornis* species were designed to distinguish these species. This is a quick and efficient method with high specificity and reproducibility that can be used in seed, seedlings and stocked wood in the management of populations in forest breeding programs.

## ACKNOWLEDGEMENTS

This work was supported by FAPESP, and the authors thank the assistance with sample collection by Suzano Pulp and Paper Company.

## REFERENCES

[1] Fonseca, S.M.; Resende, M.D.V.; Alfenas, A.C.; Guimarães,

L.M.S.; Assis, T.F. and Grattapaglia, D. 2010. Manual prático de melhoramento genético do eucalipto. Ed. UFV, Viçosa, MG.

[2] Boland, D., M.H. Brooker, G.M.; Chippendale, N.; Hall, B.P.M.; Hyland, R.D.; Johnson, D.A.; Kleining, M.W.; McDonald, and Turner, J. D. 2006. Forest trees of Australia. Melbourne: CSIRO 5: 736p.

[3] Brooker, M.I.H. 2000. A new classification of the genus *Eucalyptus* L'Her (Myrtaceae). *Aust Syst Bot* 13:79–148.

[4] Brooker, M.I.H. and Kleining, D.A. 2004. Field guide to eucalypts, vol 3, 2nd edition. Bloomings books, Melbourne, Austrália.

[5] Ladiges, P.Y.; Udovicic, F.; Nelson, G. (2003) Australian biogeographic connections and the phylogeny of large genera in the plant family Myrtaceae. *J Biogeogr* 30:989–998.

[6] Steane, D.A.; Nicolle, D.; McKinnon, G.E.; Vaillancourt, R.E.; Potts, B.M. 2002. Higher-level relationships among the eucalypts are resolved by ITS-sequence data. *Aust Syst Bot* 15:49–62.

[7] Resende, M.D.V. and Assis, T.F. 2008. Reciprocal Recurrent Selection between Synthetic Multi Species Populations (RRS-SMSP) of eucalypt. *Pesquisa Florestal Brasileira* 57:57-60.

[8] Zobel, B. and Talbert, J. 1984. Applied forest tree improvement. John Wiley and Sons, New York, NY.

[9] Kochieva, E.Z. 1999. Molecular markers of potato and tomato species and cultivar genome. Pan European Network on the genetic indicators of biodiversity (PEN/GIB): Descriptions of diversity and exchange of information by network web, Florence, 25-27.

[10] Kochieva, E.Z.; Ryzhova, N.N.; Molkanova, O.I.; Kudryavtsev, A.M.; Upelnik, A.P. and Okuneva; I.B. 2004. The genus *Syringa*: molecular markers of species and cultivars. *Russian Journal of genetics* 40(1): 30-32.

[11] Rosseto, M.; Lucarotti, F.; Hopper, S.D. and Dixon, K.W. 1997. DNA fingerprint of *Eucalyptus graniticola*: a critically endangered relict species or a rare hybrid. *Heredity* 79: 310-318.

[12] Scheef, E.; Casler, M.D. and Jung, G. 2003. Development of species-specific SCAR markers in bentgrass. *Crop Sci* 43:345–349.

[13] Balasaranavan, T. P.; Chezian, R.; Kamalakannan, R.; Yasodha, M.; Varghese, K. Gurumurthi and M. Ghosh. 2005. Identification of Species-diagnostic ISSR makers for six *Eucalyptus* species. *Silvae genetica* 55,3:119-122.

[14] Hosokawa K.; Minami M.; Kawahara K.; Nakamura I.; Shibata T. 2000. *Planta Med.*, 66, 270—272.

[15] Mandal, A. B.; Thomas, V. A.; Elanchezhian, R. 2007. *Curr. Sci.*, 93, 369— 373.

[16] Paran, I.; Michelmore, R. W. 1993. *Theor. Appl. Genet.*, 85, 985—993.

[17] Michelmore, R.; Paran, I.; Keselli, V. 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions using segregating populations. *Proc Natl Acad Sci USA* 88:9828–9832.

- [18] Steve Rozen, Helen J. Skaletsky (1998) *Primer3*. Code available at [http://www-genome.wi.mit.edu/genome\\_software/other/primer3.html](http://www-genome.wi.mit.edu/genome_software/other/primer3.html).
- [19] Brondani, R.P.V.; Brondani, C. and Grattapaglia, D. 2002. Towards a genus-wide reference linkage map for *Eucalyptus* based exclusively on highly informative microsatellite markers. *Mol. Genet. Genom.* 267(3): 338–347.
- [20] Byrne, M.; Marquez-Garcia, M.I.; Uren, T.; Smith, D.S. and Moran, G.F. 1996. Conservation and genetic diversity of microsatellite loci in the genus *Eucalyptus* Aust. J. Bot. 44: 331–341.
- [21] Marques, C.M.; Brondani, R.P.V.; Grattapaglia, D. and Sederoff, R. 2002. Conservation and synteny of SSR loci and QTLs for vegetative propagation in four *Eucalyptus* species. *Theor. Appl. Genet.* 105(2–3): 474–478.
- [22] Poke, F.S.; Vaillancourt, R.E.; Potts, B.M. and Reid, J.B. 2005. Genomic research in *Eucalyptus*. *Genetica* 125:79–101.
- [23] Ferreira, M.E.; Grattapaglia, D. 1998. Introdução ao uso de marcadores moleculares em análise genética. Brasília-DF: Embrapa – CENARGEN.