Electronic Journal of Biotechnology ISSN: 0717-3458
http://www.ejbiotechnology.info
DOI: 10.2225/vol15-issue2-fulltext-3 *RESEARCH ARTICLE*

Microsatellite markers in candidate genes for wood properties and its application in functional diversity assessment in *Eucalyptus globulus*

Cintia V. Acuña[1] · Pamela V. Villalba[1] · Martín García[1] · Pablo Pathauer[1] · H. Esteban Hopp[1,2] · Susana N. Marcucci Poltri[1] ✉

1Instituto Nacional de Tecnología Agropecuaria, Instituto de Biotecnología e Instituto de Recursos Biológicos, CNIA, Castelar, Argentina

2 Universidad de Buenos Aires, Facultad de Ciencias Exactas y Naturales, Buenos Aires, Argentina

✉ Corresponding author: smarcucci@cnia.inta.gov.ar
Received August 21, 2011 / Accepted December 21, 2011
Published online: March 15, 2012

Abstract

Background: Functional genetic markers have important implications for genetic analysis by providing direct estimation of functional diversity. Although high throughput sequencing techniques for functional diversity analysis are being developed nowadays, the use of already well established variable markers present in candidate genes is still an interesting alternative for mapping purposes and functional diversity studies. SSR markers are routinely used in most plant and animal breeding programs for many species including *Eucalyptus*. SSR markers derived from candidate genes (SSR-CG) can be used effectively in co-segregation studies and marker-assisted diversity management. **Results:** In the present study, eight new non reported SSRs were identified in seven candidate genes for wood properties in *Eucalyptus globulus*: cinnamoyl CoA reductase (CCR), homocysteine S-methyltransferase (HMT), shikimate kinase (SK), xyloglucan endotransglycosylase 2 (XTH2), cellulose synthase 3 (CesA3), glutathione S-transferase (GST) and the transcription factor LIM1. Microsatellites were located in promoters, introns and exons, being most of them CT dinucleotide repeats. Genetic diversity of these eight CG-derived SSR-markers was explored in 54 unrelated genotypes. Except for XTH2, high levels of polymorphism were detected: 93 alleles (mean of 13.1 sd 1.6 alleles per locus), a mean effective number of alleles (Ne) of 5.4 (sd 1.6), polymorphic information content values (PIC) from 0.617 to 0.855 and probability of Identity (PI) ranging from 0.030 to 0.151. **Conclusions:** This is the first report on the identification, characterization and diversity analysis of microsatellite markers located inside wood quality candidate genes (CG) from *Eucalyptus globulus*. This set of markers is then appropriate for characterizing genetic variation, with potential usefulness for quantitative trait loci (QTL) mapping in different eucalypts genetic pedigrees and other applications such as fingerprinting and marker assisted diversity management.

Keywords: functional markers, genetic diversity, lignin pathway, SSR, wood density

INTRODUCTION

Eucalyptus tree species are among the most planted hardwoods in the world. They are long-living, evergreen species belonging to the predominantly southern-hemisphere endemic angiosperm family *Myrtaceae*. *Eucalyptus* is predominantly out crossing, highly heterogeneous and genetically diverse.

Within the genus, *Eucalyptus globulus* species is native of Tasmania and coastal regions of south-eastern Australia. Several forest plantations were successfully established in southern Europe, northern Africa and southern America (Myburg et al. 2007).

Most eucalypt domestication and breeding programs are focused on increasing the volume of produced wood, as well as improving its quality properties.

Wood is essentially composed of cellulose, hemicelluloses, lignin, and extractives, each of them contributing to fiber properties which ultimately impact product properties.

Principal traits for improving pulping procedures include quantity and quality of extractives and lignin, which affect directly the economic and/or environmental cost of pulping (Raymond and Apiolaza, 2004) as well as the use of forest residues for bioenergy purposes like ethanol production. For solid timber applications wood density as well as microfibril angle (MFA) have been considered as the most important factors affecting wood properties such as stiffness, strength, and shrinkage behaviour of solid wood (Evans and Ilic, 2001).

Results from genetic variation analyses carried out within *E. globulus* subraces (Strackpole et al. 2011) suggest that selection increasing wood density tends to decrease lignin S/G relationship. However, they did not find additive genetic relationship between density and extractives, and suggest that selection for increased pulp yield would result in increased cellulose content and S/G but reduced lignin and extractives content.

As with other forest tree genera with long generation times, eucalypt domestication and breeding programs will benefit tremendously from molecular technologies which can contribute to quantify genetic diversity and relationships, breeding systems analysis, gene flow, fingerprinting and clone identification, QTL detection and molecular breeding through marker- or gene-assisted selection (Myburg et al. 2007). Also, the possible improvement by using marker-assisted selection aimed by breeders requires genomic resources publicly available involving putative candidate genes that control wood properties (Rengel et al. 2009). Good candidates for wood properties include genes involved in lignin and carbohydrate biosynthesis (Hertzberg et al. 2001).

Many genomic studies have reported the analysis of genes expressed during wood formation and xylogenesis (Hertzberg et al. 2001; Moran et al. 2002; Israelsson et al. 2003; Kirst et al. 2004; Paux et al. 2004; Paux et al. 2005; Foucart et al. 2006) and some important metabolic pathways are now well known. In *Eucalyptus,* several structural and regulatory candidate genes involved in lignin biosynthesis were identified, including those encoding components of the common phenylpropanoid pathway: phenylalanine ammonia-lyase (PAL), 4-coumarate-3-hydroxylase (C3H), Cinnamic acid 4-hydroxylase (C4H), caffeic acid 3-O-methyltransferase (COMT), caffeoyl-CoA Omethyltransferase (CCoAOMT), and 4-coumarate: CoA ligase (4CL) (Gion et al. 2000; Thamarus et al. 2002), and those of the monolignol specific pathway like cinnamoyl CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD) (Poke et al. 2003) as well as lignin regulatory genes such as MYB transcription factors (Goicoechea et al. 2005). All of these genes seem to be good candidates for QTL co-localization studies with wood-quality and lignin-content/quality QTLs.

In the past ten years a great number of studies have reported genetic maps for different *Eucalyptus* species, built from combinations of random-amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), restriction fragment length polymorphism (RFLP) markers, isozymes, EST and gene sequence derived markers (mainly by the analysis of RFLP and single strand conformation polymorphisms (SSCP) techniques) as well as microsatellites (SSR and EST-SSR) (Grattapaglia and Sederoff, 1994; Byrne et al. 1996; Verhaegen and Plomion, 1996; Marques et al. 1998; Bundock et al. 2000; Gion et al. 2000; Brondani et al. 2002; Thamarus et al. 2002; Poke et al. 2005; Faria et al. 2010, Faria et al. 2011). In Brondani et al. (2006), a set of 230 new microsatellites has been developed and a consensus map has been assembled, covering at least 90 percent of the recombining genome of *Eucalyptus*.

Recently Single Feature Polymorphisms (SFP) (Neves et al. 2011) and Single Nucleotide Polymorphism (SNP) (Novaes et al. 2008; Grattapaglia et al. 2011) analysis have been reported.

There are only few studies in *Eucalyptus* where candidate genes underlying wood quality traits have been analyzed using QTL approaches (Gion et al. 2000; Gion et al. 2001; Freeman et al. 2009; Thumma et al. 2010; Gion et al. 2011; Sexton et al. 2012). These studies have shown the co-segregation of QTLs simultaneously involved both in chemical as well as physical properties, thus suggesting that genomic regions may be organized in clusters of genes with independent effects on

different wood properties or that there are some key genes with pleiotropic effects (Gion et al. 2011). The increasing use of larger sets of interspecific transferable markers and consensus mapping information will allow faster and more detailed investigations of QTLs synteny among species, validation of QTLs and expression-QTLs across variable genetic backgrounds, and positioning of a growing number of candidate genes co-localized with QTLs, to be tested in association mapping experiments (Grattapaglia and Kirst, 2008). The first association mapping studies in forest trees were based on the analysis of SNPs allelic diversity which was carried out in fewer than 20 selected candidate genes. These studies have revealed that, although some candidate gene associations could be detected (Thumma et al. 2005; Sexton et al. 2012), it turned out very difficult to predict an appropriate set of candidate genes for any trait of interest. Instead, a much larger set of genes and regulatory sequences must be screened and interrogated for allelic diversity and trait association (Myburg et al. 2007). Recent advances in high-throughput marker technologies allowed mapping and association studies in pine (Eckert et al. 2009) and poplar (Wegrzyn et al. 2010) based on candidate genes.

Although high throughput sequencing techniques for functional diversity analysis are being developed nowadays, the use of already well established variable markers present in candidate genes is still an interesting alternative for mapping purposes and functional diversity studies. SSR markers are routinely used in most plant and animal breeding programs for many species, including *Eucalyptus* species. SSR markers derived from candidate genes (SSR-CG) can be used effectively in co-segregation studies and marker-assisted diversity management. It is also a low-cost technology that may provide an efficient way to survey natural variation in *Eucalyptus.*

In the present study, eight new non reported SSRs were identified in seven candidate genes for wood properties: CCR (catalyzes the conversion of cinnamoyl CoA esters to their corresponding cinnamaldehydes, being the first specific step in the synthesis of lignin monomers); HMT (participates in the S-adenosylmethionine biosynthesis, a substrate for the phenylpropanoid pathway); SK (takes part in phenylalanine, tyrosine and tryptophan biosynthesis); XTH2 (implicated in the modification of cell wall structure); CesA3 (involved in the coordinated synthesis of glucose chains); GST (functions in normal plant development and plant stress responses) and the transcription factor LIM1.

Diversity analyses of them were carried out in 54 unrelated genotypes to estimate their potential usefulness for QTL mapping in different eucalypts genetic pedigrees and diversity analysis. Then, frequency of alleles, observed heterozygosity (Ho), PIC and PI were calculated for other applications such as fingerprinting and marker assisted diversity management.

MATERIALS AND METHODS

Plant material and DNA extraction

A total of 54 non related trees, each from a different open-pollinated (OP) family of *E. globulus,* were analyzed for their variability. These non selected trees represented major geographical races of the species' natural distribution (Dutkowski and Potts, 1999) and land race from Portugal that were grown in a field trial in the Province of Buenos Aires, Argentina, between 1995 and 1997 (Table 1).

Total DNA was extracted from young leaves by the CTAB method (Hoisington et al. 1994) with minor modifications to avoid oil precipitation: instead of using isopropanol in the precipitation step, samples were diluted two times in 10 mM Tris-HCl, 1 mM EDTA buffer and precipitated with 2.5 volumes of ethanol and 300 mM Na acetate (Marcucci Poltri et al. 2003).

Target accession numbers of candidate genes from GenBank

Microsatellites within CGs were identified mainly by two different methods: -SSRs detection from non redundant ESTs of *E. globulus* from GenBank: Annotations of these SSR-ESTs were based on the Gene Ontology (GO) using Blast2GO (Conesa et al*.* 2005; Acuña et al*.* 2011). The selection of CGs for wood quality was made from that function assignment.

Table 1. *Eucalyptus* genotypes used in this study and their geographic locations. Details of origin/provenance latitude (Lat., S); longitude (Long., E); and number of open-pollinated (OP) families are included (one individual each).

Race	Origin/ Provenance	Lat.	Long.	N
Eastern_Otways	Otways State Forest	38°36´	143° 53´	5
Western_Otways	Otways National Park	38°47´	143°35´	10
NE_Tasmania	Seymour	41°40´	148° 17´	3
S_Tasmania	Cradoc Hill	32°04´	138°29´	2
	Glendevie	43°14´	147°02´	2
SE_Tasmania	Moogara	42°47´	146° 53´	4
Strzelecki	Jeeralang	38°20´	146° 31´	16
Furneaux	Flinders Island	40°04´	148°00´	7
King_Island	King Island (TAS)	39°53´	143°59´	3
	Portugal OP			2
total				54

-SSR Mining software (GDR Server, http://www.rosaceae.org/bio/content?title=&url=/cgi-bin/gdr/gdr_ssr) on selected candidate gene genomic sequences identified at the GenBank.

Selection was restricted to those microsatellites displaying a minimum of five repeats for di-, 4 for tri-, and 3 for tetra- and penta-nucleotide arrays. Primers were designed flanking the repeated motifs using PRIMER 3.0 software (http://www.broadinstitute.org/genome_software/other/primer3.html) (Table 2).

PCR amplification of microsatellite loci

PCR conditions were as follows: A final volume of 12 µl containing 20 ng of genomic DNA, 0.25 µM of each primer, 2 mM $MgCl_2$, 0.2 mM of each dNTP, 1X reaction buffer and 1 U Platinum Taq polymerase (Invitrogen). Amplifications were performed with the following conditions: Denaturation step of 5 min at 94ºC, 35 cycles of 30 sec at 94ºC, 30 sec at annealing temperature and 45 sec at 72ºC. The final extension step was of 10 min at 72ºC. Details of primers sequences, SSR location, amplification condition, and product size are described in Table 2.

Samples were mixed with denaturing loading buffer, treated for 5 min at 95ºC and separated in a 6% polyacrylamide gel. Amplification products were silver-stained following the DNA silver staining system procedure (Promega).

Loci with poor amplification products in standard PCR conditions were amplified with fluorescent forward primer. PCR products were detected in the ABI3100 Genetic Analyzers (Applied Biosystems, USA) and the allele assignments were made by size comparison with the standard allelic ladders and using the GenMapper ID software provided by Applied Biosystems. To minimise scoring errors associated with base pair shifts between runs, at least four control samples were also included in each run.

Table 2. Description of SSR markers derived from candidate genes for wood quality in *E. globulus.* Included are primer names, accession numbers, repeat motif within the sequence, SSR location, forward (F) and reverse (R) primers, annealing temperature (AT), expected and observed product size (bp).

Primer name	Accession number	Repeat motif	SSR location	Primer sequences (5´to 3´)	AT (Cº)	Expected product size (bp)	Observed product size (bp)
CCR	AJ132750	$(CT)_{15}$	promoter	F: ATGGAAAATAAGGGC R: ACTTTTGGACGCTTTTTGAC	42	222	182-216
CesA3_a	EU165713	$(GA)_{10}$	promoter	F: TTGGAAGTAATGGAGCATTT R: AGTTTGTTCTTCCTTCCCTC	54	292	285-313
CesA3_b	EU165713	$(CT)_{11}$	promoter	F:TCCATTTCTTATGATCCGTC R:CCAAGTGGAGTATCCTCGTA	54	340	341-357
LIM1	AB208710	$(CT)_{22}$	promoter	F: GCTTCCCTTTCTTATCCTCCA R: CTGTGAGCTTGTCCACCAGA	58	178	159-179
HMT	FJ492059	$(CT)_{17}$	intron	F: TAGGGAGGGTCTTCCCCTTA R: GATCTCCTCGGTCGATGGTA	57	307	417-453
SK	BF942502	$(CT)_{6}$	exon	F: TGGCATATGTTGTGGTTAAA R: GATGGGAAGAGAGACAGACA	54	336	326-382
XTH2	CT989247	$(CGAAA)_{3}$	exon	F: ATCAAGACATCACGAGCAA R: CTCGCGATTCTAGAGAAGAA	54	132	133-138
GST	CT988111	$(TATG)_{9}$	exon	F: TTTGATCTCCTGCTTTCTGT R: CGAAGAACGTGCTATTAGGT	54	118	87-124

Statistical analyses

The number and frequency of alleles as well as the effective number of alleles Ne ($Ne = 1 / \sum_{i=1}^{n} p_i^2$), observed heterozygosity (H_o), and probability of identity (PI, which shows the probability of two unrelated individuals displaying the same genotype) ($PI = 2(\sum_{i=1}^{n} p_i^2)^2 - \sum_{i=1}^{n} p_i^4$) were determined with the GenAlEx 6.4 program (Peakall and Smouse, 2006).

The PIC of each marker was calculated according to Botstein et al. (1980) using the following formula,

$$PIC = 1 - \sum_{i=1}^{n} p_i^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} 2 p_i^2 p_j^2$$

where pi is the frequence of the ith allele.

Null allele frequencies were estimated with INEST software (Inbreeding/Null allele Estimation) (Chybicki and Burczyk, 2009) which is adequate for this species subject to a mixed mating model (Faria et al. 2011) under the IIM (individual inbreeding model).

Distances between individuals based on the microsatellite data set were carried out according to the shared allele distance (DAS) (Chakraborty and Jin, 1993) and was computed with POPULATIONS 1.2.28 software (Langella, 2002). Cluster analyses were implemented based on distances matrices using the unweighted pair group method arithmetic average (UPGMA), and the corresponding phenogram was constructed.

RESULTS AND DISCUSSION

Screening for microsatellite sequences in candidate genes for wood quality

Microsatellite sequences that were identified *in silico* were confirmed by PCR amplification using *E. globulus* genomic DNAs. Some of them were corroborated by automatic nucleotide sequencing. After homology reconfirmation of the flanking sequences using BLAST software, microsatellite structure was corroborated by direct visualization of their DNA sequence profiles. Eight new microsatellite sequences involved in lignin and cell-wall polysaccharide biosynthesis were, thus, characterized in seven candidate genes: *Eucalyptus gunnii* cinnamoyl-CoA reductase promoter (GenBank: AJ132750), *E. grandis* cellulose synthase 3 (CesA3) (GenBank: EU165713, 2 microsatellite regions), transcription factor LIM1 (GenBank: AB208710, *E. grandis* EST CB967988 similar to *Arabidopsis thaliana* homocysteine S-methyltransferase (HMT), *E. globulus* EST GenBank: BF942502 similar to *A. thaliana* shikimate kinase (SK), a putative *E. globulus* XTH2 (GenBank: DQ100338), *E. globulus* EST CT988111 similar to *Vitis vinifera* glutathione S-transferase and *E. globulus*).

In *E. grandis* EST CB967988, similar to *Arabidopsis thaliana* homocysteine S-methyltransferase (HMT), the microsatellite was identified after automatic sequencing of fragments which were higher than the expected size. Sequence data from this genomic fragment was annotated in the GenBank under accession number FJ492059.

For diversity analysis samples of 54 non selected unrelated trees were assayed for each SSR.

Polymorphism in CCR gene

A CT microsatellite located in the promoter of the *Eucalyptus gunnii* CCR (GenBank: AJ132750) was found between positions -164 and -135. This region is close to that described by Lacombe et al. (2000) at position -119 to -77 (GenBank: AJ132750) as necessary and sufficient for the expression in vascular tissues of stems. Twelve different alleles (ranging between 182 and 216 bp size) were detected, indicating a remarkably high level of allelic variability. Like most of SSR-CG described here, the number of alleles (Na) was higher than the effective number of alleles (Ne), which was 2.7 for this

locus. This result shows that only a small number of equally frequent alleles are present (*i.e.* one allele of 182 bp with a frequency of 0.588). A relatively high PIC of 0.617 and low Ho as well as PI values of 0.275 and 0.15 respectively were found (Table 3). The highly estimated null frequency alleles (0.217) possibly explain the low observed heterozygous level and prevents its usefulness for fingerprinting analysis.

Table 3. Descriptive statistics of the eight microsatellite markers. Include N: Sample size; Na: Number of alleles; Ne: effective number of alleles; PI: probability of identity; Ho: Observed Heterozygosity; PIC: Polymorphism Index Content; Null Freq: Null allele frequency.

	N	Na	Ne	PI	Ho	PIC	Null Freq
CCR	51	12	2.722	0.150	0.275	0.617	0.217
CesA3_a	52	12	4.382	0.078	0.519	0.746	0.132
CesA3_b	52	13	5.821	0.050	0.692	0.807	0.079
LIM1	52	16	7.564	0.030	0.769	0.855	0.054
HMTi	51	13	5.335	0.056	0.745	0.791	0.047
SK	53	15	4.777	0.068	0.604	0.766	0.095
XTH2	46	3	1.566	0.459	0.370	0.310	0.054
GST	54	12	7.317	0.033	0.815	0.849	0.036
mean		12	4.936		0.599	0.718	
s.d.		3.9	2.074		0.196	0.181	

CCR locus has already been found to co-localize with QTLs for both lignin content and the syringil-guaiacyl monomers ratio in other *Eucalyptus* species (*E. grandis* x *E. urophylla* hybrids) by Gion et al. (2001) and Gion et al. (2011).

In *A. thaliana,* 10 putative *CCR* genes have been reported (Jones et al. 2001). However, for the genus *Eucalyptus*, only one CCR gene has been found in *E. gunnii, E. globulus* and *E. tenuiramis* (Lacombe et al. 1997; Poke et al. 2003).

Polymorphism for this gene was previously reported in *E. globulus* (Poke et al. 2003; Poke et al. 2004) and Thumma et al. (2005) found specific polymorphisms in CCR associated with variation in microfibril angle (MFA) with linkage disequilibrium (LD) mapping analysis using SNP-marker, haplotype analyses and family-based studies in *E. nitens*.

In this study, we identified a CT microsatellite within the promoter of the gene, which might be involved in transcription regulation. As mentioned, it is close to a position described by Lacombe et al. (2000) essential to drive xylem-localized gene expression.

Presence of microsatellites in 5'UTR has been described already by other authors (Kumpatla and Mukhopadhyay, 2005). For instance, in *Populus*, 1.2% of total SSR found after genome sequence (around 1510) was discovered in 5'UTR (Arnaud et al. 2007).

CT elements have also been described in *A. thaliana* core promoters (Molina and Grotewold, 2005). Furthermore in *Arabidopsis* and *Brassica*, analysis *in silico* of known *cis*-regulatory elements showed that light responsive elements were clustered in the region of CT/GA repeats (Zhang et al. 2006).

Polymorphism in CesA3 gene

Higher plants contain a family of cellulose synthase catalytic subunit (CesA) genes that encode components of an enzyme complex embedded in the cell membrane.

Creux et al. (2008) showed that CesA promoters from *Eucalyptus* trees are functional in *Arabidopsis* plants, despite the low amounts of sequence similarity between orthologous CesA promoters. This functional conservation is made possible by *cis*-elements that have been retained during evolution along with the corresponding transcription factors.

Rengel et al. (2009) released a collection of *Eucalyptus* ESTs originating from one normalized differentiating xylem cDNA library and four subtractive libraries, named EUCAWOOD. They found five unigenes homologous to members of the CesA multigene family that correspond to the EgCesA1-EgCesA5 genes recently described in *E. grandis* (Ranik and Myburg, 2006).

In this study, we identified two microsatellite regions in the promoter zone of CesA3 in *E. grandis* (GenBank: EU165713). One of these, with a GA repeat motif, was found between positions -804 and -787 whereas a repeat motif CT was found between positions -248 and -226. Both genomic regions were denominated CesA3a and CesA3b, respectively (Table 2) and, like the enzyme described before, these motifs could be transcription regulation elements.

Diversity analyses revealed a high number of alleles: 12 (285 to 307 bp) and 13 (342 to 367 bp) for CesA3a and CesA3b microsatellites, respectively. As it was mentioned before, Ne values were of 4.38 and 5.82 for both loci respectively and lower than their corresponding Na.

High PIC values of 0.746 and 0.807 were found for CesA3a and CesA3b. Ho was 0.519 and 0.692 and PI was 0.078 and 0.05 respectively (Table 3). Estimated null frequency alleles was higher for CesA3a (0.132) than CesA3b (0.079), possibly explaining the lower observed heterozygous level for CesA3A.

Polymorphism in LIM1 transcription factor gene

The AC-rich motif, Pal-box, is an important cis-acting element for gene expression in phenylpropanoid biosynthesis. LIM transcription factor can bind to the Pal-box *cis* element in the promoters of monolignol biosynthetic genes leading to repression of the LIM gene and inhibition of lignin biosynthesis in transgenic tobacco plants. A reduction of lignin content in the woody plant *E. camaldulensis* by the suppression of gene expression of the LIM domain transcription factor was reported by Kawaoka et al. (2006). In poplar, the distributionof ESTs in different wood tissuesindicates a rather high expression of an LIM protein homologuein tension wood (Arnaud et al. 2007).

Rengel et al. (2009) found in their libraries 3 unigenes which resemble LIM transcription factors *i.e.* contig named Cg2892 was similar to *Ec*LIM1 from *E. camaldulensis*, sharing 86% homology with *Nicotiana tabacum* (*Nt*LIM1).

We identified a GA repetitive motif between positions -43 and -1 in GenBank: AB208710 sequence. High level of variability was found in the sample assayed: 16 alleles were identified, ranging from 159 and 194 bp in size and the highest Ne (7.56) of tested loci. This locus showed the highest PIC value (0.855), a relatively high Ho value (0.769), the lowest PI value (0.03) and a relative low null allele frequency estimates (0.054) (Table 3), making it useful for fingerprinting analysis.

Polymorphism in homocysteine S-methyltransferase gene

In this study, we identified a microsatellite in *E. grandis* GenBank EST CB967988 similar to *Arabidopsis thaliana* homocysteine S-methyltransferase (HMT) gene. However, the resulting amplification products were higher than the expected size for *E. grandis.* In consequence, the presence of the microsatellite was confirmed after automatic sequencing. Sequence data from this genomic fragment was annotated in the GenBank under accession number FJ492059.

A repeat motif CT was found within intron region, after multiple alignment and comparison with EST CB967988 similar to HMT (Figure 1).

Intronic SSRs can affect gene transcription, mRNA splicing, export to cytoplasm, and can also induce heterochromatin-mediated-like gene silencing (Li et al. 2002). These SSRs may provide a molecular basis for fast adaptation to environmental changes in both prokaryotes and eukaryotes (Li et al. 2002).

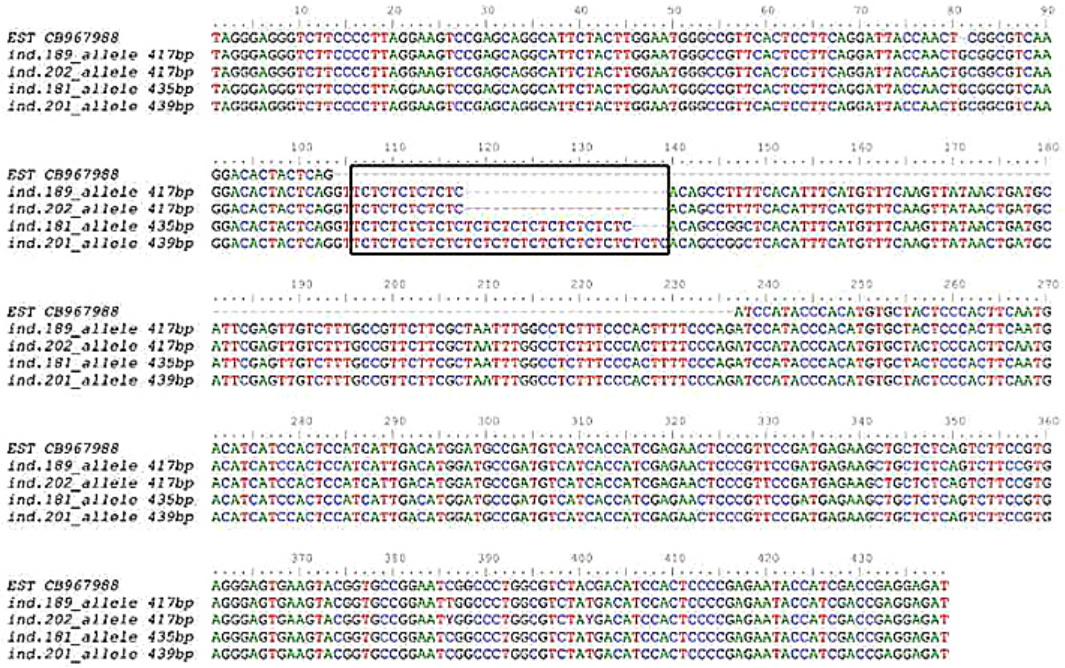

Fig. 1 Alignment of HMT gene sequences. First sequence corresponds to the *E. grandis* EST CB967988. Second to fourth sequences correspond to the genomic fragments amplified in four individuals named 189, 202, 181 and 201 of *Eucalyptus globulus*. Alignment of sequences denotes the presence of an intron, which extends from the positions 104 to 236. Microsatellite region CT (black box) is located between positions 106 to 139.

It is not known yet if the here detected differences in SSR length are actually responsible for such changes in *Eucalyptus*. Functional studies using molecular biology techniques for each of the candidate genes are needed to establish if this is happening here.

We found 13 alleles (415 bp-465 bp), generating a Ne of 5.33, a PIC of 0.791, a Ho of 0.745, a PI of 0.056 and null allele frequency estimates of 0.047 (Table 3), being this locus one of the most informative in this study.

Polymorphism in SK gene

The shikimate pathway links metabolism of carbohydrates to biosynthesis of aromatic compounds. In a sequence of seven metabolic steps, phosphoenolpyruvate and erythrose 4-phosphate are converted to chorismate, the precursor of the aromatic amino acids and many aromatic secondary metabolites (Herrmann and Weaver, 1999).

SK enzyme belongs to the family of transferases, specifically those transferring phosphorus-containing groups (phosphotransferases) with an alcohol group as acceptor.

We have found an *E. globulus* EST (GenBank: BF942502) similar to *A. thaliana* SK, harbouring a microsatellite region of CT motif.

Fifteen alleles (between 331 bp and 386 bp) and a Ne of 4.8 were revealed, whereas PIC and Ho presented values of 0.766 and 0.604, respectively. PI of 0.068 and null alleles frequency estimates of 0.095 were calculated (Table 3), showing its usefulness for fingerprinting analysis.

Polymorphism in XTH2 gene

Xyloglucans consist of a backbone of (1, 4)-βD-glucan substituted with xylosyl, galactosyl, and fucosyl residues (Hrmova et al. 2007). XTHs are enzymes involved in the modification of cell wall structure by cleaving and, often, also re-joining xyloglucan molecules in primary plant cell walls (Genovesi et al. 2008).

We found a pentanucleotide (CGAAA) microsatellite motif in a putative *E. globulus* XTH2 mRNA (GenBank: DQ100338). No sequence similarity was found when BLASTN of the SSR region was performed through sequences coding for XTH1, XTH3 and XTH4 (GenBank: DQ100337, DQ100339 and DQ100340), suggesting that SSR motif is only present in XTH2 gene.

In contrast with the above described dinucleotide motif microsatellites, only three alleles (Na) were found, with 122, 129 and 134 bp of size in the sample evaluated. While the allele of 129 bp was predominant in most individuals (n = 44), the allele of 134 bp was present in 17 while allele of 122 bp was present in just two, giving a low (Ne) of 1.57. PIC and Ho presented values of 0.310 and 0.370, respectively. PI of 0.46 and null alleles frequency estimates of 0.054 were calculated (Table 3), being this locus the less informative in this study.

Polymorphism in glutathione S- transferase gene

Plant glutathione S-transferases (GSTs) comprise a heterogeneous super family of multifunctional proteins (Dean et al. 1995; Dean and Devarenne, 1997; Edwards et al. 2000; Kampranis et al. 2000; Loyall et al. 2000; Mueller et al. 2000; Thom et al. 2002; Kilili et al. 2004).

We amplified a microsatellite region in an *E. globulus* GenBank: EST CT988111 highly similar to *Populus trichocarpa* tau class GST sequence containing a tetranucleotide motif TATG. Twelve different alleles (varying from 83 bp to 120 bp in size) and a Ne of 7.32 were found. PIC and Ho values were 0.849 and 0.815, respectively (Table 3). This high number of alleles, the lowest PI value (0.033) and low null allele frequency estimates (0.036), plus the tetranucleotide motif makes this SSR potentially very useful for identification purposes.

Genetic diversity of the putatively functional markers

A sample of 54 representative genotypes of *Eucalyptus globulus* was evaluated to accurately estimate genetic information. Table 3 shows the results obtained from the statistical analysis of the microsatellite allele pattern distribution found in the seven candidate genes analyzed.

In this report, most of the SSR found were dinucleotides repeats and ranged from 132 bp to 453 bp. Although longer motives of repeats are easily and accurate for scoring and may give better resolution patterns, usually, these are less polymorphic than shorter repeats (Tang et al. 2008; Faria et al. 2011). However, clear patterns were found in all SSR with dinucleotides motives reported here and stutter profiles could be defined adequately (Figure 2).

Ho was moderated to high except for the XTH2 and CCR loci. PIC values ranged from 0.310 to 0.855, PI of a locus varied from 0.033 to 0.151, except for XTH2 locus that showed a value close to 0.5. The resulting probability value for the combined eight loci was very low and of $1.04\ 10^{-9}$ (computed as the product of individual PI) making this set of loci mainly useful for fingerprinting analyses. This was corroborated when an analysis of genetic distances between all 54 individuals showed a mean of 0.704 (sd 0.129) and only five pairs of trees showed distances between 0.3 and 0.273 (the lowest value), reflecting the high resolution capacity of microsatellites (Figure 3).

Null alleles may bias estimates of allele and genotype frequencies, thereby hindering population genetic analyses and individual identification (Nascimento de Sousa et al 2005; Chybicki and Burczyk 2009). As no progeny tests were carried out, we cannot discard nucleotide sequence variability of the annealing site for the primer which may lead to null alleles. There are some few reports on the frequency of microsatellite null alleles like that described in Brondani et al. (2006), where they could find about an 8% overall occurrence of null alleles when mapping a segregating population between *E.*

grandis and *E. urophylla*. In *E. camaldulensis*, Butcher et al. (2009) estimated that the frequency of null alleles was less than 5% (averaged over 97 populations) for the 15 loci explored. Faria et al. (2010) and Faria et al. (2011) described that most populations that deviated from HWE displayed null allele frequencies of around 0.1-0.3.

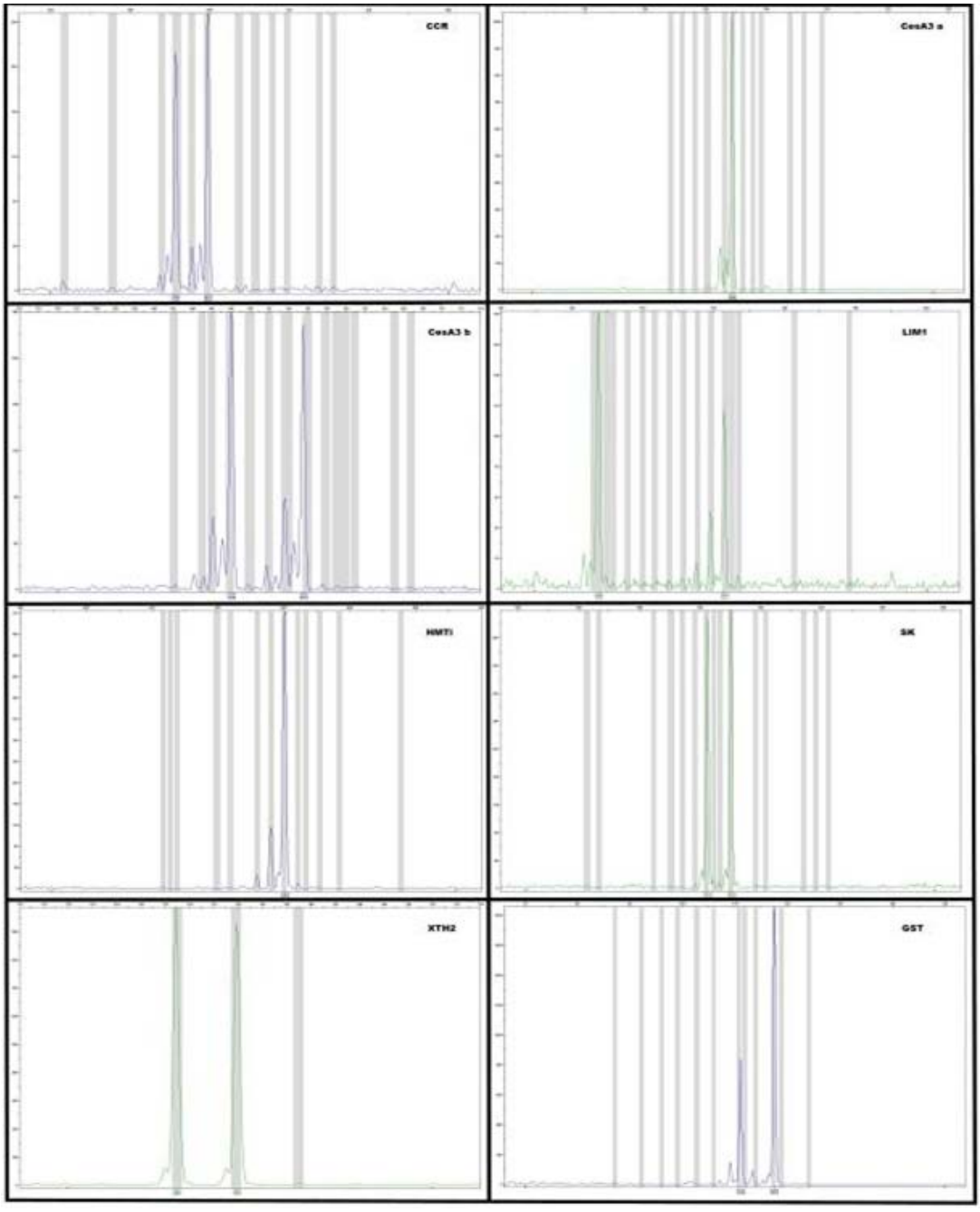

Fig. 2 Electropherogram of SSR markers from Candidate genes. Blue and green lines describe size alleles present in one individual. Grey patterns show different size alleles of the analyzed population.

In this study, low frequencies of null alleles were estimated for each loci ranging from 0.036 (GST) to 0.217 (CCR), and 6 had a maximum of 0.10 (Table 3).

In *Eucalyptus*, the probability of maintaining low-frequency alleles or rare alleles in the population is high (Grattapaglia and Kirst, 2008).

In this study, the SSR allowed to detect a total of 21 trees with low frequencies alleles. We found different proportions of alleles per locus that were present in just one individual (frequency close to 1%) for all the loci except XTH2: CCR (33.3%), CesA3a (33.3%), CesA3b (23.1%), LIM1 (31.2%), HMT (23.1%), SK (46.7%) and GST (8.3%). These results are reflected by a relatively low number of Ne (mean of 4.9 alleles per locus) which shows the number of equally frequent alleles, respect to the high number of alleles (mean A = 12 alleles per locus). SK gene had the highest percentage (46.7%) of alleles present in only one tree, and on the other extreme, GST only showed 8.3%. Besides, their classification as rare alleles is still preliminary because of the relatively limited number of trees (54 individuals) analyzed here.

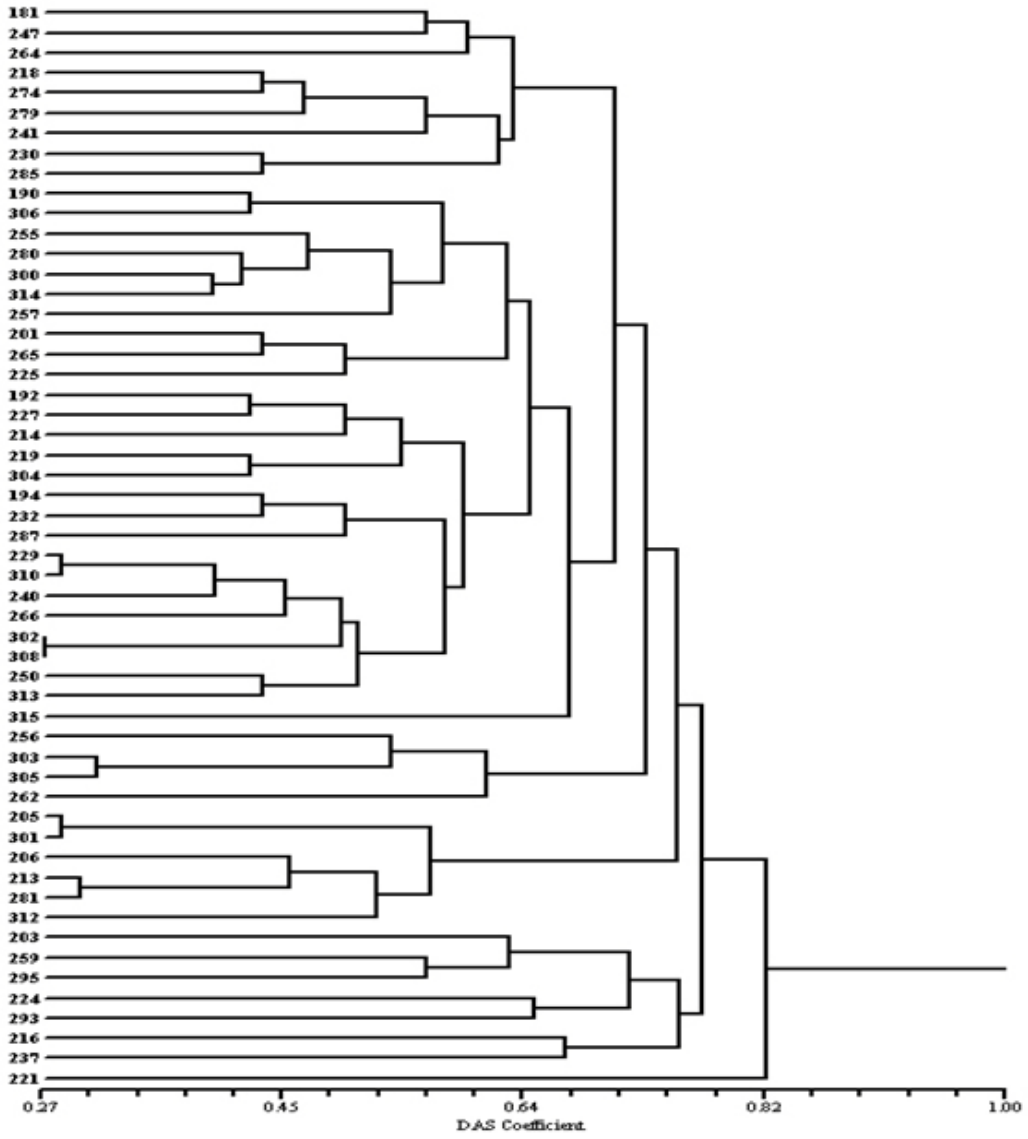

Fig. 3 Cluster analysis. Phenogram of a *Eucalyptus globulus* sample obtained by UPGMA cluster analyses based on shared allele distance (DAS) calculated with the SSR-CG.

These results are in good agreement with those of Steane et al. (2006), where they could find many rare alleles across the whole population of *E. globulus*, showing a relatively low Ne number (average 6.1 alleles per locus) and a high number of alleles (mean A = 19.4 alleles per locus).

Functional perspectives of SSR

Simple sequence repeat variation within genes could be very critical for normal gene activity because repetition motif expansion or contraction could directly affect the corresponding gene expression and even cause phenotypic changes like codon frame shift, fluctuation of gene expression, inactivation of gene activity, and/or change of function and, eventually, phenotypic changes (Li et al. 2004). Genetic microsatellite markers often have known or 'putative' functions when they are gene targeted, with the potential of representing functional differences in those cases where polymorphisms in the repeat motifs affect the function of the gene in which they reside (Andersen and Lubberstedt, 2003). Although this association discovery has not been carried out yet for *Eucalyptus*, the characterization of new markers located in candidate genes for wood quality demonstrated to be useful for genetic studies such as QTLs mapping and marker-assisted breeding in *Eucalyptus*.

It is not unusual to find SSRs mainly, but not exclusively, in the 5' and 3' untranslated regions (UTRs). From a previous analysis of public *E. globulus's* EST (Acuña et al. 2011), 72% of the SSR sequences discovered were within ORF sequences, while 28% were not. Although more SSR are usually found in UTRs than in ORF (Morgante et al. 2002) this unexpected result might be caused by an overestimation of the ORF length.

It is expected that tri- and hexanucleotide repeats would occur more frequently than other motifs in coding sequences (Metzgar et al. 2000). Such dominance of triplets over other repeats in coding regions may be explained on the basis of the selective disadvantage of non-trimeric SSR variants in coding regions, possibly causing frame-shift mutations (Metzgar et al. 2000).

In this work, we did not find trinucleotide repeats inside exons. However, two out of three non-trinucleotide repeats were multiples of three, which did not modify the reading frame (*i.e*. three motifs of five nucleotides = 15 bp = 5 codons) (Table 2).

In order to find out if any of the candidate gene alleles differing in SSR motif repetition numbers are related to wood qualities, genetic association studies are in progress. In those cases where linkage disequilibrium is confirmed, qPCR and/or differential RNAseq studies will be carried out in order to investigate if transcriptional regulation of any of these genes has a key role in these traits.

CONCLUDING REMARKS

This study contributes to previously reported analyses of nucleotide diversity in candidate genes in *Eucalyptus* (Poke et al. 2003; Thumma et al. 2005). This is the first report on the identification and diversity analysis of microsatellite markers located inside putative wood quality candidate genes. These eight microsatellites belong to seven genes involved in lignin, cellulose and hemicellulose biosynthesis: CCR, HMT, SK, GST, transcription factor LIM1, CesA3 and XTH. The characterized microsatellites were located in promoters, introns and exons and show high levels of polymorphism (mean PIC:0,769).

Potentially, these functional markers may contribute to the validation of gene and QTL positions in multiple pedigrees in *Eucalyptus globulus*.

ACKNOWLEDGMENTS

The authors thank Dr. Andrea Puebla, Verónica Nishinakamasu and Pablo Vera for gene sequencing and their help during genotyping. We are grateful to Corina Fusari for their scientific assistance during the first part of this work. The first author thanks her colleagues from IRB and IB of INTA for their continuous help and support.

Financial support: This research was supported by the ANPCyT/FONCYT, BID 1728 OC/AR, PICT-2008-00118 and INTA-PE 041120 and BiotecSur UE 127118.

REFERENCES

ACUÑA, C.; FERNANDEZ, P.; VILLALBA, P.; GARCÍA, M.; HOPP, E. and MARCUCCI POLTRI, S. (2011). Discovery, validation and in silico functional characterization of EST-SSR markers in *Eucalyptus globulus*. *Tree Genetics and Genomes,* Online first. **[CrossRef]**

ANDERSEN, J.R. and LUBBERSTEDT, T. (2003). Functional markers in plants. *Trends in Plant Science*, vol. 8, no. 11, p. 554-560. **[CrossRef]**

ARNAUD, D.; DEJARDIN, A.; LEPLE, J-.C.; LESAGE-DESCAUSES, M.C. and PILATE, G. (2007). Genome-wide analysis of *LIM* gene family in *Populus trichocarpa, Arabidopsis thaliana,* and *Oryza sativa. DNA Research*, vol. 14, no. 3, p. 103-116. **[CrossRef]**

BOTSTEIN, D.; WHITE, R.L.; SKOLNICK, M. and DAVIS, R.W. (1980). Construction of genetic linkage map in man using restriction length polymorphisms. *The American Journal of Human Genetics*, vol. 32, no. 3, p. 314-331.

BRONDANI, R.P.; BRONDANI, C. and GRATTAPAGLIA, D. (2002). Towards a genus-wide reference linkage map for *Eucalyptus* based exclusively on highly informative microsatellite markers. *Molecular Genetics and Genomics,* vol. 267, no. 3, p. 338-347. **[CrossRef]**

BRONDANI, R.P.; WILLIAMS, E.R.; BRONDANI, C. and GRATTAPAGLIA, D. (2006). A microsatellite-based consensus linkage map for species of *Eucalyptus* and a novel set of 230 microsatellite markers for the genus. *BMC Plant Biology,* vol. 6, no. 20. **[CrossRef]**

BUNDOCK, P.C.; HAYDEN, M. and VAILLANCOURT, R.E. (2000). Linkage maps of *Eucalyptus globulus* using RAPD and microsatellite markers. *Silvae Genetica*, vol. 49, no. 4-5, p. 223-232.

BUTCHER, P.A.; MCDONALD, M.W. and BELL, J.C. (2009). Congruence between environmental parameters, morphology and genetic structure in Australia's most widely distributed eucalypt, *Eucalyptus camaldulensis*. *Tree Genetics & Genomes,* vol. 5, no. 1, p. 189-210. **[CrossRef]**

BYRNE, M.; MARQUEZGARCIA, M.I.; UREN, T.; SMITH, D.S. and MORAN, G.F. (1996). Conservation and genetic diversity of *Microsatellite loci* in the genus *Eucalyptus*. *Australian Journal of Botany*, vol. 44, no. 3, p. 331-341. **[CrossRef]**

CHAKRABORTY, R. and JIN, L. (1993). A unified approach to study hyper-variable polymorphisms: Statistical considerations of determining relatedness and population distances. In: PENA, S.D.J.; CHAKRABORTY, R.; EPPLEN, J.T. and JEFFREYS, A.J. eds., Basel, Birkhauser, *DNA fingerprinting: State of the Science,* p. 153-175.

CHYBICKI, I.J. and BURCZYK, J. (2009). Simultaneous estimation of null alleles and inbreeding coefficients. *Journal of Heredity,* vol. 100, no. 1, p. 106-113. **[CrossRef]**

CONESA, A.; GÖTZ, S.; GARCIA-GOMEZ, J.M.; TEROL, J.; TALON, M. and ROBLES, M. (2005). Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, vol. 21, no. 18, p. 3674-3676. **[CrossRef]**

CREUX, N.M.; RANIK, M.; BERGER, D.K. and MYBURG, A.A. (2008). Comparative analysis of orthologous cellulose synthase promoters from *Arabidopsis, Populus* and *Eucalyptus*: Evidence of conserved regulatory elements in angiosperms. *New Phytologist,* vol. 179, no. 3, p. 722-737. **[CrossRef]**

DEAN, J.V.; DEVARENNE, T.P.; LEE, I.S. and ORLOFSKY, L.E. (1995). Properties of a maize glutathione *S*-transferase that conjugates coumaric acid and other phenylpropanoids. *Plant Physiology*, vol. 108, no. 3, p. 985-994.

DEAN, J.V. and DEVARENNE, T.P. (1997). Peroxidase-mediated conjugation of glutathione to unsaturated phenylpropanoids. Evidence against glutathione S-transferase involvement. *Physiologia Plantarum*, vol. 99, no. 2, p. 271-278. **[CrossRef]**

DUTKOWSKI, G.W. and POTTS, B.M. (1999). Geographic patterns of genetic variation in *Eucalyptus globulus* ssp. *globulus* and a revised racial classification. *Australian Journal of Botany*, vol. 47, no. 2, p. 237-263. **[CrossRef]**

ECKERT, A.J.; PANDE, B.; ERSOZ, E.S.; WRIGHT, M.H.; RASHBROOK, V.K.; NICOLET, C.M. and NEALE, D.B. (2009). High-throughput genotyping and mapping of single nucleotide polymorphisms in loblolly pine (*Pinus taeda* L.). *Tree Genetics & Genomes,* vol. 5, no. 1, p. 225-234. **[CrossRef]**

EDWARDS, R.; DIXON, D.P. and WALBOT, V. (2000). Plant glutathione *S*-transferases: Enzymes with multiple functions in sickness and in health. *Trends in Plant Science*, vol. 5, no. 5, p. 193-198. **[CrossRef]**

EVANS, R. and ILIC, J. (2001). Rapid prediction of wood stiffness from microfibril angle and density. *Forest Products Journal,* vol. 51, no. 3, p. 53-57.

FARIA, D.A.; MAMANI, E.M.C.; PAPPAS, M.R.; PAPPAS Jr. G.J. and GRATTAPAGLIA, D. (2010). A selected set of EST-derived microsatellites, polymorphic and transferable across 6 species of *Eucalyptus*. *Journal of Heredity*, vol. 101, no. 4, p. 512-520. **[CrossRef]**

FARIA, D.A.; MAMANI, E.M.C.; PAPPAS G.J. and GRATTAPAGLIA, D. (2011). 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*, vol. 7, no. 1, p. 63-77. **[CrossRef]**

FOUCART, C.; PAUX, E.; LADOUCE, N.; SAN-CLEMENTE, H.; GRIMA-PETTENATI, J. and SIVADON, P. (2006). Transcript profiling of a xylem vs phloem cDNA subtractive library identifies new genes expressed during xylogenesis in *Eucalyptus*. *New Phytologist*, vol. 170, no. 4, p. 739-752. **[CrossRef]**

FREEMAN, J.S.; WHITTOCK, S.P.; POTTS, B.M. and VAILLANCOURT, R.E. (2009). QTL influencing growth and wood properties in *Eucalyptus globulus*. *Tree Genetics & Genomes,* vol. 5, no. 4, p. 713-722. **[CrossRef]**

GENOVESI, V.; FORNALÉ, S.; FRY, S.C.; RUEL, K.; FERRER, P.; ENCINA, A.; SONBOL, F.M.; BOSCH, J.; PUIGDOMÉNECH, P.; RIGAU, J. and CAPARRÓS-RUIZ, D. (2008). ZmXTH1, a new xyloglucan endotransglucosylase/hydrolase in maize, affects cell wall structure and composition in *Arabidopsis thaliana*. *Journal of Experimental Botany*, vol. 59, no. 4, p. 875-889. **[CrossRef]**

GION, J.M.; RECH, P.; GRIMA-PETTENATI, J.; VERHAEGEN, D. and PLOMION, C. (2000). Mapping candidate genes in *Eucalyptus* with emphasis on lignification genes. *Molecular Breeding*, vol. 6, no. 5, p. 441-449. **[CrossRef]**

GION, J.M.; BOUDET, C.; GRIMA-PETTENATI, J.; PICHAVANT, F.H.; PLOMION, C.; BAILLERES, H. and VERHAEGEN, D. (2001). A candidate genes approach identifies CCR, PAL and C4H as loci for syringyl/guaiacyl ratio in an interespecific hybrid between *E. urophylla* and *E. grandis*. In: *IUFRO – International Union of Forest Research Organizations Conference: Developing the eucalypt of the future.* (10th - 15th September, 2001, Valdivia, Chile). Abstract.

GION, J.M.; CAROUCHE, A.; DEWEER, S.; BEDON, F.; PICHAVANT, F.; CHARPENTIER, J.P.; BAILLERES, H.; ROZENBERG, P.; CAROCHA, V.; OGNOUABI, N.; VERHAEGEN, D.; GRIMA-PETTENATI, J.; VIGNERON, P. and PLOMION, C. (2011). Comprehensive genetic dissection of wood properties in a widely-grown tropical tree: *Eucalyptus*. *BMC Genomics*, vol. 12, no. 301. **[CrossRef]**

GOICOECHEA, M.; LACOMBE, E.; LEGAY, S.; MIHALJEVIC, S.; RECH, P.; JAUNEAU, A.; LAPIERRE, C.; POLLET, B.; VERHAEGEN, D.; CHAUBET-GIGOT, N. and GRIMA-PETTENATI, J. (2005). EgMYB2, a new transcriptional activator from *Eucalyptus* xylem, regulates secondary cell wall formation and lignin biosynthesis. *The Plant Journal*, vol. 43, no. 4, p. 553-567. **[CrossRef]**

GRATTAPAGLIA, D. and SEDEROFF, R.R. (1994). Genetic linkage map of *Eucalyptus grandis* and *Eucalyptus urophylla* using a Pseudo-Testcross: Mapping strategy and RAPD markers. *Genetics*, vol. 137, no. 4, p. 1121-1137.

GRATTAPAGLIA, D. and KIRST, M. (2008). *Eucalyptus* applied genomics: From gene sequences to breeding tools. *New Phytology,* vol. 179, no. 4, p. 911-929. **[CrossRef]**

GRATTAPAGLIA, D.; SILVA-JUNIOR, O.; KIRST, M.; DE LIMA, B.; FARIA, D. and PAPPAS, G. (2011). High-throughput SNP genotyping in the highly heterozygous genome of *Eucalyptus*: Assay success, polymorphism and transferability across species. *BMC Plant Biology*, vol. 11, no. 65. **[CrossRef]**

HERRMANN, K.M. and WEAVER, L.M. (1999). The shikimate pathway. *Annual Review of Plant Physiology and Plant Molecular Biology*, vol. 50, p. 473-503. **[CrossRef]**

HERTZBERG, M.; ASPEBORG, H.; SCHRADER, J.; ANDERSSON, A.; ERLANDSSON, R.; BLOMQVIST, K.; BHALERAO, R.; UHLEN, M.; TEERI, T.T.; LUNDEBERG, J.; SUNDBERG, B.; NILSSON, P. and SANDBERG, G. (2001). A transcriptional roadmap to wood formation. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, no. 25, p. 14732-14737. **[CrossRef]**

HOISINGTON, D.; KHAIRALLAH, M. and GONZALEZ-DE-LEON, D. (1994). *Laboratory protocols: CIMMYT. Applied Molecular Genetics Laboratory.* 3rd edition. CIMMYT, Mexico, D.F. 102 p. ISBN 968-6923-30-6.

HRMOVA, M.; FARKAS, V.; LAHNSTEIN, J. and FINCHER, G.B. (2007). A Barley xyloglucan xyloglucosyl transferase covalently links xyloglucan, cellulosic substrates, and (1,3;1,4)-*β*-D-glucans. *The Journal of Biological Chemistry,* vol. 282, no. 17, p. 12951-12962. **[CrossRef]**

ISRAELSSON, M.; ERIKSSON, M.E.; HERTZBERG, M.; ASPEBORG, H.; NILSSON, P. and MORITZ, T. (2003). Changes in gene expression in the wood-forming tissue of transgenic hybrid aspen with increased secondary growth. *Plant Molecular Biology*, vol. 52, no. 4, p. 893-903. **[CrossRef]**

JONES, L.; ENNOS, A.R. and TURNER, S.R. (2001). Cloning and characterization of *irregular xylem4 (irx4)*: A severely lignin-deficient mutant of *Arabidopsis*. *The Plant Journal*, vol. 26, no. 2, p. 205-216. **[CrossRef]**

KAMPRANIS, S.C.; DAMIANOVA, R.; ATALLAH, M.; TOBY, G.; KONDI, G.; TSICHLIS, P.N. and MAKRIS, A.M. (2000). A novel plant glutathione *S*-transferase/peroxidase suppresses Bax lethality in yeast. *The Journal of Biological Chemistry,* vol. 275, no. 38, p. 29207-29216. **[CrossRef]**

KAWAOKA, A.; NANTO, K.; ISHII, K. and EBINUMA, H. (2006). Reduction of lignin content by suppression of expression of the LIM domain transcription factor in *Eucalyptus camaldulensis. Silvae Genetica*, vol. 55, no. 6, p. 269-277.

KILILI, K.G.; ATANASSOVA, N.; VARDANYAN, A.; CLATOT, N.; AL-SABARNA, K.; KANELLOPOULOS, P.N.; MAKRIS, A.M. and KAMPRANIS, S.C. (2004). Differential roles of tau class glutathione *S*-transferases in oxidative stress. *The Journal of Biological Chemistry*, vol. 279, no. 23, p. 24540-24551. **[CrossRef]**

KIRST, M.; MYBURG, A.A.; DE LEÓN, J.P.G.; KIRST, M.E.; SCOTT, J. and SEDEROFF, R. (2004). Coordinated genetic regulation of growth and lignin revealed by quantitative trait locus analysis of cDNA microarray data in an interspecific backcross of *eucalyptus*. *Plant Physiology*, vol. 135, no. 4, p. 2368-2378. **[CrossRef]**

KUMPATLA, S.P. and MUKHOPADHYAY, S. (2005). Mining and survey of simple sequence repeats in expressed sequence tags of dicotyledonous species. *Genome*, vol. 48, no. 6, p. 985-998. **[CrossRef]**

LACOMBE, E.; HAWKINS, S.; DOORSSELAERE, J.V.; PIQUEMAL, J.; GOFFNER, D.; POEYDOMENGE, O.; BOUDET, A.M. and GRIMA-PETTENATI, J. (1997). Cinnamoyl CoA reductase, the first committed enzyme of the lignin branch biosynthetic pathway: Cloning, expression and phylogenetic relationships. *Plant Journal*, vol. 11, no. 3, p. 429-441. **[CrossRef]**

LACOMBE, E.; VAN DOORSSELAERE, J.; BOERJAN, W.; BOUDET, A.M. and GRIMA-PETTENATI, J. (2000). Characterization of *cis*-elements required for vascular expression of the *cinnamoyl CoA reductase* gene and for protein-DNA complex formation. *The Plant Journal*, vol. 23, no. 5, p. 663-676. **[CrossRef]**

LANGELLA, O. (2002). Populations 1.2.32. Logiciel de génétique des populations. [cited 5 June 2011]. Available from Internet: http://www.bioinformatics.org/project/?group_id=84.

LI, Y.C.; KOROL, A.B.; FAHIMA, T.; BEILES, A. and NEVO, E. (2002). Microsatellites: Genomic distribution, putative functions and mutational mechanisms: A review. *Molecular Ecology,* vol. 11, no. 12, p. 2453-2465. **[CrossRef]**

LI, Y.C.; KOROL, A.B.; FAHIMA, T. and NEVO, E. (2004). Microsatellites within genes: Structure, function, and evolution. *Molecular Biology and Evolution*, vol. 21, no. 6, p. 991-1007. **[CrossRef]**

LOYALL, L.; UCHIDA, K.; BRAUN, S.; FURUYA, M. and FROHNMEYER, H. (2000). Glutathione and a UV light-induced glutathione *S*-transferase are involved in signaling to chalcone synthase in cell cultures. *The Plant Cell*, vol. 12, no. 10, p. 1939-1950. **[CrossRef]**

MARCUCCI POLTRI, S.N.; ZELENER, N.; RODRIGUEZ TRAVERSO, J.; GELID, P. and HOPP, H.E. (2003). Selection of a seed orchard of *Eucalyptus dunnii* based on genetic diversity criteria calculated using molecular markers. *Tree Physiology,* vol. 23, no. 9, p. 625-632.

MARQUES, C.M.; ARAUJO, J.A.; FERREIRA, J.G.; WHETTEN, R.; O'MALLEY, D.M.; LIU, B.H. and SEDEROFF, R.R. (1998). AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis. TAG Theoretical and Applied Genetics,* vol. 96, no. 6-7, p. 727-737. **[CrossRef]**

METZGAR, D.; BYTOF, J. and WILLS, C. (2000). Selection against frameshift mutations limits microsatellite expansion in coding DNA. *Genome Research*, vol. 10, no. 1, p. 72-80.

MOLINA, C. and GROTEWOLD, E. (2005). Genome wide analysis of *Arabidopsis* core promoters. *BMC Genomics*, vol. 6, no. 25. **[CrossRef]**

MORAN, G.F.; THAMARUS, K.; RAYMOND, C.A.; QIU, D.; UREN, T. and SOUTHERTON, S.G. (2002). Genomics of *Eucalyptus wood* traits. *Annals of Forest Science*, vol. 59, no. 5-6, p. 645-650. **[CrossRef]**

MORGANTE, M.; HANAFEY, M. and POWELL, W. (2002). Microsatellites are preferentially associated with non-repetitive DNA in plant genomes. *Nature Genetics*, vol. 30, no. 2, p. 194-200. **[CrossRef]**

MUELLER, L.A.; GOODMAN, C.D.; SILADY, R.A. and WALBOT, V. (2000). AN_9, a petunia glutathione *S*-transferase required for anthocyanin sequestration, is a flavonoid-binding protein. *Plant Physiology*, vol. 123, no. 4, p. 1561-1570. **[CrossRef]**

MYBURG, A.A.; POTTS, B.M.; MARQUES, C.M.; KIRST, M.; GION, J.M.; GRATTAPAGLIA, D. and GRIMA-PETTENATI, J. (2007). *Eucalyptus*. In: CHITTARANJAN, K. ed. *Genome mapping and molecular breeding in plants*. Forest Trees, New York, NY, USA: Springer, vol. 7, p. 115-160.

NASCIMENTO DE SOUSA, S.; FINKELDEY, R. and GAILING, O. (2005). Experimental verification of microsatellite null alleles in Norway spruce (*Picea abies* [L.] Karst.): Implications for population genetic studies. *Plant Molecular Biology Reporter*, vol. 23, no. 2, p. 113-119. **[CrossRef]**

NEVES, L.G.; MAMANI, E.M.C.; ALFENAS, A.C.; KIRST, M. and GRATTAPAGLIA, D. (2011). A high-density transcript linkage map with 1,845 expressed genes positioned by microarray-based Single Feature Polymorphisms (SFP) in *Eucalyptus. BMC Genomics*, vol. 12, no. 189. **[CrossRef]**

NOVAES, E.; DROST, D.; FARMERIE, W.; PAPPAS, G.; GRATTAPAGLIA, D.; SEDEROFF, R. and KIRST, M. (2008). High-throughput gene and SNP discovery in *Eucalyptus grandis*, an uncharacterized genome. *BMC Genomics*, vol. 9, no. 312. **[CrossRef]**

PAUX, E.; TAMASLOUKHT, M.; LADOUCE, N.; SIVADON, P. and GRIMA-PETTENATI, J. (2004). Identification of genes preferentially expressed during wood formation in *Eucalyptus. Plant Molecular Biology*, vol. 55, no. 2, p. 263-280. **[CrossRef]**

PAUX, E.; CAROCHA, V.; MARQUES, C.; MENDES DE SOUSA, A.; BORRALHO, N.; SIVADON, P. and GRIMA-PETTENATI, J. (2005). Transcript profiling of *Eucalyptus* xylem genes during tension wood formation. *New Phytology*, vol. 167, no. 1, p. 89-100. **[CrossRef]**

PEAKALL, R. and SMOUSE, P.E. (2006). GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, vol. 6, no. 1, p. 288-295. **[CrossRef]**

POKE, F.S.; VAILLANCOURT, R.E.; ELLIOTT, R.C. and REID, J.B. (2003). Sequence variation in two lignin biosynthesis genes, cinnamoyl CoA reductase (*CCR*) and cinnamyl alcohol dehydrogenase 2 (*CAD2*). *Molecular Breeding*, vol. 12, no. 2, p. 107-118. **[CrossRef]**

POKE, F.S.; RAYMOND, C.A.; REID, J.B. and VAILLANCOURT, R.E. (2004). The effect of a single amino acid substitution in a lignin biosynthesis enzyme on wood properties in *Eucalyptus globules*. In: *Eucalyptus in a Changing World-International IUFRO Conference*, (11^{th} - 14^{th} October 2004, Aveiro, Portugal), p. 388-394.

POKE, F.S.; VAILLANCOURT, R.E.; POTTS B.M. and REID, J.B. (2005). Genomic research in *Eucalyptus*. *Genetica,* vol. 125, no. 1, p. 79-101. **[CrossRef]**

RANIK, M. and MYBURG, A.A. (2006). Six new cellulose synthase genes from *Eucalyptus* are associated with primary and secondary cell wall biosynthesis. *Tree Physiology*, vol. 26, no. 5, p. 545-556. **[CrossRef]**

RAYMOND, C.A. and APIOLAZA, L.A. (2004). Incorporating wood quality and deployment traits in *Eucalyptus globulus* and *Eucalyptus nitens.* In: WALTER, C. and CARSON, M. eds. *Plantation Forest Biotechnology for the 21^{st} Century*, vol. 5, p. 87-89.

RENGEL, D.; CLEMENTE, H.S.; SERVANT, F.; LADOUCE, N.; PAUX, E.; WINCKER, P.; COULOUX, A.; SIVADON, P. and GRIMA-PETTENATI, J. (2009). A new genomic resource dedicated to wood formation in *Eucalyptus*. *BMC Plant Biology*, vol. 9, no. 36. **[CrossRef]**

SEXTON, T.R.; HENRY, R.J.; HARWOOD, C.E.; THOMAS, D.; MCMANUS, L.J.; RAYMOND, C. and HENSON, M. and SHEPHERD, M. (2012). Pectin methylesterase genes influence solid wood properties of *Eucalyptus pilularis. Plant* Physiology, vol. 158, no. 1, p. 531-541. **[CrossRef]**

STACKPOLE, D.J.; VAILLANCOURT, R.E.; ALVES, A.M.; RODRIGUES, J. and POTTS, B.M. (2011). Genetic variation in the chemical components of *Eucalyptus globulus* wood. *G3 Genes Genomes Genetics*, vol. 1, no. 2, p. 151-159.

STEANE, D.; CONOD, N.; JONES, R.; VAILLANCOURT, R. and POTTS, B. (2006). A comparative analysis of population structure of a forest tree, *Eucalyptus globulus* (Myrtaceae), using microsatellite markers and quantitative traits. *Tree Genetics & Genomes*, vol. 2, no. 1, p. 30-38. **[CrossRef]**

TANG, J.; BALDWIN, S.J.; JACOBS, J.M.; LINDEN, C.G.; VOORRIPS, R.E.; LEUNISSEN, J.A.A.; VAN ECK, H. and VOSMAN, B. (2008). Large-scale identification of polymorphic microsatellites using an *in silico* approach. *BMC Bioinformatics*, vol. 9, no. 374. **[CrossRef]**

THAMARUS, K.A.; GROOM, K.; MURRELL, J.; BYRNE, M. and MORAN, G.F. (2002). A genetic linkage map for *Eucalyptus globulus* with candidate loci for wood, fibre, and floral traits. *TAG Theoretical and Applied Genetics,* vol. 104, no. 2-3, p. 379-387. **[CrossRef]**

THOM, R.; CUMMINS, I.; DIXON, D.P.; EDWARDS, R.; COLE, D.J. and LAPTHORN, A.J. (2002). Structure of a tau class glutathione *S*-transferase from wheat active in herbicide detoxification. *Biochemistry*, vol. 41, no. 22, p. 7008-7020. **[CrossRef]**

THUMMA, B.R.; NOLAN, M.F.; EVANS, R. and MORAN, G.F. (2005). Polymorphisms in *Cinnamoyl CoA reductase* (*CCR*) are associated with variation in microfibril angle in *Eucalyptus* spp. *Genetics*, vol. 171, no. 3, p. 1257-1265. **[CrossRef]**

THUMMA, B.R.; SOUTHERTON, S.G.; BELL, J.C.; OWEN, J.V.; HENERY, M.L. and MORAN, G.F. (2010). Quantitative trait locus (QTL) analysis of wood quality traits in *Eucalyptus nitens*. *Tree Genetics & Genomes*, vol. 6, no. 2, p. 305-317. **[CrossRef]**

VERHAEGEN, D. and PLOMION, C. (1996). Genetic mapping in *Eucalyptus urophylla* and *E. grandis.* RAPD markers. *Genome*, vol. 39, no. 6, p. 1051-1061. **[CrossRef]**

WEGRZYN, J.L.; ECKERT, A.J.; CHOI, M.; LEE, J.M.; STANTON, B.J.; SYKES, R.; DAVIS, M.F.; TSAI, C.J. and NEALE, D.B. (2010). Association genetics of traits controlling lignin and cellulose biosynthesis in black cottonwood (*Populus trichocarpa*, Salicaceae) secondary xylem. *New Phytologist,* vol. 188, no. 2, p. 515-532. **[CrossRef]**

ZHANG, X.; YAZAKI, J.; SUNDARESAN, A.; COKUS, S.; CHAN, S.W.; CHEN, H.; HENDERSON, I.R.; SHINN, P.; PELLEGRINI, M.; JACOBSEN, and ECKER, J.R. (2006). Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis. Cell,* vol. 126, no. 6, p. 1189-1201. **[CrossRef]**

How to reference this article:

ACUÑA, C.V.; VILLALVA, P.V.; GARCÍA, M.; PATHAUER, P.; HOPP, H.E. and MARCUCCI POLTRI, S.N. (2012). Microsatellite markers in candidate genes for wood properties and its application in functional diversity assessment in *Eucalyptus globulus*. *Electronic Journal of Biotechnology*, vol. 15, no. 2, p. 1-17. http://dx.doi.org/10.2225/vol15-issue2-fulltext-3