Nucleotidic variability of the *Eucalyptus urophylla* CAD2 gene in populations along an altitudinal gradient in Timor island

B. Favreau, J.-M. Gion, D. Verhaegen, J.-M. Bouvet
CIRAD BIOS, UPR39 Research Unit "Genetic Diversity and Breeding of Forest Tree Species"
TA A39/C, 34398 Montpellier Cedex 5, France.
Contacts: benedicte.favreau@cirad.fr
jean-marc.bouvet@cirad.fr

Introduction

Single nucleotide polymorphism (SNP) in candidate genes offers new perspectives to analyze the impact of natural selection on gene variation along environmental gradient and to understand the molecular basis of adaptation (Garcia-Gil et al, 2003; Pot et al, 2005).
The aim of our study was to describe the nucleotidic variability within the Cinnamyl Alcohol Dehydrogenase (CAD) gene, a structural gene of the lignin biosynthesis pathway, within a representative sample of *Eucalyptus urophylla* distributed in the natural population at various altitudes in the Timor island. Previous studies have shown that the altitudinal gradient affects strongly the growth of this species, the individuals from lower altitude exhibiting a better growth (Tripiana et al, 2007).

Material and Methods

• Genetic material : Seeds of *Eucalyptus urophylla* (*Symphyomyrtus* subgenus) from Timor island were collected in 1975 and preserved in the CIRAD Genetic Forestry Laboratory. For this study, some seeds were planted and DNA from 87 individuals was extracted from leaves. This sample represents 2 provenances of the East of Timor Island (Remexio, Maubisse), on an altitudinal gradient from 500 m to 1760 m (Figure 1). From these 87 genotypes, 10 individuals have been used for the CAD2 gene sequencing and the SNPs identification.

• Gene sequencing and nucleotide variation determination: 11 overlapping primers were designed according to the CAD2 conserved regions of *E. gunnii* (X75480), *E. botryoides* (D16624), *E. globulus* (AF038561) and *E. saligna* (AF294793). The OLIGO EXPLORER v1.1.0 software (Kuulasmaa, 2000-2002) has been used for primer design. The 11 fragments (≈ 600 bp) were PCR amplified on 10 individuals and the 110 amplicons were sequenced on both sides. The gene was rebuilt by alignment of the fragment sequences and the SNPs identified by comparison of the 10 genotype sequences.

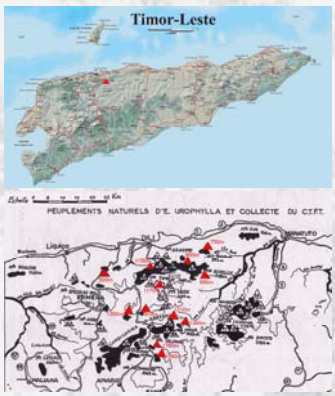

Figure 1 : Localization of the *E. urophylla* genotypes from Remexio and Maubisse (East of Timor Island)

Results & Discussion

• Description of the nucleotidic variability in the CAD2 gene

100% of the *E. urophylla* CAD2 gene was sequenced (Figure 2):
- 5 exons and 4 introns
- total length of 5395 bp
- 93% similarity with the *E. gunnii* CAD2 gene

Main results of the nucleotidic variability analysis on 10 individuals (Table 1). :
- 160 SNPs and 16 INDELs identified
- high frequency of the polymorphism but the heterogeneous distribution in the gene with 1 SNP per 74 bp in the exonic region and 1 SNP per 21 bp in the intronic region
- 4 non synonymous SNPs and 1 SNP translated in polar/non polar aminoacids (Table 2).

These results when compared with 3 nucleotidic variability studies on CAD and CCR from *E. globulus* and *E. urophylla* show that, in spite of a smaller number of sequenced genotypes, the SNP and INDEL frequencies are in the same range (Table 3).

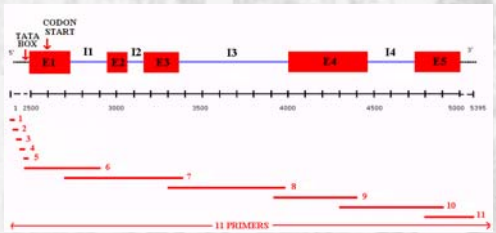

Figure 2 : Sequencing of the *E. urophylla* CAD2 gene with 11 primers

	SNP Frequency	INDEL Frequency	Transversion number Tv	Transition number Ts	Ts/Tv
5'	1/36	1/360	21	49	2
Exon 1	1/104	1/104			
Exon 2	1/57	0			
Exon 3	1/74	0	5	11	2
Exon 4	1/110	0			
Exon 5	1/39	0			
Intron 1	1/20	1/219			
Intron 2	1/11	0	20	39	2
Intron 3	1/24	1/317			
Intron 4	1/23	1/151			
3'	1/30	1/228	6	9	2

Table 1: Nucleotid variability detected in the CAD gene on 10 *E. urophylla* genotypes

	Number of sequenced SNPs	Non synonymous mutation	Aminoacid identities	Aminoacids properties
Exon 1	2 before ATG	----------	----------	----------
Exon 2	2	0	----------	----------
Exon 3	3	1	Methionine/Valine	Non-polar / Non-polar
Exon 4	4	0	----------	----------
Exon 5	5	3	Isoleucine/Methionine Thréonine/Méthionine Alanine/Valine	Non-polar / Non-polar Polar / Non polar Non-polar / Non-polar

Table 2 : Mutation characteristics in the CAD2 exonic region on 10 *E. urophylla* genotypes

• SNP typing

To identify the SNPs on the 87 genotypes, we use the TDI-FP assay (Template-directed Dye-terminator Incorporation with Fluorescent Polarization detection). To set up the assay, 1 SNP was analyzed on 7 sequenced genotypes. The SNP, identified by TDI-FP and sequencing, is compared (Table 4). This result shows that there is a perfect correlation of the SNP identification between the 2 methods and that the TDI-FP can be use for high-throughput typing.

Indivivuals	SNP sequenced	SNP typing 1	SNP typing 2
Blank		no SNP	no SNP
1	G/A	G/A	G/A
2	G/A	G/A	G/A
3	A	A	A
4	A	A	A
5	no sequence	A	A
6	G/A	G/A	G/A
7	A	A	no SNP

Table 4 : Comparison of 1 SNP analyzed by TDI-FP and sequencing

		E. globulus CCR (Poke et al, 2003)	*E. globulus* CAD2 (Poke et al, 2003)	*E. urophylla* CCR (Champurney et al, 2005)	*E. urophylla* CAD2
	Number sequenced genotypes	23	23	22 à 36	**10**
	% sequenced gene	65	45	82	**100**
EXONS	% sequenced region	100	100	72	**100**
	SNP frequency	1/48	1/147	1/27	**1/74**
	SNP synonymous	9	6	17	**10**
	SNP non synonymous	12	2	6	**4**
	INDEL number	0	0	0	**2**
INTRONS	% sequenced region	48	0	92	**100**
	SNP frequency	1/33	---	1/22	**1/21**
	INDEL number	6	---	9	**5**

Table 3 : Comparison of the nucleotidic variability between different studies : CAD2 vs CCR and *E. urophylla* vs *E. globulus*

Perspectives

- High-throughput SNPs typing of the 87 *E. urophylla* genotypes
- Haplotypic reconstruction with the Phase software (Stephens et al, 2003)
- Study of the CAD2 gene polymorphism structuration on the SNP and haplotype level
- Association studies with progeny/provenance trial established in Congo

Bibliography

Champurney N. et al, 2005. *IUFRO*; Poke F.S. et al, 2003. *Molecular Breeding*, 12:107-118; Garcia-Gil M. et al, 2003. *Molecular Ecology*, 12:1195-1206; Pot D. et al, 2005. *New Physiologist*, 167(1):101-112; Stephens M. et al 2003. *American Journal of Human Genetics,* 73:1162-1169; Tripiana V. et al, 2007. *Canadian Journal of Forest Research*, 37(4):773-785

CIRAD BIOS, UPR39 Genetic Diversity and Breeding of Forest Tree Species
TA A39/C, 34398 Montpellier Cedex 5, FRANCE.
Web site: www.cirad.fr