

Diplôme d'études approfondies
de Biologie Cellulaire et Moléculaire
Végétale

***Clonage et cartographie de
gènes impliqués dans la
biosynthèse des lignines chez
l'Eucalyptus***

Philippe RECH

Responsables de stage ;

Jacqueline GRIMA-PETTENATI et Daniel VERHAEGEN

Résumé

Eucalyptus, première espèce de reboisement au monde, est largement utilisé en industrie papetière. Toutefois, les lignines qui représentent un quart de la biomasse totale du bois constituent un produit indésirable qui limite les rendements en pâte à papier. L'extraction de ces composés phénoliques est réalisée par des processus chimiques onéreux et très polluants. Il apparaît donc souhaitable, d'un point de vue économique et environnemental, de développer des variétés d'*Eucalyptus*, présentant une réduction de teneurs en lignines et/ou des lignines plus facilement extractibles.

Une stratégie envisageable est d'exploiter la variabilité génétique qui contrôle l'hétérogénéité naturelle des lignines dans des programmes de sélection génétique. Dans cet objectif, on peut envisager d'utiliser des marqueurs moléculaires de fonction connue a priori directement impliqués dans le contrôle génétique de la quantité et/ou de la qualité des lignines. Les enzymes impliquées dans la chaîne de biosynthèse des lignines constituent de tels gènes candidats et pourraient servir de marqueurs "diagnostics" dans un programme de sélection.

Dans le cadre de cette approche, l'objectif de mon travail était de cloner des gènes impliqués dans la voie de biosynthèse des lignines chez *Eucalyptus* puis de les positionner sur des cartes génétiques chez *E.grandis* et *E.urophylla*, en utilisant une technique PCR-SSCP (Single Chain Conformation Polymorphism). Ainsi, un ADNc codant pour la caféoyl CoA O-méthyl transférase (CCoAOMT) a été cloné puis positionné sur ces cartes génétiques. De plus, deux gènes codant pour l'acide caféique O méthyl transférase (COMT) ont été localisés sur la carte d'*E urophylla*. Enfin, un fragment du gène de la 4-coumarate CoA ligase (4CL) a été isolé et séquencé.

Abstract

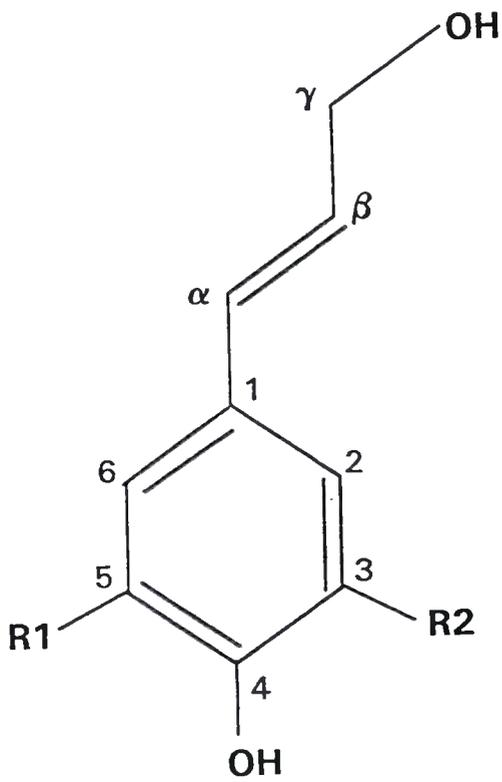
Eucalyptus is the first woody genus planted in the world and is largely used for pulp and paper industries. The abundance of lignins in wood is a limiting factor of pulp yield and quality. The removal of lignin in paper manufacture is achieved by toxic and costly processes. Therefore, from both economical and environmental point of view, it would be beneficial to select trees with less lignin or with lignin that is more extractible.

One possible strategy is to take advantage of the genetic variability which controls the natural lignin heterogeneity. With this aim, the use of molecular markers with known function directly involved in the genetic control of lignin quantity and/or quality would be of great interest. The genes encoding enzymes involved in the lignin biosynthetic pathway seem the best candidates and could potentially be used in markers-assisted selection.

The goal of the work presented here was to clone some of these genes in *Eucalyptus* in order to map them on the genetic maps of *E.grandis* et *E.urophylla*, using the PCR-SSCP (Single Chain Conformation Polymorphism) technique. A cDNA encoding caffeoyl CoA O methyl transferase (CCoAOMT) was cloned allowing the mapping of the corresponding gene. In addition, two genes encoding caffeic acid O methyl transferase (COMT) were mapped on the genetic map of *E urophylla*. Finally, a fragment of the Eucalyptus gene. 4-coumarate CoA ligase (4CL) was isolated and sequenced.

Abréviations

CAD	Cinnamyl Alcohol Dehydrogenase
CCoAOMT	Caffeoyl CoA O-methyltransferase
CCR	Cinnamoyl CoA reductase
4CL	4 coumarate CoA ligase
cM	Centimorgan
COMT	Acide cafeique O-méthyltransférase
PVP	Polyvinylpyrrolidone
RAPD	Random Amplified Polymorphism DNA
SDS	Sodium dodecyl sulfate
SSCP	Single Strand Conformation Polymorphism



$R1=R2=H$ *p*-coumaryl alcohol \rightarrow H units

$R1=H$ $R2=OCH_3$ coniferyl alcohol \rightarrow G units

$R1=R2=OCH_3$ sinapyl alcohol \rightarrow S units

Fig 1. Unités monomériques de lignines (Monolinols)

Introduction

Du fait de son excellente adaptation à différentes conditions écologiques, de sa croissance très rapide, et de la qualité de ces fibres, l'*Eucalyptus* est largement utilisé pour les reboisements industriels. Ainsi, les superficies reboisées actuellement sont estimées à 15 millions d'hectares répartis à travers le monde, dont une grande partie est destinée aux industries papetières (Verhaegen, 1996). De part l'énorme enjeu économique que représentent les *Eucalyptus*, des programmes d'améliorations génétiques, basés sur quelques caractères cibles, se sont développés. Les objectifs de ces programmes, comme celui mené par le CIRAD-Forêt, sont essentiellement orientés vers une forte productivité en bois et une multiplication végétative facile (Vigneron, 1992).

Si l'objectif de ces programmes reste encore d'augmenter le rendement en bois, les efforts de sélection s'orientent aujourd'hui vers d'autres paramètres, tel que la qualité du bois. Les lignines qui représentent un quart de la biomasse totale du bois, conditionnent ses propriétés technologiques et papetières. Elles constituent notamment le facteur limitant majeur de l'exploitation des espèces ligneuses destinées à l'industrie papetière.

Ces biopolymères complexes s'associent avec les composants polysaccharidiques pariétaux (Wallace et Fry, 1994), et confèrent aux cellules spécialisées dans les fonctions de conduction et de soutien, des propriétés particulières, telles la rigidité et l'hydrophobicité. Ainsi, elles jouent un rôle clé dans l'établissement d'un système vasculaire fonctionnel, une résistance aux contraintes mécaniques, et dans la défense des plantes vis à vis des agressions de l'environnement.

Les lignines se caractérisent par une grande variabilité au niveau quantitatif et une grande hétérogénéité de leur composition monomérique. Elles résultent, en effet, de la polymérisation oxydative de trois unités monomériques, les monolignols, qui diffèrent par leur degré de méthylation (Fig. 1) et peuvent être reliées par différents types de liaisons intermonomériques. La nature de ces liaisons conditionnent les propriétés physicochimique de ce polymère.

Dans l'industrie papetière, les lignines, qui sont intimement combinées aux celluloses et hémicelluloses, constituent un produit indésirable dont l'extraction est réalisée par des processus chimiques onéreux et très polluants. Il serait donc

souhaitable, d'un point de vue économique et environnemental, de développer rapidement des variétés d'*Eucalyptus*, présentant des teneurs et/ou des compositions en lignines mieux adaptées aux applications industrielles.

Dans ce but, deux types de stratégies, basées sur les connaissances acquises de la voie biosynthétique des lignines (Boudet *et al.*, 1995), sont actuellement envisagées.

L'ingénierie génétique est une voie pertinente pour contourner les temps de génération relativement longs des croisements conventionnels (Bouvet, 1995), qui retardent l'amélioration génétique des *Eucalyptus*. Ainsi, la modulation de l'expression de gènes impliqués dans la voie de biosynthèse des lignines, pourrait résulter dans un changement phénotypique immédiat et spécifique. Des résultats récents, obtenus notamment chez le tabac et le peuplier, ont démontré la faisabilité d'une telle approche (Boudet *et al.*, 1996). Cependant, les difficultés liées à la transformation génétique des *Eucalyptus* d'importances économiques reste une limitation de sa faisabilité à court terme (Teulières *et al.*, 1994).

La seconde stratégie s'appuie sur l'hétérogénéité naturelle des teneurs et/ou des compositions des lignines entre individus d'une même espèce, qui permet de penser qu'une variabilité génétique en est à l'origine. Cette variabilité pourrait être exploitée dans le programme d'amélioration génétique de l'*Eucalyptus*. Cependant le coût des centaines voire des milliers d'analyses concernant la quantité ou la qualité des lignines, est un frein incontestable à l'utilisation d'un tel critère de sélection. L'alternative est de trouver un caractère associé, simple à mesurer et peu coûteux. Des marqueurs moléculaires, situés au voisinage des gènes majeurs ou QTL (quantitative trait loci, pour revue Tanksley, 1993) contrôlant une partie de la variation phénotypique de la qualité du bois, pourrait jouer ce rôle. Une sélection assistée par marqueurs moléculaires, permettrait non seulement d'envisager l'incorporation des critères de qualité dans ce programme mais aussi d'améliorer l'efficacité de la sélection par rapport à une sélection purement phénotypique (Lande et Thompson, 1990).

Afin de détecter une association marqueur-QTL au niveau de la population d'amélioration (Strauss *et al.*, 1992), une stratégie envisageable est d'utiliser des marqueurs moléculaires de fonction connue a priori directement impliqués dans le contrôle génétique de ces caractères. Les enzymes de structure de la chaîne de biosynthèse des lignines constituent de tels gènes candidats. Si on parvient à mettre en

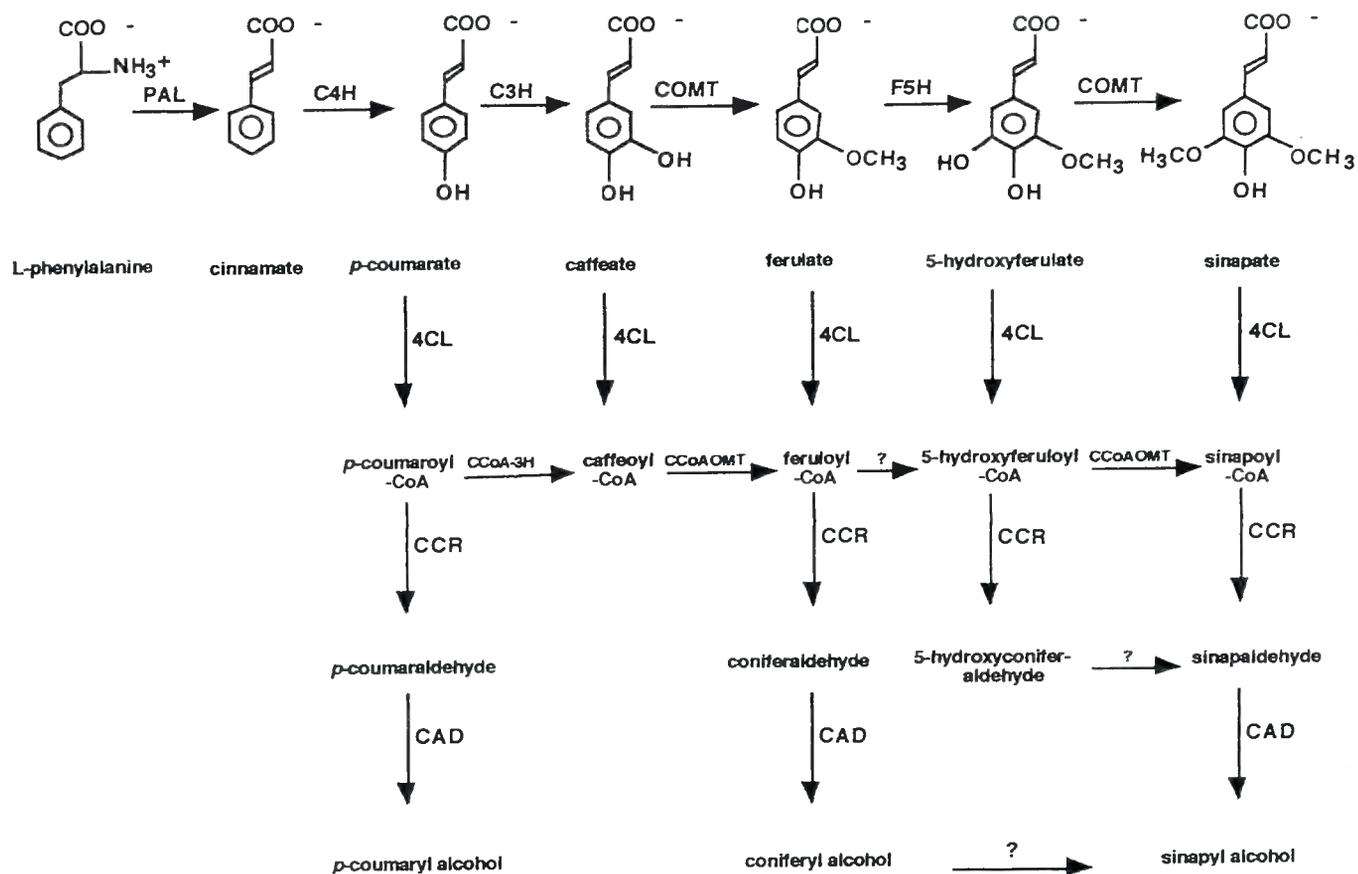


Fig 2. Voies de synthèse des lignines à partir de la phénylalanine.

Les enzymes de la voie commune des phénylpropanoïdes comprennent la phénylalanine ammonia lyase (PAL), la cinnamate-4-hydroxylase (C4H), la coumarate-3-hydroxylase (C3H), l'acide caféique 3-O-méthyltransférase (COMT), la férulate hydroxylase (F5H) et la 4-coumarate CoA ligase (4CL). Des voies alternatives d'hydroxylation et de méthylation ont été démontré au niveau des cinnamoyl CoA, via la coumaroyl coenzyme A 3-hydroxylase (CCoA-3H) et la cafeoyl coenzyme A O-méthyltransférase (CCoAOMT). Les deux étapes spécifiques de réduction des cinnamoyl CoAs en alcools cinnamyls (monolignols) sont respectivement catalysées par la cinnamoyl CoA réductase (CCR) et l'alcool cinnamyls déhydrogénase (CAD).

évidence une colocalisation entre ces gènes et des QTL liés à la qualité papetière du bois, ces gènes pourraient alors servir de marqueurs "diagnostics" dans un programme de sélection.

Mon sujet de DEA s'intègre dans cette approche complémentaire de l'ingénierie génétique des arbres. Dans ce contexte, l'objectif de mon étude est de cloner des gènes impliqués dans la voie de biosynthèse des lignines chez *Eucalyptus*, puis de les positionner sur des cartes génétiques établies chez *E.grandis* et *E.urophylla* (Verhaegen et Plomion, 1996)

La voie de biosynthèse des monomères des lignines peut être divisée en deux grandes parties : la voie commune des phénylpropanoïdes qui conduit aux esters de cinnamoyl-CoA, et une voie plus spécifique, impliquant la cinnamoyl CoA reductase (CCR) et la cinnamoyl alcool deshydrogenase (CAD) (Fig. 2) (pour revue, Boudet *et al.*, 1995). Les gènes codant pour ces enzymes (CAD2 et CCR) ont été clonés au laboratoire (Equipe expression et modulation de la lignification UMR 5546) chez *E.gunnii*, et positionnés sur les cartes génétiques d'*E.grandis* et d'*E.urophylla* (Gion, 1996. CIRAD-Forêt, Montpellier).

Pour ma part, je me suis intéressé à la dernière enzyme de la voie générale des phénylpropanoïdes, la 4 coumarate coA ligase (4CL)(EC 6.2.1.12), qui convertit les acides hydroxycinnamiques en esters de coA correspondants, ainsi qu'à deux enzymes intervenant dans le contrôle de la composition monomérique des lignines, l'acide caféique-orthodiphénol-O-méthyl transférase (COMT) (EC 2.1.1.68) et la caféoyl-coA O-méthyl transférase (CCoAOMT) (EC 2.1.1.104). La COMT qui catalyse la méthylation de l'acide caféique et de l'acide 5-hydroxyfêrulique, est étroitement associée à la lignification (Bugos *et al.*, 1991). Le(s) gène(s) codant pour cette enzyme a (ont) été caractérisé (s) dans un grand nombre d'espèces (pour une revue Boudet *et al.*, 1995), dont *Eucalyptus gunnii* (Poeydomenge *et al.*, 1994). La CCoAOMT, quant à elle, catalyse la méthylation du caféoyl coA et son implication dans le processus de lignification n'a été suggéré que plus récemment (Ye et Varner, 1995).

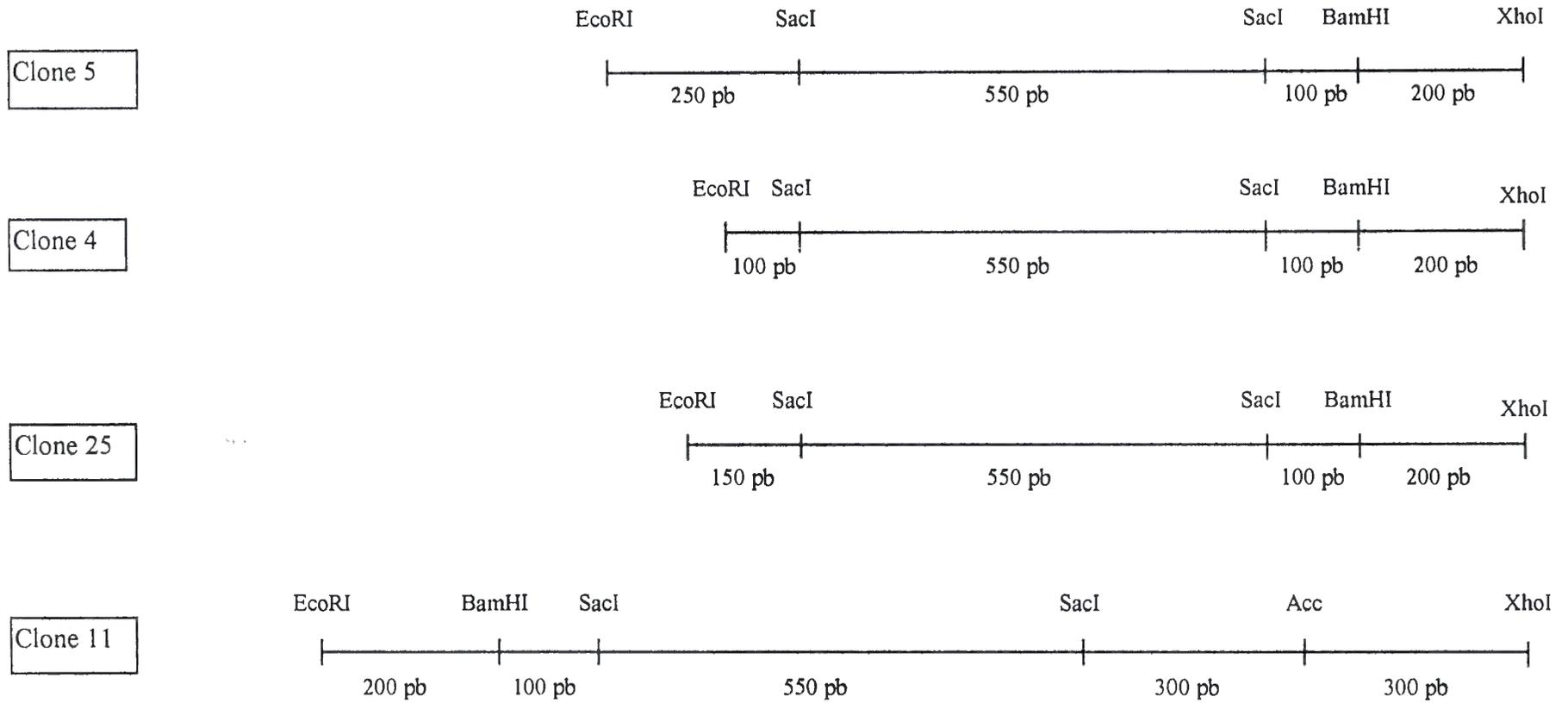


Fig 3. Carte de restriction des différents clones obtenus au dernier tour de criblage

Résultats

Afin de positionner les gènes de la 4CL et de la CCoAOMT sur les cartes génétiques d'*E.grandis* et d'*E.urophylla*, il est nécessaire au préalable d'obtenir des séquences "homologues" ou tout au moins très proches des séquences homologues. La comparaison de séquences nucléotidiques des exons du gènes CAD2 entre *E.gunnii* (Feuillet *et al.*, 1994) et *E.botryoïdes* (Hibino *et al.*, 1994). suggèrent que les régions exoniques sont extrêmement conservées dans le genre *Eucalyptus* (98% d'identité). Cette observation nous a conduit à envisager le clonage des ADNc codant pour ces enzymes à partir d'une banque d'ADNc de xylème en différenciation d'*E.gunnii*, disponible au laboratoire (UMR 5546).

Clonage

Clonage et caractérisation d'un ADNc codant de CCoAOMT d'*E.gunnii*

Une banque d'ADNc de xylème d'*E.gunnii*, construite dans λ ZAPII, a été criblée en utilisant comme sonde, un clone d'ADNc de *Nicotiana tabacum* codant pour une CCoAOMT (Martz, 1997). Après un criblage de 6×10^5 phages recombinants, 4 clones potentiels de CCoAOMT (nommés clone 4, clone 5, clone 11, clone 25) présentant respectivement des inserts de 950 pb, 1100 pb, 1450 pb, et 1000 pb, ont été purifiés. Afin d'éliminer d'éventuels faux positifs, les clones obtenus ont été analysés en Southern blot. La sonde hétérologue de tabac s'hybridant à forte stringence avec tous les clones, suggère ainsi l'existence d'un important degré d'homologies entre tous ces clones et la sonde.

Le degré de parenté entre ces 4 clones a été estimé, dans un premier temps, par l'établissement d'une carte de restriction exhaustive. Ainsi, les profils de restrictions, apparaissant similaires (le clone 11 ne se distinguant des autres clones que par l'apparition d'un site Acc), suggèrent l'existence d'un fort degré d'homologie entre tous ces clones (Fig. 3).

Toutefois, afin de ne pas éliminer précocément des clones alléliques ou de gènes différents ne présentant que de faibles variations de séquences nucléotidiques non détectées par digestion enzymatique, tous les clones ont été séquencés

partiellement au niveau de leurs extrémités sur 300 pb. L'alignement des séquences nucléotidiques, des clones 4, 5 et 25 n'a pas révélé de différence. De plus, les séquences déduites en acides aminés montrent de fortes homologues à plusieurs CCoAOMT disponibles dans les banques de données.

Le clone 11, présentant de fortes homologues avec plusieurs CCoAOMT à son extrémité 5' et avec plusieurs aquaporines végétales à son extrémité 3' (84% d'identité), s'est avéré être un ADNc chimérique. Aucune différence de séquence n'a été observée entre son extrémité 5' et les clones 4, 5 et 25.

Nous avons donc choisi de séquencer entièrement le clone 5, qui présente la taille la plus importante. Sa séquence de 1068 pb, enregistrée dans la banque de séquence nucléique EMBL sous le numéro d'accession y12228, contient une phase de lecture ouverte codant pour une protéine de 249 acides aminés dont le poids moléculaire estimé est de 27,9 kDa (Fig. 4).

L'alignement et la comparaison de cette séquence en acides aminés à différentes CCoAOMT existantes dans les bases de données, révèle 93,7% d'identité avec la séquence protéique de la CCoAOMT de *P.tremuloides*, et 88,8% d'identité avec la séquence protéique de *N.tabacum*, dont l'ADNc a servi de sonde de criblage (Martz, 1997) (Tab. 1). La production d'une protéine recombinante fonctionnelle apporterait la preuve définitive de l'identité du clone d'*Eucalyptus*. Néanmoins, les fortes homologues de séquences avec les CCoAOMT de *Petroselinum crispum* (86,3% d'identité) (Schmitt *et al.*, 1991), *Nicotiana tabacum* (88,8% d'identité) (Martz, 1997) et *Zinnia elegans* (85,3% d'identité) (Ye *et al.*, 1994), dont l'activité a été démontrée, suggèrent qu'il s'agit bien d'un clone d'ADNc codant pour une CCoAOMT d'*Eucalyptus*.

De plus, certains résidus en acides aminés impliqués dans la liaison de la S adénosyl méthionine (SAM) récemment identifiés chez la catéchol-OMT de rat (Vidgren *et al.*, 1994), sont conservés dans la séquence en acides aminés de la CCoAOMT d'*E.gunnii*. Notamment, les résidus situés entre la position 84 à 91 sont similaires à une région consensus impliquée dans la fixation de nucléotides (Ye *et al.*, 1994) (Fig. 4). Ces résidus conservés, se retrouvant notamment dans les CCoAOMT de *Nicotiana tabacum* (Martz, 1997) et *Zinnia elegans* (Ye *et al.*, 1994) constituent un élément supplémentaire de l'identité de ce clone.

Cette comparaison suggère l'existence de similarités structurales et d'éléments consensus entre les CCoAOMT et d'autres OMT SAM-dépendante.

% d'identité % similarité <i>E.gunnii</i>	<i>P.crispum</i>	<i>E.gunnii</i>	<i>P.tremuloïdes</i>	<i>V.vinifera</i>	<i>N.tabacum</i>	<i>M.sativa</i>	<i>Z.elegans</i>	<i>S.longipes</i>	<i>P.kitakamiensis</i>
	86.3 9.5								
<i>P.tremuloïdes</i>	89.2 8.3	93.7 4.5							
<i>V.vinifera</i>	88 9.5	91.3 5.4	94.2 4.1						
<i>N.tabacum</i>	86.3 11.2	88.8 7.4	90.5 7.9	91.3 6.6					
<i>M.sativa</i>	84.6 9.5	86.6 7.3	88.7 6.5	90.5 5.4	87.6 7.9				
<i>Z.elegans</i>	85.5 9.5	85.3 9.8	85.3 10.6	86.4 10.3	84.7 12.4	84.5 10.2			
<i>S.longipes</i>	53.5 20.3	56.4 18.7	55.2 18.3	54.8 19.9	53.9 21.2	55.6 19.1	52.3 21.2		
<i>P.kitakamiensis</i>	54.9 17.9	54.5 18.7	53.2 19.2	53.2 19.1	52.3 21.3	52.3 21.3	54 18.3	52.3 20.4	
<i>A.thaliana</i>	55.7 17.5	54.7 17.9	54.7 17.9	56.6 17.5	52.8 21.7	55.2 20.3	56.6 18.9	52.8 19.8	69.3 12.7

Tab 1. Homologie protéiques entre CCoAOMT de différentes familles.

Les comparaisons de séquences ont été réalisés deux à deux grâce à l'utilisation du programme Clustal. Les différentes espèces représentées sont :

Petroselinum crispum: z54183 , *Eucalyptus gunnii*; y12228 , *Populus tremuloïdes*; u27116 , *Vitis vinifera*; z54233 , *Nicotiana tabacum*; u62735 , *Medicago sativa*; u20736 , *Zinnia elegans*; u13151 , *Stellaria longipes*; l22203 , *Populus kitakamiensis*; ab000408 , *Arabidopsis thaliana*; l40031

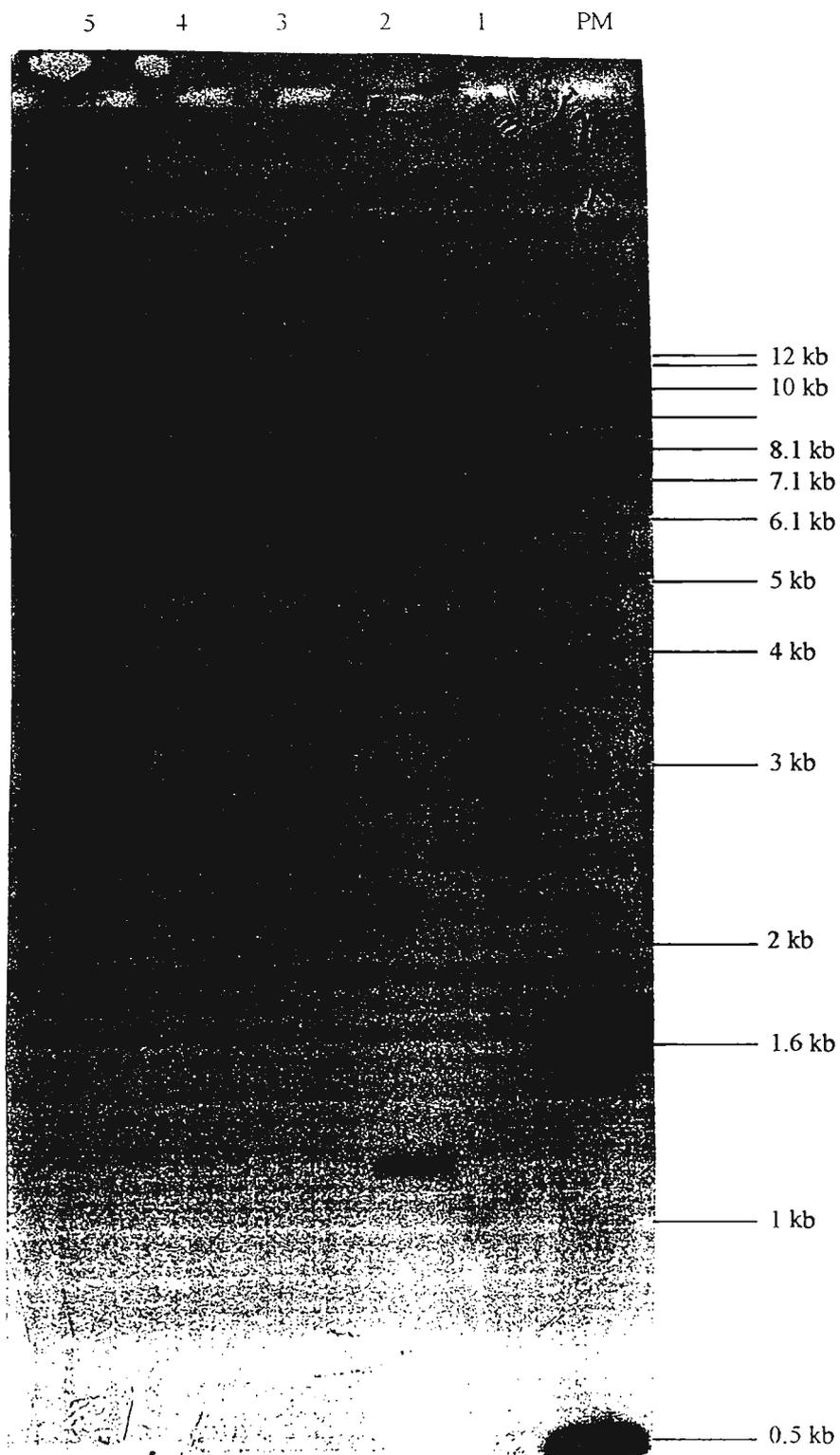


Fig 5. Analyse en Southern Blot, réalisé avec une sonde radiomarquée, synthétisée à partir de l'ADNc du clone 5, et de l'ADN génomique d'*E.gumii* digéré par différentes enzymes de restrictions :

piste 1 = PstI ; 2= HindIII ; 3= BamHI ; 4 = Xho ; 5 = XbaI ; PM = échelle de poids moléculaire (kb ladder)

O. sativa	SEDEASPLAGDEDEDKVFDPDDVVALPYSSGTTGLPKGVMLTHRSLSTSVAQQVDGENPNIGLHAGDVILCALPMFHIYSLNTIMMCGLRVGAIVVMRR	280
S. tuberosum	QSDEHE-----IPDVKIQDDVVALPYSSGTTGLPKGVMLTHKGLVTSVAQQVDGENANLYMHSDDVLMCVLPLFHIYSLNSVLLCALRVGAAILIMQK	263
N. tabacum	QANEHD-----IPEVEIQDDVVALPYSSGTTGLPKGVMLTHKGLVTSVAQQVDGENPNLYIHSDDVLMCVLPLFHIYSLNSVLLCALRVGAAILIMQK	260
P. crispum	EADESE-----MPEVVINSDDVVALPYSSGTTGLPKGVMLTHKGLVTSVAQQVDGDNPNLYMHSDDVLMCVLPLFHIYSLNAVLCCLRAGVITILIMQK	261
L. erythrorhizon	NSDETT-----LPDVEIRPDDVVALPYSSGTTGLPKGVMLTHKGLVTSVAQQVDGDNANLYMHEDVVMCTLPLFHIYSMNSILLCGLRVGAAILLMHK	263
A. thaliana	QSTTEAS---EVIDSVEISPDVVALPYSSGTTGLPKGVMLTHKGLVTSVAQQVDGENPNLYFHSDVILCVLPMFHIYALNSIMLCGLRVGAAILIMPK	281
P. taeda	EADETQC-----PAVKIHPDDVVALPYSSGTTGLPKGVMLTHKGLVTSVAQQVDGENPNLYIHSDDVLMCVLPLFHIYSLNSVLLCALRAGAATLIMQK	260
O. sativa	DADEGS-----VPEVAISPDDPVALPFSSGTTGLAKGVVLTTHRSVSVGAHEVDGENPNLHMGAGDVALCVLPLFHIYSLNSVLLSRVPAPAVLMPR	290
G. max	EANESD-----VPEVEIHPDDAVAMPFSSGTTGLPKGVILTHKSLTTSVAQQVDGENPNLYLTEDVLLCVLPLFHIYSLNSVLLCALRAGSAVLLMQK	139
** **.*.....**.*.....**.*.....**.*.....**.*.....**.*.....**.*.....**.*.....**.*.....**.*.....**.*.....	
	1 →	
O. sativa	FDLAAMMDLVERHRVTIAPLVPPIVVAVAKSEAAAARDLSSVRMVLSGAAPMGKDIEDAFMAKLPGAVLGGQGYGMTEAGPVLMSCLAFAKEPFKVKSGAC	380
S. tuberosum	FDIAQFLELIPKHKVTIGPFVPPIVLAIKSPVLDNYDSSVVRTVMGSAAPLKGKELEDVRAKFPNAKLGQGYGMTEAGPVLAMCLAFAKEPFDIKSGAC	363
N. tabacum	FDIVSFLELIQRYKVTIGPFVPPIVLAIKSPMVDYDLSSVVRTVMGSAAPLKGKELEDVRAKFPNAKLGQGYGMTEAGPVLAMCLAFAKEPFKVKSGAC	360
P. crispum	FDIVPFLELIQKYKVTIGPFVPPIVLAIKSPVVDKYDLSSVVRTVMGSAAPLKGKELEDVRAKFPNAKLGQGYGMTEAGPVLAMCLAFAKEPFYKIKSGAC	361
L. erythrorhizon	FEIVTFLELIQRYKVTIGPFVPPIVLAIKSNVVDQYDLSTVVRTVMGSAAPLKGSELEDVRAKFPNAKLGQGYGMTEAGPVLAMCLAFAKEPFKVKSGAC	363
A. thaliana	FEINLLELIQRCKVTVAPMVPPIVLAIAKSSSETEKYDLSSIRVVKSGAAPLKGKELEDVNAKFPNAKLGQGYGMTEAGPVLAMSLGFAKEPFKVKSGAC	381
P. taeda	FNLTCLELIQKYKVTIAPIVLVDITKSPIVSQYDVSSVRIMSGAAPLKGKELEDALRERFPKAI FGQGYGMTEAGPVLAMNLAFAKNPFPVKSGSC	360
O. sativa	FEMGAMLGAIERWRVTGAVVPLVVALAKNPFVERHDLSSIRIVLSGAAPLKGKELEDALRRLPQAI FGQGYGMTEAGPVLAMCPAFAKEPTPAKSGSC	390
G. max	FEIGTLLELIQRHRVSVAMVVPPLVVALAKNPMVADFDLSSIRIVLSGAAPLKGKELEALRNRMPQAVLGGQGYGMTEAGPVLAMCLGFAKQPFQTKSGSC	239
**.*.....**.*.....**.*.....**.*.....**.*.....**.*.....**.*.....**.*.....**.*.....**.*.....**.*.....	
	2 ←	
O. sativa	GTVVRNAELKIIDPDTGKSLGRNLRGECIRGQQIMKGYLNNPEATKNTIDAEGWLHTGDIGYVDDDDDEIFIVDRLKEIKYKGFQVAPAELEALLNTHP	480
S. tuberosum	GTVVRNAEMKIVDPDTGCSLPRNQPGECIRGDQIMKGYLNDPEATARTIEKEGWLHTGDIGFIDDDDELFI VDRLKELIKYKGFQVAPAELEALLINHP	463
N. tabacum	GTVVRNAEMKIVDPKTNLSLPRNQSGECIRGDQIMKGYLNDPEATARTIDKEGWLHTGDIGYIDDDDELFI VDRLKELIKYKGFQVAPAELEALLNHP	460
P. crispum	GTVVRNAEMKIVDPETNASLPRNQSGECIRGDQIMKGYLNDPESTRTTIDEEGWLHTGDIGFIDDDDELFI VDRLKELIKYKGFQVAPAELEALLTHP	461
L. erythrorhizon	GTVVRNSEMKIIDTETGASLPRNQSGECIRGDQIMKGYLNDPEATERTIDKEGWLHTGDIGYIDDDDELFI VDRLKELIKYKGFQVAPAELEALLVPHP	463
A. thaliana	GTVVRNAEMKIVDPDTGDSLRLNQPGECIRGHQIMKGYLNNPAATAETIDKDGWLHTGDIGLIDDDDELFI VDRLKELIKYKGFQVAPAELEALLIGHP	481
P. taeda	GTVVRNAQIKILDTEGSLPHNQAGECIRGPEIMKGYINDPESTAATIDEEGWLHTGDVGYIDDDDEIFIVDRVKEIKYKGFQVAPAELEALLVAHP	460
O. sativa	GTVVRNAELKVDPDTGFSLGRNLPGEICIRGPQIMKGYLNDPEATAATIDVEGWLHTGNIGYVDDDDDEVFIVDRVKEIKYKGFQVPPAELESLLIAHP	490
G. max	GTVVRNAELKVDPETGRSLGYNQPGECIRGQQIMKGYLNDPEATAASTIDSEGLHTGDVGYVDDDDDEIFIVDRVKEIKYKGFQVPPAELEGLLVSHP	339
	*****...*.*** **.*.....**.*.....**.*.....**.*.....**.*.....**.*.....**.*.....**.*.....**.*.....**.*.....**.*.....	
	GTVVRNAEMKIVDPDTGCSLPRNQSGECIRGQQIMKGYLNDPEATAATIDEEGWLHTGDVGYVDDDDDEIFIVDRVKEIKYKGFQVPPAELESLLIAHP	

Fig 6. Alignement multiple de séquences peptidiques partielles de 4Cl, de *Oriza sativa*: x52623 , *Solanum tuberosum*: M62755 , *Nicotiana tabacum*: u50846 , *Petroselinum crispum*: x13324 , *Lithospermum erythrorhizon*: d49366 , *Arabidopsis thaliana*: u18675 , *Pinus taeda*: u39405 , *Oryza sativa*: l43362, *Glycine max*: x69955. Cet alignement a été réalisé à l'aide du programme Clustal. Les symboles (*) et (.) montrent respectivement l'identité et l'homologie entre acides aminés des différentes 4CL. La position des amorces utilisées pour la PCR est noté par des flèches rouges. La séquence en acides aminés du fragment PCR «encadré» par les amorces PCR, 1 et 2 est indiqué en vert.

Les oligonucléotides dégénérés :

Oligonucléotide 1 sens : 5'GGNCARGGNTAYGGNATGAC3' (dégénérescence = 256)

Oligonucléotide 2 antisens : 5'CGDATRCADATRTCNC3' (dégénérescence = 144)

avec R=(A,G) ; Y=(C,T) ; D=(G,A,T) ; N = (A,T,G,C)

Cependant, la CCoAOMT d'*E.gunnii* ne montre que 23% d'identité (25,3 % de similarité) avec la COMT, SAM dépendante, de cette même espèce (Poeydemenge *et al.*, 1994). Les régions conservées ne montrant pas de similarité significative avec la région de fixation de la SAM de la catéchol de rat, suggère une évolution différente de la COMT pour l'utilisation de la SAM (Ye *et al.*, 1994).

Organisation génomique de la CCoAOMT d'*E.gunnii*

Afin d'estimer le nombre de gènes codant pour la CCoAOMT dans le génome d'*E.gunnii*, des extraits d'ADN génomique ont été digérés par des enzymes de restriction ne clivant pas (PstI, HindIII, Xho, XbaI), ou clivant une seule fois (BamHI) l'ADNc du clone 5. Le profil d'hybridation montre la présence d'un faible nombre de bandes après digestion par PstI, HindIII, BamHI et d'une seule bande après digestion par Xho et XbaI, suggérant l'existence d'un seul gène (Fig. 5).

Clonage d'un fragment du gène de la 4CL

Ne possédant pas de sonde hétérologue pour cribler la banque d'ADNc d'*E.gunnii*, nous avons décidé de réaliser le clonage d'un fragment partiel d'ADN de 4CL, par une approche PCR (polymerase chain reaction). L'objectif étant d'obtenir une sonde spécifique et homologue, afin de cribler la banque et d'isoler l'ADNc pleine longueur.

La comparaison de séquences protéiques de 4CL de plusieurs espèces, à permis de localiser sur les protéines des régions conservées qui ont servi à définir un couple d'amorces dégénérées (nommées 1 et 2), prenant en compte un degré de dégénérescence minimum et un pourcentage de GC élevé (Fig. 6).

Ce couple d'amorces a été testé par la suite sur un extrait d'ADN génomique d'*E.gunnii*. Après avoir déterminé les conditions optimales d'amplification, un seul fragment d'ADN de 200 pb, a été amplifié (Fig. 7). Le clonage et le séquençage de ce clone, a permis de montrer une forte homologie de séquence protéique avec la 4CL de *Nicotiana tabacum* (92%d'identité) (u50846) (Fig. 6). En outre, ce fragment d'ADN ne contient aucune séquence intronique. Possédant maintenant une sonde homologue, le criblage de la banque d'ADNc peut être réalisé.



Fig 7. Migration électrophorétique des produits PCR obtenus, avec le couple d'amorces dégénérées 1 et 2, sur de l'ADN génomique d'*E.gunnii*.
les pistes 1, 2 et 3 : PCR réalisée avec 1 μ M d'amorces
les pistes 4, 5 et 6 : PCR réalisées avec 0,5 μ M d'amorces
les pistes 7, 8 et 9 : PCR réalisées avec 0,25 μ M d'amorces

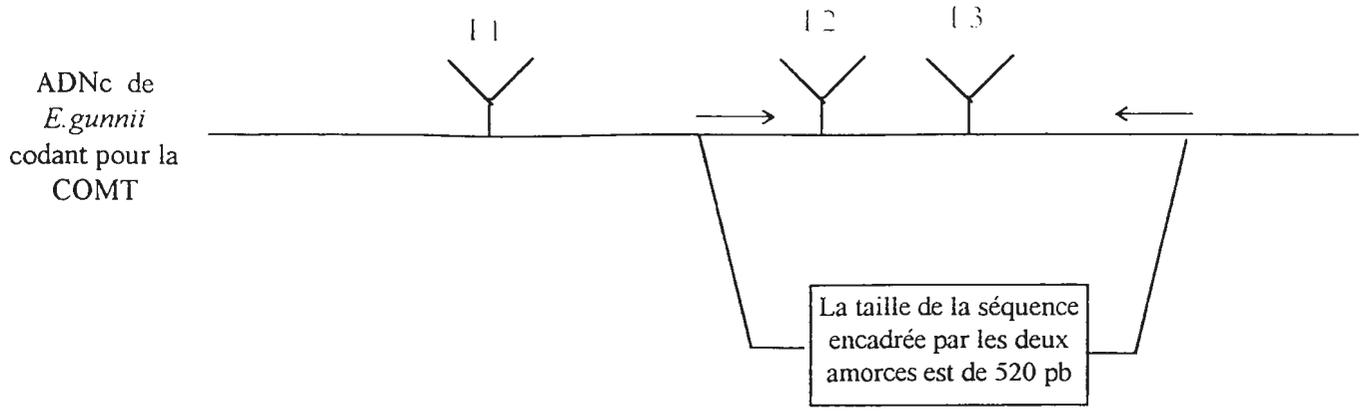


Fig 9. Position schématique des amorces PCR (flèches bleues) sur l'ADNc codant pour la COMT d'*E. gunnii* (x74814). La position des introns I1, I2 et I3 est conservée entre les gènes de *Populus tremuloïdes* (u13171) et *Populus kitakamiensis* (D49710).

Amorce sens : 5'CGCTCCACCCCTTCCT3'
 Amorce antisens : 5'GGCTCCACGACCCTTTC3' ← *faux*.

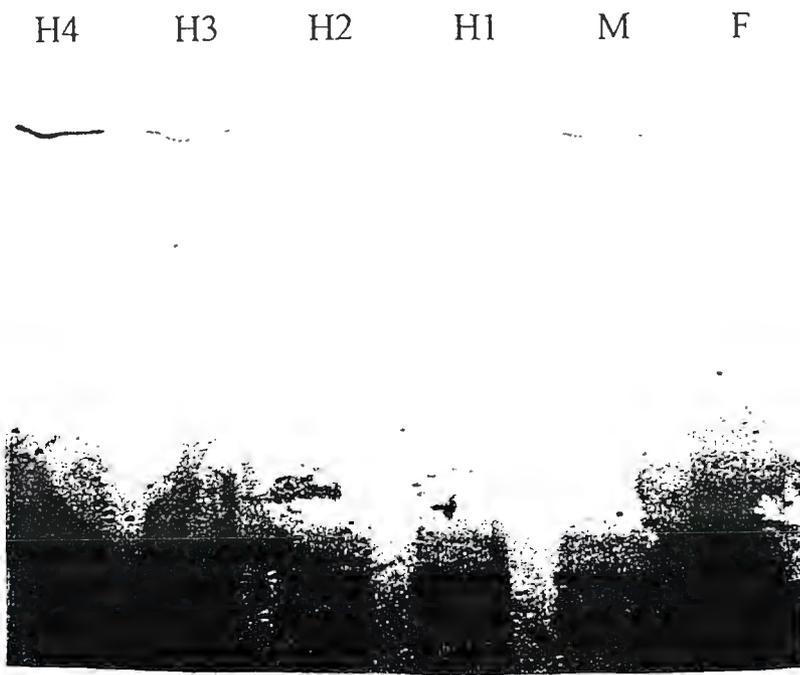


Fig 10. Profils SSCP obtenus pour la CCoAOMT pour les parents (géniteur : profil F ; géniteur : profil M) et 4 hybrides (H1, H2, H3, H4)

	Parents		4 combinaisons Hybrides			
	♂	♀	H1	H2	H3	H4
<u>profils SSCP :</u>	1 ———		————	————		
		4 ———	————		=====	=====
	2 ———	5 =====		=====	=====	=====
		6 =====	=====	=====	=====	=====
	3 ———	7 ———	=====	=====	=====	=====
<u>Génotypes :</u>	(1,3) (2,3)	(4,7) (5,6)	(1,3) (4,7)	(1,3) (5,6)	(2,3) (4,7)	(2,3) (5,6)

Fig 11. Représentation schématique des profils SSCP obtenus pour la CCoAOMT, pour les deux parents et leurs descendants hybrides.

Positionnement sur les cartes génétiques des gènes codant pour la CCoAOMT et pour la COMT

A partir des deux séquences nucléotidiques des ADNc clonés précédemment et codant pour la CCoAOMT (y12228) et pour la COMT (x74814) (Poeydomenge *et al.*, 1994), nous avons pu envisager de les localiser sur les cartes génétiques existantes. La technique choisie (PCR-SSCP, Single Strand Conformation Polymorphism, pour revue Hayashi *et al.*, 1993) doit permettre de détecter un polymorphisme de séquence entre allèles présent dans un organisme diploïde, comme l'*Eucalyptus*. Le produit d'amplification d'un fragment d'un gène, obtenu par PCR à l'aide d'amorces spécifiques, est dénaturé, puis soumis à une électrophorèse en conditions non dénaturante. Dans ces conditions les molécules d'ADN simple brin prennent une structure secondaire spécifique. Le produit PCR correspondant à un allèle est séparé en deux bandes distinctes correspondant à chacun des deux brins complémentaires. Cette technique initialement utilisée sur le génome humain (Orita *et al.*, 1989), est particulièrement sensible et facile d'utilisation, et permet de mettre en évidence des polymorphismes dus à des modifications mineures de la séquence d'ADN, comme des mutations ponctuelles. En effet, la substitution d'une ou de plusieurs bases au niveau de la molécule d'ADN se traduit par des conformations particulières qui migreront différemment sur un gel polyacrylamide non dénaturant (Fig. 8).

Choix des amorces PCR

De manière à faciliter la détection d'un éventuel polymorphisme allélique, des couples d'amorces exoniques sont choisis pour chacun des ADNc de façon à encadrer des régions introniques. En effet, les séquences des régions introniques sont généralement moins conservées que celles des régions codantes.

La comparaison de plusieurs gènes de COMT et de CCoAOMT, clonés à partir de différentes espèces végétales, a permis d'observer une bonne conservation du nombre et de la position des régions introniques au sein des angiospermes dicotylédones. La position et la séquence des amorces choisies sont représentées sur la figure 4 et 9.

Eucalyptus est un organisme diploïde :
à chaque locus il existe 2 allèles

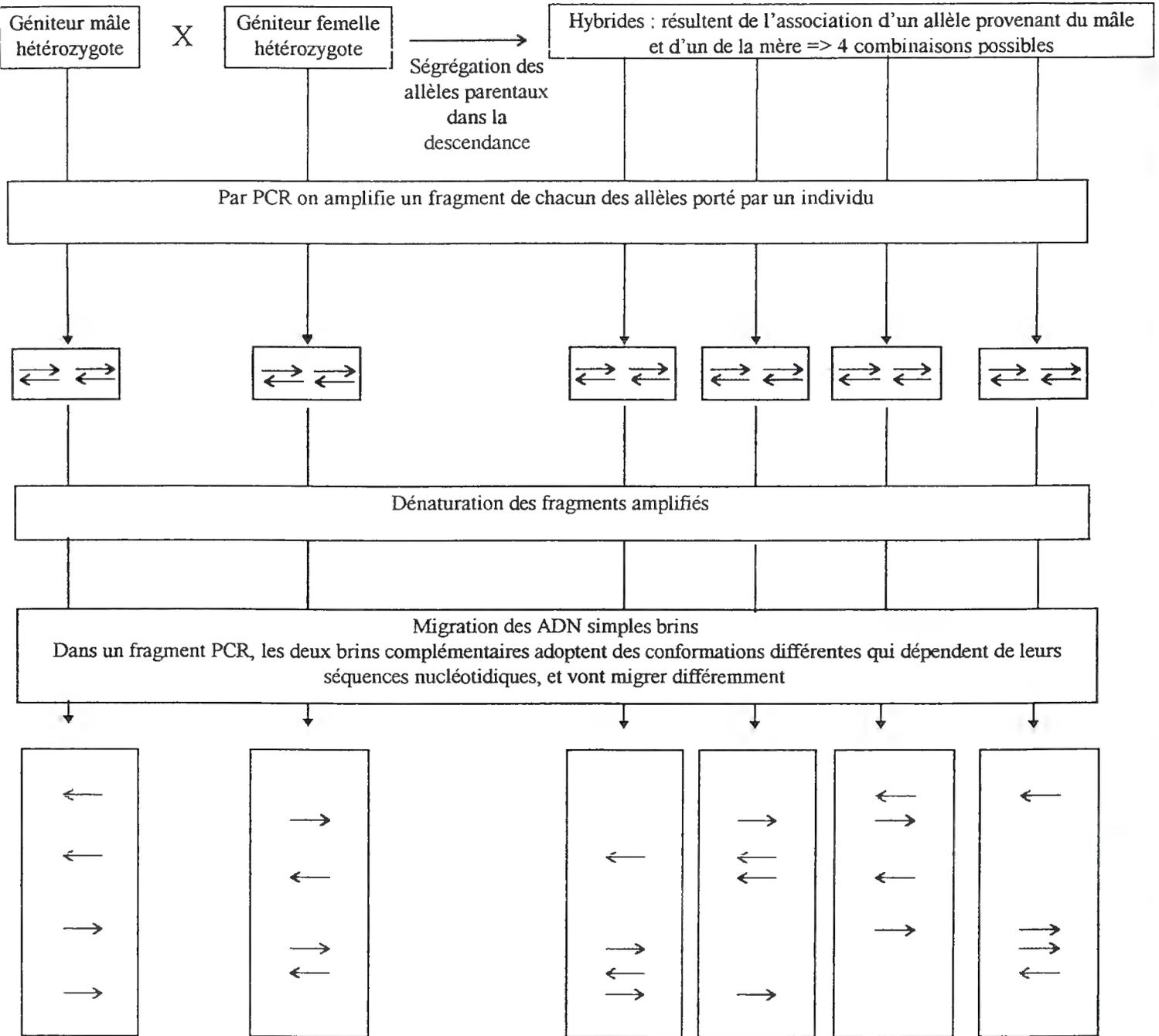


Fig 8. Schématisation de la ségrégation des allèles portés par un locus dans un organisme diploïde, et exemple de profil SSCP, obtenu après amplification par PCR d'un fragment de chaque allèle, présent dans les parents et les descendants.

Le clone partiel de la 4CL, ne possédant pas d'introns, n'a pas pu être utilisé dans cette seconde partie de mon stage

Optimisation des conditions d'amplifications par PCR

Ces couples d'amorces ont été utilisés pour des réactions d'amplification par PCR à partir d'ADN génomique extrait des géniteurs (*E.urophylla* clone 14.144 (femelle) et *E.grandis* clone 9.21(mâle)) et de leur descendance (constituée de 92 individus). Après avoir déterminé les conditions optimales d'amplification, un seul fragment d'ADN par couple d'amorces, a été amplifié, démontrant ainsi la spécificité de ces couples d'amorces et la conservation des régions codantes entre espèces d'*Eucalyptus*.

La taille de ces fragments est de 500 pb et de 800 pb pour les couples d'amorces définis respectivement à partir de l'ADNc codant pour la CCoAOMT et pour la COMT. Ces tailles sont supérieures à celles des séquences des ADNc bordées par les couples d'amorces (respectivement supérieures de 250 pb et de 450 pb), suggérant qu'une ou plusieurs régions introniques sont également amplifiées.

Optimisation des conditions de migrations en gel SSCP

Dans un premier temps, nous avons optimisé la détection de ces polymorphismes et amélioré la résolution des profils SSCP en utilisant différentes concentrations d'acrylamide, de ratio acrylamide/bis-acrylamide, de glycérol, d'intensités de voltage ainsi que de température de migration. Comme l'ont mentionné plusieurs auteurs (Sheffield *et al.*, 1993, Hayashi *et al.*, 1993), ces facteurs peuvent influencer la séparation des différentes bandes. Nous avons, en effet, constaté que les modifications des conditions SSCP ont souvent changé les profils électrophorétiques. Ainsi, nous avons obtenu, dans les conditions électrophorétiques optimales 1, 3 ou 4 bandes dans chacun des échantillons (Fig. 10 et 12). L'analyse de ces fragments a révélé l'existence d'un polymorphisme allélique présent chez les parents et ségrégeant dans la descendance.

Héritabilité mendélienne des fragments SSCP

Afin de positionner sur les cartes génétiques les gènes codant pour la CCoAOMT et la COMT, nous avons analysé dans un deuxième temps leur

M F H4 H1 H2 H2 H2 H3 H1



Fig 12. Profils SSCP obtenus pour la COMT pour les parents (géniteur femelle : profil F; géniteur mâle : profil M) et 7 hybrides (H1, H2, H3, H4)

		Parents		4 combinaisons Hybrides			
		♂	♀	H1	H2	H3	H4
<u>profils SSCP :</u>	1	████████		████████	████████	████████	████████
	2		████████	████████	████████		████████
	3		████████		████████		████████
	4		████████	████████			████████
<u>Génotypes :</u>	(1,1)	Gène A { (2,3) (0,0)		(1,1) (2,4) (0',0')	(1,1) (2,3) (0,0)	(1,1) (0,0) (0',0')	(1,1) (2,3) (2,4)
		Gène B { (2,4) (0',0')					

Fig 13. Représentation schématique des profils SSCP obtenus pour la COMT, pour les deux parents et leurs descendants hybrides.

ségrégation dans les descendants issus du croisement *E.urophylla* (clone 14.144) x *E.grandis* (clone 9.21).

* Cas de la CCoAOMT

On observe en SSCP, 3 bandes différentes (nommées 1, 2, 3) présentes dans le parent mâle et 4 (nommées 4, 5, 6, 7) dans le parent femelle (Fig. 11). Tous les descendants possèdent 4 bandes, mais dont les combinaisons sont différentes de celles des géniteurs (Fig. 11). Ainsi, nous obtenons pour les 92 individus de la descendance F1 analysée, 27 montrant le profil H1 [1, 3, 4, et 7]; 18 le profil H2 [1, 3, 5, et 6]; 25 le profil H3 [2, 3, 4, et 7] et 22 le profil H4 [2, 3, 5, et 6]. Ainsi, il apparaît que les bandes [1 et 3], [2 et 3], [4 et 7], [5 et 6] sont toujours associées dans les descendants (Tab. 2).

Sachant d'une part que chaque parent diploïde transmet un de ces deux allèles à chacun des descendants et d'autre part que les deux brins d'ADN correspondant à un allèle donné apparaissent en gel SSCP comme deux bandes distinctes (Sheffield *et al.* 1993), nous en concluons que les parents sont hétérozygotes au locus de la CCoAOMT. D'ailleurs, les proportions des différents profils observés ne sont pas significativement différents d'une ségrégation mendélienne de type $\frac{1}{4} : \frac{1}{4} : \frac{1}{4} : \frac{1}{4}$ (Tab. 2).

Ainsi, le géniteur femelle possède un allèle correspondant aux bandes 4, 7 et un allèle aux bandes 5, 6, et le géniteur mâle, un allèle correspondant aux bandes 1, 3 et un allèle aux bandes 2, 3. Dans ce dernier géniteur, la bande 3 serait constitué de deux ADN simples brins migrant au même niveau. En effet, la sensibilité de détection du polymorphisme est souvent plus grande pour un brin comparé au brin complémentaire (Michaud *et al.*, 1992). Ainsi, les génotypes hétérozygotes peuvent montrer 3 ou 4 bandes.

* Cas de la COMT

Nous constatons la présence d'1 seule bande (nommée 1) dans le parent mâle et de 3 bandes (nommées 2, 3 et 4) dans le parent femelle. Bien que l'unique bande observé dans le parent mâle ait pu être séparée en deux bandes sous différentes conditions de migration, nous avons été incapables de révéler un polymorphisme dans la descendance. En raison de ce monomorphisme, nous avons choisi des conditions de

Tab 2. Test de ségrégation des profils observés en SSCP pour la CCoAOMT et la COMT

fragment	Génotypes des parents		Génotypes des hybrides				Effectif observé pour chaque génotype	Test χ^2 (a)	ddl
	mère	/ père							
CCoAOMT	4,7 : 5,6	/ 1,3 : 2,3	4,7; 1,3 / 5,6 : 1,3 / 4,7 : 2,3 / 5,6 : 2,3				27 / 18 / 25 / 22	2	3
COMT	gène A : 2,3 : 0,0(b)		2,3	0,0	2,3	0,0			
	gène B : 2,4 : 0',0'(b)		/ 1,1	: 1,1 /	: 1,1 /	: 1,1 /	: 1,1	46 / 42 / 2 / 2	0,18

(a) valeur critique du χ^2 avec $p=0,01$ (3 dl: 11,3)
(1 dl: 6,63)

(b) 0,0 et 0',0' indiquent les allèles nuls

migration en fonction de la résolution de lecture des bandes maternelles ségrégeant dans la descendance. Ainsi, parmi les 92 descendants, nous avons observé 46 individus présentant le profil H1 (1, 2 et 3), 42 le profil H2 (1, 2 et 4), 2 le profil H3 (1, 2, 3 et 4) et 2 le profil H4 (1) (Fig. 13).

Si l'on considère uniquement le parent femelle, on peut supposer, comme précédemment, l'existence d'un locus hétérozygote. Toutefois, l'obtention d'individus présentant soit aucune soit toutes les bandes provenant de la mère, nous fait penser à l'existence de deux gènes liés. Il faut dans ce cas admettre que pour chaque locus, un seul des allèles soit amplifié. En effet, l'absence du produit d'amplification de ces deux allèles pourrait rendre compte de la ségrégation observée. Les individus présentant aucune ou toutes les bandes maternelles seraient alors les produits d'évènements de recombinaison.

Ainsi, nous pouvons décomposer le génotype de la mère en; (gène A constitué de l'allèle 2,3 et de l'allèle nul 0,0 et le gène B constitué de l'allèle 2,4 et de l'allèle nul 0',0'). Selon cette hypothèse, la ségrégation des génotypes sur chacun des loci, si l'on ne tient pas compte des recombinants, n'est pas significativement différente de celle attendue sous une loi mendélienne de type $\frac{1}{2} : \frac{1}{2}$ (Tab. 2). Les allèles 2,3 du gène A et 2,4 du gène B sont liés en trans, c'est à dire situées sur des chromosomes homologues (d'où la faible probabilité d'obtenir des recombinant de type H3 et H4)

Localisation des gènes de la CCoAOMT et de la COMT sur les cartes génétiques

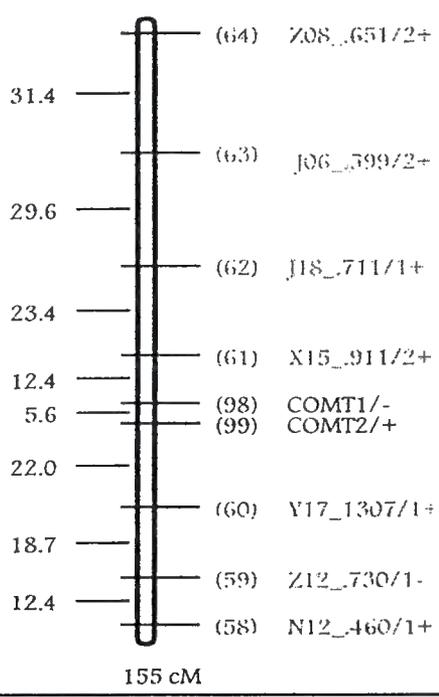
Le logiciel MapMaker a été utilisé pour positionner ces gènes sur les deux cartes génétiques parentales (Fig. 14) établies à partir de marqueurs RAPD dominants ségrégeant dans les proportions mendéliennes $\frac{1}{2} : \frac{1}{2}$ (Grattapaglia et al., 1994; Verhaegen et Plomion, 1996). Le principe de la stratégie de cartographie adoptée pour construire la carte génétique d'un parent, à partir d'une descendance pleins frères, était le suivant :

Si ce parent est hétérozygote pour un marqueur donné (présence de bande RAPD) alors que l'autre parent est homozygote nul pour ce même marqueur (absence de bande RAPD), la descendance hybride ségrégera dans les proportions $\frac{1}{2} : \frac{1}{2}$ pour un tel marqueur (la moitié des individus présentant l'allèle "bande présente", l'autre présentant l'allèle "bande absente"). L'analyse de la ségrégation de tous les marqueurs

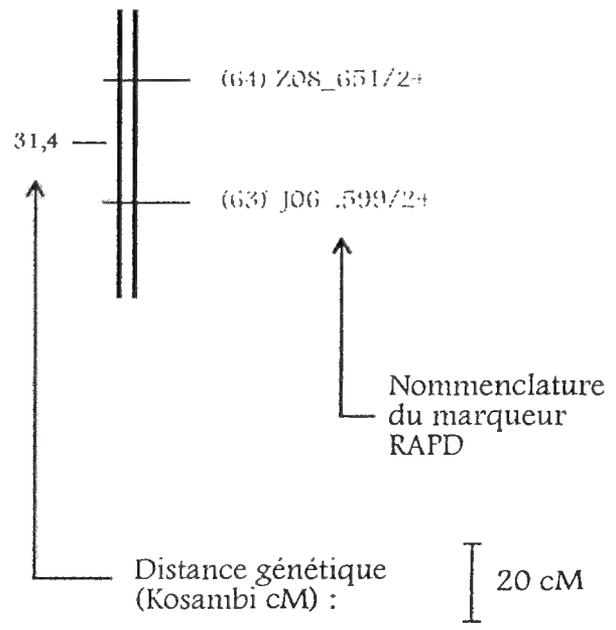
Fig 14. Localisation des gènes CCoAOMT et COMT sur les cartes génétiques d'*E.urophylla* et d'*E.grandis*

Groupe 7

E.urophylla

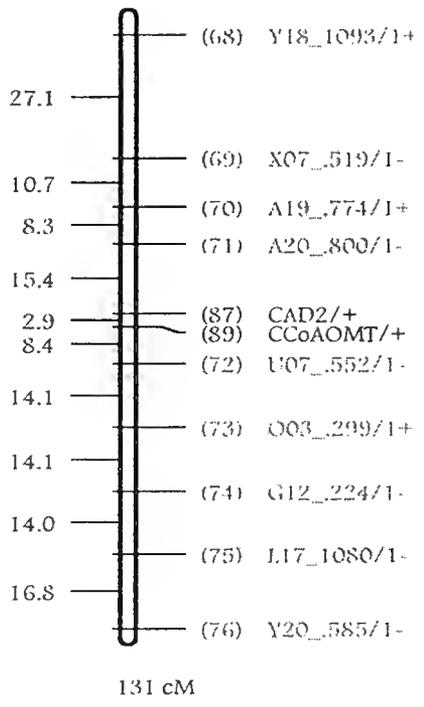


Légende

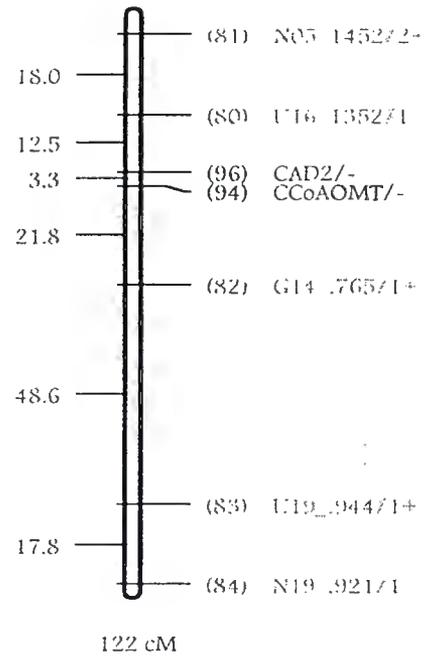


Groupe 10

E.grandis



E.urophylla



uniquement présents dans ce parent permet la construction de sa carte. (Le principe reste le même pour construire la carte du deuxième parent).

Pour positionner les gènes de la CCoAOMT et de la COMT sur les cartes génétiques parentales, nous avons dans un premier temps défini la ségrégation de nos marqueurs de façon dominante.

* CCoAOMT

Dans ce but, nous avons décomposé la ségrégation du gène de la CCoAOMT, de type ($\frac{1}{4} : \frac{1}{4} : \frac{1}{4} : \frac{1}{4}$) en deux ségrégations $\frac{1}{2} : \frac{1}{2}$, pour le père (1) et pour la mère (2).

(1) Après avoir choisi un des deux allèles paternel, les 2 génotypes des descendants, sont représentés par sa présence et par son absence.

(2) Après avoir choisi, un des deux allèles maternel, les 2 génotypes des descendants, sont représentés par sa présence et par son absence.

Ainsi, la CCoAOMT a été localisée, sur la carte génétique de la mère et sur celle du père, sur le groupe de liaison 10, à proximité (3 cM) du gène de la CAD2 (Gion, 1996) (Fig. 14).

* COMT

Puisqu'un seul allèle est amplifié pour chaque locus A et B, la ségrégation observée dans la descendance de ces gènes est similaire à une ségrégation de deux marqueurs RAPD. Ainsi, nous avons traité ces deux gènes comme deux marqueurs dominants.

Le monomorphisme détecté pour la COMT chez le géniteur mâle, ne permet pas de localiser sa position sur la carte génétique de ce géniteur. Toutefois, les deux gènes liés (5,6 cM) ont été localisés sur le géniteur femelle, sur le groupe de liaison 7 (Fig. 14).

Discussion

Le travail, réalisé lors de mon stage de DEA, a permis, dans un premier temps, de cloner et de caractériser un ADNc codant pour la CCoAOMT, ainsi qu'un fragment génomique codant pour la 4CL d'*E.gummii*. Dans un second temps, les gènes codant pour la COMT et pour la CCoAOMT ont été positionnés sur les cartes génétiques d'*E.urophylla*. et d'*E.grandis* (Verhaegen et Plomion, 1996).

Récemment, différentes techniques de détection efficace de variation génétique ont été mises au point (pour revue Hayashi *et al.*, 1993). Cependant, l'analyse de polymorphisme de conformation d'ADN simple brin (SSCP) reste, de part sa rapidité et sa simplicité, la technique la plus couramment utilisée. Ainsi des molécules d'ADN simple brin différant par une simple substitution nucléotidique adoptent des conformations différentes et peuvent être distinguées après migration électrophorétique en conditions non dénaturantes (Orita *et al.*, 1989). Toutefois, cette structure secondaire est hautement dépendante des conditions expérimentales, telles que la concentration ionique et la température (Scheffield *et al.*, 1993). Afin de mettre en évidence un polymorphisme allélique pour les gènes de la COMT et de la CCoAOMT, il a donc été nécessaire dans un premier d'optimiser les conditions de migration des fragments amplifiés par PCR.

Dans le cas de la CCoAOMT, nous avons pu mettre en évidence un polymorphisme allélique chez chacun des deux parents, permettant, après analyse de sa ségrégation dans la descendance, la localisation de ce gène sur les cartes génétiques d'*E.urophylla* et d'*E.grandis*. Ce gène est positionné sur le groupe de liaison 10, à proximité (3 cM) d'un autre gène impliqué dans la voie de biosynthèse des lignines, le gène de la CAD2 (Gion, 1996).

Les résultats de cartographie suggèrent donc la présence d'un seul gène, ce qui conforte l'analyse réalisée en Southern blot. Toutefois, on ne peut pas exclure l'existence d'autres gènes dont les séquences seraient suffisamment divergentes pour ne pas être révélées dans les conditions d'hybridation stringentes que nous avons utilisées. Aussi, une analyse en Southern blot réalisée à des conditions moins stringentes pourrait apporter un élément de réponse. En effet, des études, réalisées sur la CCoAOMT bien que peu nombreuses, suggèrent l'existence d'une famille

multigénique chez *Nicotiana tabacum* (Martz, 1997) et chez *Zinnia elegans* (Ye *et al.*, 1994).

Concernant la COMT, la situation apparaît plus complexe. Nous avons été incapable de révéler un polymorphisme allélique chez *E.grandis*. Pour expliquer ce monomorphisme, deux hypothèses peuvent être avancées, soit le parent mâle est homozygote pour ce gène, soit les amorces choisies pour la PCR-SSCP ne permettent pas de détecter une variation de séquence nucléotidique entre allèle dans les conditions utilisées. Pour tenter de trancher entre ces deux hypothèses, on peut envisager d'utiliser d'autres couples d'amorces positionnées différemment.

Chez le géniteur femelle (*E.urophylla*) en revanche, l'analyse de la ségrégation du polymorphisme allélique dans la descendance suggère l'existence de deux gènes liés, et de plus, pour chaque locus un seul des allèles est amplifié. Ces allèles nuls résultent probablement d'une absence d'amplification PCR. Il est possible qu'aux fortes conditions de stringences utilisées, des variations de séquence sur le site de fixation d'une amorce empêchent sa fixation, et par conséquent l'amplification du fragment de l'allèle concerné. Pour vérifier cette hypothèse, on peut envisager d'utiliser d'autres couples d'amorces pour les réactions PCR-SSCP et pour confirmer l'existence de deux gènes réaliser une analyse en Southern blot, en utilisant comme sonde d'hybridation l'ADNc codant pour la COMT d'*E.gunnii*. La présence de deux gènes codant pour la COMT chez *E.urophylla* semble en accord avec le clonage récent de deux gènes chez *Populus kitakamiensis* (Hayakawa *et al.*, 1995) et l'existence d'une famille multigénique chez *Nicotiana tabacum* (Pellegrini *et al.*, 1993).

En résumé, les gènes de la COMT ont pu être positionnés seulement sur la carte génétique d'*E.urophylla*. Ils se localisent sur le groupe de liaison 7 et sont distants de 5,6 cM. Cette localisation est à la fois différente de celle du gène codant pour la CCR (groupe 6, Gion, 1996) et de celles des gènes de la CCoAOMT et de la CAD2 (groupe 10). Ainsi, cette disparité de localisation des gènes de la voie de biosynthèse des lignines, suggère l'absence d'un "cluster de gènes de lignification".

Ce travail constitue une première étape d'un programme plus large visant à étudier la colocalisation éventuelle des gènes impliqués dans la biosynthèse des lignines avec des QTL liés à la teneur et/ou à la qualité de ces biopolymères. A coté des gènes de structure de la lignification, d'autres gènes candidats, contrôlant la

quantité ou la qualité de la lignine, sont à rechercher. Ainsi, des facteurs de transcription spécifiquement exprimés dans le xylème en différenciation constitue de tels gènes candidats (Martin, 1996).

Cette approche intégrée constitue un axe de recherche important dans le cadre de la sélection assistée par marqueurs de ces caractères coûteux à mesurer dans un programme d'amélioration génétique, et constitue une alternative à la transformation génétique. Cependant ces approches sont complémentaires: en effet, des gènes candidats colocalisés avec des QTL (teneur et/ou qualité des lignines), pourraient être utilisés dans le cadre de la transformation génétique. Inversement, la modification des lignines obtenue par transformation génétique d'un ou plusieurs gènes devra être considéré par les généticiens dans leur approche sélection assistée par marqueurs.

Un exemple qui illustre cette complémentarité est la découverte d'un allèle non exprimé d'un gène codant pour la CAD (allèle CAD "nul") chez *Pinus taeda* (MacKay *et al.*, 1995). Ces variants génétiques sont par ailleurs intensivement utilisés comme géniteurs élités dans les programmes de croisements conventionnels et cela depuis une trentaine d'années. Les homozygotes pour cet allèle "nul" présentent des lignines dont la composition monomérique est altérée (présence d'aldéhydes) et qui sont plus facilement dégradées lors des processus de délignification. Ces propriétés sont similaires à celles retrouvées chez des plantes transgéniques chez lesquelles l'activité CAD est très fortement diminuée (Halpin *et al.*, 1994).

Ces études démontrent ainsi que l'action de gènes individuels peut avoir une influence critique sur la composition et les propriétés des parois cellulaires et soulignent la contribution de ces gènes dans la variation génétique des propriétés du bois manipulées au travers de programmes de croisements conventionnels.

Matériels et Méthodes

Clonage

Matériel.

* Banque d'ADNc

La banque utilisée pour les différents criblages a été construite dans le bactériophage λ ZAPII, à partir d'ARNm extraits de xylème en cours de différenciation d'*Eucalyptus gunnii* (génotype 867). La construction de cette banque d'ADNc, cloné dans un site de clonage EcoRI, XhoI, a été réalisée par Clontech (USA) (Poeydomenge *et al.*, 1994).

* Souches bactériennes utilisées

E.coli souche XL1-Blue (*supE44 hsdR17 recA1 endA1 gyrA46 thi relA1 lac*⁻, *F'[proAB⁺ lacI^f lacZ Δ M15 Tn10(*tet*)]*)

E.coli souche DH5 α (*supE44 Δ lacU169 (Φ 80 lacZ Δ M15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1*)

Techniques d'usage courant en biologie moléculaire

* Digestion par des enzymes de restriction et analyses.

2 à 5 μ g d'ADN sont digérés par les enzymes de restriction adéquates (1-2U/ μ g d'ADN) dans le tampon fourni avec l'enzyme (Eurogentec), pendant une durée minimum de 2 heures à 37°C. Les produits de digestions sont déposés sur gel d'agarose (0,8% à 1%) après ajout du tampon de charge 1x (tampon 6x : 0,25% bleu de bromophénol, 0,25% xylène cyanol, 15% ficoll). Après migration électrophorétique à voltage constant (50V au minimum), l'ADN est coloré 15' dans une solution de BET (Bromure d'éthidium, 0,5 μ g/ml) puis visualisé sous lumière UV.

* Synthèse de sondes radioactives

Le marquage des fragments d'ADN est réalisé par élongation d'amorces aléatoires à l'aide du kit de "random priming" d'Appligene, en suivant les instructions du fabricant. Les sondes sont séparées des nucléotides non incorporés par

chromatographie d'exclusion sur une colonne Sephadex G50 (Sambrook *et al.*, 1989). La radioactivité, incorporée dans la sonde, est mesurée par effet Cerenkoff.

*** Séquençage et analyse des séquences**

Les réactions de séquençage ont été réalisées par Eurogentec par une méthode de terminaison de chaîne par des dideoxynucléotides (Sanger *et al.*, 1977). Les recherches de similarités entre séquences protéiques ont été traitées par un programme BLAST (Altschul *et al.*, 1990), et les alignements multiples ont été réalisés par un programme CLUSTAL.

Criblage d'une banque d'ADNc de xylème et isolement d'un ADNc codant pour la CCoAOMT

*** Infection des bactéries hôtes par les phages**

Le premier tour de criblage est réalisé sur environ 600000 phages recombinants répartis sur 3 boîtes de Petri 23x23cm. On incube pendant 30' à 37°C sous agitation, 1 ml de XL1-Blue compétentes (Sambrook *et al.*, 1989) avec 20µl d'un milieu SM (Sambrook *et al.*, 1989) contenant 200000 phages. On ajoute alors 35ml de Top Agarose (Sambrook *et al.*, 1989) à 50°C et le mélange obtenu est coulé sur une boîte contenant du milieu LB Agar (Sambrook *et al.*, 1989) et 12,5 µg/ml de tétracycline. Les boîtes sont ensuite incubées à 37°C pendant 12 heures. Le second et le troisième tour de criblage sont réalisés sur des boîtes de Pétri Ø 90 mm, de manière à obtenir au final des plages de lyse bien isolées et pures

*** Empreintes sur membranes de nylon neutre**

Après formation des plages de lyses, les boîtes sont placées à 4°C pendant 1 heure. Une membrane de nylon neutre (Appligene) (20 cm/ 20 cm) est alors appliquée pendant 1' sur la surface de chaque boîte, puis mis en contact successivement pendant 5' dans chacune des solutions suivantes : (1) Une solution de dénaturation contenant, 1,5M de NaCl et 0,5 M de NaOH, (2) une solution de neutralisation contenant 1 M de Tris-HCl (pH7,4) et (3) une solution de lavage contenant un tampon 2xSSC (Sambrook *et al.*, 1989). L'ADN est fixé sur des membranes par un traitement à 80°C pendant 2 heures sous vide.

*** Préhybridation, hybridation et lavage des membranes.**

Les membranes sont préhybridées pendant 3 heures à 60°C dans une solution contenant 5xSSPE (Sambrook *et al.*, 1989), 0,1% SDS, 0,1% BSA, 0,1% Ficoll400, 0,1% PVP, et 100 µg/ml d'ADN de sperme de saumon.

Cette étape étant réalisée, la sonde radioactive, synthétisée à partir d'un clone d'ADNc codant pour la CCoAOMT de *Nicotiana tabacum* (ref U62735) (Martz, 1997), est alors rajouté au liquide de préhybridation, et l'hybridation est réalisée pendant 12 heures à 60°C. Les membranes hybridées subissent par la suite 4 lavages successifs de 30' chacun à 60°C : 2 lavages contenant 2xSSC et SDS 0,2%, puis 2 lavages 1xSSC, SDS 0,1%. Puis elles sont séchées sur du papier filtre et mises au contact d'un film autoradiographique pendant 24 heures à -80°C. Les films sont révélés et les plages de lyses, ayant donné un signal positif, sont prélevées, puis transférées dans 1 ml d'une solution SM et conservées à 4°C.

*** Excision in-vivo**

Dans un tube de 50 ml, 200 µl de bactéries XL1-Blue compétentes (DO=1) sont mis en contact avec 20 µl du stock de phages purifiés (contenant au moins 1×10^5 particules phagiques) et 1 µl de phage filamenteux "helper" R408 (contenant au moins 1×10^6 pfu/ml). Le mélange est incubé à 37°C pendant 15', puis 3 ml de milieu LB sont rajoutés, et le tout est incubé pendant 3 heures à 37°C sous agitation. Après un traitement à 70°C pendant 20', le mélange est centrifugé 5' à 4000 rpm, et le surnageant, contenant les phagemides pBluescript SK⁺ encapsidés, est récupéré dans un tube stérile.

La transformation des bactéries XL1-Blue par les phagemides est réalisée en incubant à 37°C pendant 15', 200µl de bactéries compétentes XL1-Blue (DO = 1) mélangés avec 50µl ou 0,1 µl de la solution de phagemides. Par la suite, 100 µl de ce mélange sont étalés sur des boîtes de Pétri contenant du LB agar et ampicilline à 50µg/ml.

*** Minipréparation d'ADN: lyse alcaline.**

Les colonies bactériennes isolées sont incubées séparément dans 2 ml de milieu LB contenant de l'ampicilline (50µg/ml) pendant 12 heures à 37°C sous agitation. 1,5 ml de culture est centrifugé pendant 3' à 12000 rpm. Le surnageant éliminé, le culot est repris dans 100 µl d'une solution de 25mM de Tris-HCl, 10 mM d'EDTA (pH8). La lyse bactérienne est réalisée par l'ajout de 200 µl d'une solution de

0,2 N NaOH, 1% SDS. Après une forte agitation (vortex), le mélange est placé dans de la glace pendant 5'. 150 µl d'acétate de sodium 3 M (pH5,2) sont alors ajoutés, et suite à une agitation légère, le mélange est replacé dans de la glace pendant 10'. Une centrifugation à 12000 rpm pendant 10' permet de récupérer le surnageant, auquel est ajouté 1 ml d'éthanol absolu froid et le tout est placé à -20°C pendant 30' avant d'être centrifugé pendant 15' à 12000 rpm. Le surnageant est éliminé, et le culot est lavé dans 200 µl d'éthanol 70% puis séché sous vide. Le culot est repris dans 50 µl d'eau distillée et l'ADN plasmidique est analysé sur gel d'agarose 1%.

*** Préparation d'ADN plasmidique sur colonne échangeuse d'anions:
Midipréparation (Wizard, Promega)**

Cette méthode est utilisée pour les préparations d'ADN destinées au séquençage. Les conditions d'extraction et de purification, réalisées à partir de cultures bactériennes de 100 ml dans un milieu LB, sont celles indiquées par le fournisseur (Promega).

*** Analyse en Southern blot**

Southern génomique

30 µg d'un stock d'ADN génomique extrait de feuilles d'*E.gunnii* (Lacombe *et al.*, 1997) sont digérés séparément avec une enzyme de restriction; PstI, HindIII, BamHI, Xho, XbaI. Les produits de la digestion sont séparés par électrophorèse sur un gel d'agarose 0,8%. Le gel est alors dénaturé 30' dans une solution NaOH 0,5M, NaCl 1,5M, puis neutralisé 30' dans une solution Tris 0,5M (pH7,4), NaCl 1,5M. Le transfert de l'ADN sur une membrane de nylon chargé positivement (Applicone) est réalisé dans un tampon 10xSSC pendant 18 heures. La membrane est alors lavée avec un tampon 2xSSC pendant 15', séchée à température ambiante 15' puis mise à 55°C pendant 20'. La préhybridation est réalisée dans une solution, contenant 5xSSPE, 0,5% SDS et 100µg/ml d'ADN de sperme de saumon, pendant 3 heures à 65°C. Une sonde radioactive, synthétisée à partir de l'ADNc codant pour la CCoAOMT d'*E.gunnii* (clone 5, cf résultats), est alors rajoutée et l'hybridation est accomplie pendant 12 heures à la même température. Les membranes sont alors lavées 3 fois 10' à 55°C dans une solution 2xSSC, 0,5% SDS. Les membranes sont alors séchées puis mise en contact avec un film autoradiographique pendant 3 jours.

Southern plasmidique

Suite à une digestion enzymatique (XhoI, EcoRI) de 3 µg d'ADN plasmidique, une migration électrophorétique, réalisée dans un gel d'agarose 1%, permet de séparer les inserts des vecteurs. Cette étape précède un transfert sur membrane, une hybridation (réalisée avec une sonde synthétisée à partir de l'ADNc de *N.tabacum*) puis une mise en contact avec un film autoradiographique. Les protocoles utilisés sont identiques à ceux utilisés pour le southern génomique

Technique d'amplification génique (PCR)

*** Amplification par PCR**

Pour chaque réaction PCR, le mélange réactionnel (volume total de 20 µl) contient 75 mM de Tris-HCl (pH9), 2,5 mM de MgCl₂, 0,01% Tween20, 20 mM de sulfate d'ammonium, 250 µM de chaque dNTP, 0,25 µM d'amorces, 1 µg d'ADN génomique (extrait de feuilles d'*Eucalyptus gunnii*), et 0,4U de Taq-polymerase (Goldstar DNA polymerase, Eurogentec). L'amplification est réalisée à l'aide d'un thermocycleur Hybaid OmniGene (modèle TR3CM220) programmé pour une dénaturation préliminaire (5', 95°C), suivit de 30 cycles comprenant chacun une étape de dénaturation (1', 94°C), une hybridation (1'30, 72°C), une élongation (45", 55°C), et enfin une élongation finale (5', 72°C). Les produits d'amplification sont alors analysés sur un gel d'agarose 1%.

*** Purification de fragments d'ADN à partir d'un gel d'agarose**

Les bandes d'intérêt sont excisées et l'ADN est purifié, sur des particules de silices, à l'aide d'un kit d'extraction (QIAEXII, QIAGEN) selon le protocole décrit par le fournisseur.

*** Ligation au vecteur de clonage**

Les segments d'ADN purifiés sont introduits dans un vecteur pGEM-T selon un protocole fourni par le fournisseur (pGEM-T vector SYSTEM 1, Promega). La réaction de ligation est réalisée pendant 3 heures à 15°C et la totalité de cette réaction a été directement utilisée pour la transformation de DH5α compétentes (Sambrook *et al.*, 1989).

* Transformation des bactéries

200 µl de DH5α compétentes sont transférées dans un tube en polypropylène au froid. La totalité de la réaction de ligation est ajoutée et le mélange est homogénéisé doucement par rotation avant d'être placé pendant 20' dans de la glace. La transformation est provoquée par un choc thermique de 90" à 42°C puis replongée immédiatement dans de la glace. 800µl de milieu LB sont alors ajoutés et le tout est placé pendant 1 heure à 37°C sous agitation. 200 µl de culture sont étalées sur des boîtes de Pétri contenant un milieu LB Agar, IPTG (Isopropylthio- β-D-galactoside) 80µg/ml, Xgal (5-Bromo-4-chloro-3-indolyl- β-D-galactoside) 40µg/ml et ampicilline (50µg/ml).

Cartographie

* Matériel végétal

Une famille de pleins frères issue d'un croisement interspécifiques *Eucalyptus urophylla* (clone 14.144) (parent femelle) x *E.grandis* (clone 9.21) (parent mâle) a été choisie dans le schéma de sélection récurrente réciproque mis en place au Congo (Vigneron, 1992). Cette famille F1, constituée de 92 individus, a été utilisée pour construire les cartes de liaisons génétiques de ces géniteurs, grâce à l'utilisation de marqueurs anonymes dominants RAPD (Verhaegen et Plomion, 1996). Ces cartes génétiques ont servi de support pour positionner les gènes de la COMT et de la CCoAOMT.

* Amplification PCR et migration

Pour chaque réaction PCR, le mélange réactionnel (volume total de 25 µl) contient 67 mM de Tris-HCl (pH7,5), 2 mM de MgCl₂, 1 ng de BSA (Serum albumine Bovine), 0,2% de β-mercaptoéthanol, 16 mM de sulfate d'ammonium, 200 µM de chaque dNTP, 0,4 µM d'amorce, 15 ng d'ADN génomique (provenant soit des parents, soit des 92 individus F1 (Verhaegen et Plomion, 1996)) et 1U de Taq-polymerase (GIBCO BRL). L'amplification est réalisée avec un thermocycleur Techne cyclogene (modèle FPHC3MD) programmé pour une dénaturation préliminaire (4', 94°C), suivit de 30 cycles comprenant chacun une dénaturation (45", 94°C), une hybridation (45", 72°C), une élongation (1', 60°C pour la COMT et 45", 62°C pour la CCoAOMT), et enfin une élongation finale (3', 72°C). Les produits PCR sont déposés

sur un gel d'agarose 1,8%, puis soumis à une électrophorèse à voltage constant (110V) pendant 3 heures, dans du tampon d'électrophorèse 0,5xTBE.

*** Protocole SSCP (Single Strand Conformation Polymorphism)**

Le protocole utilisé est basé sur celui élaboré par Bodénès (1996). Pour l'analyse SSCP, 8µl de produit PCR sont ajoutés à 15µl d'une solution contenant 9,2% de formamide, 10mM de NaOH, 0,05% de xylène cyanol, et 0,05% de bleu de Bromophénol. Les échantillons sont placés à 94°C pendant 4', puis plongés immédiatement dans de la glace. Ils sont par la suite chargés sur un gel acrylamide non dénaturant (1mm x 16 cm x 18 cm). Suite à plusieurs essais de migration, nous avons utilisé, pour la CCoAOMT, un gel contenant TBE 6X (6%), acrylamide/bis, 37/1 (10%), 5% glycérol, TEMED (N,N,N,N'-tetramethylethylenediamine) (0,065%) et APS (persulfate d'ammonium) (0,065%). Pour la COMT, nous avons utilisé un gel identique mais ne contenant pas de glycérol.

La migration électrophorétique est réalisée à une température constante de 15°C, à 190 V pendant 20h, dans un tampon de migration de 0,6X TBE.

*** Coloration des gels SSCP**

Les gels ont été fixés pendant 20' dans une solution d'acide acétique à 10% et lavés 3 fois 2' dans de l'eau distillée. L'étape de coloration a eu lieu pendant 30' dans une solution contenant 0,1% de nitrate d'argent et 0,056% de formaldéhyde. Les gels ont été rincés à l'eau distillée pendant 20" , puis ont été placés pendant 5 à 10' dans une solution de révélation (30g/l de carbonate de sodium, 0,056% de formaldéhyde et 2mg/l de thiosulfate de sodium). La coloration a été arrêtée en plongeant les gels dans une solution d'acide acétique 10%. Après un dernier rinçage à l'eau, les gels ont été photographiés.

*** Analyse de ségrégation et positionnement sur les cartes génétiques**

Les deux cartes génétiques ont été construites à l'aide du logiciel MAPMAKER (Lander *et al.*, 1987). Les gènes ont été localisés sur les cartes à l'aide de la fonction "PLACE" qui permet de placer de nouveaux marqueurs dans une carte existante.

Références Bibliographiques

- Altschul S. F, Gish W., Miller W., Myers E. W., Lipman D. J., 1990. Basic local alignment search tool. *J.Mol.Biol.*, **215**, 403-410
- Bodénès C., Laigret F., Kremer A., 1997. Inheritance and molecular variations of PCR-SSCP fragments in Pedunculate oak (*Quercus robur* L.). *Theor. Appl. Genet.* (sous presse)
- Boudet A. M., Lapierre C., Grima-Pettenati J., 1995. Biochemistry and molecular biology of lignification. *New Phytol.*, **129**, 203-236
- Boudet A. M., Grima-Pettenati J., 1996. Lignin genetic engineering. *Molecular Breeding*, **2**, 25-39
- Bouvet J. M., 1995. Evolution de la variabilité avec l'âge et corrélation juvénile-adulte dans les population d'*Eucalyptus*, *Thèse de l'Institut National Agronomique Paris Grignon*, 236 p
- Bugos R. C., Chiang V. L. C., Campbell W. H., 1991. cDNA cloning, sequence analysis and seasonal expression of lignin-bispecific caffeic acid/5-hydroxyferulic acid O-methyltransferase of aspen. *Plant Molecular Biology*, **17**, 1203-1215
- Feuillet C., Lauvergeat V., Deswarte C., Pilate G., Boudet A.-M., and Grima-Pettenati J., 1995. Tissue and cell specific expression of a cinnamyl alcohol dehydrogenase promoter in transgenic poplar plants. *Plant. Mol. Biol.* **27**, 651-667.
- Gion J. M., 1996. Etude de l'architecture génétique de caractères quantitatifs chez l'*Eucalyptus*, *DEA Option génétique, Adaptation et Production Végétale ENSAR* 32p
- Grattapaglia D., Sederoff R., 1994. Genetic linkage maps of *Eucalyptus grandis* and *E. urophylla* using a pseudo-testcross mapping strategy and RAPD markers. *Genetics*, **137**, 1121-1137
- Halpin C., Knight M.E., Foxon G.A., Campbell M.M., Boudet A.M., Boon J.J., Chabbert B., Tollier M.T., and Schuch W., 1994. Manipulation of lignin quality by down-regulation of cinnamyl alcohol dehydrogenase. *Plant J.*, **6**, 339-350.
- Hayakawa T., Nanto K., Kawai S., Katayama Y. and Morohoshi N., 1996. Molecular cloning and tissue-specific expression of two genes that encode caffeic acid O-methyltransferase from *Populus kitakamiensis*. *Plant Sci.*, **113**, 157-165
- Hayashi K., Yandell D. W., 1993. How sensitive is PCR-SSCP?. *Human Mutation*, **2**, 338-346

- Hibino T., Chen J., Shibata D., Higushi T., 1994.** Nucleotide sequence of a *Eucalyptus botryoides* gene encoding cinnamoyl alcohol dehydrogenase. *Plant Physiol.*, **104**, 305-306
- Lacombe E., Hawkins S., Van Doorselaere J., Piquemal J., Goffner D., Poeydomenge O., Boudet A. M., Grima-Pettenati J., 1997.** Cinnamoyl coA reductase, the first committed enzyme of the lignin branch biosynthetic pathway: cloning, expression and phylogenetic relationships. *The Plant Journal*, **11**, 429-441
- Lande R. and Thompson R., 1990.** Efficiency of marker-assisted selection in the improvement of quantitative traits, *Genetics* **124**, 743-756
- Lander E.S., Green P., Abrahamson J., Barlow A., Daly M.J., Lincoln S.E., Newburg L., 1987.** MAPMAKER : an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics* **1**, 174-181
- MacKay J. J., Liu W., Whetten R., Sederoff R. R., O'Malley D. M., 1995.** Genetic analysis of cinnamoyl alcohol deshydrogenase in loblolly pine: single inheritance, molecular characterization and evolution. *Mol. Gen. Genet.*, **247**, 537-545
- Martin C., 1996.** Transcription factors and the manipulation of plant traits. *Curr. Opinion Biotech.*, **7**, 130-138
- Martz F., 1997.** Modification de l'activité O-méthyltransférase dans des tabacs transgéniques: conséquences sur la lignine et la résistance au virus de la mosaïque du tabac. *Thèse de l'Université Louis Pasteur, Strasbourg*
- Michaud J., Brody L. C., Steel G., Fontaine G., Martin L. S., Valle D., Mitchell G., 1992.** Strand separating conformational polymorphism analysis: efficacy of detection of point mutations in the human ornithine δ -aminotransferase gene. *Genomics*, **13**, 389-394
- Orita M., Iwahana H., Kanazawa H., Hayashi K., Sekiya T., 1989.** Detection of polymorphisms of human DNA by gel electrophoresis as single-strand conformation polymorphisms. *Proc. Natl. Acad. Sci. USA*, **86**, 2766-2770
- Pellegrini L., Geoffroy P., Fritig B., and Legrand M., 1993.** Molecular cloning and expression of a new class of ortho-diphenol-O-methyltransferase induced in tobacco (*Nicotiana tabacum* L.) leaves by infection or elicitor treatment. *Plant Physiol.*, **103**, 509-517
- Poeydomenge O., Boudet A. M., Grima-Pettenati J., 1994.** A cDNA encoding S-adenosyl-L-methionine: caffeic acid 3-O-methyl transferase from *Eucalyptus*. *Plant Physiol.*, **105**, 749-750
- Sambrook J., Fritsch E. F., Maniatis T., 1989.** Molecular cloning, a laboratory manual, *11nd edition, CSH laboratory press*

- Sanger F., Nicklen S., Coulson A. R., 1977.** DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467
- Schmitt D., Pakusch A. E., Matern U., 1991.** Molecular cloning, induction, and taxonomic distribution of caffeoyl-coA 3-O-methyltransferase, an enzyme involved in disease resistance. *The Journal of Biological Chemistry*, **266**, 17416-17423
- Sheffield V. C., Beck J. S., Kwitek A. E., Sandstrom D. W., Stone E. M., 1993.** The sensitivity of single strand conformation polymorphism analysis for detection of single base substitutions. *Genomics*, **16**, 325-332
- Strauss S. H., Lande R., and Namkoong G., 1992.** Limitations of molecular-marked-aided selection in forest tree breeding. *Can. J. For. Res*, **22**, 1050-1061
- Tanksley S. D., 1993,** Mapping Polygenes. *Annu. Rev. Genet*, **27**, 205-233
- Teulières C., Marque C. and Boudet A.-M., 1994.** Genetic transformation of *Eucalyptus*, in Y.P.S. Bajaj, Springer-Verlag (eds.) Biotechnology in Agriculture and Forestry, Vol. 29 "Plant protoplasts and genetic engineering V", p. 289-306.
- Verhaegen D. et Plomion C., 1996.** Genetic mapping in *Eucalyptus urophylla* and *Eucalyptus grandis* using RAPD markers, *Genome* **39**, 1051-1061
- Verhaegen D., 1996.** Prédiction des valeurs génétiques, cartographie génétique et recherche de QTL dans un croisement interspécifique *Eucalyptus urophylla* x *Eucalyptus grandis*, *Thèse de doctorat, Université Paris XI Orsay*, 125 p
- Vidgren J., Svensson L. A., Lijjas A., 1994.** Crystal structure of catechol O-methyltransferase. *Nature*, **368**, 354-358
- Vigneron P., 1992.** Création et amélioration de variétés d'hybrides d'*Eucalyptus* au Congo. *Bois et Forêts des tropiques*, **234**, 29-42
- Wallace G., Fry S. C., 1994.** Phenolic components of the plant cell wall. *International Review of cytology*, **151**, 229-267
- Ye Z. H., Kneusel R. E., Matern U., Varner J. E., 1994.** An alternative methylation pathway in lignin biosynthesis in *Zinnia*, *The Plant Cell*, **6**, 1427-1439
- Ye Z. H., Varner J. E., 1995.** Differential expression of two O-methyltransferases in lignin biosynthesis in *Zinnia elegans*. *Plant Physiol.*, **108**, 459-467

