

SOMMAIRE

Introduction et synthèse bibliographique

p 1-8

- 1. Généralités sur l'eucalyptus**
- 2. Programme d'amélioration de L'*Eucalyptus* au Congo**
- 3. Les principaux problèmes de la sélection chez les arbres forestiers**
- 4. Les biotechnologies au service de la sélection**
- 5. L'approche gènes candidats appliquée à la voie de biosynthèse des lignines**

Matériel et méthodes

p 8-16

- 1. Matériel végétal**
 - 1.1. Descendance hybride étudiée
 - 1.2. L'extraction d'ADN
- 2. Gènes étudiés**
- 3. PCR dirigée pour chaque gènes fonctionnels**
 - 3.1. Choix des amorces
 - 3.2. Conditions d'amplification PCR
- 4. Séquençage des produits PCR**
- 5. Révélation du polymorphisme par SSCP**
 - 5.1. Principe
 - 5.2. Conditions SSCP
- 6. Cartographie génétique**
- 7. Caractères mesurés**
- 8. Détection de QTL**
 - 8.1. Analyse de variance (ANOVA)
 - 8.2. Maximum de vraisemblance (Interval Mapping)

Partie I : Cartographie des gènes

- 1. Optimisation des conditions d'amplification PCR**
- 2. Résultats du séquençage**
- 3. Optimisation des conditions SSCP**
- 4. Héritabilité mendélienne des fragments SSCP**
- 5. Causes moléculaires du polymorphisme SSCP**
- 6. Localisation des gènes sur les cartes parentales**

Partie II : Colocalisation gène/QTL

- 1. Les caractères étudiés**
 - 1.1. Etude de la variabilité des caractères
 - 1.2. Corrélations phénotypiques des caractères
- 2. Mise en évidence de QTL**
 - 2.1. Détection par ANOVA et Interval mapping
 - 2.1.1. *Caractères de vigueur*
 - 2.1.2. *C -ractères de forme*
 - 2.1.3. *Le pylodin*
 - 2.1.4. *Propriétés mécaniques du bois*
 - 3.5.3 *Propriétés chimiques du bois*
 - 2.1. Test de permutation
- 2. Colocalisation gène-QTL**
- 3. Cas des gènes CAD et CCoAOMT**

Introduction - Synthèse bibliographique

1. Généralités sur l'eucalyptus

Le genre *Eucalyptus* appartient à la famille des Myrtacées qui compte environ 700 espèces, nombre qui varie selon les auteurs (revue par Vigneron, 2000). Les aires naturelles de répartition s'étendent des Philippines jusqu'au sud de l'Australie (Jacobs, 1954). Leur évolution ayant suivi celle du super continent Gondwana, la plupart des espèces ont des caractéristiques d'adaptation assez remarquables qui leur permettent de pousser dans des conditions pédoclimatiques très diverses (Eldridge *et al.*, 1993). Outre ces caractères d'adaptation, certaines espèces sont capables de se développer dans des zones impropres à l'agriculture ou à tout autre plantation forestière. Enfin, les eucalyptus possèdent des propriétés sylvicoles intéressantes en plantation : une croissance rapide, une possibilité de multiplication végétative par bouturage, et la propriété de rejeter permettant plusieurs rotations sur une même plantation.

Actuellement, environ 15 millions d'hectares sur la planète sont consacrés à la culture de l'eucalyptus, ce qui en fait l'essence feuillue la plus plantée dans le monde (Vigneron et Bouvet, 1997). Le bois d'eucalyptus est utilisé aussi bien comme bois d'énergie (charbon de bois et bois de chauffage), bois de construction (poteaux, bois de sciage), bois pour l'industrie papetière. Au Congo, l'eucalyptus est essentiellement utilisé comme bois de pâte et bois d'énergie.

2. Programme d'amélioration de l'Eucalyptus au Congo

En 1955, plusieurs espèces d'*Eucalyptus* ont été introduites au Congo en vue de comparer leurs performances et de repérer les mieux adaptées aux conditions pédoclimatiques du Congo. Ensuite, des hybrides naturels sont apparus avec des performances de croissance supérieures à celles des espèces introduites. A partir des années 70, ces hybrides ont été utilisés en plantations industrielles après la mise au point des techniques de multiplication végétative. Le programme d'amélioration génétique a débuté dans les années 80 avec comme objectif d'« obtenir des eucalyptus capables de dépasser la limite de production atteinte par les hybrides

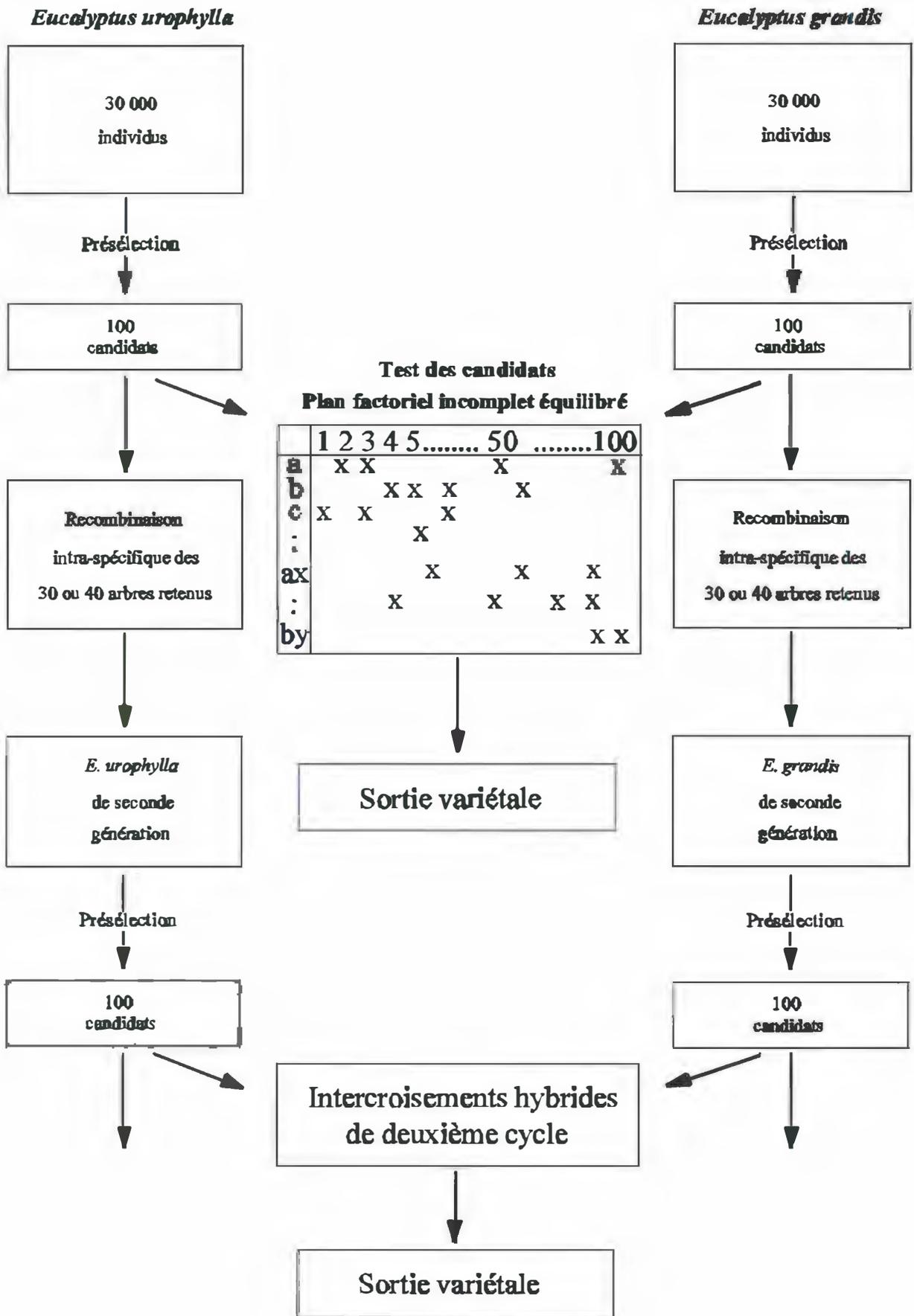


Figure 1 : Sélection Récurrente Réciproque pour *E. urophylla* x *E. grandis*. Schéma général des deux premiers cycles (Vigneron, 19911)

naturels *Eucalyptus PF1* et *E. 12ABL** *E. saligna* (Vigneron,1992). La maîtrise des techniques de pollinisation contrôlée (Martin et Quillet, 1974) a permis de tester différentes combinaisons hybrides. Une des formules hybrides la plus prometteuse s'est avérée être le croisement entre *E. urophylla* et *E. grandis*. Les caractéristiques de ces deux espèces sont complémentaires : *E. urophylla* présente une bonne adaptation au Congo, une bonne forme et une assez bonne production alors qu'*E. grandis* est mal adapté, présente une assez bonne forme et une production potentielle importante. Dans la mesure où les populations de base sont très divergentes, que les populations parentales sont fortement complémentaires et que le taux de consanguinité de l'hybride est nul, la sélection récurrente réciproque (SRR) est apparue comme la meilleure solution pour l'amélioration génétique des eucalyptus au Congo (Vigneron, 1991).

Dans son principe la SRR vise à améliorer de façon conjointe et orientée deux groupes d'individus l'un par rapport à l'autre. Ceci à pour but d'obtenir des hybrides intergroupes recombinants différents caractères présents séparément dans chacun des deux groupes (Gallais, 1978). Le schéma de la SRR est représenté à la figure 1. La SRR comporte plusieurs étapes :

- ① la constitution des populations parentales
- ② la présélection d'une centaine de géniteurs
- ③ la réalisation d'un plan factoriel incomplet pour les géniteurs sélectionnés. Il permet d'estimer la valeur de chaque géniteur en croisement (l'aptitude générale à la combinaison : AGC). Les hybrides interspécifiques de première génération constituent la base de la sortie variétale qui peut prendre plusieurs formes : clones, mélange de clones ou famille...
- ④ une recombinaison intraspécifique entre géniteurs sélectionnés sur la base de leur AGC permet d'obtenir les populations parentales de seconde génération. Cette étape de recombinaison intraspécifique évitent une perte trop rapide de la variabilité génétique, tout en augmentant la fréquence des allèles favorables.

3. Les principaux problèmes de la sélection chez les arbres forestiers

Pour de nombreuses espèces forestières un minimum de 20 à 30 ans de recherches et d'évaluation est nécessaire avant de pouvoir produire les premières variétés. En effet, l'amélioration génétique des arbres forestiers est confrontée à trois contraintes majeures :

➤ Une évaluation tardive des caractères de croissance et de qualité du bois. Il faut attendre, pour les espèces à croissance rapide 6 à 12 ans chez le pin radiata (Cotterill et Dean, 1988), le pin taeda (McKeand, 1988) et trois ans chez l'eucalyptus (Bouvet, 1991). L'évaluation est beaucoup plus tardive pour les espèces à croissance lente comme le chêne, le teck, l'acajou.

➤ Les caractères intéressants pour le sélectionneur (croissance, forme, adaptation, qualité du bois) sont soumis à de forts effets du milieu. L'héritabilité (*sensu stricto*) de ces caractères est généralement faible à moyenne variant de 0,10 à 0,40 (Kremer, 1994). Le choix des meilleurs individus dans les meilleurs familles n'est donc pas optimal.

➤ Les sélectionneurs sont contraints d'attendre la maturité sexuelle avant de pouvoir effectuer des croisements. Cette maturité sexuelle apparaît tardivement chez les arbres forestiers : de 3 à 5 ans chez l'eucalyptus, de 5 à 10 ans chez le peuplier, de 15 à 25 ans chez l'épicéa.

4. Les biotechnologies au service de l'amélioration génétique

Même si les sélectionneurs sont en mesure d'offrir des variétés forestières améliorées à l'aide des méthodes de sélection classique, des techniques visant à augmenter les gains génétiques par unité de temps (sélection précoce) et à améliorer l'efficacité de la sélection (choix des meilleurs géniteurs) sont recherchées. Dans l'optique de développer rapidement des variétés d'*Eucalyptus* à forte productivité et mieux adaptées aux applications industrielles, deux stratégies sont envisageables :

La transformation génétique combinée à la production en masse du matériel *via* l'embryogenèse somatique (Harvengt *et al.*, 2000), permet de « créer » puis de multiplier rapidement des génotypes répondant aux demandes des utilisateurs

(sylviculteurs, industriels...). Des résultats récents, obtenus notamment chez le tabac et le peuplier, ont montré qu'il était possible de modifier la composition chimique du bois au niveau des lignines et de la cellulose (Grima-Pettenati et Goffner, 1999). Cette modification permettrait d'optimiser le procédé industriel de fabrication de la pâte à papier tout en le rendant moins polluant et moins coûteux (Boudet, 1997). Chez l'eucalyptus, les problèmes liés à la régénération rendent pour le moment la multiplication par embryogenèse somatique impossible.

D'autre part, des questions restent posées : sur plusieurs années, comment se comportent ces arbres transformés en conditions naturelles ? Quelles sont les éventuelles répercussions sur l'environnement de l'introduction de transgènes (Jouanin, 2000) ? Enfin, les craintes de l'opinion publique relatives aux OGM rendent délicate l'utilisation d'arbres transgéniques pour l'instant.

La seconde stratégie s'appuie sur l'exploration et la mise à profit de la variabilité génétique naturelle. La sélection assistée par marqueurs (SAM) permettrait de sélectionner rapidement à moindre coût les individus les plus performants : sélection des individus qui possèdent le ou les allèles favorables des gènes qui contrôlent les caractères quantitatifs d'intérêt. Ces loci sont appelés QTL pour Quantitative Trait Locus (Geldermann, 1975). En 1923, Sax a été le premier à utiliser des marqueurs morphologiques pour mettre en évidence les zones chromosomiques impliquées dans les variations quantitatives. Il trouva une association entre la pigmentation des téguments de la graine de haricot (caractère qualitatif contrôlé par un seul gène) et la taille du grain (caractère quantitatif contrôlé par de nombreux gènes). Depuis, le développement des techniques de marquage moléculaire a ouvert de nouvelles perspectives pour la détection de QTL : la construction de cartes génétiques saturées et la recherche de QTL sur ces cartes. Des cartes génétiques saturées ont été construites pour différentes espèces d'*Eucalyptus*, à l'aide de marqueurs anonymes : RAPD (Grattapaglia et al., 1994 ; Verhaegen et Plomion, 1996), AFLP (Marques et al., 1998). L'existence de ces cartes a permis la mise en évidence de QTL pour des caractères de qualité du bois, de croissance et de multiplication végétative (Grattapaglia et al., 1995 ; Verhaegen et al., 1997 ; Byrne et al. 1997 ; Marques et al. 1999).

Cependant par rapport aux espèces annuelles, les analyses QTL sont encore peu développées chez les arbres forestiers (Plomion, 2000). En effet, ce sont

généralement des espèces allogames caractérisées par un haut niveau de diversité génétique et un déséquilibre de liaison faible. Or, un déséquilibre de liaison entre le locus marqueur et le QTL est indispensable pour la détection. L'utilisation des gènes sous jacent aux QTL pourrait permettre de s'affranchir du faible déséquilibre de liaison, des phénomènes de recombinaison et de caractériser ces QTL.

5. L'approche gènes candidats appliquée à la voie de biosynthèse des lignines

L'approche gènes candidats fonctionnels a été utilisée chez différentes espèces végétales.

Le premier exemple de gène de fonction connue qui pourrait correspondre à un QTL est celui de *Shrunken-2*, qui code pour une des sous-unités de l'ADP-glucose pyrophosphorylase (ADPGp-ase), enzyme clef de la synthèse de l'amidon dans le grain de maïs. En effet, Goldman *et al.*, 1993 ont trouvé qu'un QTL à très fort effet pour le contenu en amidon se localisait précisément au niveau de ce gène. Chez le pin maritime, il a été mis en évidence qu'un QTL de la croissance juvénile était localisé à proximité du gène de la glutamine synthase, enzyme intervenant dans l'assimilation de l'azote (Plomion, 2000).

Dans le cadre du programme d'amélioration de l'*Eucalyptus au Congo*, les efforts de sélection ont été centrés, entre autre, sur la qualité du bois. Dans cette optique, il apparaît pertinent de tester si le polymorphisme des gènes impliqués dans la voie de biosynthèse des lignines a une influence sur les qualités chimiques et rhéologiques du bois. Par exemple, pour l'utilisation du bois comme source d'énergie, une teneur accrue en lignines, plus énergétique que la cellulose est recherchée (Brown, 1985). Par contre, au niveau des industries papetières, l'élimination des lignines de la cellulose (constituant de la pâte à papier) est très coûteuse et polluante.

Au total, les lignines représentent 20 % de la biomasse terrestre (Higuchi, 1990) et constituent le biopolymère végétal le plus abondant après la cellulose. Les teneurs en lignines peuvent varier naturellement dans une large gamme de valeurs selon le matériel végétal considéré. Ainsi, la teneur en lignines varie de 15% à 37% chez les espèces arborescentes et de 3 à 15% chez les plantes herbacées. Chez l'eucalyptus, la teneur en lignines dépend de l'espèce considérée. Ainsi, pour deux espèces d'intérêt industriel *E. globulus* et *E. grandis* les teneurs moyennes sont

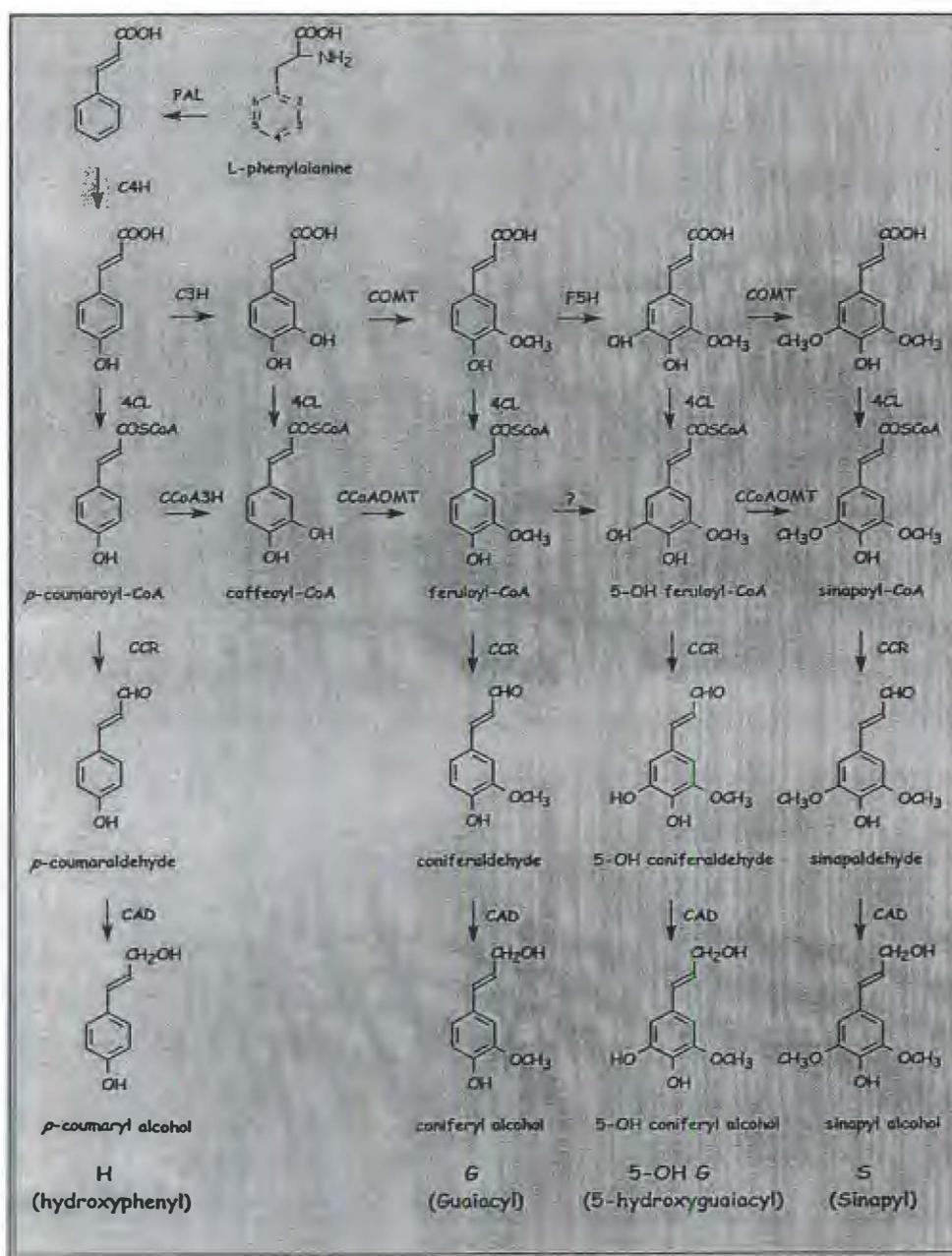


Figure 2 : Voie de biosynthèse des lignines à partir de la phénylalanine. Les enzymes de la voie commune des phénylpropanoïdes comprennent la phénylalanine ammonia lyase (PAL), la cinnamate-4-hydroxylase (C4H), la coumarate-3-hydroxylase (C3H), l'acide caféique 3-O-méthyltransférase (COMT), la férulate 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 cinnamyliques (monolignols) sont respectivement catalysées par la cinnamoyl CoA réductase (CCR) et l'alcool cinnamylique déshydrogénase (CAD).

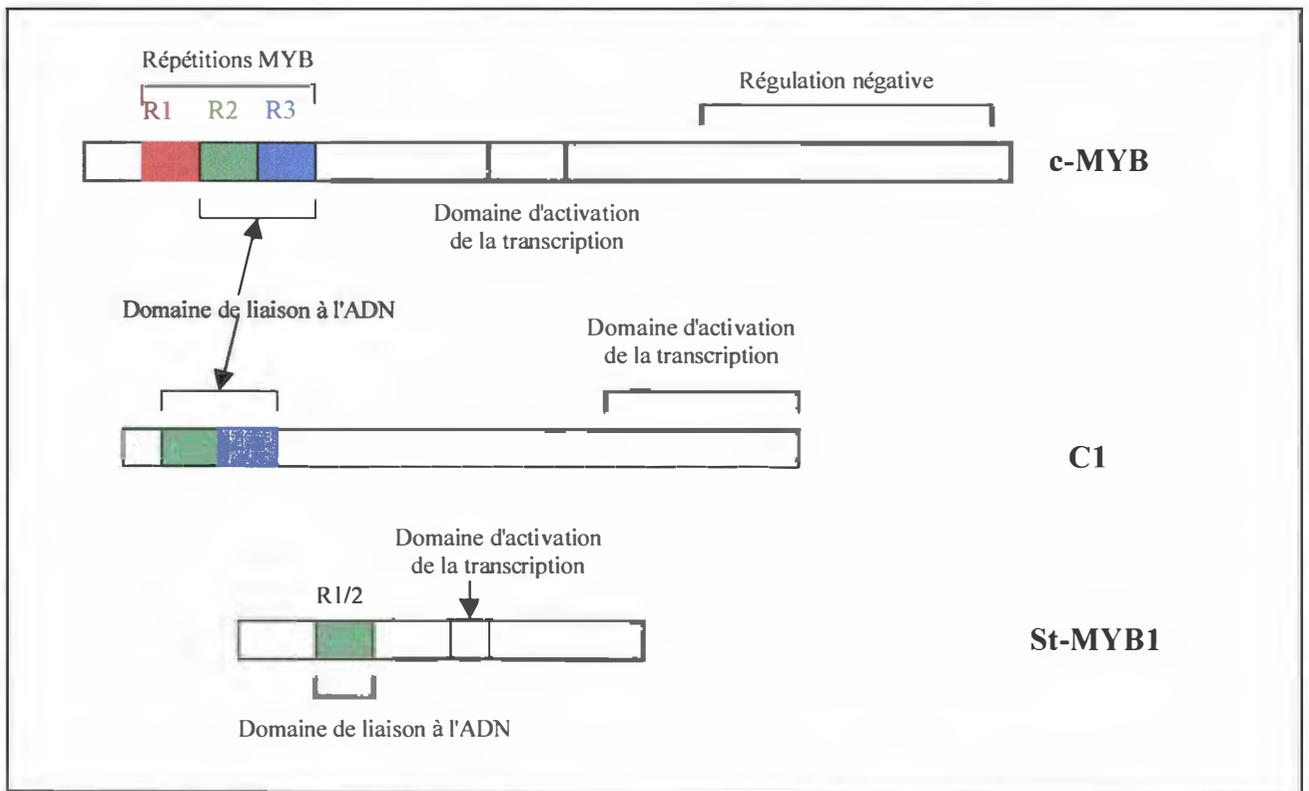
respectivement de 19,5 et de 24,5% (lignines de klason). L'approche gènes candidats devrait donc pouvoir s'appuyer sur l'hétérogénéité naturelle, entre individus, des teneurs en lignines et/ou leurs compositions monomériques.

La biosynthèse des lignines implique le fonctionnement de deux voies métaboliques distinctes (Figure 2):

- La voie générale des phénylpropanoïdes avec notamment la phénylalanine ammonia-lyase (PAL) qui est responsable de la transformation de la phénylalanine en acide cinnamique. La PAL est la première enzyme dont l'activité a été réduite dans des tabacs transgéniques. Son inhibition entraîne une nette diminution de la teneur en lignine, ainsi qu'un effet néfaste au niveau de la croissance et du développement (Elkind *et al.*, 1990). Dernière enzyme de la voie des phénylpropanoïdes, la 4CL (4 coumarate CoA ligase) catalyse la réaction permettant la formation des thioesters CoA à partir des acides cinnamiques. La réduction de son activité (via la stratégie antisens) chez arabidopsis (Lee *et al.*, 1997) et le tabac (Kajita *et al.*, 1997) conduit à une réduction modeste de la teneur en lignines mais n'a pas d'influence sur la morphologie. Chez le peuplier, on observe une diminution substantielle de la quantité en lignines et plus surprenant, une augmentation de 15 % du taux de cellulose (Hu *et al.*, 1998).

- La voie spécifique de synthèse des lignines aboutissant à la polymérisation oxydative des monolignols. La première enzyme à intervenir est la cinnamoyl CoA réductase (CCR) qui convertie les esters cinnamoyl CoA en leurs cinnamaldéhydes correspondants. Une baisse de son activité induit une diminution spectaculaire (50%) de la teneur en lignines et une diminution de l'indice Kappa (indice de délignification) des pâtes à papier (Piquemal *et al.*, 1998). La cinnamyl alcool déhydrogénase (CAD) est impliquée dans la réduction des hydroxycinnamyl aldéhydes (sinapyl, paracoumaryl et coniféryl aldéhydes) en leur alcool correspondant qui sont les précurseurs directs des lignines. L'obtention de tabacs et de peupliers ayant une réduction de l'activité CAD montre une teneur en lignines inchangée. Cependant, les lignines sont enrichies en cinnamaldéhydes impliquant une extractabilité facilitée des lignines par le traitement alcalin (Halpin, 1994).

Figure 3 : Schéma illustrant les différentes classes de protéines MYB se différenciant par le nombre de répétitions dans le domaine de liaison à l'ADN : trois répétitions MYB (c-MYB), deux répétitions MYB (C1) et une répétition MYB (StMYB1) (d'après Hailing et Martin, 1999).



Le polymère lignine est donc composé de trois unités structurales (monolignols) de types H (p-hydroxyphénol), G (guaiacyl) et S (syringyl) (Freudenberg, 1959) se différenciant les unes des autres par leur degré de méthylation.

Grâce à une collaboration avec le CNRS et la présence dans les banques de biomolécules des séquences de plusieurs gènes de la voie de biosynthèse des lignines, la localisation de six gènes de structure a été réalisée (sur les cartes génétiques établies chez *E. urophylla* et *E. grandis* (Gion et al., 2000).

Toutefois les gènes de structure ne sont pas les seuls à être potentiellement intéressants, les facteurs de transcription apparaissent aussi comme de bons gènes candidats potentiels. En effet, chez le maïs, un activateur de transcription *p 1* contrôlerait 58% de la variation de la teneur en maysine, flavonoïde à effet répulsif pour une espèce de noctuelle (*Helicoverpa zea* Boddie) (Byrne et al., 1996). Le rôle prépondérant des gènes régulateurs par rapport aux gènes de structure peut-être fortuit, mais il semble cohérent avec la théorie des flux métaboliques (Kacser et Burns, 1980), qui montre que les variations d'efficacité des gènes de structure codant pour des enzymes impliquées dans des voies métaboliques complexes ont en générale des conséquences plus faibles que celles de gènes de régulation.

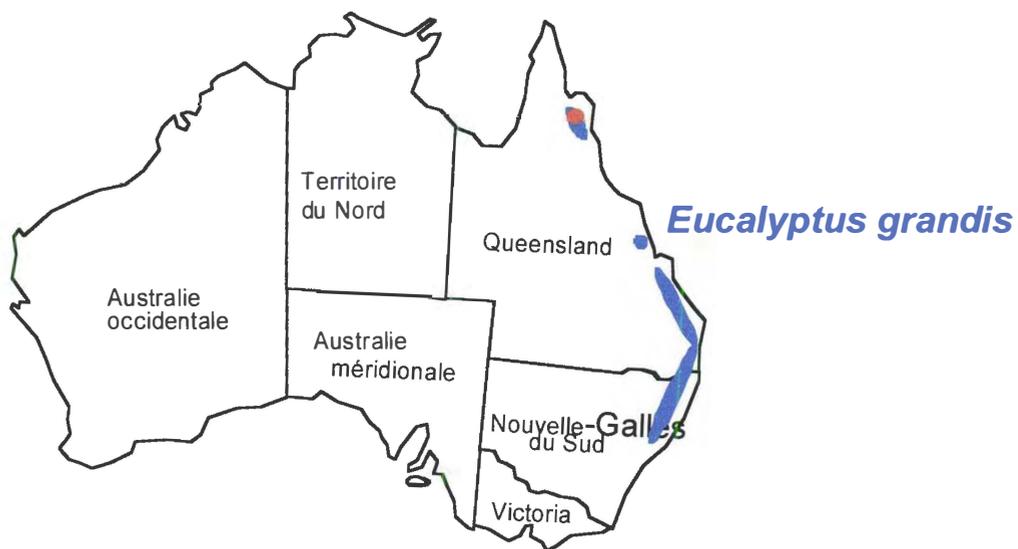
Le clonage chez *E. gunnii*, de trois séquences nucléiques codant des facteurs MYB a été réusit au CNRS à Toulouse. Il est fort probable que ces facteurs MYB régulent l'expression de certains gènes de la voie de biosynthèse des lignines (Grima-Pettenati, communication personnelle) et la voie générale des phénylpropanoïdes (Sablowski et al., 1994). Les membres de cette famille se caractérisent par la présence d'un domaine très conservé au niveau de leur structure : le domaine de liaison à l'ADN ou domaine MYB. Différentes catégories de protéines MYB ont pu être mises en évidence par le nombre de répétitions situées dans le domaine MYB (figure 3) (Hailing, 1999).

Ce stage s'inscrit dans le cadre du projet MENRT dont la finalité est d'identifier les gènes qui contrôlent la qualité du bois chez les arbres forestiers. C'est un projet multidisciplinaire, qui réunit différents partenaires : AFOCEL, CIRAD forêt, CNRS, INRA, Institut du Pin, Laboratoire de rhéologie du bois et Smurfit worlwide research.

Dans un premier temps mon travail a consisté principalement à cartographier des gènes de régulation et des gènes de structure impliqués dans la voie de biosynthèse des lignines, il a été étendu à d'autres gènes fonctionnels dont la séquence chez *Eucalyptus sp.* était disponible dans les banques de biomolécules. Ceci a nécessité la recherche et la mise au point d'amorces spécifiques pour les gènes ciblés et la révélation d'un polymorphisme à l'intérieur du fragment de chaque gène spécifique. Dans un second temps, nous avons étudié la colocalisation entre les gènes fonctionnels et les caractères de qualité du bois.

Matériel et Méthodes

 *Eucalyptus urophylla*



- Meilleures provenances testées au Congo

Figure 4 : Aires naturelles des deux espèces d 'Eucalyptus utilisées dans le programme d 'amélioration au Congo.

1. Matériel végétal

1.1. Descendance hybride étudiée

Nous avons utilisé pour notre étude un croisement interspécifique *E. urophylla***E. grandis*. Le géniteur *E. urophylla* (clone 14.144) est utilisé comme parent femelle, il provient de Monte Lewotobi (Iles de Flores) et le géniteur mâle, *E. grandis* (clone 9.21) provient d'Atherton (Australie). Les origines précises des deux géniteurs sont représentées à la figure 4. Au total 201 descendants pleins-frères sont disponibles pour notre étude. Les arbres ont été plantés en 1993 à une densité de 3m × 5m.

1.2. L'extraction d'ADN

L'extraction a été réalisée selon le protocole de Saghai-Marroof *et al.*, (1984). Chez les eucalyptus, la présence d'un grand nombre de métabolites secondaires et d'hydrates de carbone rend l'extraction d'ADN difficile. De plus au niveau de la qualité de l'ADN, seules les feuilles d'un vert tendre, prélevées sur des pousses terminales à proximité des feuilles juvéniles, donnent des résultats satisfaisants. L'ADN est extrait à partir de feuilles séchées à l'air libre et il est quantifié à l'aide d'un fluorimètre (TD-700) puis dilué dans l'eau stérile à 3ng/μl. Cette solution de travail est conservée à -20°C.

2. Gènes étudiés

Huit gènes de fonction connue ont été étudiés et parmi eux trois gènes de régulation appartenant à la famille des facteurs de transcription MYB clonés chez *E. gunnii* (MYB1, MYB2, MYB4). Il a été montré que ces gènes régulent différentes branches du métabolisme des flavonoïdes chez les plantes (Mol *et al.*, 1998) et sont supposés jouer un rôle important dans la régulation du métabolisme générale des phénylpropanoïdes (Tamagnone *et al.*, 1998). Trois gènes de structure de la voie de biosynthèse des lignines ont aussi été étudiés : C4H (cinnamate 4-hydroxylase), F5H (fêrulate 5-hydroxylase), LAC (laccase).

➤ Le gène C4H code pour une enzyme qui catalyse l'hydroxylation des acides cinnamiques. Des tabacs transgéniques sous-exprimant la C4H montrent une

réduction substantielle des teneurs en lignines et une diminution du ratio S/G (Sewalt *et al.*, 1997).

➤ Le gène F5H catalyse l'oxydation des acides féruliques, réaction dépendante du cytochrome P₄₅₀. Un mutant d'*Arabidopsis thaliana* (mutant *fah1*) est incapable d'accumuler les unités S des lignines. Or, le locus *fah1* code pour une F5H et sachant que les lignines riches en unités S sont plus facilement extractibles (au niveau du procédé de fabrication de la pâte à papier) la F5H apparaît comme une enzyme importante pour le contrôle du degré de méthylation des unités monomériques des lignines (Boudet, 1997).

➤ LAC : les laccases sont impliquées dans la polymérisation oxydative des monolignols au même titre que les peroxydases qui ont longtemps été considérées comme les seules enzymes intervenant dans cette dernière étape. Chez deux espèces, *Liriodendron* (Dean *et al.*, 1998) et le peuplier, l'activité des laccases a été diminuée à l'aide de la stratégie antisens. Dans les deux cas, il n'y pas eu de différences significatives observées au niveau des teneurs en lignines.

L'étude s'est également portée sur les gènes LEAFY (ELF), isocitrate déshydrogénase (EglCDH). LEAFY est impliqué dans l'initiation et le développement floral chez les végétaux. Chez *E. globulus*, deux gènes, ELF1 et ELF2, ont été clonés et possèdent des homologues de séquences avec LEAFY d'*Arabidopsis* (Southerton *et al.*, 1998). EglCDH catalyse la conversion de l'isocitrate en 2-oxoglutarate au niveau du cycle de Krebs.

3. PCR dirigée pour chaque gène fonctionnel

3.1. Choix des amorces

La définition des amorces a été effectuée à partir des séquences d'ADNc d'*E. gunnii* pour les trois facteurs MYB, et sur des séquences d'ADNc d'*E. globulus* pour les gènes ELF et EglCDH à l'aide du logiciel OLIGO.4 (Rychlick *et al.*, 1990).

En ce qui concerne les gènes F5H, C4H et LAC, les amorces ont été réalisées en effectuant un alignement multiple avec des ADNc et des séquences d'ADN génomique disponibles dans les banques de gènes. Les amorces ont été ciblées dans les exons et elles comportent des nucléotides dégénérés.

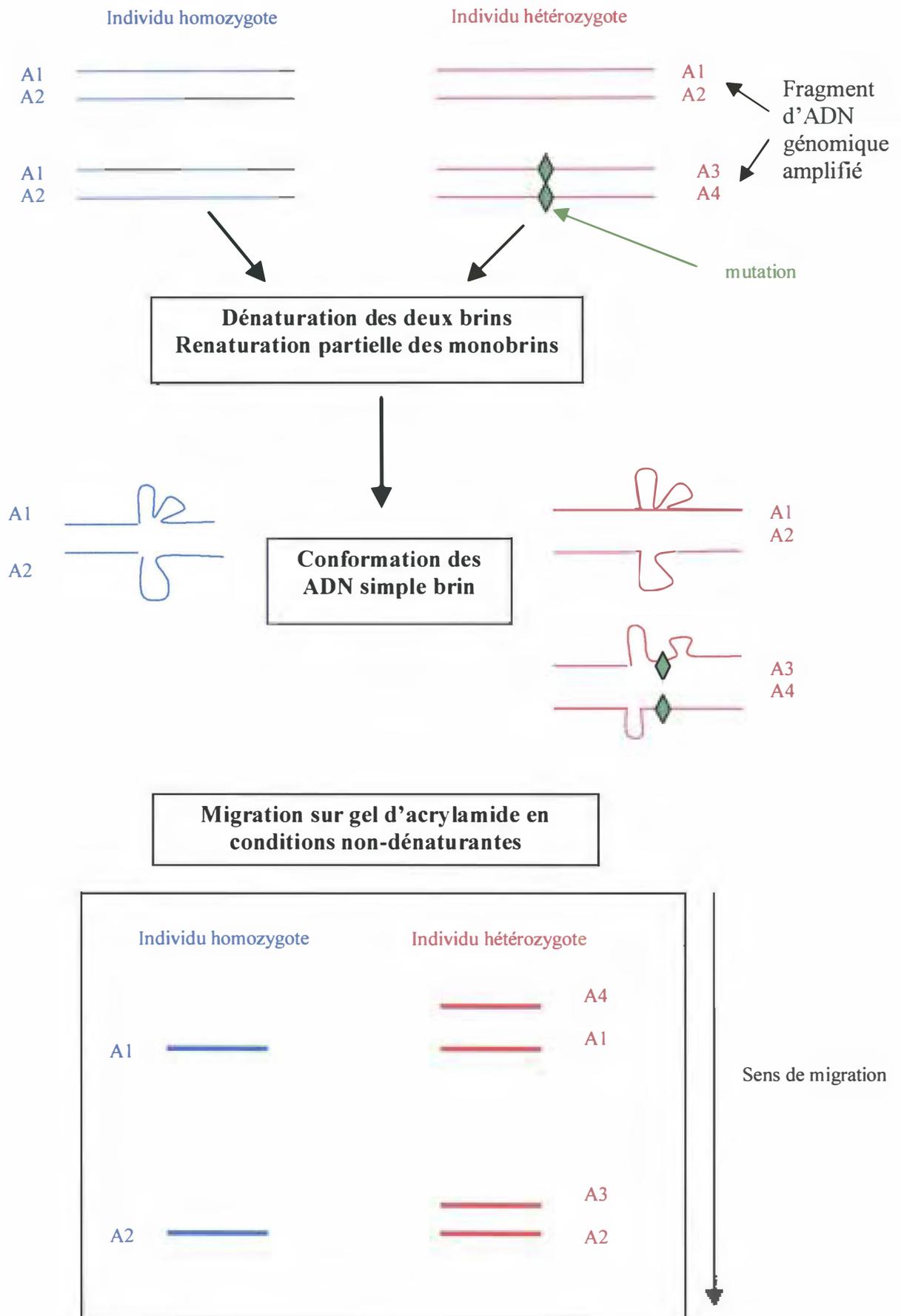


Figure 5 : Principe de la technique SSCP.



3.2. Conditions d'amplification PCR

Pour chaque amorce la température d'hybridation a été optimisée à l'aide d'un Robocycleur gradient 96 (Stratagene). Pour chaque réaction PCR le mélange réactionnel (volume total de 20 µl) contient 1µl de chaque primer, 67 mM Tris-HCl, pH 7,5, 2mM MgCl₂, 1ng de BSA, 0.2% b-mercaptoéthanol, 16 mM acétate d'ammonium, 200 µM de chaque dNTP et 1 U de Taq (Gibco BRL). A ce mélange réactionnel est ajouté 5 µl d'ADN en concentration finale de 15ng. Les conditions d'amplification sont les suivantes : une dénaturation préliminaire (4 min à 94°C), suivie par 30 cycles comprenant une dénaturation (40 à 94°C), une hybridation (45 s à une température d'hybridation spécifique à chaque couple d'amorces) et une élongation (1 min à 72°C), et une élongation finale (10 min à 72°C). Les produits PCR sont séparés sur gel d'agarose à 1,5% et colorés avec le bromure d'éthidium.

4. Séquençage des produits PCR

Le séquençage a été réalisé par Genome Express (S.A. Grenoble). Le produit PCR chez les deux parents a été séquencé à partir de l'extrémité 5' et de l'extrémité 3'. Les séquences sont ensuite comparées avec des séquences nucléiques existant dans les banques de gènes en utilisant le programme BLASTn (Blast version 2.0) (Altschul *et al.*, 1990)

5. Révélation du polymorphisme par SSCP

5.1. Principe

La technique SSCP (Single-Strand Conformation Polymorphism), développée par Orita *et al.*, (1989) est une méthode rapide et simple pour détecter du polymorphisme au sein d'un fragment d'ADN. Elle est basée sur le comportement électrophorétique d'un fragment d'ADN simple brin dans un gel d'acrylamide non dénaturant (Figure 5). En conditions non dénaturantes, l'ADN double brin ne possède qu'une seule conformation car la structure est stabilisée par des interactions intramoléculaires. Dans le cadre de la SSCP, l'ADN est dénaturé entraînant un repliement des simples brins différents selon la séquence. Ces différences de repliement des brins entraînent des différences de migration.

En général, chez un individu homozygote, on observe deux bandes, car les deux molécules d'ADN simple brin complémentaires prennent des conformations secondaires différentes. Chez un individu hétérozygote, quatre conformations secondaires sont observées. Dans quelques cas, des bandes surnuméraires peuvent apparaître en raison de l'existence de plusieurs conformations semi-stables pour un même brin (de Vienne, 1998).

5.2. Conditions SSCP

Le protocole SSCP utilisé est celui décrit par Bodénès *et al.*, (1996) . Aux 5 μ l de produit PCR, on ajoute 10 μ l d'une solution de bleu dénaturant, contenant 95% de formamide, 10 mM de NaOH, 0.05% de xylene cyanol et 0.05% de bleu de bromophénol. Les échantillons sont chauffés à 94°C pendant 4 minutes puis immédiatement refroidis dans la glace. Ils sont alors déposés sur des gels d'acrylamide non dénaturant (0.75mm \times 16 cm \times 18 cm), pour une migration dans un tampon TBE 0,6 X (annexe 1). La concentration du gel, le voltage ainsi que la température influencent la migration des fragments. Les conditions optimales SSCP de migration sont donc à rechercher pour chaque fragment.

Les gels sont révélés selon la technique de Bassam *et al.* (1991). Ils sont fixés dans une solution d'acide acétique à 10% pendant 20 minutes puis, lavés dans trois bains d'eau ultra pure. Les gels sont ensuite placés dans une solution de coloration contenant 0,1% de nitrate d'argent et 0,0056% de formaldéhyde pendant 30 minutes. Un rinçage de 20 secondes dans de l'eau ultra pure est effectué avant la révélation faite dans un bain contenant 30 g/l de carbonate de sodium et 2 mg/l de thiosulfate de sodium. La réaction de révélation est stoppée avec l'acide acétique à 10%. Enfin, les gels sont rincés dans l'eau ultra pure et séchés entre du papier cellophane.

6. Cartographie génétique

Une carte génétique est une représentation des liaisons génétiques mises en évidence entre différents locus polymorphes (les marqueurs). Les distances entre marqueurs sont exprimées en centimorgan (cM), distance qui dépend de la fréquence de recombinaison entre locus dans la population en ségrégation. La carte génétique est dite saturée quand tout point du génome est génétiquement lié à au

moins un marqueur, on observe alors autant de groupes de liaison que de chromosomes.

Chez les végétaux, les conditions sont en général favorables pour la construction de cartes génétiques : des croisements contrôlés peuvent être réalisés et des descendance de grand effectif peuvent être manipulées.

Généralement les descendance utilisées pour la cartographie sont des F2, des populations de lignées recombinantes (LR), des haploïdes doublés.... Dans notre cas, nous avons à faire à une espèce pérenne à cycle long où la constitution de telles populations peut prendre un temps considérable. La population utilisée pour la cartographie génétique des arbres forestiers est principalement une descendance F1 de pleins frères ou de demi-frères.

Dans notre cas, la mise en évidence de polymorphisme pour les gènes étudiés est de type codominant. Afin d'intégrer les gènes dans les cartes génétiques construites à partir de marqueurs RAPD, ces marqueurs sont codés comme des marqueurs dominants. Chez chaque parent on va fixer un allèle et les individus possédant l'allèle seront notés 1, ceux pour lesquels l'allèle est absent seront notés 0.

La ségrégation mendélienne de chaque marqueur a été vérifiée par un test du χ^2 , en comparant les différentes classes génotypiques dans les descendance à celles attendues sous l'hypothèse de ségrégation mendélienne.

Le logiciel MAPMAKER pour Macintosh (Lander et al, 1987) a été utilisé pour positionner les gènes sur les cartes génétiques parentales. Le test de liaison utilisé par le logiciel est un rapport de vraisemblance appelé Lod score (Log10 of the odds ratio). On teste en fait le maximum de vraisemblance des données θ (valeur observées) par rapport au maximum de vraisemblance des données $\theta = 0,5$. Donc, ce rapport permet de tester la force de liaison entre marqueurs. Le Lod minimal choisi est de 3 et le taux de recombinaison θ maximal est de 0,4. Ce seuil de Lod correspond à une liaison entre marqueurs alors 1000 fois supérieure à celle de l'indépendance entre marqueurs. Pour ordonner les marqueurs entre eux, la fonction "RIPPLE", test 3 points de MAPMAKER a été utilisés. Elle permet de permuter les triplets de marqueurs adjacents et, seul est retenu l'ordre du triplet dont sa probabilité est 10000 fois plus grande que tout autre ordre relatif. Enfin, les pourcentages de

Tableau 1 : Définition des variables étudiées. La hauteur totale, la circonférence, le pylodyn, la DRLM ont été mesurés sur le terrain.

Code	Variables	Définition des caractères
HT	Hauteur totale (cm)	X1
C	Circonférence (cm)	X2
V	Volume (dm ³)	$X3 = (1/12\pi * (X2)^2 * (X1 - 130) + 1/4 \pi * (X2)^2 * 130) / 1000$
DM	Décroissance métrique (cm/m)	$X4 = (100 * X2) / (X1 - 130)$
Pil	Pénétration sous écorce du pilodyn	
DRLM	Déformation Résiduelle Longitudinalement de Maturation	
IE1	Index d'Eclatement moyen à un jour	X5= longueur des fentes/diamètre
IES	Index d'Eclatement moyen une semaine après l'abattage	X6= idem X5
IE2m	Index d'Eclatement moyen deux mois après l'abattage	X7= idem X5
Tx ext		X8= (mb- mhe)/mhe ^(a)
Klason	Teneur en lignines totale	
G	Teneur en unités G	
S	Teneurs en unités S	
S/G	Rapport entre les unités S et les unités G	

(a) : mb : masse de bois

mhe : masse de bois hors extrait (résidus pariétaux)

recombinaison sont convertis en distances génétiques à l'aide de la fonction de cartographie de KOSAMBI (Kosambi, 1944).

7. Caractères mesurés

La famille a été abattue en 1998 et différents types de caractères ont été mesurés à 62 mois, âge d'exploitation des eucalyptus au Congo. Pour cette étude 10 variables ont été étudiées (Tableau 1) concernant :

- ♦ Les caractères de vigueur : la hauteur totale, la circonférence à 1 m30, le volume calculé à l'aide de la formule du tronc de cône.

- ♦ Un caractère de forme, le coefficient de décroissance métrique (cm/m) qui correspond à un rapport entre une hauteur totale et une circonférence à 1 m30.

- ♦ La densité du bois mesurée indirectement par la mesure de pénétration du pilodyn. Cet appareil placé contre le tronc de l'arbre libère brusquement dans le bois une aiguille maintenue sous la tension d'un ressort. Plus l'aiguille pénètre dans le tronc, plus la densité du bois est faible. Cette mesure a été réalisée à 1 m30 sur l'arbre sous écorce (Pil), c'est-à-dire que l'appareil est posé à même le tronc après avoir retiré l'écorce.

- ♦ Les propriétés mécaniques du bois :
 - Les mesures de déformation résiduelle longitudinale de maturation (DRLM) ont été réalisées à 1 m30. A la fin de sa différenciation le bois tend à se déformer, lors de la formation de la paroi secondaire. Cette déformation aussi appelée DRLM permet d'expliquer les efforts internes importants, généralement nommés "contraintes de croissance" supportées par le bois des arbres et directement liées à la structure chimique du bois. Ces mesures de DRLM ont été réalisées à l'aide de la méthode du trou unique développé au CIRAD-Forêt (Mariaux, 1982).
 - L'index d'éclatement correspond au rapport entre la longueur totale des fentes observées sur le billon et le diamètre du billon. Il a été calculé à différentes dates après l'abattage : 1 jour, une semaine et deux mois.

♦ Les propriétés chimiques du bois :

- Le taux d'extrait qui correspond aux résidus pariétaux est obtenu à partir d'une rondelle de bois prélevée à mi-hauteur. Une fois l'écorce enlevée, le bois est réduit en poudre. Puis des extractions successives (toluène/éthanol 1/1 ; éthanol ; eau ; acétone) permettent d'éliminer les phénols solubles, les sucres et les autres composés extractibles non liés aux parois.
- La teneur en lignines déterminée par la méthode gravimétrique de Klason. Elle est réalisée à partir des résidus pariétaux. Suite à une hydrolyse sulfurique, tous les résidus pariétaux solubles sont éliminés, il ne reste plus que les lignines insolubles (exprimé en % du résidu pariétal).
- La composition monomérique des lignines a été obtenue par une méthode chimique, la thioacidolyse (Lapierre *et al.*, 1986). Cette technique permet de séparer les monomères syringyl (S) et guaicyl (G) des fractions non condensées des lignines par coupure spécifique de la liaison β -O-4 alkyl aryl éther. Les pourcentages d'unités S et d'unités G sont déterminés par chromatographie en phase gazeuse. Notons que les teneurs en S et G sont établies à partir des lignines non condensées et ne constituent donc pas la fréquence absolue de ces monomères dans les lignines totales.

8. La détection de QTL.

La détection de QTL consiste à mettre en évidence une liaison statistique entre le polymorphisme d'un marqueur (moléculaire, izozymes...) et la variation d'un caractère quantitatif. Pour tester si cette liaison est significative deux méthodes sont utilisées.

8.1. Analyse de variance (ANOVA)

Il s'agit d'une analyse marqueur par marqueur (Soller *et al.*, 1976). Les allèles parentaux ségrégeant dans la descendance, les descendants hybrides peuvent être regroupés en différentes classes en fonction des allèles parentaux qu'ils possèdent. Ensuite une simple analyse de moyenne par un test F permettra de dire si les moyennes sont significativement différentes. Si c'est le cas il est fort probable que le

marqueur soit lié à une région chromosomique impliquée dans la variation quantitative du caractère.

8.2. Maximum de vraisemblance (Interval Mapping)

Cette méthode nécessite une carte génétique. Elle a été proposée par Lander et Botstein (1989) et suppose que deux marqueurs sont liés avec un taux de recombinaison défini, et qu'un QTL au plus est situé entre ces marqueurs. Cette méthode teste, en tout point de la carte, la présence d'un QTL dans un intervalle entre deux marqueurs. La puissance de détection est ainsi augmentée et précise la localisation des QTL. Le test de liaison est réalisé par un rapport de vraisemblance appelé "LOD score". Au delà d'un LOD seuil on considère qu'un QTL est détecté. Nous avons choisi un LOD seuil faible, 1,7 pour éviter de rejeter des vrais QTL. Pour vérifier si des faux QTL avaient été mis en évidence avec le LOD seuil de 1,7 des tests de permutation ont été réalisés.

L'IM se fait à l'aide des logiciels MAPMAKER/QTL (Lincoln *et al.*, 1990) et QGENE (Nelson, 1997) qui sont basés sur des algorithmes de calculs différents. Le logiciel MAPMAKER/QTL permet d'estimer le pourcentage de variation phénotypique expliqués par chaque QTL. Lorsque plusieurs QTL sont détectés pour un même caractère, une estimation du pourcentage de variation expliquée par l'action conjointe des QTL est également obtenue. QGENE a permis de déterminer les probabilités associées à la valeur du test F ("p-value") pour chaque QTL. L'effet de substitution allélique au marqueur le plus proche du QTL est calculé. Il correspond à la différence de la valeur moyenne du caractère exprimée en unité d'écart type. C'est une estimation de la part de variance phénotypique expliquée par les QTL plus parlante pour le sélectionneur.

Résultats -Discussion

Rapport-Gratuit.com

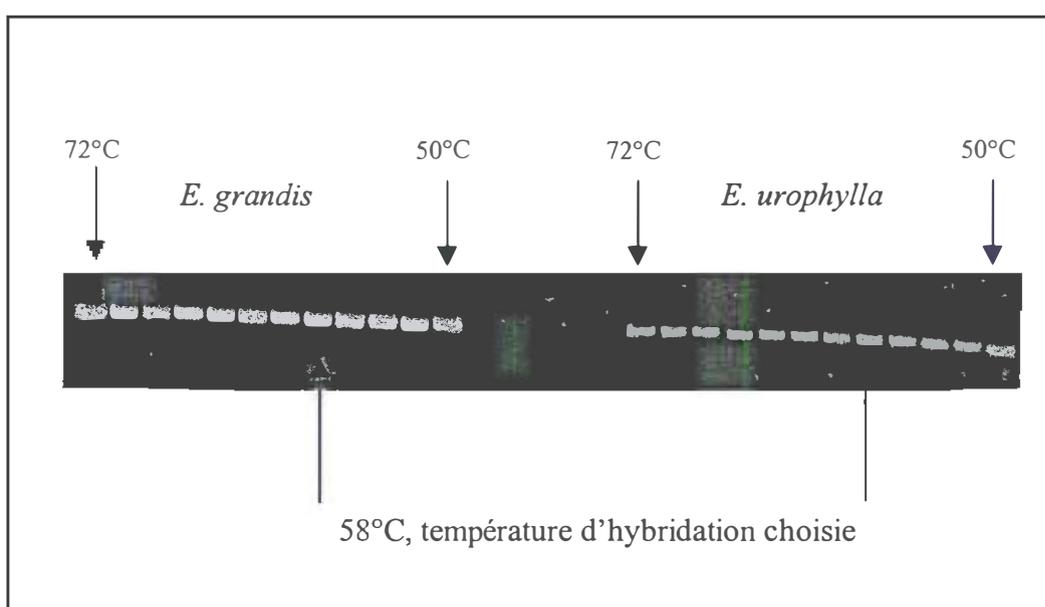
Tableau 2 : Conditions d'amplifications PCR pour les différents fragments de gènes étudiés

Gènes	Numéro d'accèsion des séquences utilisées pour l'alignement (a)	Séquences des amorces (5' vers 3') (b)	Fragment attendu	T°C hybridation	Fragment observé
MYB 1	design sur ADNc d' <i>E. gunnii</i>	Upper :ACA CGC ACA TAA GGA GGA AGC Lower :CAA GCC CAA GAA GTC ATA CCC	411 pb	58°C	400 pb
MYB 2	design sur ADNc d' <i>E. gunnii</i>	Upper :TCC AAT CCA CAA GAC ATA GC Lower: GTG GGG GAA CAG AAA ACT CG	297 pb	58°C	300 pb
MYB 4	design sur ADNcd' <i>E. gunnii</i>	Upper : CAC CAT TTC TTT TGC TGC TAA Lower : AAT CTC TGC CTC ATC TCT TCG	397 pb	58°C	400 pb
F5H	AFO68574, AJ249093, AJ010324	Upper GCGATCAGCTACCTCACG Lower <i>CCAAACATACRTCCAT</i>	608 pb sans intron 1095 pb avec intron(129 et 346 pb)	45 à 51°C	Plusieurs fragments
C4H	L38898, U71080, U119922, X92437 D82812, Z32563, L07634	upper <i>GGHCARGAYATGGTITTCAC</i> lower <i>GCRGCVACATTGATGTTTC</i>	1264 pb avec intron de 682 pb	53°C	700 pb
LAC	U43542, Y13771, U12757	Upper <i>ATCAAAYGCWRYYTYAATGA</i> Lower <i>GYMBYTCIARATGRCAATGCCA</i>	886-901pb sans intron	44°C	Plusieurs fragments
ELF	AFO56550, AFO34806	Upper GCGAGCGGTACGGCATCAAGG Lower GGTCGTAGAGATGGAAGAGGT	475pb avec intron de 62 pb	65°C	400-500 pb
EGICDH	Y16126, U80912, X77944, AB000261	Upper ATGAGTCAATCCGTTCTTT Lower TAGCCGATAAGCCAGGATA	220pb	50°C	400pb

(a) : Les séquences correspondant aux numéros d'accèsion sont indiquées en annexe 2.

(b) : Les séquences nucléiques en italique comportent des nucléotides dégénérés dont le code est le suivant : **H**=A+T+C, **Y**=C+T, **M**=A+c, **B**=T+C+G, **J**=déoxynosine, **N**=A+C+T+G, **R**=A+G, **K**=T+G, **S**=C+G, **W**=A+T, **D**=A+T+G, **V**=A+C+G.

Figure 6 : Gradient de température réalisé pour la recherche de la température optimale d'hybridation pour le MYB2. Les températures de 50 à 72°C ne montrent pas d'amplifications non spécifiques.



Partie I : La cartographie des gènes

1. Optimisation des conditions d'amplification PCR

Le Tableau 2 résume les différentes conditions d'amplification des fragments de gènes étudiés. Les températures optimales d'hybridation pour tous les couples d'amorces utilisés varient de 50 à 65°C et la taille des fragments observés est comprise entre 300 et 700 pb. Aux températures optimales un seul fragment d'ADN a été amplifié pour MYB1, MYB2, MYB4, ELF et C4H pour les deux parents et la descendance. La figure 6 montre le résultat de l'amplification du fragment MYB2 chez les deux géniteurs pour des températures d'hybridation de l'amorce comprises entre 72 et 50°C. Dans le cas du gène *EgLCDH* l'amplification d'un fragment a été possible uniquement chez le géniteur *E. urophylla*.

Les résultats des PCR dirigées n'ont pas été satisfaisants pour F5H et LAC. Plusieurs fragments ont été obtenus et l'augmentation de la température n'a pas augmenté la spécificité de l'amplification. Le degré de dégénérescence des amorces, utilisées pour ces deux gènes était, sans doute trop élevé. Nous n'avons pu réaliser d'autres amorces car trop peu de séquences étaient disponibles dans les banques de biomolécules pour effectuer un design différent du premier. Les travaux ont été temporairement abandonnés.

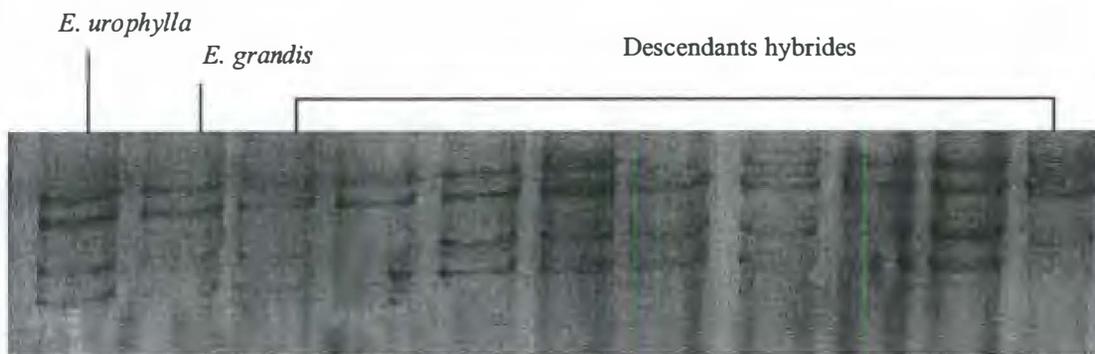
2. Résultats du séquençage

Le programme BLASTn (Blast version 2.0) (Altschul *et al.*, 1990), avec les paramètres par défaut proposés par NCBI, a été utilisé pour comparer les séquences des fragments avec celles des gènes disponibles dans les banques de biomolécules (GENEBANK, EMBL,). Le séquençage des produits PCR des deux parents pour MYB1, MYB2, MYB4 et ELF confirme la nature des fragments amplifiés. Pour ELF une homologie de séquence de 97% a été mise en évidence avec le gène LEAFY

Tableau 3 : Conditions optimales obtenues en SSCP pour la révélation de polymorphisme au sein des fragments de gènes.

Gènes	Voltage SSCP	Température de migration SSCP	durée de migration	Acrylamide (ratio 37:1)
MYB1	130V	15°C	15 h00	10%
MYB2	130V	15°C	15 h00	10%
MYB4	130V	15°C	15 h00	10%
C4H	130V	15°C	16h00	12%
ELF	130V	25°C	15h30	8%

A : On ne distingue pas de polymorphisme dans la descendance



B : Deux classes de descendants se différenciant par un des allèles d '*E. grandis*

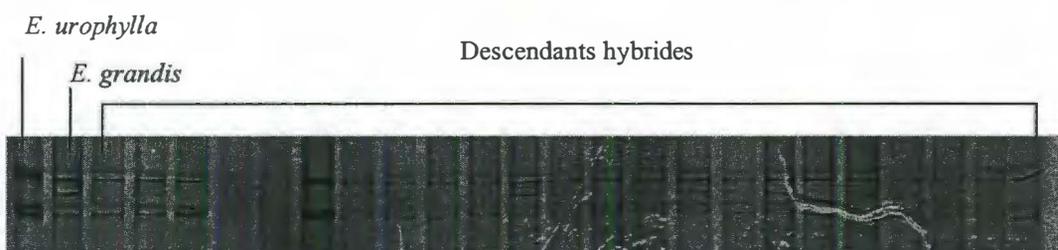


Figure 7 : Optimisation des conditions de migration SSCP pour la révélation de polymorphisme au sein du fragment de gène C4H amplifié par PCR. Le voltage, le temps et la température de migration et le type d 'acrylamide sont fournis. Deux conditions de migration sont testées : migration à température ambiante pour un gel d 'acrylamide à 10% (A) et une autre migration à 15°C et à 12% d 'acrylamide (B). Le voltage et le temps de migration sont identique dans chacun des cas (130V et 16h).

d'*E. globulus*. Pour les gènes C4H et EglCDH les résultats de séquençage n'ont pas encore été obtenus.

3. Optimisation des conditions SSCP

Différentes conditions SSCP ont été testées pour optimiser la révélation d'un polymorphisme au sein des fragments de gènes étudiés. Plusieurs paramètres ont été testés : la température de migration, la présence ou l'absence de glycérol, le temps de migration et le type (MDE ou acrylamide ratio 37:1) et la concentration en acrylamide. D'autres travaux ont montré une influence de ces facteurs sur la mise en évidence de polymorphisme (Orita *et al.*, 1989 ; Spinardi *et al.*, 1991 ; Glavac et Dean, 1993). Le tableau 3 résume les conditions optimales trouvées pour chaque gène.

Pour la C4H, le changement de température de migration et le pourcentage d'acrylamide ont permis de mettre en évidence du polymorphisme dans la descendance hybride dû aux allèles d'*E. grandis* (Figure 7). De plus, les gels ont été chargés à sec c'est-à-dire sans tampon de migration car il a été montré que ceci influençait la netteté des profils (Gion *et al.*, 2000). Enfin, nous avons montré que sous des conditions de migration identiques, les profils SSCP sont reproductibles, résultat également obtenu par Bodénès *et al.*, (1996).

Pour EglCDH, nous avons testé différentes conditions mais sans succès, aucun polymorphisme n'a été détecté au niveau du fragment amplifié. Il conviendrait de changer de couple d'amorces PCR, permettant d'amplifier une autre partie du gène.

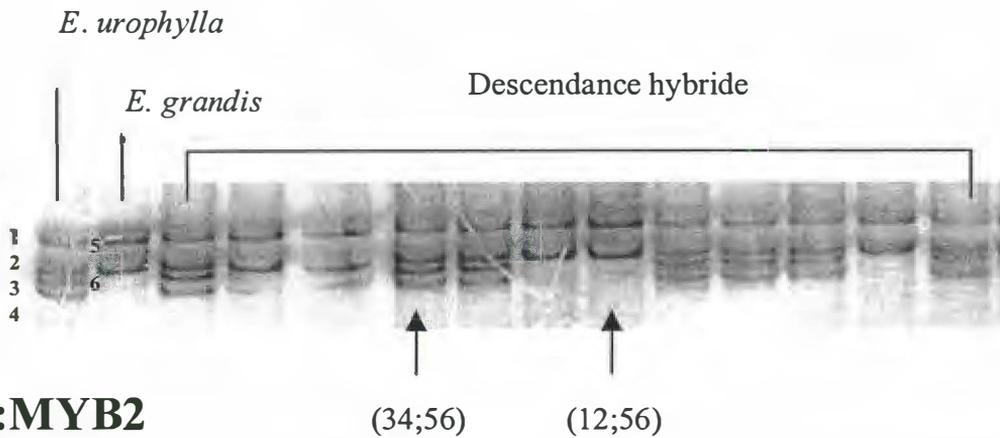
La SSCP est une technique simple, sensible (une mutation ponctuelle est visible) et peu coûteuse pour détecter du polymorphisme à l'intérieur de fragments PCR de tailles différentes et ainsi « visualiser » la ségrégation des allèles parentaux. Elle apparaît très appropriée pour cartographier des gènes de fonction connue (Plomion *et al.*, 1999). De plus, la SSCP génère des fragments de type codominant. Ainsi on peut envisager de positionner un gène chez les deux parents et par la suite confirmer l'homologie entre les groupes de liaison ou bien étudier la synténie.

Tableau 4 : Interprétation allélique des profils SSCP pour les gènes MYB2, MYB1, MYB4, C4H, et C4H

Gènes	Génotypes parentaux E. urophylla / E. grandis	Génotypes des hybrides	Nombre d'individus pour chaque génotype	χ^2	ddl ^(b)
MYB2	(12;34) / (34;56)	(12;34) / (12;56) / (34;56) / (34;34)	45 / 42 / 46 / 58	1,5	3
MYB1	(13;24) / (56;56)	(13;56) / (24;56)	92 / 103	0,62	1
MYB4	(12;34) / (56;56)	(12;56) / (34;56)	96 / 98	0,02	1
C4H	(12;12) / (34;35)	(12;34) / (12;35)	91 / 92	0,005	1

(a) : valeurs seuil du χ^2 pour $p = 0.05$: 1ddl : 3,84 ; 3ddl : 7,81.

A : MYB4



B : MYB2

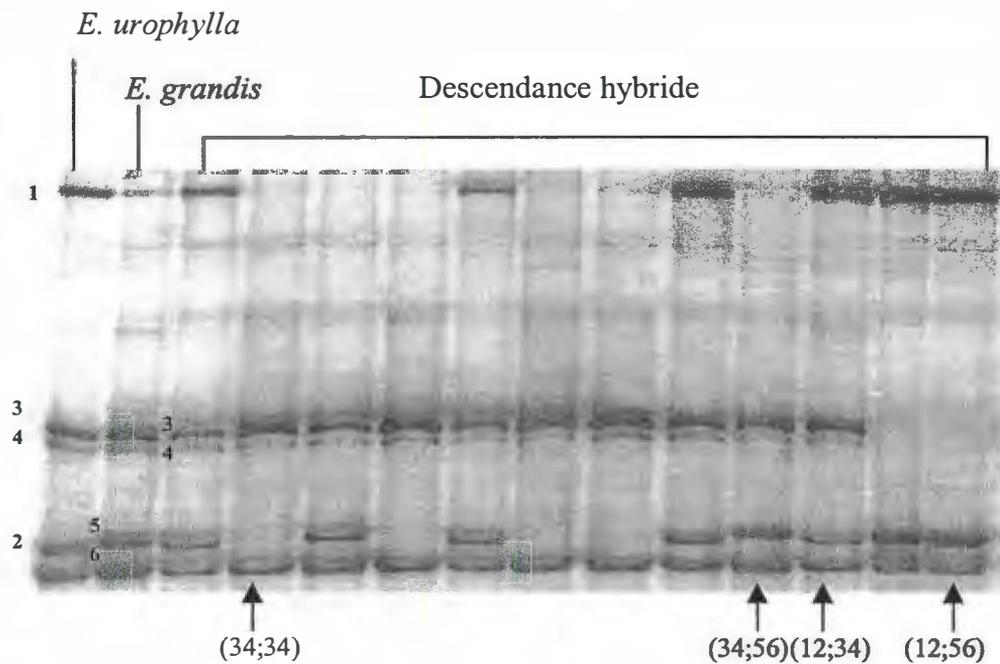


Figure 8: Profils SSCP illustrant le polymorphisme du MYB4 et du MYB2 dans le croisement interspécifique *E. urophylla***E. grandis*. Les deux premiers puits contiennent l'ADN amplifié du parent femelle (*E. urophylla*) et du parent mâle (*E. grandis*), respectivement. Les autres puits contiennent l'ADN amplifié d'une partie des descendants hybrides. A. Exemple d'un profil SSCP du MYB4. Les génotypes *E. urophylla* et *E. grandis* sont respectivement (12;34) et (56;56). La descendance F1 présente deux classes génotypiques : (12;56) et (34;56). B. MYB2. Les génotypes d'*E. urophylla* et d'*E. grandis* sont (12;34) et (34;56) respectivement. Quatre classes génotypique sont observées dans la descendance : (12;34), (34;56), (12;56) et (34;34).

4. Héritabilité mendélienne des fragments SSCP

Sur les six gènes testés en SSCP, quatre sont hétérozygotes (MYB1, MYB4, C4H, ELF) chez un parent, un seul (MYB2) est hétérozygote chez les deux parents, et un gène (EglCDH) n'a pas montré de polymorphisme dans la descendance étudiée.

Les proportions des différents génotypes des hybrides sont en accord avec les proportions d'une ségrégation mendélienne 1 :1 :1 :1 ou 1 :1 (Tableau 4).

Pour le gène MYB4, l'observation des profils SSCP de la descendance hybride a révélé deux classes de descendants se différenciant par rapport aux allèles d'*E. urophylla* (Figure 8A). Ainsi des descendants sont caractérisés par les bandes (12;56) et d'autres par les bandes (34;56). Les profils SSCP des géniteurs donnent les génotypes (12;34) et (56;56) respectivement pour les géniteurs *E. urophylla* et *E. grandis*. Un second exemple est celui de la ségrégation observée pour le gène MYB2 (Figure 8B). Chez les descendants, quatre classes génotypiques sont observées correspondant aux combinaisons alléliques suivantes : (12;34), (12;56), (34;56) et (34;34). Pour ce gène, *E. urophylla* et *E. grandis* sont hétérozygotes avec comme génotype respectif : (12;34) et (34;56). La ségrégation du gène ELF dans la descendance hybride est plus complexe. Cette ségrégation suggère l'existence de deux gènes : ELF1 et ELF2. L'analyse des profils en vue de l'interprétation allélique n'a pas permis de définir correctement les allèles parentaux. L'existence de deux gènes ELF est confortée par les résultats obtenus par Southerton *et al.* (1998) qui ont cloné deux gènes leafy chez *E. globulus*.

Sur l'ensemble des gènes de la voie de biosynthèse des lignines étudiés en SSCP, il semble que le clone 14.144 soit légèrement plus hétérozygote que le clone 9.21. En effet, six gènes sur sept étudiés (CCR, CAD, COMT, CCoAOMT, PAL, 4CL) sont hétérozygotes chez *E. urophylla* contre quatre (CCR, CAD, CCoAOMT, C4H) chez *E. grandis*.

Figure 9 : Séquence partielle d'ADN de MYB4 pour *E. urophylla* (U) et *E. grandis* (G). Dans la séquence *E. urophylla* six indéterminations en position 137, 300, 301, 354, 357, 358 et deux insertions/délétions d'une base en position 359, 394 sont décelées indiquant son hétérozygotie. Seule la séquence d'*E. urophylla* est donnée, la séquence d'*E. grandis* est identique sauf pour au niveau des nucléotides notés en rouge.

(1) : N=T ou N ; (2) : N=A ou T ; (3) : N=G ou A ; (4) : N=G ou - ; (5) : N=T ou G ; (6) : N=A ou C .

```

1                               50
U   ATTTCTTTTG CTGCTAATTC TAAGGAGATC AAAGAGATGA AGAACAACGC
G
51                               100
U   AGAGCTCAAT TTCATGTGCA ACTTAGAAGA GTCGGCAGAC GTGGCATCGT
G
101                              (1) 150
U   CGGCTCGAGA AAGGTGTCCT GACCTGAATC TCGAGCTN(1)GG AATCAGCCCT
G                               C
151                              200
U   CCTTCTCATC AACTGCATCA GCCTGAGCCA CTCTTGAGAT TCACTGGTAG
G
201                              250
U   GAAAAGTGAT TTGTGTCTGG AGTGTAATTT GGGGTTGAAA AATAGCCAAA
G
251                              (2) 300
U   ATTGCAGATG CAGTGTGGG GTGATCGAGA GTGAAACTAG TGTTGGGTAN(2)
G                               T
301 (3)                          350
U   NACTTCTTGG GCTTGAAGGC AAGTGTTTTG GATTATAGGA GCTGAATTTT
G   G
351 (4) (5,6)                    400
U   GGTNA(4)ANN-G ATGGATAATT GTGCAGCGAA GAGATGAGGC AGAG(5,6)GATTAC
G   G GAA

```

Figure 10 : Séquence partielle d'ADN du MYB2 pour *E. urophylla* (U) et *E. grandis* (G) indiquant respectivement la position des quatre et trois indéterminations. Ceci mettant en évidence l'hétérozygotie des deux géniteurs à ce locus.

(1) : N=A ou C ; (2) : N= C ou G ; (3) : N=G ou - ; (4) : N=G ou A ; (5) : N=T ou G ; (6) : N=T ou C.

```

1                               50
U   TCCAATCCAC AAGACATAGC TGCAGGA(1)AATT ATGCCCAGTT TTCATGCACA
G                               N(1)
51                              (2) 100
U   ATATGACGTC TTAG(2)N(2)TACAT GCATGGACTC GT(3)N(3)CTCCCGC ACC(4)N(4)TTCCCC
G                               C G
101                              (5) 150
U   CCGATGGACA ATATATCTGC GCCAAACCAN(5) TTCGATCCAT TCCC(6)CACACT
G                               N N(6)
151                              200
U   CAACAACCGC TGTGACACAT GGAAGGAGT CGGGTTTTTC ACCTTCCCTT
G
201                              250
U   CAGGCATTGC ACCGGTCAGT ATGGGAGATG ATCAAGTTA TTTGAATTTG
G
251                              297
U   GAGCATGCAA AAGTGGGACT TTTAGGAAGC GAGTTTTCTG TT(6)CCCCA
G

```

Tableau 5 : Résultats du séquençage pour tous les gènes avec ou sans indétermination dans les séquences nucléiques.

Gènes	Résultats du séquençage		Génotypes parentaux déduit des profils SSCP	
	<i>E. urophylla</i> / <i>E. grandis</i>		<i>E. urophylla</i> / <i>E. grandis</i>	
MYB2	i	i	het.	het.
MYB1	i	ni	het.	hom
MYB4	i	ni	het.	hom.
ELF	i	i	het	het.

i : nucléotide indéterminé dans la séquence

ni : pas de nucléotides indéterminés dans la séquence

het. : hétérozygote au locus amplifié

hom. : homozygote au locus amplifié

5. Causes moléculaires du polymorphisme SSCP

Les différences de migration observées en SSCP sont dues à des différences de séquences. L'eucalyptus étant diploïde, les produits PCR sont un mélange de deux allèles. Pour chaque gène une comparaison entre les séquences des deux géniteurs a été effectuée. Au niveau de la séquence du gène MYB4 d'*E. urophylla*, on observe six indéterminations et deux insertions/délétions d'une base (% de changement nucléotidique de 8/400). Ceci confirme la présence de deux allèles différents chez *E. urophylla* alors qu'*E. grandis* serait homozygote pour ce fragment (aucune indétermination dans la séquence) (Figure 9). Plusieurs indéterminations sont présentes dans les séquences nucléiques du MYB2 d'*E. urophylla* et *E. grandis*. Les deux géniteurs seraient donc hétérozygotes à ce locus (Figure 10).

Ainsi, généralement la ségrégation observée en SSCP est confirmée par les résultats du séquençage (Tableau 5). Pour les quatre gènes séquencés, on distingue deux cas :

- La séquence nucléique comporte des nucléotides indéterminés pour lesquels deux nucléotides sont possibles à une même position. L'hétérozygotie supposée du fragment amplifié est donc confirmée. C'est le cas pour le MYB1, MYB4, ELF1 et ELF2 chez *E. urophylla* ; et pour le MYB2 chez les deux géniteurs.
- Le fragment amplifié ne comporte pas d'indéterminations. L'individu est soit homozygote (MYB1 et MYB4 chez *E. grandis*), soit hétérozygote avec un allèle nul (mutation dans le site de fixation des amorces conduisant à une absence d'amplification).

Les concordances entre résultats de séquençage et le polymorphisme SSCP ont également été obtenues dans le cas des gènes CAD, COMT, CCoAOMT, 4CL, PAL, EgHypar, EgTubA1 (Gion *et al.*, 2000).

6. Localisation des gènes sur les cartes parentales.

Pour tous les gènes étudiés une liaison significative (LOD>12) avec les marqueurs RAPD des cartes génétiques parentales a été établie. La position des gènes sur ces cartes génétiques est représentée à la Figure 11.



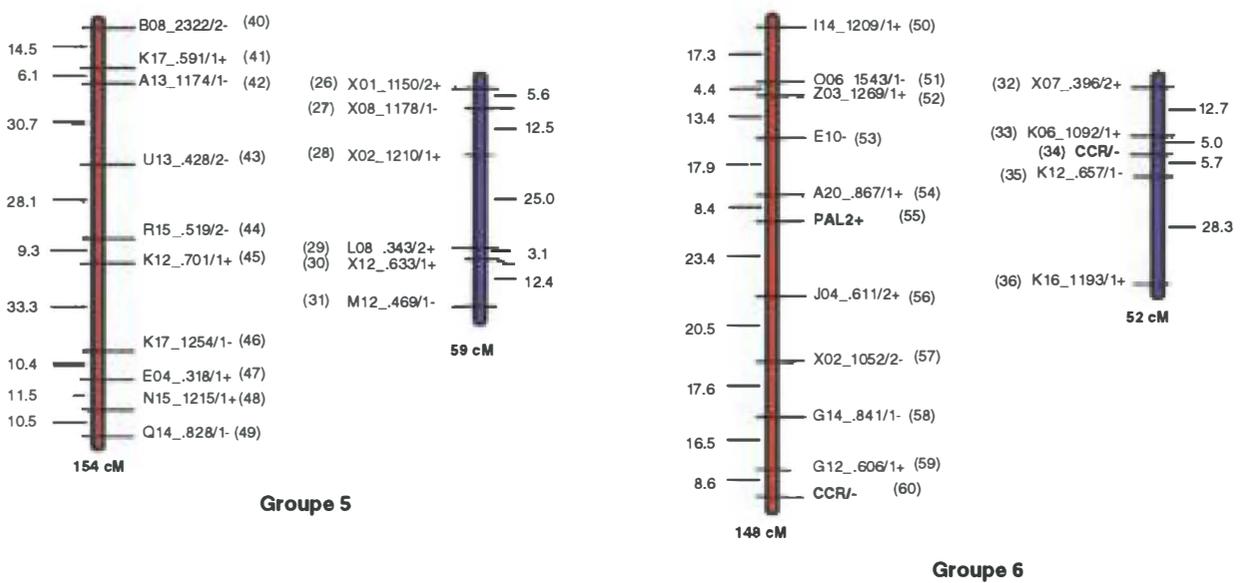
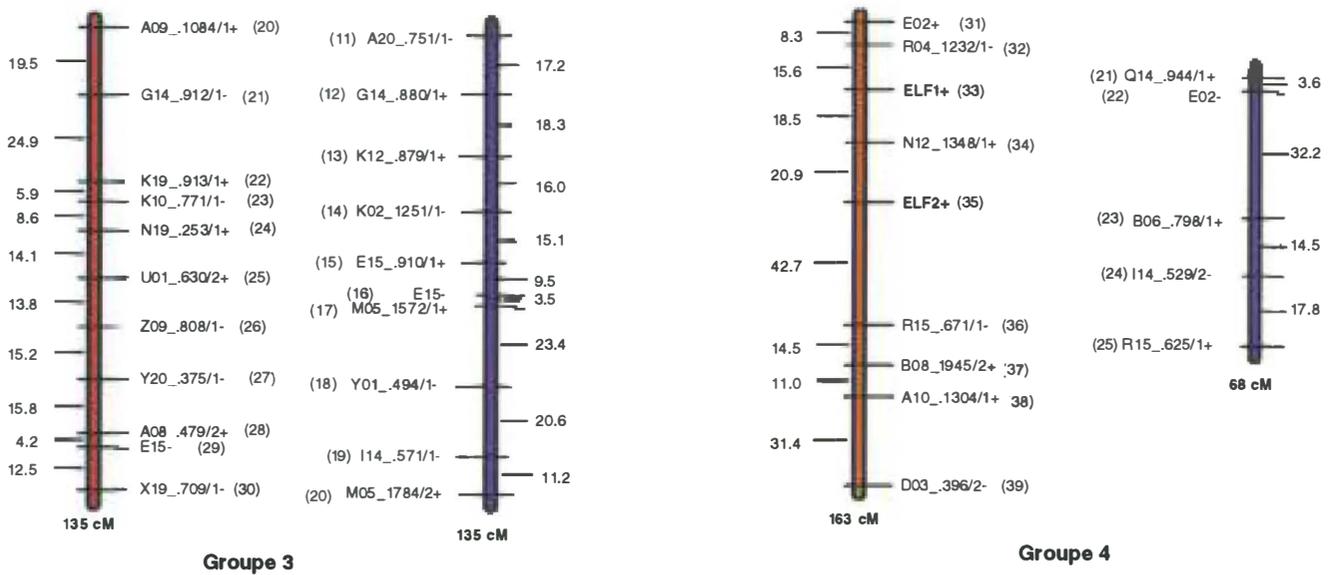
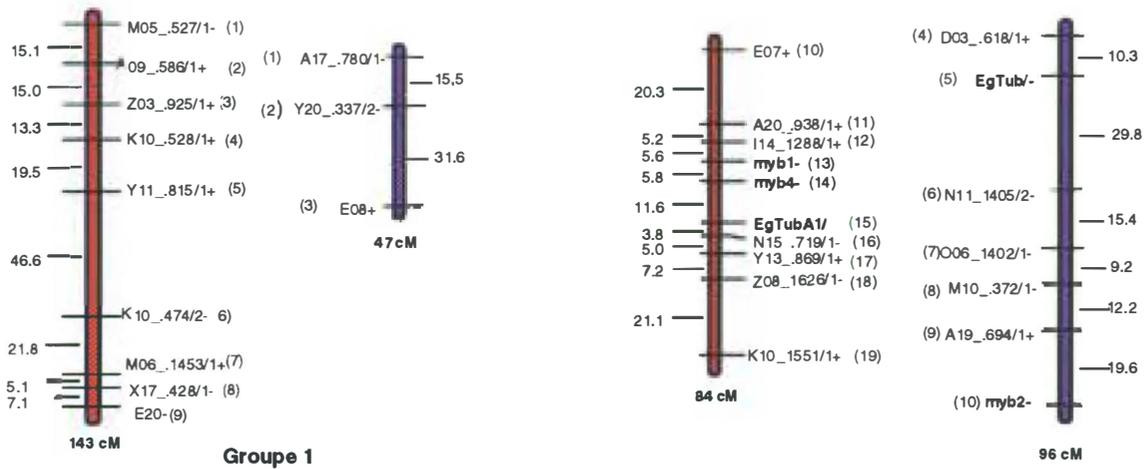
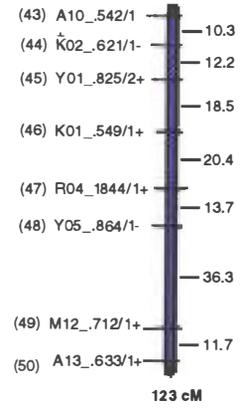
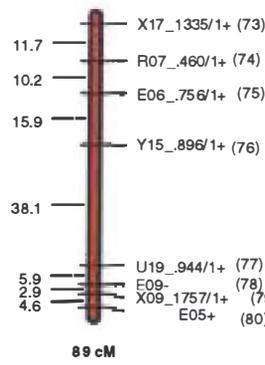
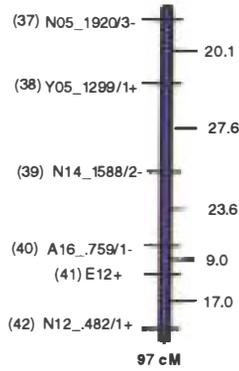
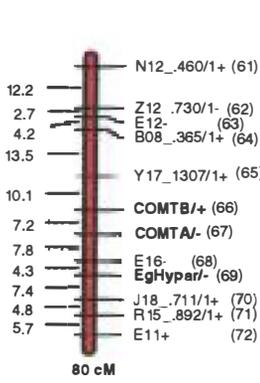
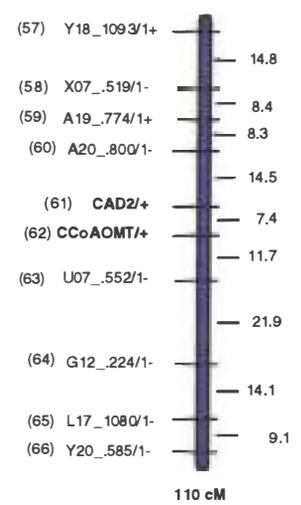
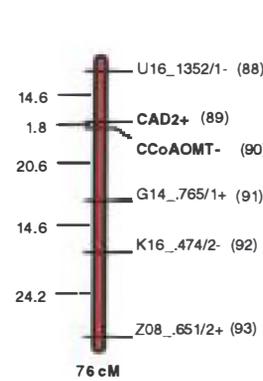
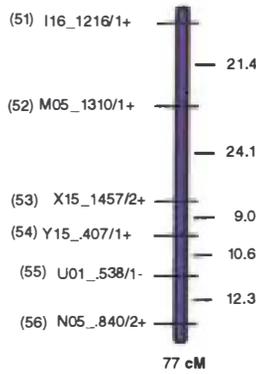
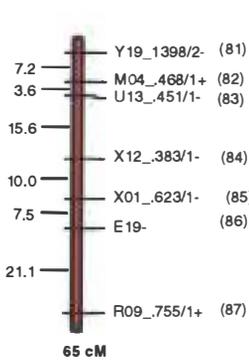


Figure 11 : Localisation des gènes de la voie de biosynthèse des lignines sur les cartes génétiques d'*E. urophylla* et d'*E. grandis*



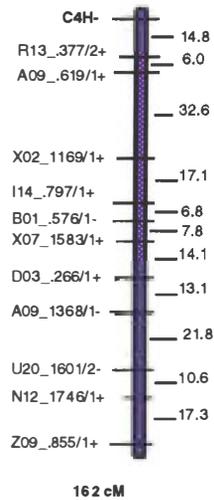
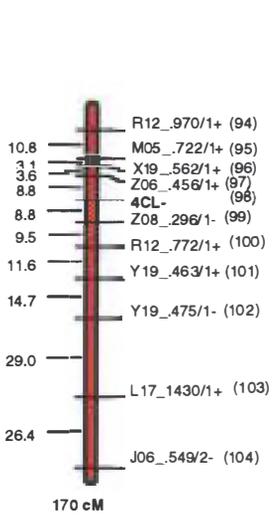
Groupe 7

Groupe 8



Groupe 9

Groupe 10



groupe 11

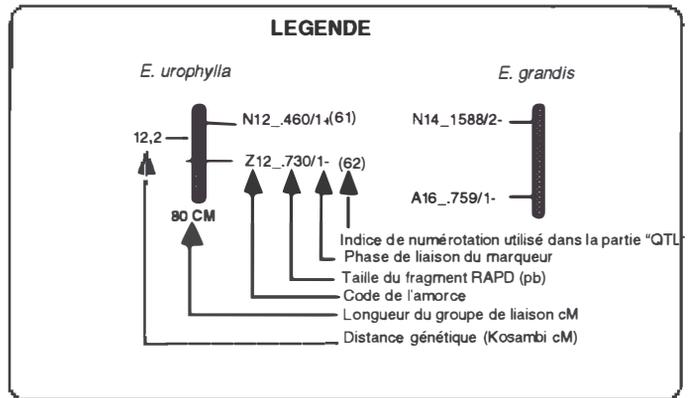


Figure 11 : Localisation des gènes de la voie de biosynthèse des lignines sur les cartes génétiques d'*E. urophylla* et d'*E. grandis* (suite).

Chez *E. urophylla*, MYB1 et MYB4 se positionnent sur le groupe 2 (GL2). Ils sont séparés l'un de l'autre par 5,2 cM (LOD=40). ELF1 et ELF2 ont été localisés tous les deux sur le GL4. ELF2 permet la liaison entre les deux sous-groupes de GL4 (GL4a et GL4b) mis en évidence par Gion *et al.*, 2000. Chez *E. grandis* la C4H et le MYB2 se localise respectivement sur le GL11 et GL2.

Chez *E. urophylla*, le MYB2 n'a pas été positionné sur la carte génétique. En effet, ce marqueur ne montre aucune liaison significative avec les marqueurs des différents groupes de liaison. Il semble donc que les cartes génétiques ne soient pas totalement saturées.

Gion *et al.*, 2000 avait localisé le gène PAL à partir d'un échantillon de 96 individus sur le GL6 d'*E. urophylla*. Le reste des descendants a été génotypé pour le gène PAL. Les résultats de cartographie précise la position de PAL sur le GL6 d'*E. urophylla*.

Pour les gènes de la lignification cartographiés, nous n'observons pas de regroupement des gènes (cluster de gènes). Ainsi, les chances de rencontrer des génotypes combinant les allèles favorables aux différents gènes sont plus importantes. Cependant, des associations entre gènes de structure intervenant à différents niveaux de la voie de biosynthèse ont été observées : PAL et CCR sur le GL6 (distants de 50 cM), CCoAOMT et CAD sur le GL10 (distants de 1,8 et 7,4 cM respectivement pour *E. urophylla* et *E. grandis*), 4CL et C4H sur le GL11 respectif d'*E. urophylla* et d'*E. grandis*.

En ce qui concerne les trois gènes de régulation MYB, on observe un cluster sur le GL2. En effet, les MYB1 et MYB4 se positionnent sur le GL2 d'*E. urophylla* et le MYB2 sur le GL2 d'*E. grandis*. L'homologie entre les GL2 d'*E. urophylla* et d'*E. grandis* a été établie à l'aide de marqueurs microsatellites par Gion (thèse en cours). On peut donc penser que le MYB2 se localise également sur le GL2 d'*E. urophylla*, même si ce dernier n'a pas été cartographié chez ce géniteur.

Chez *E. gunnii*, le MYB1 et le MYB4 sont définis comme deux MYB différents mais ils possèdent une région fortement conservée (Grima-Pettenati, communication personnelle). La définition des amorces pour le MYB1 et le MYB4 a été réalisée dans cette région très conservée des deux MYB. On pourrait s'attendre à voir un mélange des allèles de MYB1 et des allèles de MYB4 lors de l'amplification PCR. Ce n'est pas

Tableau 6 : Statistiques descriptives (moyenne, écart-type, coefficient de variation(Coef.var.), normalité des distributions avec g1 et g2) des caractères étudiés

Variables	Moyenne	Ecart-type	Coef. Var.	g1	g2
HT	2415	238.17	0.10	-2.7*	9.21*
C	54.02	8.54	0.16	-0.84*	1.36
V	216.30	71.54	0.33	-0.18	0.84
DM	2.53	0.24	0.16	-0.13	-0.02
Pil	22.26	2.42	0.11	-0.1	-0.54
DRLM	72.43	15.57	0.22	2.39*	12.96*
IE1	1.19	0.48	0.40	0.58*	0.53
IES	0.96	0.49	0.51	1.13*	2.13*
IE2m	1.06	0.48	0.45	2.12*	6.10*
Tx ext	3.70	0.62	0.17	0.42*	0.02
Klason	24.61	0.85	0.03	0.1	-0.39
G	6.42	0.78	0.12	0.14	-0.16
S	27.55	4.18	0.15	0.81*	0.18
S/G	4.03	0.53	0.13	0.48*	0.22

* : valeur significative au seuil 1%

l'amplification PCR. Ce n'est pas le cas puisque aux vues des résultats de ségrégation et de séquençage nous avons bien amplifié deux MYB différents.

Partie II : Etude des Colocalisations gène-QTL

Après avoir mis en évidence une variabilité allélique des gènes étudiés au sein de notre descendance, l'étude des colocalisations entre gènes et QTL devrait permettre de voir si cette variabilité allélique est responsable en partie de la variation des caractères étudiés.

1. Les caractères étudiés

1.1. Etude de la variabilité des caractères

Le Tableau 6 résume les statistiques descriptives des variables étudiées dont les valeurs g1 et g2 correspondent respectivement au coefficient d'asymétrie et d'aplatissement d'une distribution. Le test g1 permet de vérifier si la moyenne et la médiane coïncident, une valeur négative correspond à une asymétrie gauche et une valeur positive à une asymétrie droite. Le test g2 décrit les proportions trouvées au centre et aux queues de la distribution par rapport aux proportions intermédiaires. Ces deux valeurs rendent compte de la normalité de la distribution.

On observe une asymétrie et un aplatissement significatifs au seuil 1% pour la hauteur, la circonférence, les DRLM, les index d'éclatement à une semaine et deux mois. Une asymétrie significative est observée pour l'index d'éclatement à un jour, le taux d'extrait, les unités S et le ratio S/G.

Une des conditions à l'analyse QTL est la normalité des caractères. Une étude réalisée par Kumar *et al.*, (2000) a montré que les résultats de détection de QTL utilisant des données transformées (pour que l'hypothèse de normalité soit respectée) comparés à ceux obtenus sans modification des données ne diffèrent pas. Des résultats similaires ont également été obtenus par Plomion *et al.*, (1996) et Emebiri *et al.*, (1998).

Tableau 7: Corrélations pour tous les caractères étudiés.

La première valeur donnée est le coefficient de corrélation de Pearson et la seconde valeur est la probabilité qui lui est associée

	HT	C	V	DM	Pi	DRLM	IEJ	IES	IEM	TXEXT	KLASS	G	S	S/G
C	0.813	1												
V	0.000	0	1											
DM	0.789	0.986	0	1										
Pi	0.000	0.000	0	0.364	0.804	0.799	1							
DRLM	0.000	0.000	0.000	0	1									
IEJ	0.147	0.319	0.321	0.347	-	1								
IES	0.039	0.000	0.000	0.000	0	0	1							
IEM	0.213	0.140	0.128	0.008	-0.048	1	0	1						
TXEXT	0.003	0.048	0.071	0.905	0.503	0	0	0.305	0.411	0.738	1			
KLASS	0.433	0.376	0.376	0.146	-0.135	0	0	0	0	0	0	1		
G	0.000	0.000	0.000	0.040	0.057	0.000	0	0	0	0	0	0	1	
S	0.362	0.294	0.295	0.122	-0.118	0.411	0.738	1	0	0	0	0	0	1
S/G	0.000	0.000	0.000	0.091	0.102	0.000	0.000	0	0	0	0	0	0	0
	0.178	0.067	0.086	0.002	-0.118	0.321	0.380	0.604	1	1	0	0	0	0
	0.014	0.359	0.237	0.975	0.105	0.000	0.000	0.000	0	0	0	0	0	0
	-0.055	0.129	0.104	0.230	-0.046	-0.167	-0.134	-0.120	-0.034	1	0	0	0	0
	0.456	0.077	0.157	0.002	0.526	0.022	0.068	0.107	0.640	0	0	0	0	0
	0.047	0.116	0.104	0.076	-0.087	-0.142	0.053	-0.091	-0.109	0.170	1	0	0	0
	0.520	0.113	0.155	0.301	0.234	0.051	0.469	0.222	0.136	0.020	0	0	0	0
	0.070	0.032	0.042	-0.021	-0.007	-0.021	-0.057	-0.050	0.094	-0.050	0.123	1	0	0
	0.338	0.666	0.567	0.770	0.922	0.775	0.434	0.504	0.197	0.499	0.091	0	0	0
	0.157	0.175	0.185	0.151	0.273	0.032	-0.184	-0.227	0.013	0.126	-0.163	0.549	1	0
	0.031	0.016	0.011	0.038	0.000	0.667	0.011	0.002	0.856	0.084	0.025	0.000	0	0
	0.114	0.175	0.175	0.199	0.322	0.059	-0.158	-0.222	-0.082	0.193	-0.286	-0.322	0.610	1
	0.118	0.016	0.016	0.006	0.000	0.418	0.030	0.003	0.262	0.008	0.000	0.000	0.000	0

Le coefficient de variation qui correspond à l'écart type sur la moyenne permet la comparaison de la variabilité entre caractères d'unités de mesure différentes. Pour chacun des caractères la variabilité observée est supérieure à la précision de la mesure. On constate que ce sont les mesures d'index d'éclatement qui possèdent les coefficients de variation les plus élevés.

1.2. Corrélations phénotypiques des caractères

Nous avons étudié les corrélations phénotypiques entre les différents caractères, grâce au coefficient de corrélation linéaire de Pearson. Il traduit la relation linéaire entre deux variables.

Au niveau des caractères de vigueur, on observe les plus fortes corrélations qui se situent entre 0,789*** et 0,986*** (Tableau 7). Ces fortes corrélations entre ces différents caractères de vigueur s'expliquent d'une part par le fait qu'une des variables (le volume) est calculée à partir des deux autres (la hauteur et la circonférence) et d'autre part, que ces trois caractères font intervenir les mêmes voies métaboliques et sont donc étroitement liés.

Pour les caractères de qualité du bois, on observe une faible corrélation négative (-0,142 au seuil de 5%) entre la DRLM et les teneurs en lignines de klason. La diminution de la DRLM moyenne associée à une augmentation de la teneur en lignines a été observée par Boyd, (1977) sur *Eucalyptus regnans* et Baillères *et al.* (1995) sur des hybrides d'eucalyptus du Congo. L'augmentation de la teneur en lignines traduit une augmentation de la proportion de la matrice de lignines et d'hémicelluloses (qui se dilate lors de la maturation cellulaire). Parallèlement à cette augmentation, la teneur (complémentaire) en cellulose cristalline diminue. Les tensions créées lors de la formation du bois diminuent donc avec la teneur en matrice.

On n'observe pas de corrélation significative entre le niveau de DRLM et le ratio S/G. Dans une étude réalisée sur des clones d'eucalyptus hybrides du Congo Baillères *et al.* (1995) montrent une corrélation positive significative au seuil de 1%. Les liaisons, entre les monomères S et les monomères G, plus ou moins fortes, peuvent être de nature différentes et entraîner des "conformations" différentes des monomères de lignines. Ainsi, des variations au niveau de la proportion en monomères peuvent induire des différences de gonflement ou de rigidité de la

Tableau 8 : Mise en évidence de QTL de vigueur (hauteur, circonférence, volume), de forme (DM), de pénétration du pilodyn (Pil), propriétés mécaniques (DRLM, indice d'éclatement), propriétés chimiques (txext., klason, G, S, S/G) par la méthode d'interval mapping (lod seuil 1,7) pour *E. urophylla*. La position du QTL est donnée en cM en partant du marqueur gauche de l'intervalle. Le pourcentage de variation phénotypique expliqué par chaque QTL (% var.), ainsi que le coefficient de détermination multiple (% var. mult.) sont donnés. L'allèle favorable au marqueur le plus proche et le coefficient de substitution allélique (σ ou Δ) sont fournis.

Caractère	Groupe de liaison	Intervalle	Distance (cM)	Marqueur flanquant	Allèle favorable	LOD max	Proba.	% de var.	Δ	% var mult.
HT	1	2-3	0	A09 .586	+	1,9	2,2.10 ⁻³	4,5	0,53	10,2
HT	7	65-66(a)	4	Y17 1307	+	2,0	1,5.10 ⁻³	5,0	0,36	
C	1	2-3	6	A09 .586	+	4,8	<10 ⁻⁴	12,2	0,41	
C	2	12-13	4	I14 1288	+	2,2	1,6.10 ⁻³	5,2	0,70	15,1
C	3	29-30	14	EMBRA15	-	2,0	2,7.10 ⁻³	4,5	0,43	
V	1	2-3	6	A09 .586	+	5,3	<10 ⁻⁴	13,3	0,73	
V	2	12-13	2	I14 1288	+	2,2	1,6.10 ⁻³	5,2	0,44	19,2
V	3	29-30	14	EMBRA15	+	2,2	1,5.10 ⁻³	5,0	0,15	
DM	1	2-3	8	A09 .586	+	4,4	4,0.10 ⁻⁴	11,2	0,60	12,5
DM	2	14-15	6	MYB 4	-	2,8	3,8.10 ⁻³	7,4	0,38	
Pil	2	14-15	0	MYB 4	-	3,3	1,0.10 ⁻⁴	7,3	0,48	
Pil	4	37-38	0	B08 1945	+	2,4	1,2.10 ⁻³	5,8	0,51	21,9
Pil	5	44-45	6	R15 .519	-	2,1	2,2.10 ⁻³	5,2	0,41	
Pil	11	98-99	0	4CL	-	2,0	3,1.10 ⁻³	4,7	0,36	
DRLM	7	64-65	6	Y17 1307	+	2,2	1,9.10 ⁻³	5,4	0,42	
DRLM	9	84-85	2	X12 .383	-	2,5	<10 ⁻⁴	5,9	0,48	18,3
DRLM	10	92-93	0	K16 .474	-	3,5	1,0.10 ⁻⁴	7,9	0,57	
IE1	7	65-66	4	Y17 1307	+	4,4	<10 ⁻⁴	11,5	0,57	16,5
IE1	10	91-92	10	G14 .765	+	2,0	3,0.10 ⁻³	5,8	0,37	
IES	7	65-66	0	Y17 1307	+	5,2	<10 ⁻⁴	13,2	0,38	
IE2m	7	66-65	4	COMTB	+	4,1	4,0.10 ⁻⁴	9,7	0,44	18,2
IE2m	10	92-93	0	K16 .474	+	4,1	8,0.10 ⁻⁴	9,4	0,61	
Tx ext	3	20-21	14	A09 1084	+	4,6	<10 ⁻⁴	13,7	0,49	
Tx ext	6	50-51	4	I14 1209	+	2,5	4,0.10 ⁻⁴	6,7	0,48	25,3
Tx ext	6	55-56	16	PAL	+	2,7	7,0.10 ⁻⁴	8,8	0,44	
Tx ext	10	92-93	16	K16 .474	-	1,9	7,0.10 ⁻⁴	6,8	0,37	
Klason	3	28-29	0	A08 .479	+	3,6	1,0. 10 ⁻⁴	8,5	0,62	
Klason	5	46-47	10	K17 1254	-	2,2	1,5.10 ⁻³	5,4	0,41	40,6
Klason	6	59-60	8	G12 .606	-	5,6	<10 ⁻⁴	13,5	0,65	
Klason	8	77-78	6	U19 .944	-	2,6	6,0.10 ⁻⁴	6,4	0,42	
G	1	4-5	12	Y11 .815	+	3,9	<10 ⁻⁴	11,9	0,57	
G	5	45-46	10	K12 .701	-	2,2	1,8.10 ⁻³	7,6	0,47	25,4
G	6	58-59	10	G14 .841	+	4,1	<10 ⁻⁴	12,7	0,55	
S	4	38-39	0	A10 1304	+	3,5	1,0. 10 ⁻⁴	8,2	0,56	16,3
S	6	55-56	14	PAL	+	1,8	1,4.10 ⁻³	5,5	0,39	
S	8	79-80	0	XO9 1757	+	1,9	3,0.10 ⁻³	4,9	0,43	
S/G	1	3-4	4	Z03 .925	-	1,8	3,2.10 ⁻³	4,6	0,41	
S/G	4	38-39	0	A10 1304	+	1,8	3,3.10 ⁻³	4,4	0,42	
S/G	6	54-55	0	A20 .867	-	4,1	<10 ⁻⁴	9,7	0,61	37,8
S/G	6	58-59	10	G14 .841	-	9,7	<10 ⁻⁴	24,7	0,86	
S/G	8	79-80	0	XO9 1757	+	2,5	8,0. 10 ⁻⁴	6,2	0,45	

Tableau 8 (suite) : Mise en évidence de QTL de vigueur (hauteur, circonférence, volume), de forme (DM), de pénétration du pilodyne (Pil), propriétés mécaniques (DRLM, indice d'éclatement), propriétés chimiques (txext., klason, G, S, S/G) par la méthode d'interval mapping (lod seuil 1,7) pour *E. grandis*. La position du QTL est donnée en cM en partant du marqueur gauche de l'intervalle. Le pourcentage de variation phénotypique expliqué par chaque QTL (% var.), ainsi que le coefficient de détermination multiple (% var. mult.) sont donnés. L'allèle favorable au marqueur le plus proche et le coefficient de substitution allélique (σ ou Δ) sont fournis.

Caractère	Groupe de liaison	Intervalle	Distance (cM)	Marqueur flanquant	Allèle favorable	LOD max	Proba.	% de var.	Δ	% var cumul
HT	4	22-23	34	EMBRA02	-	3,0	2,0.10 ⁻⁴	8,5	0,27	
C	2	8-9	0	M10_.372	-	2,05	2,5.10 ⁻³	4,6	0,43	
C	5	30-31	4	X12_.633	+	2,7	7,0.10 ⁻⁴	7,5	0,48	19,5
C	6	35-36	10	K12_.657	-	2,4	2,0.10 ⁻⁴	7,9	0,51	
V	2	5-6	2	EgTub	+	1,9	3,0.10 ⁻³	4,8	0,39	
V	2	8-9	0	M10_.372	+	2,0	2,4.10 ⁻³	4,6	0,43	27,3
V	4	22-23	26	EMBRA02	-	2,5	1,0.10 ⁻³	9,1	0,29	
V	5	30-31	4	X12_.633	+	2,6	7,0.10 ⁻⁴	7,0	0,48	
V	6	35-36	14	K12_.657	-	2,1	1,3.10 ⁻³	6,5	0,45	
DM	5	30-31	2	X12_.633	+	4,1	<10 ⁻⁴	7,5	0,53	17,9
DM	6	35-36	8	K12_.657	-	3,4	1,0.10 ⁻⁴	10,5	0,55	
Pil	2	6-7	16	M10_.372	+	2,2	1,5.10 ⁻³	5,5	0,30	
Pil	6	35-36	16	K12_.657	-	4,1	<10 ⁻⁴	14,8	0,52	21,1
Pil	10	64-65	0	G12_.224	+	1,9	1,9.10 ⁻³	4,7	0,48	
DRLM	11	76-77	18	A09_.1368	-	2,5	4,0.10 ⁻⁴	6,6	0,44	
IE1j	7	39-40	18	N14_.1588	-	2,2	4,3.10 ⁻³	7,6	0,42	
IE1s	8	48-49	22	Y05_.864	-	1,7	2,2.10 ⁻³	7,0	0,05	
IE2m	2	7-8	0	006_.1402	-	1,7	1,9.10 ⁻³	4,2	0,44	
Tx ext	4	21-22	0	Q14_.944	-	2,0	3,0.10 ⁻³	4,9	0,46	8,9
Tx ext	9	53-54	0	X15_.1457	-	2,1	2,1.10 ⁻³	5	0,42	
Klason	2	9-10	20	A19_.694	-	1,7	3,3.10 ⁻³	4,4		
Klason	3	18-19	0	Y01_.494	+	1,9	3,1.10 ⁻³	4,5	0,44	
Klason	4	22-23	22	EMBRA02	+	2,7	6,0.10 ⁻⁴	10,6	0,41	
Klason	7	39-40	14	N14_.1588	-	2,8	5,0.10 ⁻⁴	9,8	0,45	48,7
Klason	8	48-49	44	Y05_.864	+	2,0	2,4.10 ⁻³	5,6	0,22	
Klason	10	57-58	8	Y18_.1093	+	1,7	5,2.10 ⁻³	4,3	0,45	
Klason	11	72-73	0	I14_.797	-	2,1	2,4.10 ⁻³	4,9	0,46	
G	-	-	-	-	-	-	-	-	-	-
S	2	6-7	12	N11_.1405	+	2,3	1,1.10 ⁻³	6,8	0,41	11,2
S	11	68-69	0	C4H	-	2,1	2,2.10 ⁻³	5,1	0,42	
S/G	6	32-33	0	X07_.396	+	3,2	1,0.10 ⁻⁴	7,6	0,54	
S/G	11	68-69	4	C4H	-	4,9	1,1.10 ⁻³	13,7	0,67	19,1

(a) : les marqueurs en couleur correspondent à un gène.

matrice de lignines et d'hémicelluloses, paramètres qui influencent directement les déformations longitudinales de maturation.

Les index d'éclatement sont corrélés positivement avec la DRLM. Comme la DRLM, ils rendent compte de l'état de nervosité du bois. Un bois à forte DRLM va, lors des opérations de transformation (de l'abattage au conditionnement), "libéré" les contraintes de croissance avec des effets néfastes, notamment les fentes en bout.

2. Mise en évidence de QTL

Un nombre total de 41 QTL, tous caractères confondus a été trouvés chez *E. urophylla*, et 29 QTL pour *E. grandis* (Tableau 8)

2.1. Détection par ANOVA et Interval Mapping

2.1.1. Caractères de vigueur (Hauteur, circonférence, volume)

Chez *E. urophylla*, 8 QTL de vigueur ont été mis en évidence contre 9 QTL chez *E. grandis* respectivement sur les groupes (GL) : GL1, GL2, GL3, GL7 et GL2, GL4, GL5, GL6. Pour le caractère de hauteur 10,2% de la variation phénotypique totale sont expliqués chez *E. urophylla* contre 8,2% chez *E. grandis*.

2.1.2. Caractère de forme : la décroissance métrique (DM)

2 QTL ont été mis évidence chez *E. urophylla* et chez *E. grandis* respectivement sur les GL1, GL2 et GL5, GL6. Ils expliquent conjointement 12,5% et 17,9% de la variation phénotypique totale.

2.1.3. Le pylodin

Chez *E. urophylla*, 4 QTL (GL2, GL4, GL5, GL11) ont été détectés expliquant 21,9% de la variation phénotypique totale. En ce qui concerne *E. grandis* 3 QTL (GL2, GL6, GL10) expliquent 21,1% de la variation phénotypique totale.

2.1.4. Propriétés mécaniques du bois (DRLM, inde d'éclatement)

Tableau 9 : Comparaison de valeurs de F obtenues par "Interval mapping" (a) et par ANOVA (b). F(1%) et F max sont les seuils par un test de permutations (5000 tirages) pour les QTL de vigueur (hauteur, circonférence, volume), de forme (DM), de pénétration du pilodyne sous écorce (Pil), propriétés mécaniques (DRLM, index d'éclatement), propriétés chimiques (tx ext, Klason, G, S, S/G) chez *E. urophylla*.

Caractère	Groupe de liaison	Marqueur flanquant	F obs.	F obs.	Test de permutation (5000)	
			au QTL (a)	au marqueur (b)	F(1%)	Fmax
HT	1	Z03_925	9,65	6,84	6,63	12,93
HT	7	Z12_730	10,41	7,13	6,83	17,27
C	1	A09_586	25,52	24,99	7,07	15,94
C	2	I14_1288	10,64	9,79	7,09	13,16
C	3	X19_709	9,20	8,78	6,72	14,98
V	1	A09_586	26,64	26,78	6,82	12,73
V	2	I14_1288	10,23	9,84	6,60	14,77
V	3	X19_709	10,34	9,91	7,13	15,43
DM	1	A09_586	22,08	19,57	7,18	17,76
DM	2	Z08_1626	12,22	12,07	6,63	14,54
Pil	2	MYB4	15,01	12,33	7,19	15,30
Pil	4	B08_1945	10,81	12,78	6,77	15,50
Pil	5	R15_519	9,6	8,46	6,40	15,62
Pil	11	R12_772	8,94	6,91	6,59	12,84
DRLM	7	B08_365	9,95	10,25	7,32	15,44
DRLM	9	X12_383	11,36	11,31	6,47	13,05
DRLM	10	K16_474	16,93	16,76	6,46	14,02
IE1	7	Y17_1307	20,10	17,51	6,73	14,40
IE1	10	K16_474	8,03	6,85	6,43	14,30
IES	7	Y17_1307	24,74	21,95	6,79	16,04
IE2m	7	COMT2	19,84	17,86	6,49	16,79
IE2m	10	K16_474	20,18	19,12	6,42	11,84
Tx ext	3	G14_912	21,96	19,79	6,51	26,61
Tx ext	6	J04_611	13,16	11,64	6,90	16,80
Tx ext	6	I14_1209	11,83	10,36	7,01	19,02
Klason	3	A20_479	17,17	17,75	6,52	15,39
Klason	5	E04_318	10,44	8,41	6,81	17,75
Klason	6	CCRB	27,56	28,93	6,69	16,32
Klason	8	EMBRA05	12,05	11,29	6,80	16,37
G	1	Y11_815	17,80	15,29	6,85	17,18
G	5	K12_701	10,05	10,81	7,13	17,08
G	6	G14_841	17,25	14,61	6,97	14,31
S	4	A10_1304	16,76	15,76	6,56	16,69
S	6	A20_867	10,5	6,43**	6,79	16,79
S	8	EMBRA05	8,55	9,34	6,93	16,53
S/G	1	Z03_925	7,95	7,93	6,34	11,87
S/G	4	A10_1304	8,62	8,11	7,06	15,10
S/G	6	G14_841	50,96	41,15	7,20	13,23
S/G	6	J04_611	34,31	26,76	6,44	14,81
S/G	8	EMBRA05	11,96	9,76	7,14	14,40

Tableau 9 (suite) : Comparaison de valeurs de F obtenues par "Interval mapping" (a) et par ANOVA (b). F(1%) et Fmax sont les seuils par un test de permutations (5000 tirages) pour les QTL de vigueur (hauteur, circonférence, volume), de forme (DM), de pénétration du pilodyne sous écorce (Pil), propriétés mécaniques (DRLM, index d'éclatement), propriétés chimiques (Tx ext, Klason, G, S, S/G) chez *E. grandis*.

caractère	Groupe de liaison	Marqueur flanquant	F obs.	F obs.	Test de permutation (5000)	
			au QTL (a)	au marqueur (b)	F(1%)	Fmax
HT	4	B06_798	14,12	12,81	6,24	11,49
C	2	M10_372	10,04	9,57	6,06	22,89
C	5	X12_633	11,78	11,68	6,62	17,95
C	6	K12_657	12,24	11,86	6,41	15,32
V62	2	M10_372	9,48	9,58	6,49	18,14
V62	4	B06_798	11,17	9,48	6,84	19,57
V62	5	X12_633	11,84	11,69	6,37	11,11
V62	6	K12_657	10,71	10,59	6,14	15,87
DM	5	X12_633	9,54	9,29	6,97	12,96
DM	6	K12_657	15,57	15,68	6,66	12,80
Pil	2	006_1402	10,42	13,63	7,15	15,31
Pil	6	N11_419	17,49	16,54	6,87	15,11
Pil	10	G12_224	9,94	9,98	7,47	14,24
DRLM	11	U20_1601	13,16	11,27	6,07	13,76
IE1	7	A16_759	9,27	8,35	6,15	12,46
IES	8	M12_712	7,96	8,55	6,57	12,17
IE2m	2	006_1402	10,19	8,55	6,57	12,17
Tx ext	4	Q14_944	9,06	9,60	6,76	11,12
Tx ext	9	X15_1457	9,75	8,86	6,46	14,34
Klason	3	Y01_494	8,78	8,83	6,70	14,65
Klason	4	B06_798	12,04	9,35	6,73	12,43
Klason	7	N14_1588	12,76	9,44	6,98	13,74
Klason	8	M12_712	9,48	9,20	6,71	13,10
Klason	10	Y18_1093	8,00	8,72	6,50	13,02
Klason	11	I14_797	9,49	10,32	6,94	14,33
G	-	-	-	-	-	-
S	2	N11_1405	10,96	8,23	6,54	12,08
S	11	C4H	9,60	7,87	6,61	16,92
S/G	6	X07_396	15,08	12,93	7,32	19,05
S/G	11	C4H	22,81	20,80	6,59	11,91

Pour la DRLM 3 QTL (GL7, GL9, GL10) ont été détectés chez *E. urophylla* expliquant 18,3% de la variation phénotypique totale. Chez *E. grandis*, un seul QTL (GL11) est mis en évidence expliquant 6,6 % de la variation phénotypique totale.

Pour l'indice d'éclatement, pour les trois périodes de mesure (un jour, une semaine et deux mois) un QTL (GL7) a été mis en évidence chez *E. urophylla* et pour l'index d'éclatement à un jour et à deux mois un QTL (GL10) a été détecté. Chez *E. grandis*, un QTL à chaque date a été mis en évidence sur des groupes de liaison différents : GL7, GL8 et GL2 respectivement pour les dates : un jour, une semaine et deux mois.

2.1.5. Propriétés chimiques du bois (taux d'extrait, teneur en lignines, composition monomérique des lignines)

Pour le taux d'extrait 4 QTL (GL3, GL6*2, GL10) ont été mis évidence chez *E. urophylla* expliquant 25,3% de la variation phénotypique totale. Chez *E. grandis*, seulement 2 QTL ont été détectés (GL4, GL9) expliquant 8,9% de la variation phénotypique totale. En ce qui concerne la teneur en lignines de klason, pour *E. urophylla*, 4QTL (GL3, GL5, GL6, GL8) ont été mis en évidence expliquant 40,6% de la variation phénotypique totale. Chez *E. grandis* 48,7% de la variation phénotypique totale sont expliqués par 7 QTL (GL2, GL3, GL4, GL7, GL8, GL10, GL11).

Trois QTL (GL1, GL5, GL6) pour les unités G chez *E. urophylla* ont été détectés expliquant 25,4% de la variation phénotypique totale. Aucun QTL pour les unités G chez *E. grandis* n'a été mis en évidence. Pour les unité S, 3 (GL4, GL6, GL8) et 2 (GL2, GL11) QTL ont été détectés respectivement chez *E. urophylla* et *E. grandis*. Chez *E. urophylla*, 5 QTL (GL1, GL4, GL6*2, GL8) pour le rapport S/G ont été mis en évidence expliquant 37,8% de la variation phénotypique totale. En ce qui concerne *E. grandis*, 2 QTL (GL6 et GL11) expliquent 19,1% de la variation phénotypique totale.

2.2. Test de permutation

Nous avons utilisé des tests de permutation qui permettent de valider l'existence de QTL détectés par IM. Ce test est fourni par le logiciel QGENE et est basé sur la publication de Churchill et Doerge (1994). Il consiste en une analyse marqueur par marqueur. Pour chaque marqueur, les génotypes "absence du fragment" ou "présence du fragment", sont redistribués au hasard, pour chaque caractère étudié, sur l'ensemble de la descendance. Cette redistribution est réalisée

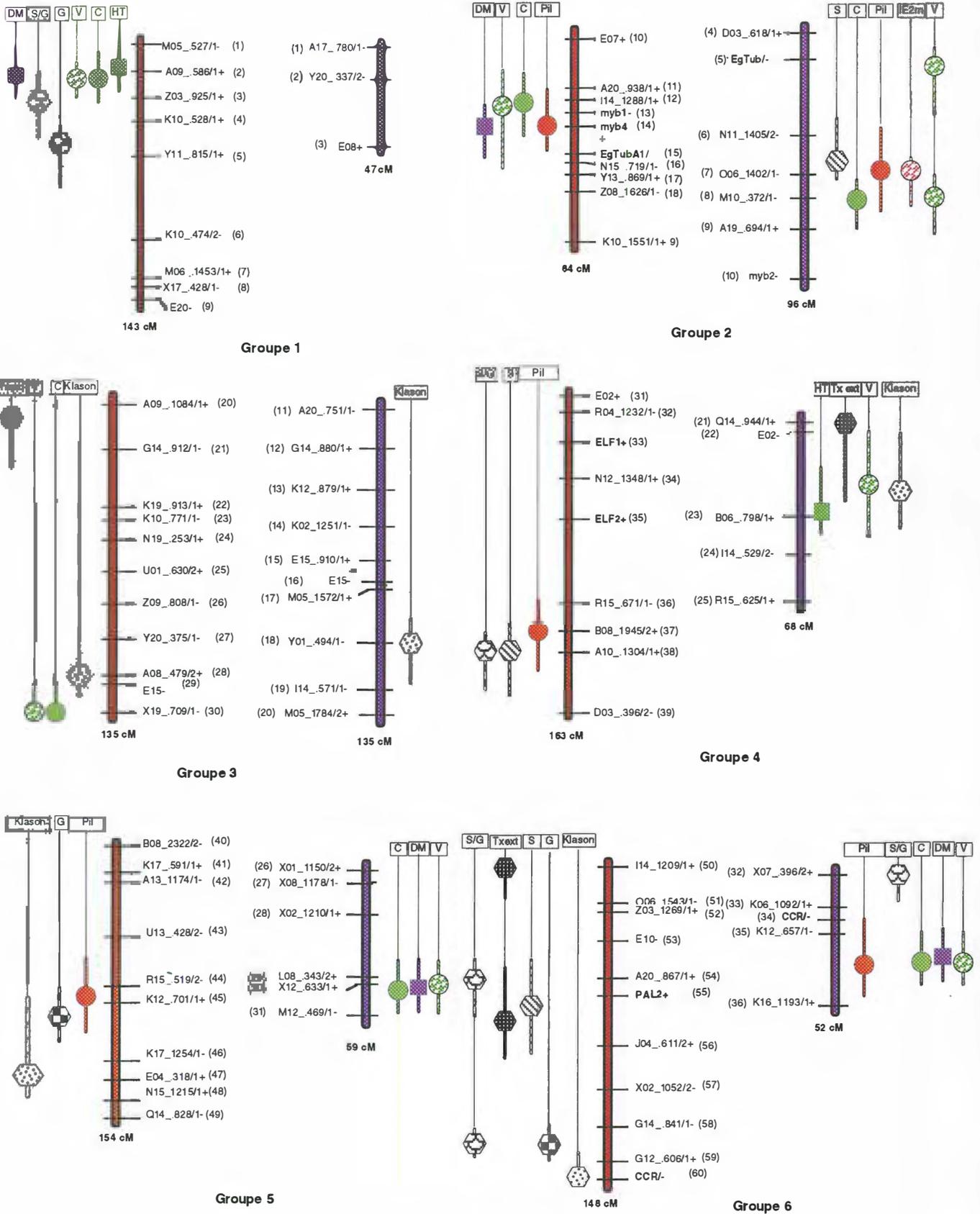


Figure 12 : Mise en évidence de QTL sur les cartes génétiques d'*E. urophylla* et d'*E. grandis*

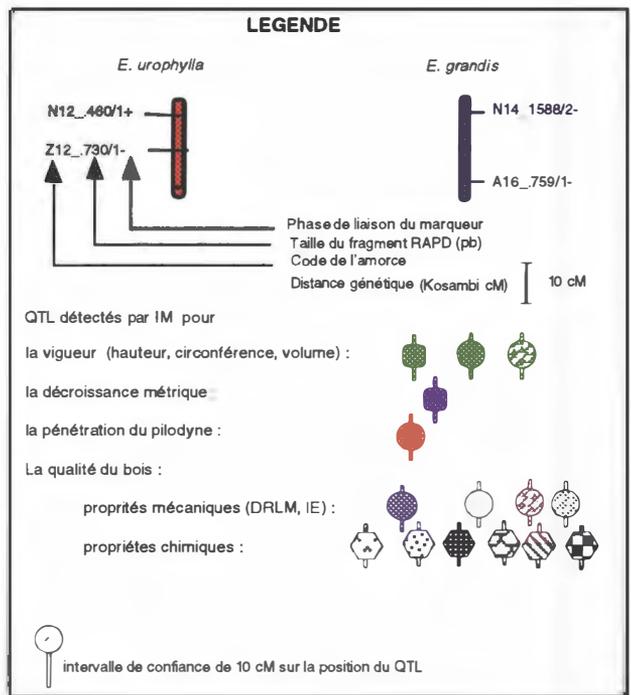
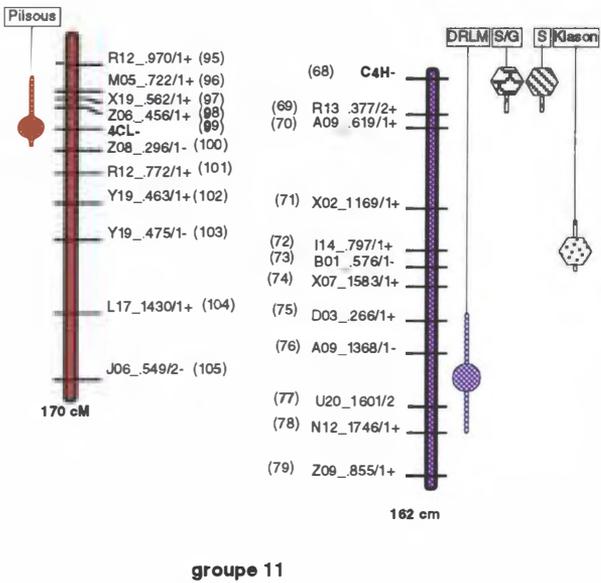
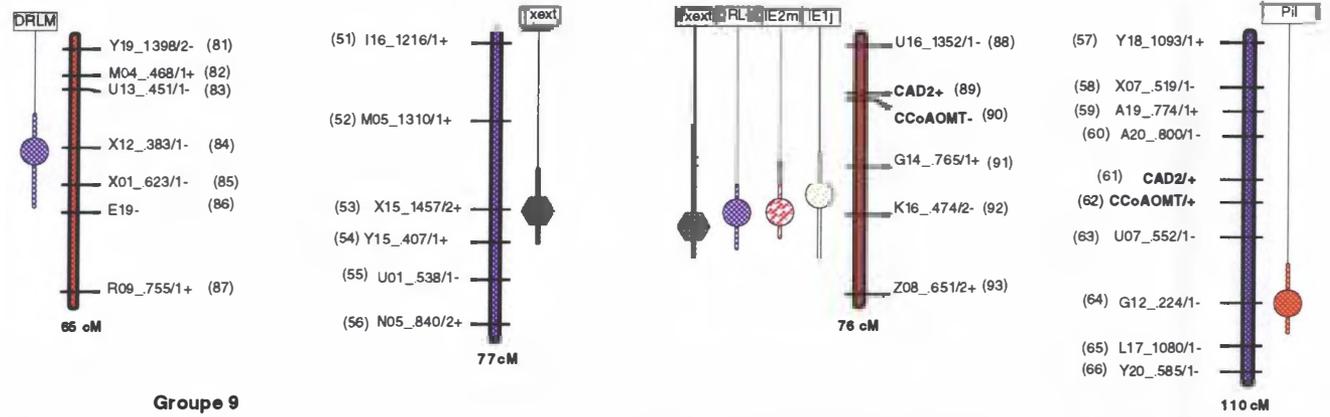
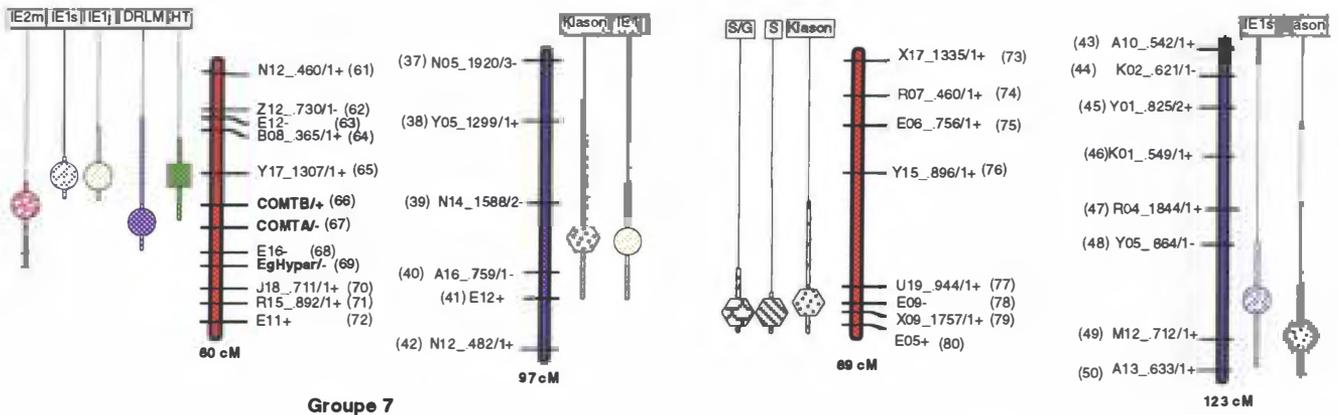


Figure 12 : Mise en évidence de QTL sur les cartes génétiques d'*E. urophylla* et d'*E. grandis* (suite).

5000 fois. Elle permet d'obtenir une nouvelle distribution de F pour l'hypothèse nulle H_0 : "absence de QTL". Ainsi, si le F observé est supérieur au F_{α} à un seuil de 1 %, l'hypothèse H_0 est rejetée. La présence d'un QTL est donc validée avec un risque d'erreur de $\alpha=0,01$.

Chez *E. urophylla*, 97,5% des QTL détectés ont un F observé au marqueur le plus proche supérieur au F(1%). Parmi ces 97,5% des QTL 33,33% ont un F observé au marqueur supérieur au Fmax. Chez *E. grandis*, tous les QTL détectés ont un F observé au marqueur le plus proche supérieur au F(1%) et 17% de tout ces QTL ont un F observé au marqueur le plus proche supérieur au F(max). De plus, si on regarde le F observé au QTL (Tableau 9), pour tous les QTL il est supérieur au F(1%). Ceci conforte la validité des QTL détectés dans notre étude.

3. Colocalisations gène-QTL

Au total 12 colocalisations entre les gènes impliqués dans la voie de biosynthèse des lignines et des caractères de qualité du bois ont été mises en évidence au niveau de quatre groupes de liaison. Afin de mieux les visualiser, ils ont été positionnés sur les cartes génétiques parentales (Figure 12).

Sur le groupe de liaison 6 d'*E. urophylla*, plusieurs colocalisations ont été détectées (Figure 13). Un QTL pour la teneur en lignines de Klason colocalise avec le gène CCR et explique 13,5% de la variation du caractère. Au niveau du marqueur G14_.841 (à 20cM du gène CCR), un QTL pour le rapport S/G a été mis en évidence. Des études réalisées sur des tabacs anti-sens CCR (activité CCR réduite de 90%) montrent que ces derniers ont une réduction sévère des teneurs en lignines (jusqu'à 50% de réduction) et une augmentation du rapport S/G (diminution des unités G) par rapport aux tabacs sauvages (Piquemal *et al.*, 1998). Les travaux de transgénèse sur CCR vont dans le sens de l'existence d'un effet de la variabilité allélique du gène CCR sur la variation de la teneur en lignines dans notre descendance. L'allèle de CCR favorable aux teneurs en lignines est lié à l'allèle favorable aux teneurs en monomères G mais défavorable au rapport S/G du marqueur G14_.841. On peut donc supposer qu'une augmentation de la teneur en

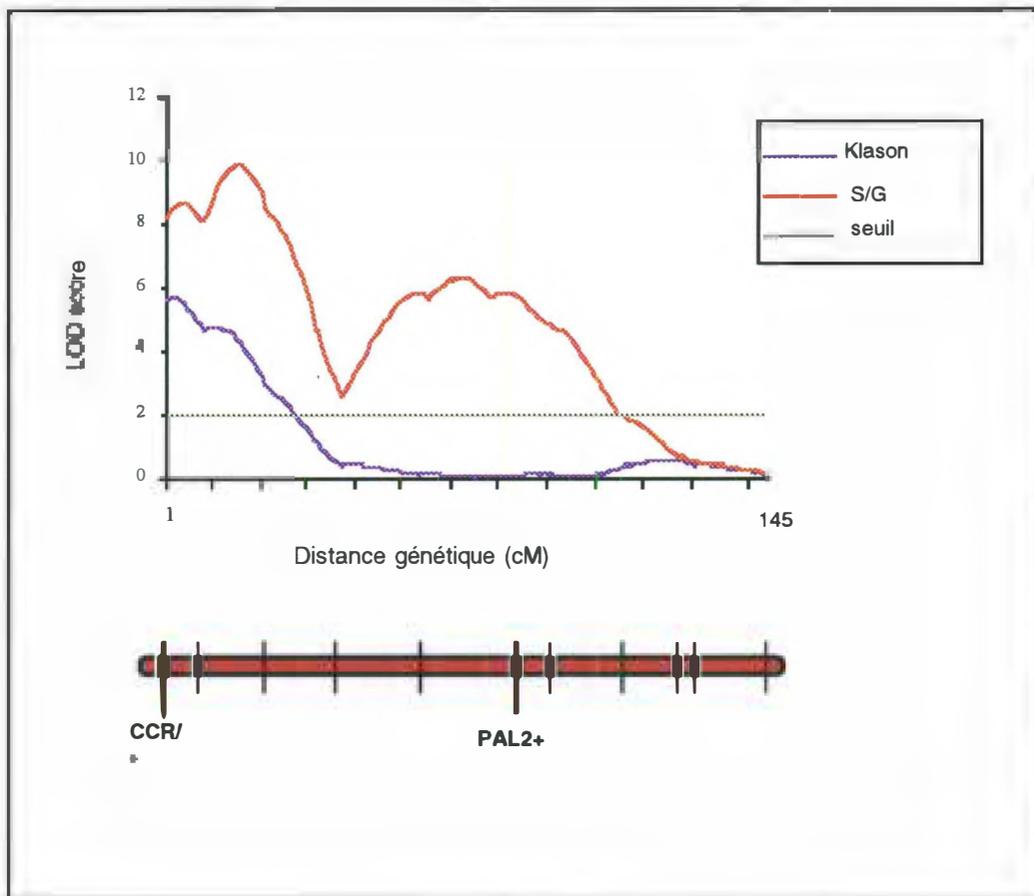


Figure 13 : Evolution des LOD sur le GL6 pour les caractères de Klason et de rapport S/G.

lignines due au gène CCR puisse s'accompagner d'une diminution du rapport S/G laquelle est due à une augmentation des monomères G.

D'autre part sur le groupe de liaison 6 d'*E. urophylla*, des colocalisations entre un QTL de S/G et un QTL de teneur en monomères S à proximité du gène PAL ont été mises en évidence, expliquant respectivement 9,7% et 5,5% de la variation des caractères. L'allèle favorable du gène PAL aux monomères S est également favorable au rapport S/G. Il semblerait donc que ce soit le même QTL. Des travaux de transgénèse réalisés sur des tabacs (Bate *et al.*, 1994) corroborent ce résultat : une réduction de l'activité du gène PAL entraîne une augmentation du rapport S/G.

Enfin une autre colocalisation entre un QTL du rapport S/G et un gène de la voie de biosynthèse des lignines, C4H est mise en évidence sur le groupe de liaison 11 d'*E. grandis*. L'allèle favorable du gène C4H pour le rapport S/G est favorable à la teneur en monomères S. Les travaux de transgénèse réalisés sur le gène C4H montrent que la sous expression de la C4H chez le tabac (Sewalt *et al.*, 1997) induit une diminution du ratio S/G. Ces travaux vont dans le sens qu'une variabilité allélique du gène C4H ait un effet sur la variation du rapport S/G. Sewalt *et al.*, ont émis l'hypothèse que les métabolites seraient peut-être dirigés vers une voie G-spécifique et une voie S-spécifique très tôt dans la biosynthèse des monolignols. Des isoformes spécifiques du gène PAL et du gène C4H pourraient être organisés dans différents complexes, peut-être en association avec des enzymes situés en aval de la voie, dirigeant les métabolites vers des produits finaux différents : soit en monomères S, soit en monomères G.

A la vue des résultats, il apparaît que dans notre descendance la variabilité allélique de trois gènes de structure de la voie de biosynthèse des lignines influence (en partie) les teneurs et la composition des lignines. Etant donné qu'il n'y a pas de cluster de gènes, il sera possible d'observer des individus présentant les différentes combinaisons alléliques favorables pour les caractères de qualité du bois et donc de sélectionner, dans notre descendance, des clones caractérisés par des teneurs en lignines plus ou moins importantes avec des compositions monomériques différentes.

Un QTL de pylodin colocalise avec le MYB4 sur le groupe de liaison 2 d'*E. urophylla*. La surexpression de deux MYB d'*Antirrhinum* (AmMYB330 et AmMYB308) chez le tabac entraîne la répression du métabolisme des acides phénoliques et la biosynthèse des lignines (Tamagnone *et al.*, 1998). Sachant que ces deux métabolismes influencent la formation du bois, la variabilité allélique du gène MYB4 influence la densité du bois. D'autre part, une colocalisation entre le MYB4 et la DM a été mise en évidence. Le pylodin et la DM sont significativement corrélés (0,342 au seuil d'1‰). Lorsque l'arbre croit, la densité du bois augmente. La colocalisation pourrait s'expliquer soit par le fait que MYB4 est un gène à effet pléiotropique, soit par la présence de deux gènes liés.

Sur le groupe de liaison 11 d'*E. urophylla*, le gène 4CL colocalise avec un QTL de pylodin. Il a été montré que plusieurs gènes de la voie générale des phénylpropanoïdes tels que la PAL, la C4H et la 4CL et des enzymes plus spécifiques de la voie de biosynthèse des lignines comme COMT et CAD contiennent des motifs conservés au niveau de leur promoteur, séquences qui sont reconnues par la famille de facteurs de transcription MYB (Bugos *et al.*, 1991 ; Douglas, 1996). De plus, Tamagnone *et al.* (1998) ont constaté que l'activité du gène 4CL est affectée par le taux de transcrits AmMYB308. On peut supposer que l'association entre un allèle X1 du gène 4CL et un allèle Y1 du gène MYB4 influence la variation de la densité du bois, alors que l'association entre un allèle X2 du gène 4CL et un allèle Y2 du gène MYB4 n'entraîne pas de variation significative de la densité de bois (effets opposés).

Sur le groupe de liaison 7 d'*E. urophylla*, nous avons mis en évidence une colocalisation entre un QTL ayant un effet sur l'index d'éclatement (à un jour, une semaine et deux mois) et le gène COMTB. Des travaux réalisés sur le peuplier (Tsai *et al.*, 1998) dont l'activité COMT a été supprimée montrent au niveau des lignines une incorporation importante de monomères G. Baillères (1994) a montré que les déformations longitudinales de maturation dont une des conséquences est l'apparition de fentes étaient influencées directement par la composition monomérique des lignines. L'effet du gène COMTB sur la nervosité du bois pourrait s'expliquer par une influence sur la composition monomérique des lignines, même si

aucun QTL n'a été détecté pour ce caractère. Le gène COMTB pourrait constituer lui aussi un gène candidat potentiel pour la sélection.

4. Cas des gènes CAD et CCoAOMT

On constate qu'aucune colocalisation avec un caractère de qualité du bois n'est détectée au niveau du gène CAD et du gène CCoAOMT cartographiés chez les deux géniteurs. Pourtant, les sous expressions de l'activité des gènes CAD et CCoAOMT chez le tabac et le peuplier entraînent respectivement un changement dans la composition monomérique des lignines et un changement en teneurs et composition des lignines. Ainsi, les allèles aux gènes mis en évidence dans cette descendance n'ont pas d'effet significatif sur la variation des caractères de qualité du bois. Soit l'effet des allèles aux gènes est trop faible et donc l'effet n'est pas détecté soit les deux allèles aux gènes possèdent des effets similaires. La détection de QTL n'est pas exhaustive car elle met en évidence l'effet des allèles sur la variation d'un caractère dans un fond génétique donné. Il faudrait envisager d'étudier la variabilité allélique des gènes CAD et CCoAOMT dans d'autres croisements.

Conclusions et perspectives

L'approche gènes candidats fonctionnels appliquée à la voie de biosynthèse des lignines s'est avérée efficace dans notre descendance hybride d'eucalyptus.

La définition d'amorces spécifiques à des espèces d'eucalyptus différentes est en pratique difficile. Dans notre cas, les hybrides interspécifiques d'eucalyptus amplifient relativement facilement. L'utilisation de la technique SSCP pour la cartographie de gènes de fonction connue est une technique simple et peu coûteuse pour révéler efficacement du polymorphisme. En effet, elle a permis de positionner quatre gènes fonctionnels de la voie de biosynthèse des lignines, sur les cartes génétiques parentales et le résultat du séquençage confirme les hypothèses émises quant aux profils SSCP.

La variabilité génétique de gènes de structure et certains gènes de régulation de la voie de biosynthèse des lignines semble avoir un effet significatif sur différents caractères de qualité du bois. En effet, six gènes colocalisent avec des caractères mécaniques et chimiques du bois et les résultats sont confirmés par les travaux de transgénèse. L'absence de cluster des gènes étudiés laisse envisager la possibilité de trouver des génotypes recombinants favorables par exemple, pour une teneur en lignines moindre.

L'approche gènes candidats fonctionnels est de plus en plus utilisée et menée avec succès chez les plantes annuelles. Chez le colza, la variabilité allélique du gène *fad3* codant pour une Δ^{15} désaturase microsomale influence la teneur en acide linoléique des huiles (Jourden *et al.*, 1996). Barret *et al.* (1998) ont montré que le gène *FAE1* (Fatty Acid Elongation 1) colocalise avec un QTL de teneur en acide érucique chez les graines de colza.

L'approche gènes candidats fonctionnels ouvre des perspectives prometteuses quant à la sélection assistée par marqueurs chez les arbres forestiers (Plomion, 2000). La caractérisation des gènes impliqués dans le contrôle des caractères d'intérêt agronomique permet de s'affranchir du problème de détection lié au déséquilibre de liaison faible et des phénomènes de recombinaison entre le marqueur et le QTL. La sélection en est facilitée par la possibilité de sélectionner les allèles désirés et donc de "construire" les meilleures combinaisons alléliques souhaitées (en théorie). Toutefois, il faut rappeler que l'effet des QTL dépend de différents facteurs (Strauss *et al.*, 1992) : le fond génétique, l'environnement et le temps (stabilité des QTL). Les colocalisations gène-QTL mises en évidence ne sont

valables que dans notre croisement. Q'en est-il dans un fond génétique différent, notamment dans la population d'amélioration ? Cependant avant de s'engager dans l'étude des colocalisations au sein de la population d'amélioration, il faudra tenter de valider le gène candidat.

Toutefois, les gènes candidats fonctionnels définis de la sorte sur la base des connaissances moléculaires et physiologiques du caractère étudié sont en nombre limité. De plus, pour des caractères complexes comme la hauteur, le choix de gènes fonctionnels impliqués dans une voie métabolique spécifique paraît peu pertinent. D'une part, ce caractère fait intervenir de nombreuses voies métaboliques dont la plupart ne sont pas décrites. D'autre part, ce caractère est trop composite et sujet aux effets de l'environnement.

Ainsi, une approche gènes candidats expressionnels serait plus adéquate. Cette dernière est basée sur des études physio-moléculaires où l'on compare l'expression de transcrits en réponse à différents traitements ou lors de différentes étapes du développement. Elle permettrait donc de réduire les possibilités de gènes candidats impliqués dans le contrôle d'un caractère très composite.

Bibliographie

- 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.
- Baillères H. (1994). Précontraintes de croissance et propriétés mécano-physiques de clones d'*Eucalyptus* (Pointe Noire – Congo): hétérogénéités, corrélations et interprétations histologiques. Thèse de Doctorat Université de Bordeaux I. 161p.
- Baillères H., Chanson B., Fournier M., Tollier M.T., Monties B. (1995). Structure, composition chimique et retraits de maturation du bois chez les clones d'*Eucalyptus*. *Ann. Sci. For.* 52 : 157-172.
- Barret P., Delourme R., Renard M., Domergue F., Lessire R., Delseny M., Roscoe T.J. (1998). A rapeseed FAEI gene is linked to the E1 locus associated with variation in the content of erucic acid. *Theor Appl Genet* 96: 177-186
- Bassam B.J., Caetano-Anolles G., Gresshoff P.M. (1991). Fast and sensitive silver staining of DNA in polyacrylamide gels. *Anal Biochem* 196 : 80-83.
- Bate N.J., Orr J., Ni W., Meromi A., Nadler-Hassar T., Doerner P., Dixon R.A., Lamb C.J., Elkind Y. (1994). Quantitative relationship between phenylalanine ammonia-lyase levels and phenylpropanoid accumulation in transgenic tobacco identifies a rate-determining step in natural product synthesis. *Proc Natl Acad Sci USA* 91: 7608-7612.
- Bodénès C., Laigret F., Kremer A. (1996). Inheritance and molecular variations of PCR-SSCP fragment in Pedunculate oak (*Quercus robur* L.). *Theor Appl Genet* 93 : 348-354.
- Boudet A. (1997). Des lignines sur mesure. *Biofutur* 172 : 58-59.
- Bouvet J.M. (1991). Geno-phenotypic regression and juvenile mature correlations : methodological tools for clonal selection of Eucalyptus hybrids in Congo. *In* : IUFRO Symp "Intensive Forestry : the role of Eucalyptus" Durban, South Africa : 13-18.
- Boyd J.D. (1977). Basic cause of differentiation of Tension Wood and Compression Wood. *Australien For. Res.* 7 (3) : 121-143.
- Brown A. (1985). Review of lignin in biomass. *J. Appl. Bioch.* 7 : 371-387.
- Bugos R.C., Chiang V.L.C., Campbell W.H. (1991). CDNA cloning, sequence analysis and seasonal expression of lignin-bispecific caffeic acid 5-hydroxymethylferulic acid O-methyltransferase of aspen. *Plant Mol. Biol.* 17 : 1203-1215.
- Byrne P.F., McMullen M.D., Snook M.E., Musket T.A., Theuri J.M., Widstrom N.W., Wiseman B.R., Coe E.H. (1996). Quantitative trait loci and metabolic pathways : genetic control of the concentration of maysin, a corn earworm resistance factor, in maize silks. *Proc. Natl. Acad. Sci. USA* 93 : 8820-8825.

- Byrne M., Murrell J.C., Owen J.V., Kriedemann P., Williams ER, Moran G.F. (1997). Identification and mode of action of quantitative trait loci affecting seedling height and leaf area in *Eucalyptus nitens*. *Theor Appl Genet* 94: 674-681.
- Churchill G.A., Doerge R.W. (1994). Empirical threshold values for quantitative trait mapping. *Genetics* 138 : 963-971.
- Cotteril P.P., Dean C.A. (1988). Changes in the genetic control of growth of radiata pine to 16 years and efficiencies of early selection. *Silvae Genet* 37 : 138-146.
- Dean J.F.D., Lafayette P.R., Rugh C., Tristram A.H., Hoopes J.T., Eriksson K.E.L., Merkle S.A. (1998). Laccases associated with lignifying vascular tissues. In *ACS Symposium Series "Lignin and Lignan biosynthesis" Washington DC (USA)* 697, chap 8 : 96-108.
- De Vienne D. (1998). Les marqueurs moléculaires en génétique et biotechnologies végétales. Coédition INRA-CNED p 200.
- Douglas C. (1996). Phenylpropanoid metabolism and lignin biosynthesis : From weeds to trees. *Trends Plant Sci.* 1 : 171-177.
- Eldridge K., Davidson J., Hardwood C., and Van Wyk G. (1993). *Eucalyptus domestication and Breeding*. Clarendon Press. Oxford. 288 p.
- Elkind Y., Edwards R., Mavandad M., Hedrick S.A., Ribak O., Dixon R.A., Lamb C.J. (1990). Abnormal plant development and down-regulation of phenylpropanoid biosynthesis in transgenic tobacco containing a heterologous phenylalanine ammonia-lyase gene. *Proc Natl Acad Sci USA* 87 : 9057-9061.
- Emebiri L.C., Devey M.E., Matheson A.C., Slee M.U. (1998). Age-related changes in the expression of QTLs for growth in radiata pine seedlings. *Theor Appl Genet* 97 : 1053-1061.
- Freudenberg K. (1959). Biosynthesis and constitution of lignin. *Nature* 183 : 1152-1155.
- Gallais A. (1978). Amélioration des populations, méthodes de sélection et création de variétés. III. Bases théoriques pour l'étude de la sélection récurrente réciproque. *Ann. Amélio. Plantes* 28 : 637-666.
- Geldermann H. (1975). Investigations on inheritance of quantitative characters in animals by gene markers. I. methods. *Theor. Appl. Genet.* 46 : 319-330.
- Gion J.M., Rech P., Grima-Pettenati J., Verhaegen D., Plomion C. (2000). Mapping candidate genes in *Eucalyptus* with emphasis on lignification genes. *Mol Breed* (sous presse).

Glavac D., Dean M. (1993). Optimization of the single-strand conformation polymorphism (SSCP) technique for detection of point mutations. *Hum. Mutat.* 2 : 404-414.

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

Grattapaglia D., Bertolucci F.L.G., Sederoff R. (1995). Genetic mapping of QTLs controlling vegetative propagation in *Eucalyptus grandis* and *E. urophylla* using a pseudo-testcross mapping strategy and RAPD markers. *Theor Appl Genet* 90 : 933-947.

Grima-Pettenati J., Goffner D. (1999). Lignin genetic engineering revisited. *Plant Sci* 145 : 51-65.

Goldam I.L., Rocheford T.R., Dudley J.W. (1993). Quantitative trait loci influencing protein and starch concentration in the Illinois long term selection maize strains. *Theor Appl Genet* 87 : 217-224.

Hailing J., Martin C. (1999). Multifunctionality and diversity within the plant MYB-gene family. *Plant Mol. Biol.* 41 : 577-585.

Halpin C., Knight M.E., Foxon G.A., Campbell M.M., Boudet A., Boon J.J., Chabbert B., Tollier M.T., Schuch W. (1994). Manipulation of lignin quality by down regulation of cinnamyl alcohol dehydrogenase. *Plant J.* 6 : 339-350.

Harvengt L., Pâques M., Philipon P. (2000). La multiplication des pins. *Biofutur* 199 : 22-25.

Higuchi T. (1990). Lignin biochemistry : biosynthesis and degradation. *Wood Sci. Technol.* 24 : 23-63.

Hu W.J., Lung J., Popko J.L., Harding S.A., Kawaoka A., Kao Y.Y., Hideki S., Stokke D.D., Rinaldi P.L., Tsai C.J., Chiang V.L. (1998). Reduced lignin quality, increased cellulose content, and enhanced growth in transgenic aspen trees. The 8th Meeting of the Conifer Biotechnology working group, june 7-11, Rutgers University, USA.

Jacob M.R. (1954). Les eucalyptus dans les reboisements. Ed. FA.O., 753 p.

Jouanin L. (2000). Arbres transgéniques : quels risques?. *Biofutur* 199 : 16-18.

Jourdren C., Barret P., Brunel D., Delourme R., Renard M. (1996). Specific molecular marker of the genes controlling linolenic acid content in rapeseed. *Theor Appl Genet* 93: 512-518

Kacser H., Burns J.A. (1981). The molecular basis of dominance. *Genetics* 97 : 639-666.

- Kajita S., Hishiyama S., Tomimura Y., Katayama Y., Omori S. (1997). Structural characterization of modified lignin in transgenic tobacco plants in which the activity of 4-coumarate:Coenzyme A ligase is depressed. *Plant Physiol* 114 : 871-879
- Kosambi D.D. (1944). The estimation of map distance from recombination values. *Ann. Eugen.* 12 : 172-175.
- Kremer A. (1994). Diversité génétique et variabilité des caractères phénotypiques chez les arbres forestiers. *Genet. Sel. Evol* 26 : 105s-123s.
- Kumar S., Spelman R.J., Garrick D.J., Richardson T.E., Lausberg M., Wilcox P.L. (2000). Multiple-marker mapping of wood density loci in an outbred pedigree of radiata pine. *Theor. Appl. Genet.* 100 : 926-933.
- 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.
- Lander E.S., Botstein D. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* 121 : 185-199.
- Lapierre C., Monties B., Rolando C. (1986). Preparative thioacidolysis of spruce lignin : isolation and identification of main monomeric products. *Holzforschung* 40 : 1215-1225.
- Lee D., Meyer K., Chapple C., Douglas C.J. (1997). Antisense suppression of 4-Coumarate:coenzyme 1 ligase in *Arabidopsis* leads to altered lignin subunit composition. *Plant Cell* 9 : 1985-1998.
- Lincoln S.E., Lander E.S. (1990). Mapping genes for quantitative traits using MAPMAKER/QTL. A Whitehead Institute for Biomedical Research Technical Report, Cambridge, Massachusetts.
- Mariaux H. (1982). Structure fine du bois de wapa en relation avec les contraintes de croissance. Colloque "Sciences et industries du bois", groupe n°1, Grenoble, France, 20p.
- Marques C.M., Araujo J.A., Ferreira J.G., Whetten R., O'Malley D.M., Liu B.H., Sederoff R. (1998). AFLP genetic maps of *Eucalyptus globulus* and *E. tereticornis*. *Theor Appl Genet* 96: 727-737.
- Marques C.M., Vasquez-Kool J., Carocha V.J., Ferreira J.G., O'Malley D.M., Liu B.H., Sederoff R. (1999). Genetic dissection of vegetative propagation traits in *Eucalyptus tereticornis* and *E. globulus*. *Theor Appl Genet* 6 : 936-946.
- Martin B., Quillet G. (1974). Bouturage des arbres forestiers au Congo : résultats des essais effectués à Pointe-Noire de 1969 à 1973. *Bois et Forêts des Tropiques* 154 : 41-57, 155 : 15-33, 156 : 61, 157 : 21-40.

McKeand S.E. (1988). Optimum age for family selection for growth in genetic tests of loblolly pine. *For Sci.* 34 : 400-411.

Mol J., Grotewold E., Koes R. (1998). How genes paint flowers and seeds. *Trends Plant Sci.* 3 : 212-217.

Nelson J.C. (1997). QGENE : software for marker-based genomic analysis and breeding. *Mol Breed* 3 : 239-245.

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.

Piquemal J., Lapierre C., Myton K., O'Connell A., Schuch W., Grima-Pettenati J., Boudet A. (1998). Down-regulation of cinnamoyl CoA reductase induces significant changes of lignin profiles in transgenic tobacco plants. *Plant J* 13 : 71-83.

Plomion C., Durel C-E., O'Malley D. (1996). Genetic dissection of height in maritime pine seedlings raised under accelerated growth condition. *Theor. Appl. Genet.* 93 : 849-858.

Plomion C., Hurme P., Frigerio J.M., Ridolfi M., Pot D., Pionneau C., Avila C., Gallardo F., David H., Neutelings G., Campbell M., Canovas F.M., Savolainen O., Bodenès C., Kremer A. (1999). Developing SSCP markers in two *Pinus* species. *Mol Breeding* 5 : 21-31.

Plomion C. (2000). Développement de la génomique pour l'étude de la variation des caractères d'intérêt chez le pin maritime. Mémoire d'habilitation à diriger des recherches, Université Paris-Sud : 125 p.

Rychlik W., Spencer W.J., Rhoads R.E. (1990). Optimization of the annealing temperature for DNA amplification in vitro. *Nucl Acids Res* 18: 6409-6412.

Sablowski R.W.M., Moyano E., Cullianez-Macia F.A., Schuch W., Martin C., Bevan M. (1994). A flower-specific Myb protein activates transcription of phenylpropanoid biosynthesis genes. *EMBO J.* 13 : 128-137.

Saghai-Marooif M.A., Soliman K.M., Jorgensen R.A., Allard R.W. (1984). Ribosomal DNA spacerlength polymorphisms in barley : Mendelian inheritance, chromosomal location, and population dynamics. *Proc. Natl. Acad. Sci. USA* 81 : 8014-8018.

Sax K. (1923). The association of size differences with seed-coat pattern and pigmentation in *Phaseolus vulgaris*. *Genetics* 8 : 552-560.

Sewalt V., Ni W., Blount J., Jung H., Masoud S., Howles C. (1997). Lamb, R. A. Dixon, Reduced lignin content and altered lignin composition in transgenic tobacco down-regulated in expression of L-phenylalanine ammonia-lyase or cinnamate 4-hydroxylase. *Plant Physiol.* 115 : 41-50.

- Soller M., Brody T., Genezi A. (1976). In the power of experimental designs for the detection of linkage between marker loci and quantitative loci in crosses between inbred lines. *Theor. Appl. Genet.* 47 : 35-39.
- Southerton SG, Strauss SH, Olive MR, Harcourt RL, Decroocq V, Zhu X, Llewellyn DJ, Peacock WJ, Dennis ES. (1998). *Eucalyptus* has a functional equivalent of the Arabidopsis floral meristem identity gene LEAFY. *Plant Mol Biol* 37 (6) : 897-910.
- Spinardi L., Mazars R., Theillet C. (1991). Protocols for an important detection of point mutations by SSCP. *Nucleic Acids Res.* 19 : 4009.
- Strauss S.H., Lande R., Namkoong G. (1992). Limitations of molecular-marker-aided selection in forest tree breeding. *Can. J. For. Res.* 22 : 1050-1061.
- Tamagnone L., Merida A., Parr A., Mackay S., Culianez-Macia F.A., Roberts K., Martin C. (1998). The AmMYB308 and AmMYB330 Transcription factors from *Antirrhinum* regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. *Plant Cell* 10 : 135-154.
- Tsai C-J., Popko J.L., Mielke M.R., Hu W-J., Podila G.K., Chiang V.L. (1998). Suppression of O-méthyltransferase gene by homologous sense transgene in quaking aspen causes red-brown wood phenotypes. *Plant Physiol.* 117 : 101-112.
- Verhaegen D., Plomion C. (1996). Genetic mapping in *Eucalyptus urophylla* and *E. grandis* using RAPD markers. *Genome* 39 : 1051-1061.
- Verhaegen D., Plomion C., Gion J.M., Poitel M., Costa P., Kremer A. (1997). Quantitative trait dissection analysis in *Eucalyptus* using RAPD markers : 1. Detection of QTL in interspecific hybrid progeny, stability of QTL expression across different ages. *Theor Appl Genet* 95: 597-608.
- Vigneron P. (1991). Création et amélioration de variétés d'hybrides d'*Eucalyptus* au Congo. In: IUFRO Symp "Intensive Forestry : the role of *Eucalyptus*" Durban, South Africa, 345-360.
- Vigneron P. (1992). Création et amélioration de variétés hybrides d'*Eucalyptus* au Congo. *Bois et Forêts des Tropiques* 234 : 29-41.
- Vigneron P. (2000). Classification des sous-genres *Symphyomyrtus* et *Telocalyptus* genre *Eucalyptus* l'Hérit. (Myrtaceae). Note interne CIRAD-Forêt 43 p.

Annexes

Annexe 1

Tampon TBE 10X (Tris Borate EDTA)

Tris HCl	08 g
Aide borique	5 g
EDTA *pH=8 0,5 M	40 ml
H2O	qsp. 1L

Ajuster le pH final à 8,3. Autoclaver.
*=éthylène diamine tétra acétique acide

Tampon TBE 0,6X pour migration SSCP

TBE 10X	60 ml
H2O	qsp. 1L

Annexe 2

Séquences correspondant aux numéros d'accèsion utilisées pour réaliser les alignements multiples

AFO68574= *Arabidopsis thaliana* (ADN)
AJ249093= *Eucalyptus gunnii* (ADNc)
AJ010324= *Populus sp.* (ADNc)
L38898= *Petroselinum crispum* (ADNc)
U71080= *Arabidopsis thaliana* (ADN)
U19922= *Zinnia elegans.* (ADNc)
X92437= *Glycine max* (ADNc)
D82812= *Populus sp.* (ADN)
Z32563= *Roseus sp.* (ADNc)
L07634= *Vigna radiata* (ADNc)
Y13771= *Populus trichocarpa.* (ADNc)
U43542= *Tabaco nicotina* (ADNc)
U12757= *Acer pseudoplatanus* (ADNc)
AF034806= *Eucalyptus globulus* (ADNc)
AF056550= *Arabidopsis thaliana* (clone partiel)
U80912= *Eucalyptus globulus* (ADNc)
Y16126= *Tomato sp.* (ADNc)
X77944= *Tabaco nicotina* (ADNc)
AB000261= *Aspergillus sp.* (ADN)

Résumé

L'objectif de ce travail est d'étudier les colocalisations gènes-QTL (Quantitative Trait Locus) pour des caractères de qualité du bois au sein d'une famille élite d'Eucalyptus. L'étude porte sur une famille pleins frères de 201 arbres issus du croisement interspécifique *E. urophylla* x *E. grandis*. L'utilisation de la technique SSCP (Single Strand Conformation Polymorphism) a permis de positionner cinq gènes fonctionnels sur les cartes génétiques parentales établies à partir de marqueurs RAPD. Parmi ces cinq gènes, quatre interviennent au niveau de la voie de biosynthèse des lignines (dont huit gènes ont déjà été cartographiés) : cinnamate 4-hydroxylase (gène de structure) et trois gènes de régulation (MYB). Un dernier gène a été positionné, le gène Leafy qui intervient dans le développement floral. Ces gènes se localisent sur différentes régions génomiques. La détection de QTL a été réalisée sur des caractères de croissance, de forme, de propriétés mécaniques et chimiques du bois. Un total de 12 colocalisations gènes-QTL a été observé. Les résultats montrent des effets importants de la variabilité allélique des gènes de la voie de biosynthèse des lignines sur la teneur en lignines et le rapport S/G. L'approche gènes candidats comme alternative à la transgénèse semble donc appropriée pour explorer et mettre à profit la variabilité naturelle des gènes contrôlant tout ou une partie des caractères agronomiques d'intérêt. Ainsi les connaissances acquises sur ces QTL permettraient de les intégrer dans le programme d'amélioration de l'eucalyptus au Congo.

Mots clés : *Eucalyptus*, approche gènes candidats, lignines, cartographie, QTL.

Abstract

The aim of this work is to investigate the colocalisations gene-QTL (Quantitative Trait Locus) for wood quality traits within the Eucalyptus elite family. The study deals with full sib family of 201 trees generated from an interspecific cross *E. urophylla* * *E. grandis*. The use of the SSCP (Single Strand Conformation Polymorphism) was allowed to map five functional genes on the parental linkage maps. Among these genes, four are involved in the biosynthesis lignins pathway (eight genes of which were already mapped) : cinnamate 4-hydroxylase (structural gene) and three regulated genes (MYB). A last gene has been mapped, the leafy gene which is involved in the floral development. These genes are mapped on different genomic regions. The detection of QTL was carried out on growth; form, mechanical and chemical property traits. A total of twelve colocalisation gene-QTL was observed. The results show the important effects of the allelic variability of gene of the biosynthesis lignin pathway on the amount of lignin and the ratio S/G. The candidate gene approach, as an alternative to the transgenese one, appears appropriate to explore and put allelic variability of genes controlling all or part of the interesting agronomic traits to good use. So the acquired knowlegde of these QTL will permit to integrate them in the breeding programme of the Eucalyptus in Congo.

Key words : *Eucalyptus*, candidate gene approach, lignin, mapping, QTL.