BEDON FRANK

STRUCTURE GÉNIQUE ET CARACTÉRISATION FONCTIONNELLE DE FACTEURS DE TRANSCRIPTION *MYB-R2R3* IMPLIQUÉS DANS LA FORMATION DU XYLÈME CHEZ LES CONIFÈRES

Thèse de doctorat en cotutelle présentée à la Faculté des études supérieures de l'Université Laval, Québec dans le cadre du programme de doctorat en Science Forestières pour l'obtention du grade de *Philosophiae Doctor* (PhD)

SCIENCES FORESTIÈRES FACULTÉ DE FORESTERIE ET DE GÉOMATIQUE UNIVERSITÉ LAVAL QUÉBEC, CANADA

et

BIOSCIENCES VÉGÉTALES ÉCOLE DOCTORALE DE BIOLOGIE, SANTÉ ET BIOTECHNOLOGIES UNIVERSITÉ PAUL SABATIER TOULOUSE III TOULOUSE, FRANCE

2007

© Frank Bedon, 2007

Résumé

Chez les arbres, la formation du bois dépend de la différenciation d'un tissu vasculaire appelé xylème secondaire. Les parois cellulaires du xylème sont riches en lignines, ce qui leur confère rigidité et imperméabilité indispensables au transport de la sève brute et au maintien du port dressé. La xylogénèse requiert la coordination de l'expression de plusieurs centaines de gènes dans le temps et dans l'espace. Cette coordination est assurée au niveau transcriptionnel par différents médiateurs protéiques, parmi lesquels il y a des facteurs de transcription de la famille MYB-R2R3. Peu de gènes de cette sous famille ont été préalablement caractérisés chez les conifères. Mon travail de thèse a consisté dans un premier temps à identifier plusieurs gènes MYB-*R2R3* potentiellement impliqués dans la formation du xylème chez les conifères, en particulier chez le pin loblolly (Pinus taeda) et l'épinette blanche (Picea glauca), et à réaliser leur caractérisation moléculaire. Différentes analyses phylogénétiques ont permis de mettre en évidence des similarités et des différences significatives dans la structure de cette famille entre les gymnospermes et les angiospermes. L'expression de ces mêmes gènes a été quantifiée par RT-QPCR, ce qui a révélé que PgMYB2, 4 et 8 sont préférentiellement exprimés dans le xylème secondaire de l'épinette et que leurs transcrits s'accumulent dans le bois de compression. La deuxième partie de mon travail a visé la caractérisation fonctionnelle de deux MYB-R2R3 de pin (PtMYB1 et 8) réalisée *via* leur surexpression constitutive chez des transformants stables d'épinette. Ces deux facteurs MYB induisent une lignification ectopique cependant plus intense avec PtMYB8, qui bloque le développement normal des racines et conduit à la mort des plantules *in vitro*. L'analyse du transcriptome des plantules transgéniques par biopuces à ADN a dévoilé entre autres, l'augmentation de l'expression des gènes de la synthèse des lignines. La dernière partie de ma thèse est consacrée à la caractérisation fonctionnelle de *PtMYB14* placé en parallèle sous le contrôle d'un promoteur constitutif et d'un promoteur préférentiel du xylème (gène codant l'alcool cinnamylique deshydrogénase). Les plantules transgéniques ont une synthèse accrue de plusieurs terpènes et produisent de nombreux transcrits potentiellement associés aux mécanismes de défense.

Abstract

In trees, wood formation depends on the differentiation of a vascular tissue refered to as secondary xylem. Xylem cell walls are rich in lignins, which confer stiffness and impermeability, properties essential for the mechanical support of the tree and for the transport of water and minerals. Xylogenesis requires the coordinated expression of several hundred genes in time and space. The regulation of gene expression is controled at the transcriptional level by different protein mediators, including transcription factors of the R2R3-MYB family. Hitherto, very few genes of this sub-family have been characterized, specially in conifers. The first part of my thesis consists in the identification and molecular characterization of several genes belonging to this family and potentially involved in xylem formation in conifers, particularly in loblolly pine (Pinus taeda) and white spruce (Picea glauca). Phylogenetic analyses highlight the major similarities as well as significant differences in the MYB family structure between gymnosperms (trees) and angiosperms (herbaceous plants and trees). I used qRT-PCR to determine transcript levels of these genes and assess their expression profiles in spruce. The data showed that the spruce transcripts for PgMYB2, 4 and 8 accumulate preferentially in secondary xylem and are upregulated in compression wood. The second part of my work aimed the functional characterization of two pine R2R3-MYB (PtMYB1 and 8) via their constitutive overexpression in stable spruce transformants. The overexpression of both of these MYB transcription factors led to ectopic lignification. However, the overexpression of PtMYB8 led to a more intense lignification; it also prevented normal root development leading to the death of the *in vitro* plantlets. Microarrays analyses of the transgenic plantlets showed that there was a significant upregulation of many lignin biosynthesis genes, among others. The final section of my thesis describes the functional characterization of *PtMYB14* based on its overexpression with either a constitutive promoter or a xylem-preferential promoter (gene coding for the cinnanyl alcohol dehydrogenase). Both types of transgenics plantlets accumulated several terpenes compounds and produced numerous transcripts potentially involved in defence mechanisms.

Avant-Propos

Voici enfin venu le moment tant attendu de consacrer le temps nécessaire à la rédaction de la dernière partie de cette thèse qui représente tant de souvenirs et de rencontres pendant ces cinq années de doctorat. La réalisation et l'accomplissement de cette thèse a été permis grâce à toutes les personnes qui m'ont aidé et encouragé pendant cette aventure effectuée à Toulouse et à Québec. Cette thèse est aussi le reflet d'une extraordinaire collaboration entre plusieurs personnes dont la contribution scientifique est à la hauteur de l'aventure humaine. A toutes ces personnes je voudrais exprimer ma pleine gratitude et mes remerciements les plus sincères

A mes directeurs, John MacKay et Jacqueline Grima-Pettenati : une aventure en cotutelle entre Québec et Toulouse qui a commencé dans le laboratoire « Régulation transcriptionnelle » de Toulouse avec Jacqueline Grima-Pettenati, déclenchée par Armand Séguin en année sabbatique à Toulouse, et concrétisée par John MacKay à Québec.

Merci Jacqueline pour avoir cru en moi et pour m'avoir fait confiance dans cette aventure !

Merci John pour m'avoir accepté au sein de ton équipe, de ton projet grandiose et pour ces années de formation, de discussion et de confiance !

Aux membres de mon jury de thèse, c'est-à-dire Normand Brisson et Gille Pilate, pour avoir accepté d'être les rapporteurs de ma thèse et pour l'avoir jugé, Armand Séguin et Christophe Roux, qui m'ont fait l'honneur de leur présence lors de ma soutenance à Québec.

Aux membres de mon comité de thèse, Jean Robert Thibault, Nathalie Isabel et Armand Séguin pour leurs conseils avisés à un moment important de ma thèse.

A Claude Bomal, mon plus proche collaborateur sur les MYB « plantules », pour ces nombreux moments passés sous la hotte aspirante à trier ces sacrés embryons transgéniques sous le binoculaire, pour nos discussions et surtout tes exclamations devenues mythiques : « les MYBs, c'est merveilleux ! ». A mes premiers mentors qui m'ont donné le goût de la recherche à travers différents stages allant du BTS au DEA : Pascal Leplatois, Jean-Marc Gion, Béatrice Teulat-Merah et Dominique This, Simon Turner, Pauline Garnier-Géré et Christophe Plomion.

Mes remerciements s'adressent aussi à toutes les personnes rencontrées dans les laboratoires et qui ont contribué à nos travaux, l'équipe Arborea du pavillon C-E Marchand de l'université Laval : Sylvie, Isabelle, Vicky, Mario, Sébastien, Janice, Nathalie, Nancy, Brian, André, Marie-José, Pierre Olivier, Philippe, ...; le Centre de Foresterie des Laurentides à Québec : Armand, Caroline, Laurence, Françoise, Marie-José, Denis, Frédéric, ...; l'équipe « RT » de Toulouse avec Nathalie, Patricia et Pierre. Et aussi Sauphie Senneville et Claude Fortin, ainsi que Axel Schmidt, Shawn Mansfield et Catherine Lapierre pour leurs collaborations.

Aux différentes personnes des services administratifs qui ont agis dans les deux universités et les deux équipes de recherche « RT » et « Arborea » pour faciliter mon intégration et mes démarches administratives, à Toulouse : Nathalie pour les nombreuses inscriptions à distance, Michèle, Nicole et Christel Chérie ; à Québec : Pascal, Marie-José, Marie-Hélène, Fabienne et Carmen Bilodeau.

A mes collègues thésards, post-docs, stagiaires et non permanents, passés et présents, rencontrés dans les différents laboratoires, Toulouse : Camille, David, Sylvain, Audrey, Anja, Nicolas (pour une correction orthographique efficace de ma thèse : merci !), . . . Québec : Caroline R, Fred, Sylvain, Vincent, André, Caroline C, Francis, Hugo, Florian, Claude, Nathalie, Pat, Benjamin, Karine, Marie-Eve, Mirella, Alain, Séb de Marseille, Julie, Julien et aussi Richard, Frédéric, Louis-Philippe, Marie-Claude, . . .

A mes amis de Québec, parfois colocataires, Greg, Sylvain, Guillaume, Fred « the drummer », Julien, Marie-Eve, Jean-Luc, Soul, et mes amis de Toulouse, Yohann, Blaze, Gibi, La mouette, AAAlex, Minouz, Nix and Nono.

A mes parents, Francine et Joseph

SOMMAIRE

Résumé	ii
Abstract	iii
Avant-Propos	iv
SOMMAIRE	vii
LISTE DES FIGURES	xiii
LISTE DES TABLEAUX	XV
LISTE DES ANNEXES	xvi
ABRÉVIATIONS	xvii

CHAPITRE I	1
INTRODUCTION GÉNÉRALE	1
I-1-La formation du bois	2
I-1-1-Origine et mise en place des tissus vasculaires chez les plantes	2
I-1-2-Mise en place des tissus vasculaire chez les ligneux pérennes	2
I-1-2-1-Le cambium vasculaire	4
I-1-2-2-Le xylème	7
I-1-2-3-Le phloème	7
I-1-2-4-Le collenchyme et le sclérenchyme	8
I-1-3-La xylogénèse	9
I-1-3-1-Élongation cellulaire et paroi primaire	9
I-1-3-2-Formation de la paroi secondaire et lignification	11
I-1-3-3-Mort cellulaire	11
I-1-4-Hétérogénéité du bois à l'intérieur de l'arbre	12
I-1-5-Les canaux résinifères	14
I-1-6-Les composés phénoliques	15
I-1-6-1-Les phénylpropanoïdes	17
I-1-6-2-La lignine	19
-Rôles de la lignine	20
-Biosynthèse des monolignols	20
-Transport et polymérisation des monolignols	21
I-1-6-3-Modulation de l'expression des gènes des phénylpropanoïo	les et
des lignines	22
I-1-7-Régulation de la formation du bois	24
I-1-7-1-Les hormones	24
I-1-7-2-Les facteurs de transcription	25
I-2- Les facteurs de transcription MYB	27
1-2-1-Le mecanisme general de la transcription des genes	27
1-2-2-Identite et fonction generale des MYB	28
1-2-3-Les MYB chez les vegetaux	31
1-2-3-1-Classifications des membres de la famille MYB	32
- Structuration par DBD et element cis fixe	52 22
-monys proteiques consensus en carboxy terminal	33 22
-Nombre et position des introns	55

I-2-3-2-Duplication génique et évolution des gènes MYB	33
I-2-4-Rôles biologiques des facteurs de transcription MYB	34
I-2-5-Régulation de l'activité des facteurs de transcription MYB	39
I-2-6-Le complexe protéique MYB/bHLH/WD40	40
I-2-7-Régulation transcriptionnelle de l'expression des gènes du	métabolisme
des phénylpropanoïdes	43
I-3-Objectifs du travail de thèse et présentation	46
I-3-1-Hypothèses et objectifs des travaux de thèse	46
I-3-2-Présentation des travaux de thèse	47
I-4-Références	48

CHAPITRE II

67

CONIFER R2R3-MYB TRANSCRIPTION FACTORS: SEQUENCE ANALYSES AND GENE EXPRESSION IN WOOD-FORMING TISSUES OF WHITE SPRUCE (*PICEA GLAUCA*)

II-1-Avant-propos	67
II-2-Résumé	68
II-3-Abstract	69
II-4-Introduction	70
II-5-Materials and Methods	72
II-1-1-Plant material and RNA isolation	72
II-1-2-DNA cloning and accession numbers	73
II-1-3-Sequence analyses and phylogenetic studies	75
II-1-4-Analysis of transcript accumulation by Q-RTPCR	76
II-6-Results	78
II-6-1-Isolation and sequence analysis of 18 <i>R2R3-MYB</i> genes from s and pine	spruce 78
II-6-2-Phylogenetic relationships and gene family structure of conifer	R2R3-
MYBs	83
II-6-3-Sequence analysis of conserved regions in the C-terminal of <i>P. g</i> MYBs	zlauca 87
II-6-4-Expression of <i>P. glauca MYB</i> genes in tissues of young and n trees	nature 88
II-6-5-Spruce MYB genes are differentially expressed in compression wo	od 91
II-7-Discussion	93
II-7-1-Sequence conservation and identification of amino acid mot spruce R2R3-MYBs	ifs in 93
II-7-2-Spruce MYB phylogeny and evolution	95
II-7-3-Potential involvement of the spruce R2R3-MYBs in the lignificat woody tissue	ion of 96
II-8-Conclusion	99
II-9-Acknowledgments	99
II-10-References	100

CHAPITRE III

DISTINCT AND OVERLAPPING EFFECTS OF *PINUS TAEDA* MYB1 AND MYB8 OVEREXPRESSION ON LIGNIFICATION AND RELATED METABOLISM IN CONIFERS

III-1-Avant-propos	104
III-2-Résumé	105
III-3-Abstract	106
III-4-Introduction	107
III-5-Materials and Methods	110
III-5-1-Pine tissues collection for <i>PtMYB1</i> and <i>PtMYB8</i> expression profile	e 110
III-5-2-Vector construction and spruce transformation	110
III-5-3-Transgenic lines and culture conditions used for somatic see	dling
production	111
III-5-4-Seedling growth quantification and tissue sampling	111
III-5-5-Histology	111
III-5-6-RNA extraction, cDNA synthesis and real-time RT-PCR analysis	112
III-5-7-Microarray experiment	113
III-5-8-Statistical analysis of microarray data	114
III-5-9-Soluble phenolic metabolite analysis	114
III-5-10-Acetyl bromide lignin determination	115
III-6-Results	116
III-6-1-Transcript abundance of PtMYB1 and PtMYB8 in various tissu	es of
Pinus taeda trees	116
III-6-2-PtMYB1- and PtMYB8-OE in spruce leads to reduced root growt	h and
to ectopic lignification	117
III-6-3-Lignin and phenolics in PtMYB8 overexpressing spruce	119
III-6-4-PtMYB1- and PtMYB8-OE have distinct but also overlapping ir	npact
on spruce transcriptome	122
III-6-5-PtMYB1- and PtMYB8-OE in spruce affected the expression leve	els of
spruce endogenous MYBs	131
III-7-Discussion	133
III-7-1-Involvement of pine MYBs in lignin deposition	133
III-7-2-Transcriptional regulation of genes linked to lignin biosynthesis	s and
secondary metabolism pathways	135
III-7-3-Different roles of MYB1 versus MYB8 in metabolism and	plant
development	137
III-7-4-Do MYBs form a regulatory network to control lignification?	139
III-9-Acknowledgments	140
III-10-References	140

104

CHAPITRE IV	145
RÉSULTATS COMPLÉMENTAIRES SUR	LA
CARACTÉRISATION PHÉNOTYPIQUE ET FONCTION	NELLE
DES JEUNES ÉPINETTES SUREXPRIMANT LE GÈNE PA	TMYB1
IV-1-Avant-propos	145
IV-2-Résumé	146
IV-3-Introduction et mise en contexte	147
IV-4-Matériels et méthodes	147
IV-5-Résultats	149
IV-5-1-Phénotypes des arbres surexprimant <i>PtMYB1</i>	149
IV-5-2-Nuveaux de transcrits de gènes reliés à la formation de la p	paroi dans
les pousses terminales	152
IV-6-Conclusions préliminaires et perspectives	154
IV-7-Références	155
CHAPITRE V	156
PINUS TAEDA MYB14. A MEMBER OF A	NEW
PHYLOGENETIC CLADE OF R2R3-MYB GENE	S. IS
IDENTIFIED AS A PUTATIVE REGULATOR IN DEL	TENCE-
RESPONSE IN PINE AND SPRIICE	
RESI GIVE IN THE AND SI RUCE	
V-1-Avant-propos	156
V-2-Resume	157
V-J-ADSTRACT	158
v-4-minouucion V-5-Materials and Methods	161
V-5-1-DNA cloning and accession number	161
V-5-2-Sequences analysis of the PtMYB14 and related sequences in	n nine and
spruce	163
V-5-3-Vectors construction and genetic transformation of white spru	ce 164
V-5-4-Histological analysis, X-gluc histochemical staining, and	nd MUG
fluorometric enzymatic assay	165
V-5-5-Starch content determination	166
V-5-6-Anthocyanin extraction and analysis	166
V-5-7-Extraction of resin terpenes	167
V-5-8-Analysis of monoterpenes, sesquiterpenes and diterpenes	167
V-5-9-RNA Extraction, cDNA Synthesis and real-time PCR analysis	168
V-5-10-Microarray experiment	169
V 5 12 Wounding and isomonic sold treatments in nine and sprus	1/0
nlantlets	<i>un vuro</i> 170
V-6-Results	170
V-6-1-Isolation of several R2R3-MYB sequences similar to PtMYB	14 in pine
and annua	172

- V-6-2-PtMYB14 belongs to a new conifer specific clade in the subgroup 4 of plant R2R3-MYBs 174
- V-6-3-RNA transcript profiles of *MYB14* sequences in pine and spruce 174
- V-6-4-Constitutive and xylem preferential overexpression of PtMYB14 in transgenic spruce 177
- V-6-5-Determination of anthocyanin and starch contents in transgenic PtMYB14-OE plantlets 182
- V-6-6-Constitutive and tissue-preferential overexpression of PtMYB14 has an overlapping impact on spruce transcriptome 182

V-6-7-PtMYB14 overexpression affects terpene production in spruce plantlets 190

V-6-8-Wounding and JA treatments induce the expression of PtMYB14 related genes and isoprenoid/defense pathway genes in pine and spruce seedlings 190

V-7-Discussion

- V-7-1-Structure and evolution of a new subgroup of R2R3-MYB genes specific to conifers pine and spruce 196
- V-7-2-Constitutive and tissue-preferential overexpression as a robust strategy for the functional analysis of transcription factors 198
- V-7-3-PtMYB14 as a putative partner in isoprenoid-oriented response leading to terpene accumulation 201 202

V-8-Références

CHAPITRE VI

ÉTUDE DE L'EXPRESSION DE PTMYB14 DANS LES TISSUS **D'ÉPINETTES** VASCULAIRES **BLANCHES** SOUS LE CONTROLE DU PROMOTEUR DU GÈNE DE L'ALCOOL CINNAMYLIOUE DESHYDROGENASE (CAD)

VI-1-Avant-propos	208
VI-2-Résumé	209
VI-3-Introduction et mise en contexte	210
VI-4-Matériels et méthodes	210
VI-4-1-Isolement de 1,2 Kb de région génomique promotrice du gène	e CAD
d'épinette blanche	210
VI-4-2-Détermination du site d'initiation de la transcription dans le pro-	moteur
<i>CAD</i> par RACE-5'	211
VI-4-3-Construction du plasmide vecteur d'expression et matériel végét	al 212
VI-4-4-Coloration histochimique et tests enzymatiques de la ß-glucuronida	lse 213
VI-4-5-Analyse de la qualité des fibres et analyses chimiques du bois	214
VI-4-6-Isolement de l'ARN et analyses de l'expression des gènes pa	r PCR
quantitative	214
VI-4-7-Expériences microarray sur des épinettes « Pro _{CAD} PtMYB1-	4 » de
troisième cycle de croissance et analyses statistiques	216
VI-4-8-Application de jasmonate et blessure d'épinettes de deuxième c	ycle de
croissance non transgéniques	217

195

208

VI-5-Résultats	218
VI-5-1-Isolement et analyse des séquences en amont (5') du gèr	ne CAD
d'épinette blanche (présumé promoteur du gène CAD, ProCAD)	218
VI-5-2-Phénotypes, niveaux d'expression et activités enzymatiques	de la β-
glucuronidase chez les arbres exprimant la construction mol	éculaire
$Pro_{CAD}GUS$	220
VI-5-3-Le promoteur du gène CAD dirige l'expression dans le xyl	ème en
différenciation	222
VI-5-4-Le gène CAD est inductible par la blessure et l'acide jasmoniqu	ie 226
VI-5-5-Caractérisation des épinettes « Pro _{CAD} PtMYB14 » en troisièn	ne cycle
de croissance	228
VI-5-6-Analyse du profil transcriptomique et identification des	gènes
deregules chez les arbres « Pro _{CAD} PtMYB14 »	230
VI-5-/-Etude de l'expression de genes lies au stress oxydatif	et aux
mecanismes de defense chez des plants non transformes ble	sses ou
traites par l'acide jasmonique	238
VI-/-Conclusions et perspectives	240
v 1-8-Keierences	241
CHAPITRE VII	243
DISCUSSION GÉNÉRALE ET PERSPECTIVES	243
VII-1-Évolution et dénombrement des facteurs de transcription MYB-	R <i>2R3</i> de
VII-1-Évolution et dénombrement des facteurs de transcription <i>MYB-R</i> conifères	2 <i>R3</i> de 243
VII-1-Évolution et dénombrement des facteurs de transcription <i>MYB-F</i> conifères VII-1-1-Les séquences codantes	2773 de 243 243
 VII-1-Évolution et dénombrement des facteurs de transcription MYB-F conifères VII-1-1-Les séquences codantes VII-1-2-Les séquences non codantes 	22R3 de 243 243 244
 VII-1-Évolution et dénombrement des facteurs de transcription MYB-R conifères VII-1-1-Les séquences codantes VII-1-2-Les séquences non codantes VII-1-3-Le cas du sous-groupe 4 de conifères (Sg4-C) 	22R3 de 243 243 244 246
 VII-1-Évolution et dénombrement des facteurs de transcription MYB-A conifères VII-1-1-Les séquences codantes VII-1-2-Les séquences non codantes VII-1-3-Le cas du sous-groupe 4 de conifères (Sg4-C) VII-1-4-Représentation de l'ensemble des MYB chez l'épinette blanch 	22R3 de 243 243 244 246 ne <i>Picea</i>
 VII-1-Évolution et dénombrement des facteurs de transcription MYB-R conifères VII-1-1-Les séquences codantes VII-1-2-Les séquences non codantes VII-1-3-Le cas du sous-groupe 4 de conifères (Sg4-C) VII-1-4-Représentation de l'ensemble des MYB chez l'épinette blanch glauca 	22R3 de 243 243 244 246 ne <i>Picea</i> 246
 VII-1-Évolution et dénombrement des facteurs de transcription MYB-R conifères VII-1-1-Les séquences codantes VII-1-2-Les séquences non codantes VII-1-3-Le cas du sous-groupe 4 de conifères (Sg4-C) VII-1-4-Représentation de l'ensemble des MYB chez l'épinette blanch glauca VII-2-Démarches expérimentales pour la caractérisation des M 	R2R3 de 243 243 244 246 ae <i>Picea</i> 246 YB de
 VII-1-Évolution et dénombrement des facteurs de transcription MYB-R conifères VII-1-1-Les séquences codantes VII-1-2-Les séquences non codantes VII-1-3-Le cas du sous-groupe 4 de conifères (Sg4-C) VII-1-4-Représentation de l'ensemble des MYB chez l'épinette blanch glauca VII-2-Démarches expérimentales pour la caractérisation des M conifères 	R2R3 de 243 243 244 246 e <i>Picea</i> 246 YB de 248
 VII-1-Évolution et dénombrement des facteurs de transcription MYB-R conifères VII-1-1-Les séquences codantes VII-1-2-Les séquences non codantes VII-1-3-Le cas du sous-groupe 4 de conifères (Sg4-C) VII-1-4-Représentation de l'ensemble des MYB chez l'épinette blanch glauca VII-2-Démarches expérimentales pour la caractérisation des M conifères VII-3-Rôles biologiques des facteurs de transcription MYB-R2R3 de conifères 	22R3 de 243 243 244 246 ae <i>Picea</i> 246 YB de 248 res 250
 VII-1-Évolution et dénombrement des facteurs de transcription MYB-R conifères VII-1-1-Les séquences codantes VII-1-2-Les séquences non codantes VII-1-3-Le cas du sous-groupe 4 de conifères (Sg4-C) VII-1-4-Représentation de l'ensemble des MYB chez l'épinette blanch glauca VII-2-Démarches expérimentales pour la caractérisation des M conifères VII-3-Rôles biologiques des facteurs de transcription MYB-R2R3 de conifères VII-3-1-Coordination transcriptionnelle de l'expression des gènes 	R2R3 de 243 244 246 ae <i>Picea</i> 246 YB de 248 res 250 pour la
 VII-1-Évolution et dénombrement des facteurs de transcription MYB-R conifères VII-1-1-Les séquences codantes VII-1-2-Les séquences non codantes VII-1-3-Le cas du sous-groupe 4 de conifères (Sg4-C) VII-1-4-Représentation de l'ensemble des MYB chez l'épinette blanch glauca VII-2-Démarches expérimentales pour la caractérisation des M conifères VII-3-Rôles biologiques des facteurs de transcription MYB-R2R3 de conifères VII-3-1-Coordination transcriptionnelle de l'expression des gènes formation de la paroi secondaire 	R2R3 de 243 244 246 ae <i>Picea</i> 246 YB de 248 res 250 pour la 250
 VII-1-Évolution et dénombrement des facteurs de transcription MYB-R conifères VII-1-1-Les séquences codantes VII-1-2-Les séquences non codantes VII-1-3-Le cas du sous-groupe 4 de conifères (Sg4-C) VII-1-4-Représentation de l'ensemble des MYB chez l'épinette blanch glauca VII-2-Démarches expérimentales pour la caractérisation des M conifères VII-3-Rôles biologiques des facteurs de transcription MYB-R2R3 de conifères VII-3-1-Coordination transcriptionnelle de l'expression des gènes formation de la paroi secondaire VII-3-2-Complexité et interdépendance des mécanismes de défens 	22R3 de 243 243 244 246 ae <i>Picea</i> 246 YB de 248 res 250 pour la 250 e et de
 VII-1-Évolution et dénombrement des facteurs de transcription MYB-R conifères VII-1-1-Les séquences codantes VII-1-2-Les séquences non codantes VII-1-3-Le cas du sous-groupe 4 de conifères (Sg4-C) VII-1-4-Représentation de l'ensemble des MYB chez l'épinette blanch glauca VII-2-Démarches expérimentales pour la caractérisation des M conifères VII-3-1-Coordination transcriptionnelle de l'expression des gènes formation de la paroi secondaire VII-3-2-Complexité et interdépendance des mécanismes de défens réponse aux stress chez les conifères : le cas de PtMYB14 	22R3 de 243 243 244 246 Picea 246 YB de 248 res 250 pour la 250 e et de 252
 VII-1-Évolution et dénombrement des facteurs de transcription MYB-R conifères VII-1-1-Les séquences codantes VII-1-2-Les séquences non codantes VII-1-3-Le cas du sous-groupe 4 de conifères (Sg4-C) VII-1-4-Représentation de l'ensemble des MYB chez l'épinette blanch glauca VII-2-Démarches expérimentales pour la caractérisation des M conifères VII-3-Rôles biologiques des facteurs de transcription MYB-R2R3 de conifère VII-3-1-Coordination transcriptionnelle de l'expression des gènes formation de la paroi secondaire VII-3-2-Complexité et interdépendance des mécanismes de défens réponse aux stress chez les conifères : le cas de PtMYB14 VII-4-Surexpression des facteurs de transcription pour étudier la for du hair e Arabus aritiere de l'expression 	22R3 de 243 244 246 246 246 YB de 248 res 250 pour la 250 e et de 252 rmation
 VII-1-Évolution et dénombrement des facteurs de transcription MYB-R conifères VII-1-1-Les séquences codantes VII-1-2-Les séquences non codantes VII-1-2-Les séquences non codantes VII-1-3-Le cas du sous-groupe 4 de conifères (Sg4-C) VII-1-4-Représentation de l'ensemble des MYB chez l'épinette blanch glauca VII-2-Démarches expérimentales pour la caractérisation des M conifères VII-3-Rôles biologiques des facteurs de transcription MYB-R2R3 de conifères VII-3-Coordination transcriptionnelle de l'expression des gènes formation de la paroi secondaire VII-3-2-Complexité et interdépendance des mécanismes de défens réponse aux stress chez les conifères : le cas de PtMYB14 VII-4-Surexpression des facteurs de transcription pour étudier la for du bois : Analyse critique de l'approche 	22R3 de 243 244 246 246 Picea 246 YB de 248 res 250 pour la 250 e et de 252 rmation 257
 VII-1-Évolution et dénombrement des facteurs de transcription MYB-R conifères VII-1-1-Les séquences codantes VII-1-2-Les séquences non codantes VII-1-3-Le cas du sous-groupe 4 de conifères (Sg4-C) VII-1-4-Représentation de l'ensemble des MYB chez l'épinette blanch glauca VII-2-Démarches expérimentales pour la caractérisation des M conifères VII-3-Rôles biologiques des facteurs de transcription MYB-R2R3 de conifér VII-3-1-Coordination transcriptionnelle de l'expression des gènes formation de la paroi secondaire VII-3-2-Complexité et interdépendance des mécanismes de défens réponse aux stress chez les conifères : le cas de PtMYB14 VII-4-Surexpression des facteurs de transcription pour étudier la for du bois : Analyse critique de l'approche VII-5-Les gènes de la famille MYB-R2R3: une nouvelle source de condidate pour l'amédiant de la canalité du bais et un nouvelle 	22R3 de 243 243 244 246 Picea 246 YB de 248 res 250 pour la 250 e et de 252 rmation 257 e gènes
 VII-1-Évolution et dénombrement des facteurs de transcription MYB-R conifères VII-1-1-Les séquences codantes VII-1-2-Les séquences non codantes VII-1-3-Le cas du sous-groupe 4 de conifères (Sg4-C) VII-1-4-Représentation de l'ensemble des MYB chez l'épinette blanch glauca VII-2-Démarches expérimentales pour la caractérisation des M conifères VII-3-Rôles biologiques des facteurs de transcription MYB-R2R3 de conifèr VII-3-1-Coordination transcriptionnelle de l'expression des gènes formation de la paroi secondaire VII-3-2-Complexité et interdépendance des mécanismes de défens réponse aux stress chez les conifères : le cas de PtMYB14 VII-4-Surexpression des facteurs de transcription pour étudier la for du bois : Analyse critique de l'approche VII-5-Les gènes de la famille MYB-R2R3: une nouvelle source de candidats pour l'amélioration de la qualité du bois et un nouv nouv nouve la conservation des farêts neturelles? 	22R3 de 243 244 246 246 246 77B de 248 77B de 248 77B de 248 77B de 248 77B de 248 750 250 pour la 250 e et de 252 rmation 257 e gènes rel outil 260
 VII-1-Évolution et dénombrement des facteurs de transcription MYB-A conifères VII-1-1-Les séquences codantes VII-1-2-Les séquences non codantes VII-1-3-Le cas du sous-groupe 4 de conifères (Sg4-C) VII-1-4-Représentation de l'ensemble des MYB chez l'épinette blanch glauca VII-2-Démarches expérimentales pour la caractérisation des M conifères VII-3-Rôles biologiques des facteurs de transcription MYB-R2R3 de conifères VII-3-1-Coordination transcriptionnelle de l'expression des gènes formation de la paroi secondaire VII-3-2-Complexité et interdépendance des mécanismes de défens réponse aux stress chez les conifères : le cas de PtMYB14 VII-4-Surexpression des facteurs de transcription pour étudier la for du bois : Analyse critique de l'approche VII-5-Les gènes de la famille MYB-R2R3: une nouvelle source de candidats pour l'amélioration de la qualité du bois et un nouv pour la conservation des forêts naturelles? 	22R3 de 243 244 246 246 246 778 de 248 res 250 pour la 250 e et de 252 rmation 257 e gènes rel outil 260 261
 VII-1-Évolution et dénombrement des facteurs de transcription <i>MYB-A</i> conifères VII-1-1-Les séquences codantes VII-1-2-Les séquences non codantes VII-1-3-Le cas du sous-groupe 4 de conifères (Sg4-C) VII-1-4-Représentation de l'ensemble des MYB chez l'épinette blanch glauca VII-2-Démarches expérimentales pour la caractérisation des <i>M</i> conifères VII-3-Rôles biologiques des facteurs de transcription MYB-R2R3 de conifère VII-3-Coordination transcriptionnelle de l'expression des gènes formation de la paroi secondaire VII-3-2-Complexité et interdépendance des mécanismes de défens réponse aux stress chez les conifères : le cas de PtMYB14 VII-4-Surexpression des facteurs de transcription pour étudier la for du bois : Analyse critique de l'approche VII-5-Les gènes de la famille <i>MYB-R2R3</i>: une nouvelle source de candidats pour l'amélioration de la qualité du bois et un nouv pour la conservation des forêts naturelles? 	22R3 de 243 243 244 246 Picea 246 YB de 248 res 250 pour la 250 e et de 252 rmation 257 e gènes rel outil 260 261 262
 VII-1-Évolution et dénombrement des facteurs de transcription <i>MYB-A</i> conifères VII-1-1-Les séquences codantes VII-1-2-Les séquences non codantes VII-1-3-Le cas du sous-groupe 4 de conifères (Sg4-C) VII-1-4-Représentation de l'ensemble des MYB chez l'épinette blanch glauca VII-2-Démarches expérimentales pour la caractérisation des <i>M</i> conifères VII-3-Rôles biologiques des facteurs de transcription MYB-R2R3 de conifère VII-3-Coordination transcriptionnelle de l'expression des gènes formation de la paroi secondaire VII-3-2-Complexité et interdépendance des mécanismes de défens réponse aux stress chez les conifères : le cas de PtMYB14 VII-4-Surexpression des facteurs de transcription pour étudier la for du bois : Analyse critique de l'approche VII-5-Les gènes de la famille <i>MYB-R2R3</i>: une nouvelle source de candidats pour l'amélioration de la qualité du bois et un nouv pour la conservation des forêts naturelles? VII-5-1-Des gènes candidats pour l'amélioration de la qualité du bois VII-5-2-Des gènes candidats pour la conservation des forêts naturelles? 	22R3 de 243 244 246 246 246 77B de 248 res 250 pour la 250 e et de 252 rmation 257 e gènes rel outil 260 261 263 264

LISTE DES FIGURES

Figure I-1. La mise en place des tissus vasculaires chez les arbres: représentation schématique et histologie illustrant les croissances primaires et
secondaires 3
Figure I-2. La polarité des divisions nucléaires des initiales du cambium vasculaire 5
Figure I-3. Schéma d'une section transversale de la zone cambiale d'un conifère. 6
Figure I-4. Représentation tridimensionnelle du bois de conifère. 8
Figure I-5. Les interactions plante-environnement et la synthèse des composés
phénoliques. 16
Figure I-6. La voie de biosynthèse des phénylpropanoïdes et ses dérivés. 18
Figure I-7. Structure tridimensionnelle d'une protéine MYB-R2R3 interagissant avec
la molécule d'ADN. 30
Figure I-8. Représentation des domaines fonctionnels des protéines de type MYB typiques chez les animaux et chez les végétaux. 30
Figure I-9. Représentation schématique des interactions moléculaires entre des
protéines de type MYB et ses partenaires facteurs de transcription bHLH
et WD40. 41
Figure II-1. Alignment of predicted MYB domain protein sequences from spruce and pine 82
Figure II-2. Phylogenetic tree of gymnosperm and angiosperm R2R3-MYB proteins 86
Figure II-3. Transcript abundance for 13 spruce MYB genes and 4CL in various
organs and tissues.
Figure II-4. Transcript abundance for MYB genes and secondary cell-wall-related
genes in differentiating secondary xylem and in primary growth (new
flush) of spruce seedlings. 90
Figure II-5. Transcript accumulation for MYB genes and secondary cell-wall-related
genes in differentiating compression wood and opposite wood. 92
Figure III-1. Expression pattern of <i>PtMYB1</i> and <i>PtMYB8</i> in <i>Pinus taeda</i> .116
Figure III-2. Phenotypes induced by PtMYB1 overexpression in spruce.118
Figure III-3. Phenotypes induced by PtMYB8 overexpression in spruce. 120
Figure III-4. Lignin content and free phenolic profiles by induced PtMYB8
overexpression in spruce. 121
Figure III-5. Real time RI-PCK analysis of transcript accumulation in wild type,
PLIVEY BI and PLIVEY B& transgenic plantiels: validation of microarray and
wall assembly 120
Figure III-6 Endogenous spruce PaMVRs and PtMVR transcript accumulation in
wild type PtMVR1 and PtMVR8 transgenic plantlets 132
Figure IV-1 Phénotypes des épinettes transgéniques surexprimant <i>PtMVR1</i> sous
contrôle du promoteur <i>ubiavitine</i> (<i>ProuvePtMYR1</i>) 150
Figure IV-2. Sections transversales des pousses terminales des éninettes
« Prouble MYB1 ». 151
Figure IV-3. Niveaux d'expression du transgène <i>PtMYB1</i> et de l'endogène présumé
<i>PgMYB1</i> chez les arbres transgéniques « Pro _{UBI} PtMYB1 ». 152

- Figure IV-4. Quantification des transcrits de gènes reliés au métabolisme de la paroi et de l'azote chez les arbres transgéniques « Pro_{UBI}PtMYB1 ». 153
- Figure V-1. CLUSTAL W alignments of the predicted partial amino acids sequences of pine and spruce close to PtMYB14. 173
- Figure V-2. Aminoacids sequences analysis of conifers and angiosperms R2R3-MYB proteins identify a new motif specific to conifers in subgroug 4. 175
- Figure V-3. Tissue survey of RNA transcripts of *PtMYB14* sequences in *Pinus taeda*and *PgMYB14a* and *14b* in *Picea glauca*.176
- **Figure V-4.** Histochemical analysis of Pro_{UBL} -, and Pro_{CAD} driven GUS expression in root and hypocotyl of spruce transgenic plantlets. 179
- **Figure V-5.** Morphological and histological phenotypes induced by *Pro_{UBI}* and *Pro_{CAD}* driven overexpression of PtMYB14 in spruce somatic plantlets. 180
- Figure V-6. Accumulation of anthocyanins and starch in wild type and PtMYB14-OE transgenic spruce. 181
- **Figure V-7.** Number and functional grouping of genes differentially expressed in *Pro*_{UBI}*PtMYB14* and *Pro*_{CAD}*PtMYB14* overexpressing spruce plantlets. 189
- Figure V-8. Accumulation of terpenoid compounds in wild type and PtMYB14-OE trasngenic spruce. 192
- Figure V-9. Expression of R2R3-MYB, MVA- and DOX pathway related genes in response to wounding and jasmonic treatments in pine seedlings. 193
- Figure V-10. Expression of R2R3-MYB, MVA- and DOX pathway related genes in response to wounding and jasmonic treatments in spruce somatic plantlets. 194
- **Figure VI-1.** Séquence nucléotidique de 1,2 Kb en amont (5') du gène codant pour l'alcool cinnamylique deshydrogénase (*CAD*) de *Picea glauca* représentant la région promotrice potentielle de ce même gène. 219
- **Figure VI-2.** Phénotypes, accumulation des transcrits et activités enzymatiques GUS chez les épinettes transgéniques exprimant la construction moléculaire *Pro_{CAD}GUS*. 221
- Figure VI-3. Niveaux d'expression du gène *CAD* de *Picea glauca* dans différents tissus et organes d'une épinette blanche de 30 ans. 222
- **Figure VI-4.** Localisation histochimique de l'activité GUS pendant la formation du bois chez des épinettes blanches transformées avec *Pro_{CAD}GUS* (sections transversales à main levée). 223
- **Figure VI-5.** Localisation histochimique de l'activité GUS pendant la formation du bois chez des épinettes blanches transformées avec *Pro_{CAD}GUS* (sections transversales "déparaffinées"). 224
- **Figure VI-6.** Localisation histochimique de l'activité GUS dans les racines et les aiguilles chez des épinettes blanches transformées avec *Pro_{CAD}GUS*. 225
- Figure VI-7. Localisation histochimique de l'activité GUS lors d'une blessure dans
le tronc d'épinette blanche « Pro_{CAD}GUS ».226
- **Figure VI-8.** Niveaux d'expression du gène *CAD* dans les pousses terminales avec aiguilles et le tronc après blessure mécanique et pulvérisation d'acide jasmonique. 227
- **Figure VI-9.** Croissance et morphologie des épinettes blanches transformées avec la construction $Pro_{CAD}PtMYB14$, en troisième cycle de croissance. 228
- Figure VI-10. Analyse par microarrays des épinettes « Pro_{CAD}PtMYB14 » en troisième cycle de croissance. 231

- **Figure VI-11.** Niveaux d'expression de plusieurs gènes dans différents tissus et organes d'épinettes blanches « Pro_{CAD}PtMYB14 » de troisième cycle de croissance. 235
- Figure VI-12. Niveaux d'expression de plusieurs gènes dans les pousses terminales d'épinettes blanches « Pro_{CAD}PtMYB14 » de troisième cycle de croissance. 236
- Figure VI-13. Niveaux d'expression de plusieurs gènes dans les pousses terminales
avec aiguilles et dans la tige principale (tronc) après 24 heures
d'induction à l'acide jasmonique et à la blessure.239

LISTE DES TABLEAUX

Tableau I-1. Les principales classes de composés phénoliques chez les plantes.19

- **Tableau I-2.** Exemples d'expériences de perte et de gain de fonction par
transformation génétique utilisant les gènes de la voie de biosynthèse
des phénylpropanoïdes et des lignines.23
- Tableau I-3. Protéines à domaine MYB-R2R3 caractérisées fonctionnellement chez

 A. thaliana.
 36
- Tableau I-4. Protéines MYB atypiques caractérisées fonctionnellement chez A.
thaliana.38
- **Table II-1.** Predicted lengths and C-terminal motifs of spruce MYB proteins.80
- Table II-2. Length of spruce MYB coding sequences and introns with their predicted splice junctions.
 85
- **Table II-3.** Pair-wise sequence amino acids identities of the DBD and full CDS of
closest spruce and pine homologs.87
- Table III-1.Selected genes differentially expressed in PtMYB1 and PtMYB8
transgenic spruce.128
- **Tableau IV-1.** Hauteurs et diamètres des épinettes transgéniques exprimant la
construction *Pro_{UBI}PtMYB1*.149
- **Table V-1.** Overlapping set of genes differentially expressed (P value < 0.01, and Log_2 ratio $\geq |0.5|$) following both Pro_{UBI} and Pro_{CAD} -driven overexpression of PtMYB14 in spruce. 188
- **Tableau VI-1.** Analyse de la longueur et de la largeur des fibres chez les arbres
transgéniques « Pro_{CAD}PtMYB14 ».229
- **Tableau VI-2.** Contenu en lignine du bois des arbres « Pro_{CAD}PtMYB14 », déterminé
par la méthode Klason.229
- Tableau VI-3. Liste des gènes les plus fortement dérégulés et communs entre les pousses terminales, le bois de compression, le bois entier écorcé d'épinettes blanches « Pro_{CAD}PtMYB14 » de troisième cycle de croissance et les plantules « Pro_{UBI} » et « Pro_{CAD}PtMYB14 ». 232

LISTE DES ANNEXES

Annexe II-1. Protocole d'extractions des ARN messagers.	270
Appendix II-2. Primers sequences of spruce MYBs used for genomic amplificat	tion
and sequencing.	271
Appendix II-3. Phylogenetic tree of MYBs from spruce, pine and nearest sequer	ices
from other species.	272
Appendix II-4. Conifer MYB phylogeny based on partial sequences using a	31
amino acids region.	273
Appendix III-1. Primer sequences used for Real time RT-PCR (QPCR.)	274
Appendix III-2. Microarray manufacturing and quality control.	275
Appendix V-1. Primers sequences generating DNA matrix for qRT-P	' CR
amplification in pine.	277
Appendix V-2. Primers sequences used for RT-qPCR in pine and spruce.	278
Appendix V-3. Expression level (RT-qPCR) of PtMYB14, genes validate	ting
microarray and R2R3-MYB genes in hypocotyls of PtMYB14-	·ОЕ
plantlets, normalized with CDC2 transcripts.	279
Appendix V-4. Expression level (RT-qPCR) of PtMYB14, genes validated	ting
microarray and R2R3-MYB genes in roots of PtMYB14-OE plantl	lets,
normalized with CDC2 transcripts.	280
Annexe VI-1. Liste des gènes dérégulés dans les expériences PT, BE et BC chez	des
épinettes blanches de troisième cycle de croissance expriman	t la
construction $Pro_{CAD}PtMYB14$.	281
Annexe VII-1. Longueurs et pourcentages d'identités de séquences des introns	de de
quelques MYB-R2R3 d'épinettes et de pin.	286
Annexe VII-2 Alignement de séquences des premiers introns de six membres	du
groupe Sg4-C	287
Annexe VII-3. Alignement de séquences des seconds introns de six membres	du
groupe Sg4-C.	288
Annexe VII-4. Alignements de séquences des introns comparant Pt/PgMYB1, 3, 4 et 8.	289

ABRÉVIATIONS

ADN, acide désoxyribonucléique ARN, acide ribonucléique ARNm, acide ribonucléique messager bHLH, « basic helix-loop-helix protein » DBD, « DNA-binding domain » EST, « expressed sequence tag » GC, gène candidat GUS, béta-glucuronidase MYB, myéloblastose NOS, nopaline synthase pb, paire de base PCR, « polymerase chain reaction » PgMYB, gène MYB de Picea glauca Pro_{CAD}, promoteur du gène codant l'alcool cinnamylique deshydrogenase de P. glauca Pro_{UBI}, promoteur du gène codant l'ubiquitine de maïs PtMYB, gène MYB de Pinus taeda RACE-5', « 5'-rapid amplification of cDNA ends » RT-qPCR, « reverse transcriptase quantitative PCR) Sg4-C, sous-groupe 4 de conifère TF, « transcription factor » UTR-3', région transcrite non traduite de l'extrémité 3'

CHAPITRE I

Introduction générale

L'étude des processus biologiques liés à la formation du bois, c'est à dire la croissance du xylème secondaire, sous-jacente à l'expansion radiale du tronc de l'arbre, comporte de nombreux aspects autant dans la différenciation des tissus vasculaires mis en place que dans la structure et la composition du bois qui en résulte. La caractérisation des gènes et des protéines reliés à ces processus a connu un développement considérable ces dernières années, grâce notamment aux avancées permises par la plante modèle herbacée Arabidopsis thaliana utilisée en génétique végétale. De part son génome séquencé (Arabidopsis Genome Initiative, 2000), son très court cycle de reproduction (quelques semaines), la mise en place de banques de mutants et la simplicité des procédures pour sa transformation génétique, entre autres, Arabidopsis est effectivement une plante modèle. De nombreuses découvertes récentes sur la différenciation du système vasculaire chez Arabidopsis peuvent être étendues aux autres espèces végétales. Néanmoins, des différences importantes existent entre Arabidopsis et les arbres forestiers, particulièrement chez les conifères. D'une part, l'anatomie de leurs tissus vasculaires est très différente, et d'autre part, les arbres se caractérisent par la formation d'une croissance très importante du xylème secondaire, minimale chez Arabidopsis. Enfin, considérant la grande distance phylogénétique entre Arabidopsis et les conifères (angiosperme versus gymnosperme), il est pertinent de poursuivre l'étude de la formation du bois chez ces derniers si l'on désire générer des connaissances plus spécifiques ou des retombées appliquées. Par ailleurs, les ressources et les méthodes requises pour l'étude moléculaire et génétique de la formation du bois chez les conifères ont connu un développement récent important, particulièrement chez l'épinette blanche *Picea glauca* [Moench] Voss de la famille des *Pinaceae*.

I-1-La formation du bois

I-1-1-Origine et mise en place des tissus vasculaires chez les plantes

Chez les plantes, la formation des organes et des tissus résulte de l'activité des méristèmes. Les méristèmes apicaux, ou méristèmes primaires, apparaissent en premier au cours de l'embryogenèse et assurent la formation de cellules plus spécialisées comme les cellules du procambium ou tissus provasculaires (Esau, 1965). Localisés aux extrémités des organes en croissance (tiges, racines), ces tissus méristématiques sont constitués de cellules indifférenciées qui donnent naissance aux tissus primaires puis aux méristèmes secondaires (cambium vasculaire et phellogène) (Figure I-1). Selon une définition structurale, le procambium, est identifiable pendant le développement de l'organe comme un feuillet continu de cellules étroites, denses en cytoplasme et avec leur axe orienté parallèlement au feuillet lui même (Nelson et Dengler, 1997). Les cellules procambiales et leurs initiales (cellules préprocambiales) se divisent préférentiellement de façon parallèle au développement du feuillet vasculaire qui lui même se positionne parallèlement à la croissance de l'organe. Les cellules préprocambiales sont de petites cellules indifférenciées polygonales et isodiamétriques qui se divisent intensément. Les sites pour l'initiation des cellules procambiales déterminent le profil de l'organisation vasculaire et leur activité contrôle la différenciation des tissus vasculaires xylème-phloème.

I-1-2-Mise en place des tissus vasculaire chez les ligneux pérennes

Le système vasculaire d'un arbre se met en place pendant la phase d'élongation des extrémités des tiges et des branches à partir des méristèmes apicaux (croissance primaire) et se poursuit pendant la croissance radiale du tronc et des branches à partir d'un méristème latéral appelé cambium vasculaire (croissance secondaire) (Figure I-1). L'organisation des tissus primaires formés au cours de l'élongation de la tige est conservée lors de cette croissance dite secondaire et est à l'origine de l'établissement du cambium vasculaire secondaire. La suite de cette synthèse bibliographique mettra l'emphase sur la croissance secondaire.



Figure I-1. La mise en place des tissus vasculaires chez les arbres: représentation schématique et histologie illustrant les croissances primaires et secondaires (adapté de Colonna, 2006).

Voir légende à la page suivate.

Figure I-1. M: moelle; 1 et 2 représentent respectivement la première et la deuxième année de croissance de l'arbre; ZC: zone cambiale; Pa: parenchyme; Ph: phloème; Xy: xylème; CR: conduit résinifère; R: rayon. Les deux photographies sont des sections transversales A) d'une pousse terminale d'épinette blanche en fin de cycle de croissance montrant l'organisation des faisceaux vasculaires xylème et phloème et B) de la base d'un tronc d'épinette blanche en fin de troisième cycle de croissance montrant une croissance radiale vers l'intérieur de la tige. Les deux sections ont été colorées selon la procédure de la safranine-orangée de Sharman (1943) pour mettre en évidence les cellules de xylème aux parois lignifiées en orange (10 μ m d'épaisseur, grossissement x40).

I-1-2-1-Le cambium vasculaire

Chez les dicotylédones (angiospermes) et les gymnospermes, la croissance en diamètre des tiges et des branches est donc assurée par un méristème secondaire, le cambium vasculaire ou zone génératrice libéro-ligneuse, qui se met en place en une assise continue entre les faisceaux de xylème et de phloème primaires. Ce cambium produit un xylème secondaire vers l'intérieur de la tige et un phloème secondaire, vers l'extérieur. Chez les plantes pérennes des régions tempérées, la croissance du cambium est inactive pendant la saison de dormance (période hivernale) et redevient fonctionnelle au printemps. Au fil des cycles de croissance l'accumulation de couches, ou cernes, de tissus de xylème lignifié ou xylème secondaire constituera le bois (Figure I-1). Le terme cambium est également utilisé pour désigner le phellogène qui est le cambium de l'écorce (Esau, 1977), mais dans cette étude je me focaliserai sur le cambium vasculaire en raison de son importance pour la formation du bois.

Le cambium vasculaire est un méristème latéral cylindrique organisé en files radiales contenant des cellules initiales de type méristématique et des cellules-mères capables de se différencier en cellules du xylème ou du phloème. L'ensemble est appelé zone cambiale, puisqu'il y a très peu de différences morphologiques qui permettent de distinguer les deux types cellulaires. Il existe trois types de divisions cellulaires à l'origine de l'activité et du renouvellement du cambium : les divisions périclines, pseudotransverses et anticlines (Figure I-2).

Les divisions périclines donnent des cellules mères du xylème ou du phloème respectivement à l'intérieur et l'extérieur du cylindre cambial. De ce fait, l'accumulation de cellules de xylème de façon centripète maintient le cambium à la périphérie du tronc et nécessite un accroissement en diamètre du cambium réalisé par des divisions pseudotransverses produisant des cellules filles identiques aux cellules-mères. Finalement, les divisions anticlines permettent au cambium d'accompagner la croissance de l'arbre en hauteur.





Division péricline: croissance en diamètre du cambium, division pseudotransverse: augmentation de la circonférence cambiale, anticline: élongation du tronc de cylindre cambial et croissance en longueur des tissus en stade primaire. La croissance anticline positionne une partie des cellules l'une devant l'autre, on parle de croissance intrusive.

Le cambium est composé de deux types cellulaires appelés respectivement : initiales fusiformes et initiales de rayon. Les initiales fusiformes de forme allongée en section tangentielle sont majoritaires. Elles se différencient dans un système axial orienté longitudinalement (vertical) aboutissant à la formation de cellules mères du xylème ou du phloème (Koslowski, 1971). Ces cellules mères vont se différencier en plusieurs types cellulaires. Dans le xylème secondaire, elles sont à l'origine du parenchyme axial, les fibres et les vaisseaux. Dans le phloème secondaire, elles vont se différencier en cellules du parenchyme axial, en fibres, en tubes criblés et en cellules-compagnes. Les initiales de rayon plus courtes et isodiamétriques se différencient en parenchyme de rayon de façon transversale dans l'axe horizontal ou radial (Chaffey, 2002) (Figure I-3). Les cellules des rayons permettent le stockage de métabolites divers et leur translocation entre le xylème et le phloème via des plasmodesmes.



Figure I-3. Schéma d'une section transversale de la zone cambiale d'un conifère (Plomion et *al.*, 2001).

Le schéma montre les cellules initiales fusiformes (F) et de rayon (R) dans la zone cambiale (ZC). X: différentiation centripète du xylème, les cellules sont en cours de croissance radiale, en maturation et mature. P: différentiation centrifuge du phloème de la croissance radiale à la cellule mature. La flèche pleine indique l'emplacement d'une paroi tangentielle nouvellement déposée (division péricline) et la flèche vide une nouvelle paroi radiale (division anticline).

I-1-2-2-Le xylème

Le xylème conduit la sève minérale, c'est à dire l'eau et les substances minérales dissoutes prélevées dans le sol, à travers les racines, le tronc et les branches vers l'apex et les sites d'évapotranspiration. De plus, le xylème transporte des hormones telles que l'auxine, l'acide abscisique et les cytokinines à l'intérieur de l'arbre ou de la plante (Hartung et *al.*, 2002 ; Haberer et Kieber, 2002). Les cellules du xylème contribuent au soutien mécanique et permettent le port dressé de l'arbre. Les deux principaux types de cellules conductrices du xylème sont les trachéides chez les gymnospermes et les éléments de vaisseaux chez les angiospermes. Les vaisseaux sont larges et courts avec des extrémités ouvertes. Les trachéides sont longues (2 à 4 millimètres chez l'arbre adulte), étroites (20 à 50 micromètres). Leurs extrémités sont closes et elles communiquent les unes avec les autres *via* des pores ponctués (Figure I-4). Le bois de conifère est une structure relativement homogène, dont la fraction longitudinale est constituée à 90% de trachéides (mortes à maturité) et de cellules vivantes de parenchyme longitudinal ainsi que des conduits résinifères dans certains cas (Figure I-4).

I-1-2-3-Le phloème

Le phloème est un tissu vasculaire complexe de l'écorce qui conduit la sève dite élaborée ou nutritive, riche en sucres, des organes photosynthétiques vers les autres organes pour assurer leur croissance ou permettre la mise en réserve (Figure I-4). Le phloème est composé de tubes criblés constitués de cellules allongées, vivantes mais ayant perdu leur noyau, dont les cloisons transversales sont perforées et au travers desquelles circule la sève. Entre ces cellules se trouvent des cellules compagnes plus petites, vivantes et nucléées, et supposées participer au contrôle des échanges entre tubes criblés et organes végétaux. On retrouve également des cellules lignifiées, les sclérides, et des cellules parenchymateuses pouvant accumuler des composés organiques (sucres, lipides, protéines). De plus, le phloème fournit la voie pour la translocation des peptides, des protéines, et des ARNm impliqués dans la croissance et le développement de la plante ainsi que dans la défense contre les pathogènes (Citovsky et Zambryski, 2000; Oparka et Cruz, 2000; Ruiz-Medrano et al., 2001; van Bel et al., 2002).





La sève minérale monte à travers le xylème et la sève nutritive descend via le phloème. La sève peut également être transportée *via* les cellules des rayons et tangentiellement par les pores ponctués.

I-1-2-4-Le collenchyme et le sclérenchyme

En ce qui concerne les tissus de soutien, outre le xylème, on trouve des tissus formés de cellules vivantes ou mortes à parois très épaissies. Il s'agit du collenchyme dans les axes jeunes, constitué de cellules vivantes avec des parois cellulaires cellulosiques, et le sclérenchyme (fibres et sclérides), des cellules mortes aux parois secondaires lignifiées situées dans les organes dont l'allongement est terminé.

I-1-3-La xylogénèse

La formation du bois ou xylogénèse dépend de la différenciation des cellules du xylème secondaire à partir du cambium vasculaire. La différenciation des cellules conductrices et de support du xylème, comporte de nombreux événements biochimiques, morphologiques et génétiques. Ainsi, quatre étapes majeures peuvent être distinguées : la division et l'expansion cellulaire, suivie de la formation de la paroi cellulaire secondaire épaisse et de la mort programmée de la cellule (Fukuda, 1996; Ye, 2002; Turner et *al.*, 2007). Les parois secondaires des cellules du xylème sont constituées de cellulose et d'hémicellulose, et de lignine, composé phénolique tridimensionnel complexe.

I-1-3-1-Élongation cellulaire et paroi primaire

L'initiation de la différenciation du xylème secondaire commence avec la division péricline des initiales cambiales vers l'intérieur du cylindre cambial donnant lieu à une cellule-mère du xylème. Ensuite les cellules filles générées se différencient, c'est à dire s'allongent et s'élargissent. A ce stade de différenciation, la paroi cellulaire est constituée de:

-la lamelle moyenne, située entre deux cellules voisines, riche en pectine acide contenant peu de cellulose et conférant un caractère hydrophile aux cellules jeunes en élongation,

-la paroi primaire, souple et extensible, avec environ 75 à 80%, d'eau qui permet l'élongation cellulaire. Sa synthèse est initiée au niveau du phragmoplaste lors de la cytokinèse (Dickison, 2000). Elle est constituée d'une structure fibrillaire de cellulose enrobée dans une matrice polysaccharidique de pectines et d'hémicelluloses de type xyloglucanes par exemple (Cosgrove, 1997). Cette paroi contient 1 à 5% de protéines structurales comme les extensines, les arabinogalactane protéines (AGP), les glycoprotéines riches en hydroxyproline (HRGP), les protéines riches en glycine (GRP) ou en proline (PRP).

La cellulose est un polymère de glucose constitué de chaînes linéaires de beta-1,4-D-glucanes synthétisées et arrangées en microfibrilles par des complexes appelés rosettes composés de six unités, contenant chacune six cellulose synthases CesA (Doblin et *al.*, 2002). Les rosettes effectuent la synthèse de la cellulose au

niveau de la membrane plasmique. Le dépôt de ces microfibrilles est fait de manière orientée et peut changer au cours de la formation de la paroi secondaire (Brett, 2000). Plusieurs gènes CesA liés à la synthèse de la cellulose ont été identifiés par l'étude de mutants déficients en cellulose (Somerville, 2006).

Les hémicelluloses sont des polysaccharides homo- ou hétéropolymériques constitués d'une chaîne centrale de résidus beta-D-1-,4-pyranosyl avec de courtes chaînes latérales de résidus glycosyl (xylose, galactose et fucose). Certaines hémicelluloses peuvent être plus abondantes dans la paroi primaire que dans la paroi secondaire comme les xyloglucanes ou bien l'inverse dans le cas des xylanes par exemple. La biosynthèse des hémicelluloses est réalisée au niveau de l'appareil de Golgi par des enzymes de type « glycane synthase » et « glycosyltransférase ». Avant d'être intégrées à la paroi, les hémicelluloses sont transportées vers la paroi *via* des vésicules de sécrétion.

La pression interne de la cellule, ou turgescence, permet l'expansion longitudinale et radiale. De ce fait, l'élongation cellulaire nécessite le relâchement du réseau pariétal entre la cellulose et l'hémicellulose. Le processus de relâchement pariétal fait intervenir des enzymes comme les expansines, qui se lient à la paroi *via* la cellulose pour rompre et déplacer les liaisons non-covalentes entre les constituants de la paroi (Cosgrove, 1997; Cosgrove et *al.*, 2002), et des enzymes de lyse pariétale de type « xyloglucane endotransglycosylases » (XET) (Nishitani, 1997; Darley et *al.*, 2001; Bourquin et *al*;, 2002) et « endoglucanases » (EG) qui clivent aussi les xyloglucanes (Darley et *al.*, 2001), ou encore des « yieldines » (Okamoto-Nakazato et *al.*, 2002). Une fois la taille définitive de la cellule atteinte la paroi primaire subit une rigidification et l'expansion s'arrête. Ce phénomène impliquerait potentiellement des modifications de la composition pariétale, une augmentation des liaisons covalentes ou encore une diminution des processus de relâchement pariétal (Cosgrove, 1997; Darley et *al.*, 2001).

I-1-3-2-Formation de la paroi secondaire et lignification

Aprés l'élongation de la cellule, la paroi secondaire se met en place de façon dynamique et séquentielle de l'extérieur vers l'intérieur de l'élément trachéaire avec le dépôt de trois sous-couches (S1, S2 et S3) (Plomion et *al.*, 2001). La paroi secondaire est formée principalement de cellulose (40 à 50% du bois), d'hémicellulose (environ 25% de la paroi secondaire) et de lignines (25 à 35% du bois), puis de composés minoritaires tels que les pectines, les protéines structurales et les extractibles (Timell, 1969; Plomion et *al.*, 2006).

Le réseau cellulosique et hémicellulosique constituant la paroi secondaire est imprégné d'un autre polymère carboné qui joue un rôle d'agent de liaison : la lignine. Il s'agit d'un hétéropolymère phénolique complexe essentiel pour la structure de la paroi secondaire et formé de trois monomères de monolignols reliés entre eux par des processus d'oxydation (Boerjan et *al.*, 2003). La lignine imperméabilise la paroi et apporte de la rigidité à la cellule.

I-1-3-3-Mort cellulaire

La fin de la formation de la paroi secondaire s'enclenche avec le dépôt de lignines dans le maillage de microfibrilles de cellulose suivi de la mort cellulaire programmée (MCP). La mort cellulaire concerne spéciquement les trachéides (ou éléments trachéaires) et suit un déroulement identique pour toutes, ce qui indique que ce processus est régulé par un programme développemental. De plus, il s'agit d'un processus actif qui conduit la cellule à élever le niveau d'expression des gènes qui conduisent à sa propre destruction. Récemment, Nakaba et *al.*, (2006) ont montré que, chez le conifère *Abies sachalinensis*, la mort cellulaire des trachéides se produit successivement et est reliée à la distance par rapport au cambium.

La MCP commence avec la rupture du tonoplaste par autolyse entraînant la libération du contenu vacuolaire dans le cytoplasme (Kuriyama, 1999). Des enzymes hydrolytiques dont des cystéines protéases, sérines protéases et nucléases sont ensuite libérées pour digérer les autres organites cellulaires à paroi simple (appareil de Golgi et réticulum endoplasmique) puis double (chloroplastes, mitochondries et noyau) (Funk et *al.*, 2002; Bozhkov et *al.*, 2005; Ye et Warner, 1996; Ito et Fukuda, 2002). La digestion partielle des parois primaires aux extrémités cellulaires crée des perforations, ce qui permet l'évacuation du contenu cellulaire et donc le transport par les trachéides lignifiées interconnectées longitudinalement.

Les mécanismes biochimiques et moléculaires à l'origine de la synthèse des enzymes hydrolytiques et leur stockage ainsi que le processus d'autolyse de la vacuole sont encore méconnus. Des études suggèrent que le calcium et l'oxide nitreux pourraient être impliqués dans le déclenchement de la MCP (Groover et Jones, 1999; Gabaldon et *al.*, 2005). Au niveau de la synthèse des protéines, l'application d'un inhibiteur de la traduction (cycloheximide) bloque la MCP des trachéides chez *Zinnia elegans*, indiquant la nécessité de la synthèse de nouvelles protéines pour accomplir la MCP (Kuriyama et *al.*, 1999). La transition vers la phase de formation de la paroi secondaire et la MCP est également bloquée par l'utilisation d'inhibiteurs de la synthèse des brassinostéroides chez *Zinnia* (Groover et Jones, 1999). De plus, il semblerait que la MCP et la formation de paroi secondaire soient deux évènements indépendants puisque chez le mutant *gpx* d'*arabidopsis* la MCP se produit sans que la synthèse de la paroi cellulaire secondaire ait été complété (Turner et Hall, 2000).

I-1-4-Hétérogénéité du bois à l'intérieur de l'arbre

La croissance de l'arbre dans son environnement est déterminé par un programme génétique de développement mais est aussi fortement influencée par des contraintes biotiques (interaction ou attaque par des micro-organismes, insectes, animaux, végétation compétitrice) et abiotiques (température, vent, eau) qui entraînent des changements physiologiques, hormonaux et morphologiques affectant le développement du xylème. La morphologie des trachéides est déterminée par le taux de division des cellules de la zone cambiale et la durée des phases d'expansion et de formation de la paroi secondaire. Les changements dans la structure des cellules et dans la composition de leur paroi conduisent à la formation d'un bois hétérogène et variable d'une région à l'autre du tronc. Aussi on peut oberver jusqu'à huit types différents de bois chez un seul arbre (Plomion et *al.*, 2001; Leprovost et *al.*, 2003) en fonction:

- de la période à l'intérieur de la saison ou du cycle de croissance annuel : au début de la saison de croissance des arbres, le printemps, se forme le bois initial (ou précoce) et à la fin de la saison, le bois final (ou tardif). La transition entre les deux types de bois peut être subite ou graduelle. Le bois final est constitué d'éléments trachéaires aux parois plus épaisses que celles du bois initial, ce qui entraîne une forte augmentation de la densité. De plus, le bois final contient légèrement moins de lignines mais plus de cellulose que le bois initial.
- du stade de développement ou de l'âge de l'arbre : pendant les premières années de croissance de l'arbre (environ 8 ans pour le peuplier, 20 ans pour l'épinette), le bois est de type juvénile puis devient mature. Les phases juvénile et mature peuvent être séparées par une longue phase de transition pouvant représenter plusieurs années. Comparé au bois juvénile, le bois mature produit des trachéides beaucoup plus longues, ses cernes sont généralement plus étroits et il y a une nette diminution de l'angle des microfibrilles de cellulose.
- de stress mécaniques ou phototropiques qui déplacent l'axe de l'arbre par rapport à la verticale (vent, poids de la neige, recherche de lumière, etc.) : la réponse physiologique de l'arbre sera de redresser son tronc en produisant un type de bois particulier, dit bois de réaction. Chez les conifères, un bois dit de compression se forme sur la face inférieure des troncs alors que chez les angiospermes, le bois de tension est mis en place sur la partie supérieure. L'angle d'inclinaison influence le taux de production de bois de réaction. Le bois de compression est composé de trachéides plus courtes apparaissant arrondies en section transversale, les parois sont plus épaisses que celles du bois normal, contiennent plus de lignine et présentent un angle des microfibrilles plus ouvert (Timmel et *al.*, 1986, 1969). L'explication de ce phénomène reste encore méconnue mais plusieurs hypothèses associent la perception du stimulus gravitropique à des cellules contenant de nombreux amyloplastes, les statocytes (Blancaflor, 2002).

I-1-5-Les canaux résinifères

Récemment, Hudgins et Franceschi (2004) ont montré que la production d'éthylène induite par le méthyl jasmonate (MJ) est responsable de la mise en place de réactions de défense dans le phloème des conifères et de la nouvelle programmation de la zone cambiale de la tige pour former des canaux résinifères (Figure I-1 et I-4). En effet les cellules mères du xylème de la zone cambiale à l'origine des trachéides peuvent se différencier en cellules épithéliales sécrétrices de résine *via* l'augmentation d'éthylène (Lachaud et *al.*, 1999; Savidge, 2001). De plus la synthèse d'éthylène à travers l'application de MJ peut aussi affecter la lignification des sclérides présentes dans le phloème (Hudgins et *al.*, 2004). Chez *Arabidopsis*, Cano-Delgado et *al.* (2003) ont mis en évidence que des mutations dans les gènes de cellulose synthèse de cellulose mais activent la synthèse de lignine et les réponses de défense *via* les voies de signalisation du MJ et de l'éthylène.

Plusieurs conifères, les genres Pinus et Picea entre autres, contiennent des conduits résinifères qui produisent de la résine dans le phloème secondaire ou dans le xylème comme moyen de défense constitutive (Figure I-1 et I-4, Langenheim, 2003). Cependant, la formation de ces structures peut être induite après l'action d'herbivores, l'invasion de pathogènes ou encore la blessure; on parle alors de canaux résinifères traumatiques (CRT) (Franceschi et al., 1998, 2000; Nagy et al., 2000; Martin et al., 2002; Hudgins et al., 2003; Mckay et al., 2003; Luchi et al., 2005). La résine qui se forme dans les CRT peut avoir des propriétés plus toxiques ou anti-fongiques que les conduits résinifères constitutifs en raison de l'addition de composés phénoliques et d'un changement dans la composition du contenu en terpènes (Nagy et al., 2000; Fäldt et al., 2003; Krokene et al., 2003). La sécrétion de la résine permet de cicatriser localement la blessure, de circonscrire l'invasion du pathogène et le tuer, de décourager les herbivores, d'éloigner les insectes ou encore d'attirer leurs prédateurs. En réponse aux attaques de pathogènes, les pins produisent de la résine dans les canaux résinifères préformés (Christiansen et al., 1987) ainsi que dans les CRT axiaux nouvellement formés dans le xylème (Lombardero et al., 2000). De plus, des composés phénoliques s'accumulent constitutivement dans les vacuoles de quelques cellules du parenchyme du phloème secondaire dans la plupart des espèces de *Pinaceae*, on parle de cellules polyphénoliques (Franchesci et *al.*, 1998; Krekling et *al.*, 2000; Hudgins et *al.*, 2003).

La résine de conifères ou oléorésine, qu'elle soit constitutive ou induite, est composée d'un mélange liquide visqueux et odorant de plusieurs métabolites secondaires de type terpènes, les volatiles mono- (C_{10}) , sesqui- (C_{15}) et non volatiles di-terpénoïdes (C₂₀) (Trapp et Croteau, 2001). Les terpénoïdes sont issus des voies de biosynthèse du mévalonate cytosolique (MEV, Chappell, 1995; Bochar et al., 1999) et du méthyl-erythritol 4-phosphate plastidique (MEP), appelée aussi 1-deoxy-Dxylulose plastidique (DOX) (Bochar et al., 1999; Lichtenthaler, 1999). Les gènes et les enzymes impliqués dans la biosynthèse des terpènes chez les conifères ont été peu étudiés et aucun n'a été caractérisé fonctionnellement in planta à part quelques prényltransférases (Keeling et Bohlmann, 2006). La plupart des terpènes connus proviennent de structures synthétisées par des terpène synthases issues d'un des trois précurseurs de type prenyl diphosphate (Chappell, 1995; Dewick, 2002). Les monoterpènes, sesquiterpènes et diterpènes proviennent respectivement du geranyl diphosphate (GPP) (Wise et Croteau, 1999), du farnesyl diphosphate (FPP) (Cane, 1999) et du geranylgeranyl diphosphate (GGPP) (Davis et Croteau, 2000; Martin et al., 2004). Les trois précurseurs de type prényl diphosphate proviennent de la condensation du diméthylallyl diphosphate (DMAPP) avec un ou plusieurs isopentyl diphosphates (IPP), catalysée par des prényltransférases (Koyama et Ogura, 1999). L'isopentyl diphosphate peut être issue des deux voies MEP et MEV.

I-1-6-Les composés phénoliques

Les composés phénoliques sont des métabolites secondaires issus de la voie de synthèse des phénylpropanoïdes dont les produits majeurs sont les lignines et les flavonoïdes (Figure I-6). Les composés phénoliques sont des molécules très variées en fonction de la plante considérée, du type tissulaire et cellulaire mais leur nature varie aussi en fonction du stade de développement ou encore en réponse à l'environnement. Quelques exemples de l'implication des composés phénoliques dans le développement de la plante en interaction avec son environnement abiotique et biotique sont présentés sur la figure I-5.



Figure I-5. Les interactions plante-environnement et la synthèse des composés phénoliques (d'après Dixon et Paiva, 1995).

I-1-6-1-Les phénylpropanoïdes

La voie de biosynthèse des phénylpropanoïdes produit des composés dotés d'un squelette carboné de base de type C6-C3 issu de la phénylalanine. La grande majorité des enzymes impliquées dans les principales étapes de cette voie de biosynthèse ont été caractérisées chez les plantes supérieures (Boerjan et *al.*, 2003). Les composés phénoliques se définissent par un ou plusieurs cycles benzéniques portant une ou plusieurs fonctions hydroxyl. C'est la complexité du squelette carboné de base qui permet de différencier les différentes classes de composés phénoliques (Tableau I-1). Ces derniers dérivent d'acides aminés aromatiques (essentiellement la phénylalanine) provenant de la voie de biosynthèse du shikimate avec la condensation du phosphoénolpyruvate et de l'érythrose –4-phosphate. La voie de biosynthèse des phénylpropanoïdes présentée à la figure I-6 est une voie commune à plusieurs autres voies de biosynthèse issues d'un même précurseur: la phénylalanine.

Parmi les voies connues et étudiées, on trouve celle de la synthèse des lignines mais aussi celles conduisant aux flavonoïdes, à l'acide salicylique, aux hydroxycoumarines et aux stilbènes entre autres. Les composés phénoliques interviennent dans des rôles aussi variés que les réactions de défense contre les herbivores et les pathogènes, le support mécanique, l'attraction des pollinisateurs et la dispersion des fruits, la compétition avec des plantes voisines (en réduisant leur taille par exemple). Les flavonoïdes sont des molécules de type phénylbenzopyrone comme les anthocyanines, les flavones, les flavonols et les isoflavones qui sont connus pour leurs activités antioxidantes. Les flavonoïdes sont présents dans la plupart des tissus végétaux, souvent dans les vacuoles sous la forme de monomères ou d'oligomères, dans le bois de cœur et l'écorce de certains arbres et ils peuvent aussi intervenir dans des mécanismes de symbiose plantebactérie. Parmi les dérivés des acides hydroxycinnamiques, les stilbènes, coumarines et acides salicyliques sont impliqués dans les réactions de défense alors que les alcools *p*-coumarylique, coniférylique et sinapylique entrent dans la composition du polymère de lignines indispensable au soutien et à la structure de la plante.



Figure I-6. La voie de biosynthèse des phénylpropanoïdes et ses dérivés (Adapté de Boerjan et *al.*, 2003).

Les enzymes mentionnées sont: la 2-déhydro-3-déoxyphosphoheptonate aldolase synthase (DAHP), la phénylalanine ammonia-lyase (PAL), la cinnamate–4-hydroxylase (C4H), la coumarate-3-hydroxylase (C3H), l'acide caféique/acide 5-hydroxyférulique 3-O-méthyltransférase (COMT), la férulate-5-hydroxylase (F5H), la 4-hydroxycinnamate CoA ligase (4CL), la *p*-coumaroyl-CoA 3-hydroxylase (CCoA-3H), la caféoyl-CoA 3-O-méthyltransférase (CCoAOMT), la cinnamoyl CoA réductase (CCR), l'alcool cinnamylique déshydrogénase (CAD), la chalcone synthase (CHS), la chalcone isomérase (CHI), la dihydroflavonol réductase (DFR), l'isoflavone synthase (IFS) et la flavanone 3 hydroxylase (F3H).

A : acide cinnamique, B : acide *p*-coumarique, C : acide caféique, D : acide férulique, E : acide 5-OH-férulique, F : acide sinapique, G : *p*-coumaroyl-CoA, H : caféoyl-CoA, I : féruloyl-CoA, J : 5-OH-férulyl-CoA, K : sinapoyl-CoA, L : aldéhyde *p*-coumarylique, M : aldéhyde coniférylique, N : aldéhyde 5-OH-coniférylique, O : aldéhyde sinapylique, P : alcool 5-OH-coniférylique

Squelette carboné	Classes	Exemples
C6	Phénols simples	Catéchol, hydroquinone
	Benzoquinones	2,6-Diméthoxybenzoquinone
C6-C1	Phénols acides	<i>p</i> -Hydroxybenzoïque
		acide salicilique
C6-C2	Acétophénones	3-Acétyl-6-méthoxybenzaldehyde
	Acides phénylacétiques	<i>p</i> -Hydroxyphénylacétique
C6-C3	Acides hydroxycinnamiques	Caféate, férulate
	Phénylpropènes	Myristicine, eugénol
	Coumarines	Scopolétine, aesculétine
	Isocoumarines	Bergénine
	Chromones	Eugénine
C6-C4	Naphthoquinones	Juglones, plumbagine
C6-C1-C6	Xanthones	Mangiférine
C6-C2-C6	Stilbènes	Acide lunularique, resvératrol
	Anthraquinones	Emodine
C6-C3-C6	Flavonoïdes	
	Flavonols	Kaempférol, quercétine
	Anthocyanes	Cyanidine, pélargonine
	Flavanols	Catéchine, épicatéchine
	Flavanones	Naringérines
	Isoflavonoïdes	Daidzéine
(C6-C3)2	Lignanes	Pinorésinol
	Néolignanes	Eusidérine
(C6-C3-C6)2	Biflavonoïdes	Amentoflavone
(C6-C3)n	Lignines	-
(C6-C3-C6)n	Flavolanes (tanins condensés)	-

Tableau I-1. Les principales classes de composés phénoliques chez les plantes (Adapté de Macheix et *al.*, 2005).

I-1-6-2-La lignine

La lignine est le second composé organique de la biosphère terrestre après la cellulose. Elle représente 15 à 36% de la masse de matière sèche du bois contre seulement 3 à 15% chez les graminées et autres plantes annuelles (Sarkanen et Hergert, 1971; Hartley et Jones, 1977). Ce polymère carboné à noyau aromatique se retrouve chez tous les végétaux, principalement au niveau des parois cellulaires secondaires, mais avec des compositions qualitatives et quantitatives différentes selon les espèces. La voie de biosynthèse de la lignine et son mode de synthèse sont très étudiés à cause de l'importance économique et écologique de la matière végétale ligneuse. Chez les arbres, la quantité de lignines augmente la teneur calorifique du bois de chauffage, elle est prépondérante pour faire des planches de bois rigide. En revanche, c'est un problème majeur pour l'industrie papetière.
-Rôles de la lignine

La lignine est un composé phénolique important qui sert d'agent de liaison entre les différents polymères de sucres (cellulose, hémicellulose et pectine) de la paroi secondaire. On lui attribue les rôles qui suivent.

- Imperméabiliser les cellules conductrices de sève et résister aux fortes pressions négatives nécessaires pour le transport de l'eau jusqu'aux feuilles (Jones et *al.*, 2001; Reina et *al.*, 2001),
- Apporter un soutien mécanique aux plantes et leur permettre un port dressé (Zhong *et al.*, 1997; Jones *et al.*, 2001),
- Protéger les plantes contre des attaques invasives de pathogènes ou vis à vis de stress abiotiques en constituant une barrière physique (Henderson et Friend, 1978; Vance et *al.*, 1980; Lawton et Lamb, 1987; Lange et *al.*, 1995; Hammond-Kosack and Jones, 1996; Franke et *al.*, 2002).

-Biosynthèse des monolignols

Les monolignols, ou alcools *p*-hydroxycinnamiques, représentent trois monomères appelés alcool *p*-hydroxycoumarylique, alcool coniférylique et alcool synapylique; ces trois précurseurs, aprés polymérisation oxydative, conduisent à un polymère de lignine constitué respectivement d'unités hydroxyphényle (H), guaiacyle (G) et syringyl (S) (Figure I-6). On trouve des unités hydroxyphényles (type H) dans les parois des cellules végétales des angiospermes, gymnopermes et plantes herbacées en très faible proportion mais elles sont plus abondantes dans les cellules formant le bois de compression, ainsi que chez plusieurs graminées. Les lignines guaiacyles (type G) sont présentes dans les lignines des angiospermes et gymnospermes. On observe une proportion significative de polymère de lignine avec des unités syringyl (type S) seulement chez les angiospermes.

Les monolignols sont produits par la voie de biosynthèse des phénylpropanoïdes. Il s'agit d'une voie commune pour la synthèse de la lignine mais aussi pour bien d'autres composés tels que les stilbènes, les hydroxycoumarines et les flavonoïdes (Figure I-6). Le métabolisme des phénylpropanoïdes débute avec la désamination de la phénylalanine par la phénylalanine ammonia-lyase (PAL) puis conduit à la synthèse des cinnamoylCoA *via* la cinnamate 4-hydroxylase (C4H) et la 4-coumarate CoA ligase (4CL). Les deux étapes terminales de la synthèse des monolignols sont catalysées successivement par les enzymes cinnamoyl-CoA réductase (CCR) et alcool cinnamylique déshydrogénase (CAD) qui réduisent respectivement les fonctions cinnamoyl-CoAs de la chaîne latérale en aldéhydes puis en alcools (Figure I-6). Le degré de méthoxylation des monomères dépend des actions successives des hydroxylases (C3H : coumarate-3-hydroxylase, F5H : férulate-5-hydroxylase et CCoA-3H : *p*-coumaroyl-CoA 3-hydroxylase) et des *0*-méthyltransférases (COMT : acide caféique/acide 5-hydroxyférulique 3-O-méthyltransférase, CCoAOMT : caféoyl-CoA 3-O-méthyltransférase) en positions 3 et 5 du cycle aromatique (Figure I-6). Finalement, les monolignols, qui sont des composés toxiques et instables, ne sont pas accumulés dans la cellule végétale. Dans le cas de l'alcool coniférylique chez les conifères, celui-ci est glycosylé sous forme de coniférine pour être stocké dans la vacuole puis excrété par des vésicules d'exocytose (Schmid et Grisebach, 1982).

-Transport et polymérisation des monolignols

Les monolignols sont synthétisés dans l'appareil de Golgi puis transportés vers la paroi pour y être polymérisés. Le processus de lignification consiste en l'assemblage, ou polymérisation, des précurseurs monolignols entre eux *via* des réactions de couplage de radicaux. Ce processus est permis après la déhydrogénation des monolignols par des enzymes telles les peroxidases, les laccases, les polyphénols oxidases et les alcools coniféryliques oxydases (Boerjan et *al*, 2003). Le dépôt des lignines s'initie au niveau des coins cellulaires dans la lamelle moyenne et la paroi primaire (Samuels et *al.*, 2002; Boerjan et *al.*, 2003). La polymérisation *in situ* des lignines est influencée par l'environnement pariétal incluant la lignine préalablement déposée. Les liaisons chimiques effectuées avec des celluloses et hémicelluloses insolubles ainsi que des protéines pariétales vont graduellement éliminer l'eau et créer un environnement hydrophobe.

I-1-6-3-Modulation de l'expression des gènes des phénylpropanoïdes et des lignines

La différenciation du xylème et la lignification sont des étapes inscrites dans un continuum spatio-temporel impliquant l'action d'un grand nombre de gènes dont ceux des enzymes de la voie des phénylpropanoïdes. Aussi, les enzymes nécessaires à la synthèse des phénylpropanoïdes et des lignines ont été caractérisés biochimiquement et génétiquement (pour revue Boerjan et al., 2003). Les résultats de ces études renseignent sur la fonction enzymatique et le rôle que joue chaque enzyme au niveau du contrôle métabolique. Ils soulignent aussi la complexité de cette voie de synthèse. Plusieurs expériences de gain de fonction et/ou de perte de fonction de gènes intervenant dans la voie de synthèse des lignines ont montré que la modulation de l'expression de ces gènes entraine des variations en terme de quantité (augmentation/réduction) et de qualité (ratio S/G) des lignines obtenues (Figure I-6, Tableau I-2). Dans certains cas, la lignine devient même plus facile à extraire (Pilate et al, 2002, Baucher et al, 2003). De plus, la diversité naturelle des végétaux offre des mutants qui présentent des phénotypes différents et rares. Dans ces cas, le gène muté code pour une protéine qui a une activité enzymatique ou une fonction moléculaire réduite voire absente. Plusieurs exemples ont été décrits concernant la voie de biosynthèse des lignines dont cad-nl (gène CAD) chez le pin à encens Pinus taeda (MacKay et al., 1997; Ralph et al., 1997; Lapierre et al., 2000), brown midrib 1 et 3 (bml, gène CAD et *bm3*, gène *COMT*) chez le maïs (Barrrière et Argillier, 1993; Vignols et *al.*, 1995; Halpin et al., 1998), fahl (gène F5H) chez Arabidopsis (Chapple et al., 1992), etc. Dans le cas du mutant cad-null, MacKay et al. (1997) ont observé une faible diminution du taux de lignines en raison de l'incorporation massive de coniféraldéhydes dans ces lignines couplée à une coloration rose du bois et à une réduction de croissance. Par ailleurs, des études de répression du gène HCT (hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase) chez Arabidopsis indiquent une diminution de la synthèse de la lignine et de la croissance. Dans ces plantes transgéniques, il y a une redirection du flux métabolique dans les flavonoïdes impliquant la chalcone synthase (Hoffmann et al, 2004; Besseau et al., 2007).

Ainsi, les gènes impliqués dans la synthèse des monolignols sont directement liés à la lignification des parois végétales (Boerjan et *al.*, 2003). Les travaux de transgénèse et les études de mutants des gènes en relation avec les phénylpropanoïdes montrent que dans le cas d'un déficit en précurseurs conventionnels de la lignine, la plasticité biochimique de la lignification permettra tout de même l'incorporation de précurseurs inhabituels pour la formation de celle-ci (Boudet et *al.*, 2003).

Gènes	Espèces	Taux de Lignine	Ratio S / G	Références
PAL (+)	Tabac	Augmenté	Inchangé	Sewalt et al., 1997
PAL (-)	Tabac	Diminué	Augmenté	Sewalt et al., 1997
C4H (-)	Tabac	Diminué	Diminué	Sewalt et al., 1997
*COMT (-)	Tabac	Inchangé	Diminué	Atanassova et al., 1995
*COMT (-)	Peuplier	Inchangé	Diminué	Van Doorsselaere et <i>al.</i> , 1995
F5H (-)	Arabidopsis	N. C.	Abscence de S	Chapple et al., 1992
F5H (+)	Tabac/Arabidopsis	N. C.	Augmenté	Meyer et al., 1998
4CL (-)	Tabac	Diminué	Diminué	Kajita et al., 1996
4CL (-)	Arabidopsis	Diminué	Augmenté	Lee et al., 1997
CCR (-)	Tabac	Diminué	Augmenté	Boudet et al., 1996
CCR (-)	Tabac	Diminué	Augmenté	Piquemal et al., 1998
**CAD (-)	Tabac	Inchangé	Diminué	Halpin et al., 1994
**CAD (-)	Peuplier	Inchangé	Inchangé	Baucher et al., 1996

Tableau I-2. Exemples d'expériences de perte et de gain de fonction par transformation génétique utilisant les gènes de la voie de biosynthèse des phénylpropanoïdes et des lignines.

(-): sous-exprimé, (+):sur-exprimé, N. C.: non communiqué, S : Unité monomérique syringyl (alcool synapylique), G : Unité monomérique guaiacyl (alcool coniférylique), *: Lignine difficilement extractible, **: Lignine facilement extractible; *PAL* : *phénylalanine ammonia-lyase; C4H* : *cinnamate 4-hydroxylase; COMT* : *acide caféique/acide 5-hydroxyférulique 3-O-méthyltransférase; F5H* : *férulate-5 hydroxylase; 4CL* : *4-hydroxycinnamate CoA ligase; CCR* : *cinnamoyl CoA réductase; CAD* : *alcool cinnamylique déshydrogénase.*

I-1-7-Régulation de la formation du bois

I-1-7-1-Les hormones

Les régulateurs endogènes de la formation du xylème peuvent être des hormones végétales (auxines, cytokinines, éthylène, gibbérellines et acide abscissique), des brassinostéroïdes, des oligosacharides ou encore des peptides (Turner et *al.*, 2007). Par exemple, Israelsson et *al.* (2005) ont suggéré que les gibbérellines régulent les étapes précoces de la formation du bois ainsi que l'élongation cellulaire. Kazama et *al.* (2004) ont pour leur part démontré que l'exposition transitoire à l'éthylène stimulait la division cellulaire et altèrait la forme et la polarité des cellules épidermiques d'hypocotyle.

Dans cette étude bibliographique, je développerai plus particulièrement le rôle de l'auxine qui est le mieux caractérisé dans la différenciation vasculaire.

Les zones de synthèse des hormones végétales, ou phytohormones, sont les aiguilles ou les extrémités des feuilles, les apex des tiges pour l'auxine et les racines pour l'acide abscisique et les cytokinines ou encore les graines pour les gibbérellines. L'effet d'une phytohormone dépends de sa concentration et de la réceptivités des cellules cibles. Entre elles, les hormones ont des modes d'actions complémentaires ou opposés (antagonistes), ce qui implique qu'un équilibre hormonal doit être maintenu dans la plante pour son développement. Les hormones peuvent être des agents promoteurs (auxines, gibbérellines, cytokinines) ou inhibiteurs (acide abscicique) de l'activité des méristèmes apicaux (Wareing, 1965; Thomas et *al.*, 1965).

Le rôle de l'auxine, acide 3-indole-acétique (AIA ou IAA en anglais), dans la différenciation des tissus vasculaire a été bien documenté (Aloni, 1987; Sachs, 1991; Sundberg et *al.*, 2000). Le système modèle *Zinnia elegans* a permis de mettre en évidence qu'une certaine proportion dans la combinaison auxine et cytokinine est suffisante pour induire la différenciation des éléments trachéaires mais insuffisante quand l'une de ces deux hormones est absente (Milioni et *al.*, 2001). L'observation de l'élévation des niveaux d'expression des gènes liés à la synthèse de l'auxine a également été corrélée avec des niveaux élevés d'auxine dans les cellules en division du cambium et des cellules mères du xylème (Moyle et al., 2002). Afin de maintenir l'état méristématique de ces cellules, l'auxine stimulerait l'expression de gènes de sa propre voie de biosynthèse (Bhalerao et Bennett, 2003). La thèse de la régulation de l'activité cambiale par l'auxine est renforcée par la distribution de l'auxine en gradient de part et d'autre de la zone cambiale de tiges de pin avec son maximum au niveau de la zone cambiale (Ugla et al., 1996, 1998; Schrader et al., 2003). De plus, chez Pinus contorta, le cambium vasculaire disparaît en l'absence d'un apport continu d'auxine exogène (Savidge et Wareing, 1981; Savidge, 1983). En ce qui concerne la formation du bois de réaction, plusieurs études ont établi que le bois de tension nécessite une différence de concentration auxinique autour de la tige et se forme dans la région déficiente en AIA. Au contraire, le bois de compression serait induit par une concentration accrue en auxine (Timell, 1986; Little et Savidge, 1987; Srivastava, 2002). Toutefois, chez le peuplier et le pin, Israelsson et al. (2005) n'ont pas détecté de changements dans la balance auxinique en lien avec le bois de réaction. Ils suggèrent que le principal rôle des auxines pendant la formation du bois est de réguler l'expansion cellulaire pendant les stades précoces de la différenciation du xylème.

I-1-7-2-Les facteurs de transcription

Les facteurs de transcription sont des protéines capables de réguler des processus physiologiques et biochimiques en modulant le taux de transcription de gènes cibles (Ptashne, 1988). Il s'agit d'une modulation par activation ou répression de l'expression des gènes au cours de l'étape de la transcription qui se déroule dans les noyaux des cellules chez les eucaryotes.

Des études d'expression de gènes lors de xylogénèse chez différents modèles végétaux ont permis de mettre en évidence l'expression coordonnée de gènes impliqués dans la différenciation du xylème ou plus spécifiquement dans la lignification (Demura et *al.*, 2002; Ko et *al.*, 2006; Demura et Fukuda, 2007). Des analyses transcriptomiques indiquent que de nombreux gènes codant des facteurs de transcription sont exprimés pendant la formation du bois chez les plantes (Demura et Fukuda, 2007). Plusieurs membres de classes différentes de facteurs de transcription ont été identifiés par l'analyse génétique de mutants *Arabidopsis* possédant des défauts dans l'organisation et le développement des tissus vasculaires (Ye, 2002; Scarpella et Meijer, 2004; Carlsbecker et Helariutta, 2005). Parmi ceux-ci, on trouve les facteurs de transcription appartenant à la famille des ARFs (Auxin Response Family), dont certains ont une expression inductible par l'auxine (Scarpella et Meijer, 2004; Badescu et Napier, 2006). On trouve également des homéodomaine-leucine zippers de la classe III et les KANADIs (famille des FT de type GARP, Eshed et al., 2001) dont les mutations affectent la polarité et l'identité des cellules xylème-phloème (Emery et al., 2003). D'autres types de facteurs de transcription comme les NAC, de la famille des gènes <u>NAM/ATAF/CUC</u>, se révèlent être préférentiellement exprimés pendant le développement du bois et en particulier dans les éléments trachéaires en différenciation. Les analyses fonctionnelles de ces gènes ont clairement montré des rôles importants dans le développement des tissus vasculaires et la transdifférenciation de vaisseaux immatures en métaxylème (Kubo et al., 2005) ainsi que dans la transdifférenciation de cellules diverses en cellules possédant des parois secondaires épaissies (Mitsuda et al., 2005; Zhong et al., 2006; Mitsuda et al., 2007). Enfin, l'analyse détaillée des promoteurs des gènes codant des enzymes de la biosynthèse des lignines a révélé la présence et la conservation de courtes régions nucléotidiques, les éléments AC ou boîte H, nécessaires à une expression dirigée dans le xylème (Rogers et Campbell, 2004). De plus, il a été montré que des facteurs de transcription de type LIM (LIN11/ISL-1/MEC-3) et MYB sont capables de se fixer à ces éléments AC et de réguler l'expression de quelques uns des gènes associés à la synthèse des lignines pour réguler la formation de la paroi secondaire (Goicoechea et al., 2005; Kawaoka et al., 2000).

I-2- Les facteurs de transcription MYB

I-2-1-Le mécanisme général de la transcription des gènes

Plusieurs points de contrôle existent entre la molécule d'ADN et l'activité de la protéine codée, mais le plus important reste celui de la régulation de la transcription. La régulation de la transcription s'effectue par la présence de protéines appelées facteurs de transcription, capables de reconnaître et de lier de courtes séquences nucléotidiques appelées « éléments cis régulateurs » qui se trouvent généralement dans la région promotrice des gènes qui s'étend de quelques centaines à quelques milliers de nucléotides en amont du site d'initiation de la transcription (Maston et al., 2006). Des travaux récents ont cependant montré que des éléments cis peuvent se retrouver en aval ou même dans la séquence codante ou les introns d'un gène. Dans la plupart des gènes eucaryotes, une boite TATA est située environ 30 nucléotides en amont du site d'initiation, ce qui permet l'attachement des acteurs moléculaires de la transcription. Le complexe de pré-initiation de la transcription est formé par la liaison de l'ARN polymérase II au site d'initiation et dépend de la présence de plusieurs complexes protéiques (TFIIB, TFIID, TFIIF et TFIIE) dont la « TATA-binding protein » qui assure la liaison à l'ADN au niveau du 'promoteur proximal (Werner, 1999).

Le mode d'action des facteurs de transcription est de destructurer le nucléosome afin d'accéder aux éléments *cis* et les fixer, puis recruter des coactivateurs ou corépresseurs ainsi que les facteurs généraux pour la transcription. D'une façon générale, le rôle fonctionnel de ces protéines régulatrices est de modifier le taux d'initiation de la transcription en exerçant un contrôle sur le taux de recrutement des partenaires protéiques du complexe de pré-initiation ou encore de réguler le taux de transition du complexe de pré-initiation vers la configuration ouverte (Nikolov et Burley, 1997).

Les facteurs de transcription peuvent être activateurs ou répresseurs du taux de synthèse des ARNm de part leurs structures protéiques. Ainsi, ils ont été répartis en différentes classes en fonction de leur site de fixation à l'ADN (« DNA binding domain » ou DBD), c'est à dire en fonction de la structure moléculaire de

leur DBD. Il existe quatre types de structures protéiques suceptibles d'interagir avec l'ADN (Patikoglou et Burley, 1997) :

- structure hélice-tour-hélice (HTH pour Helix-Turn-Helix, MYB par exemple), les hélices alpha se fixent dans le sillon majeur de l'ADN pour former un angle.
- structure doigt de gant (protéines « zinc finger » pour doigt de zinc), le pouvoir de lier des atomes de zinc confère à ces protéines une conformation permettant la liaison à l'ADN.
- structure en fermeture éclair à leucine (protéines « leucine zippers », bZip), constitués d'une série de leucines positionnées tous les sept acides aminés autorisant les protéines à se dimériser. Adjacents à ces structures se trouvent plusieurs acides aminés chargés positivement qui sont impliqués dans la fixation à l'ADN.
- structure hélice-boucle-hélice (HLH pour Helix-Loop-Helix, MYOD et MYB chez les animaux, Murre et *al.*, 1989), les hélices présentent une face avec des acides aminés hydrophobes et l'autre face avec des acides aminés chargés. Ce motif moléculaire permet la dimérisation entre protéines identiques et c'est la région basique proche de ce motif qui est en contact avec l'ADN.

I-2-2-Identité et fonction générale des MYB

« MYB » est un acronyme dérivé du mot « MYeloBlastome », décrivant un virus responsable de la leucémie chez le poulet. Le premier gène *MYB* identifié était l'oncogène *v-myb* issu du virus du MyeloBlastome aviaire (AMV) intégré dans le génome du poulet (Gonda et *al.*, 1982; Klempnauer et *al*, 1982). Introna et *al.* (1994) ont montré que la protéine virale v-MYB était impliquée dans la transformation cancéreuse des cellules hématopoïétiques *via* la dérégulation de gènes importants pour le développement cellulaire.

Les facteurs de transcription de type MYB reconnaissent de façon spécifique les séquences YAAC(G/T)G *via* un motif protéique de type hélice-tour-hélice (HTH) (Biedenkapp et *al.*, 1988). Les protéines MYB se définissent par la présence d'un domaine MYB constitué d'environ 50 acides aminés contenant trois tryptophanes conservés et régulièrement espacés de 18 ou 19 acides aminés qui sont à l'origine de la formation d'un cœur hydrophobe nécessaire à la fixation de l'ADN (Peters et *al.*, 1987)

(Figure I-7, Ogata et *al.*, 1992). Plusieurs copies de ce domaine MYB peuvent être répétées dans une seule protéine et permettre l'attachement à l'ADN (Klempnauer et Sippel, 1987). Dans les cas de répétitions imparfaites, les tryptophanes peuvent être remplacés par des résidus phénylalanine ou tyrosine et des acides aminés supplémentaires peuvent être observés (Hovring et *al.*, 1994; Morrow et *al.*, 1993). Le domaine de fixation des protéines MYB présent en N-terminal est conservé chez les animaux, les plantes et les levures alors qu'en en C-terminal se trouvent des régions capables de moduler l'expression génique (Lipsisk et *al*, 1996) (Figure I-8). Chez les animaux, il n'existe que des protéines MYB à trois répétitions (MYB-R1R2R3), alors que chez les plantes et les levures le nombre de répétitions est variable (MYB-R1 à une répétition, deux répétitions MYB-R2R3 et MYB-R1R2R3 à trois répétitions). De ce fait, les protéines MYB ont été arrangées en trois sous classes possédant 1, 2 ou 3 répétitions (Jin H et Martin C, 1999).

Chez les aninaux, les protéines de type MYB se présentent sous trois formes et on distingue uniquement trois membres chez l'Homme: A-MYB, B-MYB et c-MYB, (Figure I-8, Kanei-Ishii et al., 1996) impliqués dans la division, la différenciation et la mort cellulaire programmée (Weston, 1998). Ces MYB-R1R2R3 possèdent la propriété moléculaire de se lier à de courtes séquences d'ADN appelées « Myb Regulatory Element » (MRE) ayant comme séquence consensusYAACNGHH (Y=C/T; H=A/C/T; N=A/C/T/G) dans les régions promotrices des gènes pour en réguler l'expression (Ording et al., 1996). La spécificité de liaison de séquence de même que leurs rôles dans le contrôle transcriptionnel ont été démontrés pour plusieurs protéines MYB-R1R2R3 (Biedenkapp et al., 1988; Watson et al. 1993; Weston et Bishop, 1989; Foos et al., 1994). Dans la partie C-terminale de la protéine c-MYB on retrouve un domaine d'activation et aussi des domaines de régulation négative (NRD) qui pourraient agir sur la transcription via l'interaction avec d'autres protéines (Dubendorff et al., 1992; Wang et al., 1999). De la même manière, la protéine B-MYB est décrite comme un répresseur (Foos et al. 1994; Watson et al., 1993), mais elle contient aussi un domaine d'activation (Ansieau et al., 1997). Les A et B-MYB sont activés par la phosphorylation du NRD via un complexe cycline A/Cdk2 (Saville et Watson, 1998).



Figure I-7. Structure tridimensionnelle d'une protéine MYB-R2R3 interagissant avec la molécule d'ADN (Ogata et *al.*, 1994).

La photographie est tiré du site web http://gibk26.bse.kyutech.ac.jp/jouhou/image/ dna-protein/all/all.html (numéros d'accession Pfam: PF00249 et PROSITE: PS50090). La répétition R3 se fixe dans le sillon majeur de la double hélice d'ADN.





Les différentes répétitions (R) du domaine MYB sont indiquées. DBD: DNA-binding domain, TAD: transcriptional activation domain, NRD: negative regulatory domain, PRD: positive regulatory domain, Z: structure leucine zipper. Les chiffres indiquent le nombre d'acides aminés constituant chaque protéine MYB. Les pourcentages d'identité (I) et de similarité (S) entre les DBD de deux protéines sont calculés avec Bioedit (matrice Blosum62). C-MYB de l'Homme (P10242), A-MYB de *Mus Musculus* (X82327), B-MYB de *Mus Musculus* (X70472) et c1 MYB de maïs (M37153).

I-2-3-Les MYB chez les végétaux

Au sein du règne végétal, le premier membre de cette famille a été découvert chez le maïs et il est appelé C1 pour *colourless1* (Paz-Arez et *al.*, 1987). Ce gène représente aussi la première étude d'un locus ciblé avec la technique de mutagénèse induite par transposons et mettant en relation un phénotype induit, ici un changement de couleur des grains de maïs, avec une modification génique (Paz-Ares et *al.*, 1987). Depuis l'identification et la caractérisation de cette protéine MYB à deux répétitions, il apparaît que les facteurs de transcription de type MYB chez les plantes représentent une superfamille de gènes de par la diversité et le nombre élevé de membres. Par ailleurs, on leur attribue la régulation de différents aspects du métabolisme et du développement spécifiques des plantes.

Le séquençage du génome de la plante modèle Arabidopsis thaliana (AGI, 2000) a permis de découvrir que, parmi les 26 000 gènes identifiés, environ 1 500 codent des facteurs de transcription, soit 5% du génome (Riechmann et al., 2000). Finalement, et après plusieurs études, la plus récente de Yanhui et al (2006) rapporte l'identification de 198 gènes appartenant à la superfamille des MYB. L'analyse complète du génome d'Arabidopsis révèle que 126 de ces gènes sont du type MYB-R2R3, 5 sont des MYB-R1R2R3, 64 des «MYB-related » et 3 sont des gènes MYB atypiques de données web AGRIS : (base http://arabidopsis.med.ohio-state.edu/, Palaniswamy et al., 2006; Davuluri et al., 2003). Chez Arabidopsis, la superfamille des gènes MYB est une des classes de facteur de transcription qui possède le plus de membres et la famille des MYB à deux répétitions est la plus représentée (Riechmann et Ratcliffe, 2000). Dans plusieurs espèces végétales différentes, d'autres membres de la famille des gènes MYB ont également été identifiés et caractérisés, et leur dénombrement n'est pas encore fini (environ 30 membres chez le pétunia selon Avila et al., 1993; 80 chez le maïs d'après Rabinowicz et al, 1999, voire plus de 200 selon Jiang et al., 2004b; environ 200 chez le cottonnier selon Cedroni et al., 2003; 85 chez le riz selon Jiang et al, 2004a; etc.).

I-2-3-1-Classifications des membres de la famille MYB

- Structuration par DBD et élément cis fixé

Les protéines de type MYB-R2R3 sont des facteurs de transcription reconnaissables par leur domaine conservé de fixation à l'ADN à deux répétitions imparfaites (Figure I-8). Les MYB-R2R3 se lient aux éléments *cis* avec leur deuxième (R2) et troisième (R3) répétitions dans le sillon majeur de l'ADN (Figure I-7). Cette fixation implique plus précisément la troisième hélice de chaque répétition qui est en contact direct avec l'ADN. La présence des tryptophanes régulièrement espacés forment un cœur hydrophobe nécessaire à une liaison spécifique de motifs nucléotidiques (Ogata et *al.*, 1992). De plus, des études d'interaction protéine-protéine ont mis en évidence la présence de motifs protéiques dans le domaine MYB permettant un partenariat avec des facteurs de transcriptions de type b-HLH (motif [DE]Lx2[RK]x3Lx6Lx3R, Zimmermann et *al.*, 2004) ou encore avec des calmodulines (motif RWLnyLRpdVRRgnItLe, Yoo et *al.*, 2005).

Les facteurs MYB végétaux ont été classés en fonction des motifs d'ADN potentiellement reconnus par leur domaine de fixation à l'ADN et ce classement peut être visualisé par un arbre phylogénétique montrant les identités de séquences protéiques entre les membres (Romero et al, 1998). Ces études d'affinité de liaison de protéines MYB avec des séquences nucléotidiques s'effectuent principalement par le biais d'expériences de type « gel retard » ou « EMSA » (« Electrophoretic Mobility Shift Assays ») (Blackwell et Weintraub, 1990), mais aussi avec des analyses par « DNA footprinting », méthylation interférence ou encore précipitation de l'ADN. Ainsi, trois familles ont été établies: les membres de la famille A se lient sur le site de type I :CNGTTR, les membres de la famille B reconnaissent aussi bien les sites de type I et II (Type II:GTTWGTTR) et les membres de la famille C préfèrent fixer le site IIG :GKTWGGTR (N=A/G/C/T, K=G/T, R=A/G, W=A/T) (Romero et al., 1998). Cette approche phylogénétique, de plus en plus utilisée, permet de regrouper et de classer des gènes MYB entre différentes espèces sur la base de leur homologie de séquence et des éléments cis régulateurs fixés.

-Motifs protéiques consensus en carboxy terminal

Des analyses phylogénétiques de séquences complètes de gènes *MYB* converties en acides aminés ont permis de regrouper les membres de la famille R2R3 en 22 sous groupes chez *Arabidopsis* (Kranz et *al*, 1998). Une telle étude considère les deux parties de la protéine MYB, le domaine de liaison d'environ 150 acides aminés et le domaine carboxy-terminal de taille variable mais généralement supérieur ou égal à 100 résidus. La phylogénie entre MYB a aussi mis en évidence que, dans la plupart des feuillets, plusieurs membres peuvent posséder dans leur région c-terminale des motifs conservés de quelques acides aminés (Kranz et *al.*, 1998; Jiang et *al.*, 2004a; Stracke et *al.*, 2001). Par exemple, le motif pdLNL^D/_Elxi^G/_S du sous-groupe 4 observé par Kranz et *al.* (1998) a été lié à la répression transcriptionnelle du gène codant pour la *cinnamate 4-hydroxylase* (*C4H*) par le gène *AtMYB4* chez Arabidopsis (Jin et *al.*, 2000).

-Nombre et position des introns

Le séquençage des régions génomiques correspondant aux gènes *MYB* chez des espèces comme *Arabidopsis* et le riz montre que 59% et 53% respectivement possèdent une structure exon-intron de trois exons et deux introns présents dans le domaine R2R3. Un seul ou les deux introns sont absents dans 19% des gènes d'Arabidopsis et 12% chez le riz (Jiang et *al.*, 2004a). De plus, la structure exon-intron diffère entre les 22 sous-groupes mais est conservée dans un même sous-groupe (Jiang et *al.*, 2004a). Le nombre, la position et la phase d'insertion des introns sont des éléments de classification des gènes *MYB*, classification qui suit de très près les résultats des analyses phylogénétiques et la structure en sous-groupes determinée d'après les comparaisons de séquences protéiques (Kranz et *al.*,1998; Jiang et *al.*, 2004a). Des approches similaires de structuration de familles multigéniques ont été appliqués pour les kinésines et les bHLH (Lawrence et *al.*, 2002; Toledo-Ortiz et *al.*, 2003).

I-2-3-2-Duplication génique et évolution des gènes MYB

De nombreuses études indiquent que la duplication des gènes est un des mécanismes majeurs à l'origine de l'émergence de gènes possédant de nouvelles fonctions (Ohno, 1970; Nei, 1969; Zhang et *al.*, 1998). Les duplications géniques

contribuent à la formation de familles de gènes, observables dans tous les génomes séquencés (Lynch et Conery, 2000; Prince et Pickett, 2002). Après l'évènement de duplication génique, les contraintes sélectives (i.e. pression de sélection) sont réduites sur les copies résultantes, ce qui permet de tolérer des modifications nucléotidiques dans une copie ou l'autre du gène. Si la nouvelle protéine générée trouve une fonction biochimique innovante et/ou avantageuse pour la plante, alors ce nouveau gène sera maintenu actif (Sankoff, 2001; Wagner, 2001). Par exemple, les gènes *MYB* homologues *C1* et *PL1* chez le maïs proviennent d'une ancienne duplication génique et diffèrent seulement de 5 acides aminés (Cone et *al.*, 1993). Ces deux gènes régulent l'expression des gènes structuraux des anthocyanines, mais dans des parties distinctes de la plante. Le gène *C1* contrôle la pigmentation de la couche d'aleurone du grain et le gène *PL1* régule la pigmentation dans les tissus végétatifs et floraux.

L'obtention d'un grand nombre de séquences de gènes *MYB* a mené à l'élaboration de deux hypothèses opposées sur l'évolution de la superfamille des MYB. La première hypothèse propose que les MYB de type R2R3 soient relativement nouveaux et auraient évolué d'un ancêtre commun à trois répétitions R1R2R3 par la perte de la répétition R1 (Lipsick, 1996; Rosinski et Atchley, 1998; Braun et Grotewold, 1999; Kranz et *al.*, 2000; Chen et *al.*, 2006). Au contraire, la seconde hypothèse suggère que les membres de type R2R3 seraient des ancêtres des membres R1R2R3 qui auraient évolué en gagnant le domaine R1 après duplication intragénique d'un *MYB-R2R3* (Jiang et *al.*, 2004b). Quoiqu'il en soit, dans les gènes *MYB*, c'est la partie fixant l'ADN (répétitions R2 et R3) qui est la mieux conservée au cours de l'évolution, faisant état du rôle de coordination indispensable pour l'organisme au sein de différents processus biologiques via le contrôle transcriptionnel de plusieurs gènes.

I-2-4-Rôles biologiques des facteurs de transcription MYB

Les fonctions moléculaires des MYB sont d'activer ou de réprimer la transcription de gènes cibles dans le but de conduire les cellules, tissus ou organes où ils agissent à un état biologique particulier. Les études de gain de fonction (e.g. sur-expression par transgénèse) et de perte-de-fonction (e.g. mutant

insertionnel) permettent de lier ou de démontrer l'utilité d'un gène à un processus biologique donné (Memelink, 2005). D'autre part, des études qui n'utilisent pas de modifications géniques héritables mais des conditions expérimentales (infections, carences nutritives, modulation de la lumière, etc.) peuvent induire l'expression de certains facteurs de transcription et ainsi révéler une réponse biologique dans lequel est impliqué un de ces gènes régulateurs.

L'utilisation de mutants chez la plante modèle Arabidopsis a permis d'associer des rôles biologiques à une partie seulement des protéines MYB. En effet, le caractère régulateur des protéines MYB fait que la perte d'expression de certaines d'entre elles conduit à des conditions physiologiques létales pour la plante. Cependant, les études de caractérisation fonctionnelle chez Arabidopsis thaliana permettent d'avoir une vision générale des processus biologiques régulés par les MYB dans une espèce végétale. Les tableaux I-3 et I-4 présentent les différents gènes MYB ainsi que les processus biologiques mis en évidence expérimentalement chez Arabidopsis. Chez les plantes, des membres de la familles MYB-R2R3 sont impliqués dans de nombreux processus biologiques : le développement de la graine et la germination (AtMYB61 :Penfield et al., 2001; GAMYB: Diaz et al., 2002), le développement des anthères (AtMYB103: Higginson et al., 2003; HvGAMYB : Murray et al., 2003; AtMYB26 : Steiner-Lange et al., 2003), la réponse à des stress (AtMYB102 : Denekamp et Smeekens, 2003; AtMYB30 : Vailleau et al., 2002; AtMYB2 : Abe et al., 2003), la détermination du devenir des cellules épidermiques (AtMYB23 : Kirik et al., 2001; WEREWOLF : Lee et Schiefelbein, 1999, 2001), la photomorphogénèse (LAF1 : Ballesteros et al., 2001, Seo et al., 2003) ou encore la biosynthèse des phénylpropanoïdes (Rogers et Campbell, 2004) et la mise en place du système vasculaire (Scarpella et Meijer, 2004). Certains MYB peuvent aussi avoir des rôles multiples en agissant sur différents tissus ou organes, par exemple en régulant l'initiation des trichomes et aussi des poils absorbants (Tableaux I-3 et I-4). Chez la gueule de loup Anthirrinum majus, le gène AmMIXTA intervient dans la formation de cellules épidermiques et dans la synthèse d'anthocyanes au niveau des fleurs (Noda et al., 1994).

Nom du gène	Locus ID	Nom générique	Processus régulé	Références
GL1	At3g27920	AtMYB000	Initiation des trichomes	Oppenheimer et al., 1991
AtMYB2	At2g47190	AtMYB002	Régulation de l'expression des gènes par l'ABA, réponses aux stress salin et à la déshydratation	Abe et <i>al.</i> , 2003
AtMYB4	At4g38620	AtMYB004	Répression du métabolisme des esters de l'acide sinapique, réponses aux UV-B, développement du pollen	Jin et <i>al.</i> , 2000
AtMYB5	At3g13540	AtMYB005	Développement des trichomes et de la graine	Li et <i>al.</i> , 1996
HOS10	At1g35515	AtMYB008	Réponses aux stress abiotiques (froid, stress osmotique et salin)	Zhu et al., 2005
AtMYB12	At2g47460	AtMYB012	Régulateur spécifique des flavonols, de la biosynthèse des phénypropanoïdes	Mehrtens et al., 2005
AtMYB13	At1g06180	AtMYB013	Morphogènèse de la tige	Kirik et al., 1998
AtMYB15	At3g23250	AtMYB015	Régulation des gènes de tolérance au froid, régulateur de la voie du shikimate en réponse à la blessure	Agarwal et <i>al.</i> , 2006; Chen et <i>al.</i> , 2006
LAF1	At4g25560	AtMYB018	Voie de signalisation du phytochrome A	Ballesteros et al., 2001
AtMYB21	At3g27810	AtMYB021	Photomorphogènèse, développement des étamines, réponses au jasmonate	Shin et <i>al.</i> , 2002; Mandaokar et <i>al.</i> , 2006
AtMYB23	At5g40330	AtMYB023	Différenciation des trichomes	Kirik et <i>al.</i> , 2001
AtMYB24	At5g40350	AtMYB024	Développement des anthères, biosynthèse des phénylpropanoïdes, réponse au jasmonate	Mandaokar et <i>al.</i> , 2006, Yang et <i>al.</i> , 2007
AtMYB26	At3g13890	AtMYB026	Déhiscence des anthères	Steiner-Lange et al., 2003
AtMYB30	At3g28910	AtMYB030	Régulateur positif de la réponse d'hypersensibilité, dépendant de l'accumulation de l'acide salicylique	Vailleau et <i>al.</i> , 2002; Raffaele et <i>al.</i> , 2006
AtMYB32	At4g34990	AtMYB032	Développement du pollen, biosynthèse des phénylpropanoïdes	Preston et al., 2004
AtMYB33	At5g06100	AtMYB033	Développement des anthères, de la graine et de la plante via la régulation par miR159	Millar and Gubler 2005, Reyes and Chua 2007
ATR1	At5g60890	AtMYB034	Régulation du métabolisme des tryptophanes et du glucosinolate indolique	Bender et Fink 1998; Celenza et <i>al.</i> , 2005
RAX1	At5g23000	AtMYB037	Initiation de la formation des méristèmes latéraux	Keller et <i>al.</i> , 2006; Müller et <i>al.</i> , 2006
RAX2	At2g36890	AtMYB038	Initiation de la formation des méristèmes latéraux	Müller et <i>al.</i> , 2006
AtMYB60	At1g08810	AtMYB060	Régulation de l'ouverture des stomates et de la tolérance à la sécheresse	Cominelli et al., 2005
AtMYB61	At1g09540	AtMYB061	Production du mucilage de la graine, lignification, rôle dans l'ouverture des stomates	Penfield et <i>al.</i> , 2001; Newman et <i>al.</i> , 2004; Liang et <i>al.</i> , 2005
AtMYB65	At3g11440	AtMYB065	Développement des anthères	Millar et Gubler 2005
WER	At5g14750	AtMYB066	Initiation des poils racinaires	Lee et Schiefelbein 1999
AtMYB68	At5g65790	AtMYB068	Développement racinaire, lignification	Feng et al., 2004

TableauI-3.ProtéinesàdomaineMYB-R2R3caractérisées

fonctionnellement chez A. thaliana (Adapté de Baudry, 2004).

Suite et légende page suivante.

Nom du gène	Locus ID	Nom générique	Processus régulé	Références
PAP1	At1g56650	AtMYB075	Métabolisme des anthocyanes	Borevitz et al., 2000
RAX3	At3g49690	AtMYB084	Initiation de la formation des méristèmes latéraux	Muller et <i>al.</i> , 2006
AtMYB88	At2g02820	AtMYB088	Développement des stomates	Lai et al., 2005
PAP2	At1g66390	AtMYB090	Métabolisme des anthocyanes	Borevitz et al., 2000
AS1	At2g37630	AtMYB091	Différenciation des cellules foliaires	Byrne et al., 2000
AtMYB98	At4g18770	AtMYB098	Développement du gamétophyte femelle	Kasahara et al., 2005
AtMYB101	At2g32460	AtMYB101	Germination des graines via la régulation de AtMYB101 par l'ABA et miR 159	Reyes and Chua 2007
AtMYB102	At4g21440	AtMYB102	Réponses à la blessure et aux stress osmotiques	Denekamp et Smeekens 2003
AtMYB103	At1g63910	AtMYB103	Développement tapetum et trichomes	Higginson et al., 2003
BOS1	At3g06490	AtMYB108	Réponses aux stress biotiques et abiotiques	Mengiste et al., 2003
TT2	At5g35550	AtMYB123	Métabolisme des proanthocyanidines (graines)	Nesi et <i>al.</i> , 2001; Baudry et <i>al.</i> , 2004
FLP	At1g14350	AtMYB124	Développement des stomates	Yang et Sack., 1995; Lai et al., 2005

TableauI-3 (suite).Protéines à domaineMYB-R2R3caractériséesfonctionnellement chez A. thaliana (Adapté de Baudry, 2004).

AS1: asymmetric leaves 1; ATR1: altered tryptophane regulation 1; BOS1: botrytis susceptible 1; FLP: four lips; GL1: glabra 1; HOS10: hight expression of osmotically responsive genes; LAF1: long after far-red light 1 ; PAP:production of anthocyanin pigment ; RAX: regulator of axillary meristem; TT2: transparent testa 2; WER: werewolf.

Nom du gène	Locus ID	Type de MYB	Processus régulé	Références
APL	At1g79430	Domaine MYB et coil-coil	Développement des tissus vasculaires	Bonke et <i>al.</i> , 2003
AtMYBL2	At1g71030	Une seule répétition MYB	Initiation des trichomes	Sawa 2002
CCA1	At2g46830	Une seule répétition MYB	Rythme circadien	Wang et al., 1997
CPC	At2g46410	Une seule répétition MYB	Production des poils racinaires et des trichomes	Wada et <i>al.</i> , 1997; Lee et Schiefelbein, 2002
EPR1	At1g18330	Une seule répétition MYB	Rythme circadien	Kuno et al., 2003
ETC1	At1g01380	Une seule répétition MYB	Production des poils racinaires et des trichomes	Kirik et <i>al.,</i> 2004
LHY	At1g01060	Une seule répétition MYB	Rythme circadien	Schaffer et al., 1998
LUX	At3g46640	Une seule répétition MYB	Rythme circadien	Hazen et al., 2005
PHR1	At4g28610	Domaine MYB et coil-coil	Voie de signalisation du phosphate inorganique	Rubio et <i>al.</i> , 2001
TRY	At5g53200	Une seule répétition MYB	Production des poils racinaires et des trichomes	Schellmann et <i>al.</i> , 2002

Tableau I-4. Protéines MYB atypiques caractérisées fonctionnellement chez

A. thaliana (Adapté de Baudry, 2004).

APL: altered phloem development; CCA1: circadian clock associated 1; CPC: caprice; EPR1: early phytochrome responsive 1; ETC1: enhancer of TRY and CPC1; LHY: late elongated hypocotyl; LUX: lux arrhythmo; PHR1:phosporus starvation response 1; TRY: triptychon.

Encore peu d'études se détachent de la plante modèle Arabidopsis pour s'intéresser à des espèces d'intérêt économique, comme les arbres par exemple. Chez les conifères, peu de gènes MYB avaient été caractérisés avant la réalisation des travaux présentés dans cette thèse, soit PtMYB1 et 4 (Newman et al., 2004; Patzlaff et al., 2003a, 2003b), et PmMBF1 (Xue et al., 2003). De plus, une dizaine de séquences codantes partielles ont été obtenues de différents conifères (Kusumi et al., 2002). Néanmoins, la multiplication des projets de séquençage ces dernières années a fait augmenter le nombres d'EST et de cDNA correspondant aux facteurs de transcription dans les bases de données de séquences publiques (NCBI, TIGR gene index, etc.). Ainsi, plusieurs bases de données avec interface web permettent l'accès à une large collection de facteurs de transcription, et parfois d'éléments cis régulateurs, dans plusieurs espèces végétales comme Arabidopsis thaliana (AGRIS et AtRegNet, Palaniswamy et al., 2006, Davuluri et al., 2003), Populus trichocarpa (peuplier), Oryza sativa (riz), et d'autres encore (PlnTFDB, Riano-Pachon et al., 2007, http://plntfdb.bio.unipotsdam.de). Récemment, l'analyse préliminaire du premier génome nucléaire séquencé d'un arbre, celui du peuplier (Populus trichocarpa), a mis en évidence la présence de 45 555 gènes probables dont 11 654 seraient orthologues avec Arabidopsis et au moins 200 sont des facteurs de transcription de type MYB (Tuskan et al., 2006; www.jgi.doe.gov/poplar; www.phytozome.net). Concernant les conifères aucun génome n'est séquencé et seulement 2 espèces sont actuellement représentées dans le « TIGRgene index » (Pinus et Picea) contre 32 chez les angiospermes incluant les espèces séquencées Arabidopsis et Populus.

I-2-5-Régulation de l'activité des facteurs de transcription MYB

Les facteurs de transcription sont la cible de nombreuses régulations allant de la synthèse de l'ARN messager (régulation de la transcription même du FT et modifications transcriptionnelles) à la synthèse d'une protéine active (modification post-transcriptionnelle), en passant par son transport dans le noyau, ou ses interactions avec d'autres protéines et avec l'ADN. Au niveau de la régulation transcriptionnelle, des éléments nucléotidiques pouvant être des *cis* éléments régulateurs ont été identifiés dans les régions non codantes comme les promoteurs mais aussi dans les introns. Chez le coton, Wang et *al* (2004) ont démontré que le premier intron des deux MYB GL1 et GaMYB2 joue un rôle dans la formation des trichomes, celui-ci aurait un rôle d'activation dans les trichomes et de répression dans les cellules non trichomes. L'activation des facteurs de transcription et donc de la transcription de gènes cibles dépend de la perception et de la transduction d'un stimulus par la cellule. Yoo et al. (2005) ont mis en évidence la présence d'un motif en acides aminés dans le domaine MYB de certaines protéines, comme AtMYB2 par exemple, pouvant interagir avec une calmoduline dont la fonction est de répondre aux variations de calcium. De même, l'état général d'oxydoréduction de la cellule peut avoir un effet sur la capacité de fixation à l'ADN des protéines. Les transcrits d'ARN de certains gènes MYB tels que AtMYB33 et AtMYB101 par exemple sont la cible des micros ARN d'environ 40 ribonucléotides complémentaires au brin transcrit, qui conduisent à la dégradation spécifique des transcrits (Reves et Chua, 2007). Enfin, d'un point de vue plus général au niveau physiologique de la plante, des blessures, des pathogènes, le froid, les privations nutritives, la lumière, le rythme circadien et encore d'autres facteurs, peuvent induire l'expression des facteurs de transcription MYB.

I-2-6-Le complexe protéique MYB/bHLH/WD40

Le mécanisme de la transcription des gènes fait intervenir de nombreux acteurs protéiques qui forment des complexes transcriptionnels. Certaines protéines de type MYB ont un ou plusieurs partenaires protéiques, obligatoires ou facultatifs, pour assurer la régulation transcriptionnelle du gène cible (Figure I-9; Serna et Martin, 2006; Guimil et Dunand, 2006; Ramsay et Glover, 2005). Les études de sur- et sous-expression des gènes chez *Arabidopsis* couplées à des analyses d'interactions protéine-protéine ont montré qu'un complexe protéique composé de MYBs à une ou deux répétitions avec des protéines bHLH et WD40 était à l'origine de l'initiation de multiples voies de différenciation cellulaire dans quelques espèces végétales (Bernhardt et *al.*, 2003; Zhang et *al.*, 2003; Quattrochio et *al.*, 1998; Spelt et *al.*, 2000).



Figure I-9. Représentation schématique des interactions moléculaires entre des protéines de type MYB et ses partenaires facteurs de transcription bHLH et WD40 (Adapté de Broun et *al.*, 2005).

La figure illustre les processus par lesquels les différents régulateurs sont impliqués et leur influence dans le devenir morphologique et métabolique des cellules chez *Arabidopsis thaliana*. Les flèches indiquent un effet positif sur l'activation des voies et les barres en T montrent un effet de répression. Les associations entre TTG1 et un partenaire MYB n'ont été démontrées que dans le cas de TT2. Les MYB de type R2R3 sont indiqués en gras. Les rôles et fonctions biologiques indiqués en ombré signifient que l'implication de la protéine dans le mécanisme biologique a été démontrée mais pas son interaction avec d'autres protéines.

Les protéines WD40 de la famille des protéines à hélice-béta se caractérisent par la présence d'une région centrale de 20 acides aminés définie par un dipeptide glycine-histidine (GH) et un tryptophane-aspartate (WD) (Smith et al., 1999). Ce motif est répété en tandem de 4 à 16 fois dans la même protéine. La plus étudiée des protéines WD40 est la sous-unité béta de la protéine hétérotrimérique G, impliquée dans la transduction de signaux entre autres. Une fonction commune des unités de répétitions WD40 est de faciliter les interactions protéine-protéine et n'ont pas de fonctions enzymatiques intrinsèques. Plusieurs protéines WD40 sont impliquées dans certains aspects de différenciation cellulaire de l'épiderme végétal et ont été identifiées chez le pétunia (Anthocyanin 11 ou AN11, anthocyanes, de Vetten et al., 1997), Arabidopsis (Transparent Testa 1 ou TTG1, trichomes, poils racinaires, anthocyanines, stomates et forme cellulaire dans l'épiderme, mucilage de la graine et pigmentation de la graine, Walker et al., 1999, Berger et al., 1998, Galway et al., 1994), Perilla (Perilla frutescens WD ou PFWD, anthocyanines, Sompornpailin et al., 2002) et le maïs (Pale Aleurone Color 1 ou PAC1, trichomes, anthocyanines, mucilage de la graine et pigmentation de la graine, Carey et al., 2004). Ces protéines appartiennent à un même groupe phylogénétique et peuvent former une surface sur laquelle peuvent interagir d'autres protéines (Carey et al., 2004; Sompornpailin et al., 2002). L'expression ectopique d'une protéine WD40 n'appartenant pas au génome d'Arabidopsis dans le génotype mutant de ttgl complémente plusieurs des phénotypes mutants, indiquant que cette protéine peut interagir avec des partenaires protéiques similaires. Des analyses génétiques et des expérimentations de double hybride chez la levure ont identifié des partenaires appartenant à la classe des MYB et des bHLH (Bernhardt et al., 2003; Zhang et al., 2003).

Les bHLH sont des protéines capables de se lier à l'hexanucléotide « Ebox » *via* leurs région basique, CANNTG (Voronova et Baltimore, 1990), et de faire des homo et hétérodimères avec leurs région HLH (Murre et *al.*, 1994; Massari et Murre, 2000). Chez les plantes, la première bHLH identifiée est Leaf color, ou Lc, impliquée dans la pigmentation de tissus spécifiques *via* des anthocyanes chez le maïs (Ludwig et *al.*, 1989). Depuis, d'autres bHLH capables d'induire des modifications phénotypiques similaires ont été identifiées chez *Arabidopsis* (GL3, EGL3, TT8), le maïs (Lc, R), le pétunia (JAF13, AN1) et *Antirrhinum* (DELILA) (Ramsay et Glover, 2005; Guimil et Dunand, 2006). Comme dans le cas des WD40, la surexpression de Lc dans le génotype mutant *ttg1* permet de complémenter les phénotypes mutants (e.g. présence d'anthocyanines), indiquant l'implication de bHLH similaires à Lc dans d'autres aspects de la différenciation cellulaire (Lloyd et *al.*, 1994). De plus, Lc agit avec une protéine de type MYB-R2R3, MYB C1 ou Colourless 1 (Cone et *al.*, 1986; Paz-Ares et *al.*, 1987), pour l'activation de l'expression des gènes de la voie de biosynthèse des anthocyanes chez le maïs (Goff et *al.*, 1992; Cone et *al.*, 1993, Hernandez et *al.*, 2004). Cette interaction entre MYB et bHLH s'effectue par l'entremise d'un motif protéique présent dans la première hélice de la répétition R3 du MYB-R2R3 (Grotewold et *al.*, 2000; Zimmermann et *al.*, 2004).

Chez les conifères, le MYB-R2R3 de *Picea marianna* PmMBF1, phylogénétiquement proche de C1, est capable de transactiver le promoteur de maïs Bronze 2 de la synthèse des anthocyanes (Bz2, Bodeau et Walbot, 1992), en combinaison avec le gène codant le bHLH R de maïs dans du tissu embryogène de mélèze et d'épinette, sans pour autant amener à l'apparition d'anthocyanes colorées (Xue et *al.*, 2003). Au contraire, la combinaison de PmMBF1 avec le bHLH R de maïs introduit dans des cellules en suspension de maïs résulte en l'apparition d'îlots de cellules rouges (Xue *et al.*, 2003). Finalement, si ce complexe protéique tripartite se révèle être conservé entre quelques espèces, il apparaît que les conséquences phénotypiques ne sont pas strictement identiques entre des angiospermes et des conifères, plus distants évolutivement. Néanmoins, il a été démontré que le rôle général des protéines MYB-R2R3 de type C1 (maïs), AN2 (pétunia) et PmMBF1(épinette) réside dans l'activation transcriptionnelle de gènes de la voie de biosynthèse des phénylpropanoïdes et sont fonctionnellement interchangeables (Quattrocchio et *al.*, 1999; Xue et *al.*, 2003).

I-2-7-Régulation transcriptionnelle de l'expression des gènes du métabolisme des phénylpropanoïdes

Des études fonctionnelles des régions promotrices des gènes de la voie de biosynthèse des phénylpropanoïdes ont montré que la présence de courts motifs

nucléotidiques (5-10 nucléotides) étaient nécessaires à une expression dans les cellules du xylème en lignification (PAL de haricot : Levya et al. 1992, Séguin et al, 1997; 4CL de Persil : Hauffe et al, 1993 ; TED3 pour Tracheary Element Differenciation de Zinnia elegans : Igarashi et al, 1998 ; CCoAOMT de Pinus taeda: Li et al, 1999; CCR d'Eucalyptus: Lacombe et al, 2000; CAD d'Eucalyptus : Lauvergeat et al, 2002). Ces motifs sont appelés des éléments AC de part leur contenu nucléotidique riche en A et C. Il s'agit de sites de fixation des protéines de type MYB-R2R3 qui agissent en tant que facteurs de transcription. Les interactions entre éléments AC et protéines MYB-R2R3 ont été mises en évidence via la technique de gel retard. Par exemple, Sablowski et al. (1994) ont montré qu'une protéine MYB d'Anthirinnum majus (MYB305) spécifique aux fleurs se liait au motif CCTACC présent sur les promoteurs des gènes de la voie de synthèse des phénylpropanoïdes. La caractérisation fonctionnelle de certains gènes tels que les facteurs de transcription MYB par mutagènèse, par sur-expression ou sous-expression a montré leur implication dans la synthèse de phénylpropranoïdes (lignines et dérivés anthocyaniques) en lien avec différents processus biologiques (Tableau I-3). Ainsi, il apparaît qu'un nombre important de facteurs de transcription de type MYB pourraient être impliqués de près ou de loin dans la régulation génétique de la formation du bois via les gènes des phénylpropanoïdes et des lignines (Rogers et Campbell, 2004).

En ce qui concerne les arbres forestiers, deux MYB-R2R3 (PtMYB1 et 4) isolés de xylème de pin sont exprimés dans les cellules en cours de lignification et sont capables de fixer des éléments AC du gène PAL (Patzlaff et *al*, 2003a, 2003b; Newman et *al*, 2003). Récemment, Giocoechea et *al*. (2005) ont mis en évidence qu'un MYB-R2R3 chez l'Eucalyptus (EgMYB2), agit comme activateur transcriptionnel de plusieurs gènes des phénylpropanoïdes et des lignines par surexpression dans le tabac et se fixe sur des éléménts AC. Au contraire, des gènes liés aux phénylpropranoïdes peuvent être réprimés par exemple chez *Arabidopsis thaliana* (AtMYB4 : Jin et *al.*, 2000), *Antirrhinum majus* (AmMYB308 : Tamagnone et *al.*, 1998), le maïs (ZmMYB31 et 42 : Fornalé et *al.*, 2006) ou encore la fraise (FaMYB1 : Aharoni et *al.*, 2001). Dans les plantes ligneuses comme les arbres, PttMYB21a (peuplier) serait impliqué dans la répression transcriptionnelle du gène *CCoAOMT* (Karpinska et *al.*, 2004),

alors que EgMYB1 (eucalyptus) réprimerait les gènes *CAD* et *CCR* (Legay et *al.*, 2007).

De façon plus exhaustive, une analyse informatique portant sur les promoteurs de gènes *d'Arabidopsis thaliana* a démontré la présence d'éléments AC dans 10 des 34 gènes impliqués dans la biosynthèse des monolignols exprimés dans les tissus du xylème en lignification (Raes et *al*, 2003). Par ailleurs, aucun élément AC n'a été détecté sur les régions promotrices des gènes impliqués dans la synthèse des lignines S synthétisées uniquement chez les angiospermes et principalement dans les fibres. La spécificité des éléments AC à cette partie de la voie de synthèse suggère que les gènes intervenant dans la synthèse des unités G dérivent d'ancêtres génétiques commun (présents chez les angiospermes et les gymnospermes) et pourraient être régulés de façon coordonnée par les MYB et par l'intermédiaire des éléments AC.

I-3-Objectifs du travail de thèse et présentation

I-3-1-Hypothèses et objectifs des travaux de thèse

La caractérisation fonctionnelle des gènes de quelques facteurs de transcription de type MYB-R2R3 chez Arabidopsis et le tabac suggère fortement un lien entre le produit de ces gènes et la régulation coordonnée de l'expression des gènes impliqués dans la formation de la paroi cellulaire secondaire, produite lors de la xylogénèse. Certains gènes MYB contrôleraient la synthèse de la lignine, alors que d'autres interviendraient dans l'organisation tissulaire. À ce jour, quelques études ont analysé des gènes MYB provenant d'espèces ligneuses (pin, épinette, eucalyptus, peuplier), et ont démontré la capacité des facteurs qu'ils codent à se lier à l'ADN et à activer la transcription. Toutefois, les analyses fonctionnelles de ces gènes ont été menées sur des espèces annuelles, donc à croissance essentiellement primaire, de la famille des angiospermes (tabac, arabette). Au début de mes travaux, trois gènes appartenant à cette famille avaient été décrits chez les conifères et analysés par expression dans des systèmes hétérologues. Par conséquent, le rôle exact des MYB-R2R3 dans la croissance vasculaire secondaire et la synthèse des composés pariétaux comme la lignine restait à préciser chez les conifères.

Mon travail de thèse est basé sur les hypothèses suivantes :

<u>Hypothèse 1.</u> Il existe de nombreux gènes *MYB* de type R2R3 chez les gymnospermes, plus précisément ches les Pinacées.

Cette hypothèse s'appuie sur plusieurs études récentes réalisées chez le groupe des angiospermes. Les génomes séquencés à ce jour contiennent tous plus de 100 gènes différents appartenant à cette classe. Il découle deux objectifs de cette hypothèse : d'abord, isoler plusieurs gènes *MYB-R2R3* chez le pin loblolly (*Pinus taeda*) et l'épinette blanche (*Picea glauca*), et ensuite, caractériser la structure de la famille en comparant les séquences isolées à celles des angiospermes, dans le but d'en évaluer l'évolution.

<u>Hypothèse 2.</u> Les gènes *MYB-R2R3* de l'épinette présentent des profils d'expression en ARNs messagers nettement diversifiés.

Un corrolaire de cette hypothèse est que les profils d'expression variés reflètent la diversité des rôles biologiques. L'objectif général relié à cette hypothèse visait à obtenir des informations sur l'expression des MYBs en lien avec le développement normal et suite à des traitements ou des stress qui modulent le développement, ou bien des stress qui induisent des réponses de défense. Dans le cadre de mes travaux, les niveaux d'expression ont été déduits d'après l'accumulation et la fluctuation des ARN transcrits déterminés par RT-PCR quantitative.

<u>Hypothèse 3.</u> Certains des facteurs de transcription de type MYB-R2R3 sont impliqués dans la formation de la paroi secondaire du xylème chez les conifères.

Ici, l'objectif a été d'analyser le rôle potentiel de certains gènes *MYB-R2R3* dans la formation de la paroi secondaire en manipulant leur expression par transformation stable chez l'épinette blanche. Mes travaux ont ciblé trois gènes MYBs différents isolés du pin (*Pinus taeda*) et surexprimés chez l'épinette, un système d'expression "quasi-homologue". Ces travaux ont été rendus possibles de part la mise en place de plateformes spécialisées en transcriptomique et en transgénèse dans le cadre d'un projet de recherche ARBOREA.

I-3-2-Présentation des travaux de thèse

Mes travaux de thèse ont consisté en l'identification de plusieurs gènes codant des facteurs de transcription de type MYB-R2R3 en lien avec la formation du bois chez les conifères et, la caractérisation fonctionnelle de trois d'entre eux en système quasi-homologue par transformation stable chez l'épinette blanche.

Je présente dans un premier temps l'identification et l'analyse de séquences de 13 *MYB-R2R3* de *Picea glauca* et 5 de *Pinus taeda* ainsi que les niveaux d'ARN messagers des *MYB* de *P. glauca* en lien avec la formation du bois chez l'épinette blanche (Chapitre II). Je montre que trois de ces gènes sont préférentiellement exprimés dans le xylème en différenciation de tige et de racine

ainsi que dans le bois de compression après 76 heures d'inclinaison de la même façon que certains gènes de la voie de biosynthèse des phenylpropanoïdes. Puis, dans une seconde partie, j'aborde la caractérisation fonctionnelle de PtMYB1 et PtMYB8 par surexpression constitutive dans de jeunes plantules d'épinettes blanches in vitro (Chapitre III), et dans des jeunes arbres transférés en sol dans le cas de *PtMYB1* (Chapitre IV). Cette étude suggère fortement l'implication de ces gènes dans la régulation de la formation de la paroi secondaire lignifiée. Elle permet aussi de proposer l'existence d'un réseau impliquant des intéractions entre différents MYB-R2R3. Finalement, je présente la caractérisation d'un nouveau groupe de MYB appartenant au sous-groupe 4 (selon Krantz et al., 1998), le sousgroupe 4-C, qui est particulier aux conifères de part sa structure et sa composition. Ce sous-groupe est représenté entre autres par PtMYB14 qui a été soumis à une analyse fonctionnelle (Chapitre V et VI). Ce travail a nécessité la caractérisation du promoteur de l'alcool cinnamylique dehydrogénase (CAD) de P. glauca dont j'ai analysé le profil d'expression préférentiel aux tissus vasculaires (Chapitre V et VI). Mes résultats impliquent PtMYB14 dans les mécanismes de réactions de défense aux stress biotiques et abiotiques chez l'épinette blanche.

I-4-Références

- Abe H., Urao T., Ito T., Seki M., Shinozaki K. et Yamaguchi-Shinozaki K. (2003). Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *Plant Cell* 15:63-78
- Agarwal M., Hao Y., Kapoor A., Dong C.H., Fujii H., Zheng X. et Zhu J.K. (2006). A R2R3 type MYB transcription factor is involved in the cold regulation of CBF genes and in acquired freezing tolerance. *J Biol Chem* 281: 37636-37645
- Agarwal U.P. (2006). Raman imaging to investigate ultrastructure and composition of plant cell walls: distribution of lignin and cellulose in black spruce wood (*Picea mariana*). *Planta* 224: 1141-1153
- Aharoni A., De Vos C.H., Wein M., Sun Z., Greco R., Kroon A., Mol J.N. et O'Connell A.P. (2001). The strawberry FaMYB1 transcription factor suppresses anthocyanin and flavonol accumulation in transgenic tobacco. *Plant J* 28: 319-332
- Aloni R. (1987). Differentiation of vascular tissues. Annu Rev Plant Physiol 38: 179-204
- Ansieau S., Kowenz-Leutz E., Dechend R. et Leutz A. (1997). B-Myb, a repressed trans-activating protein. *J Mol Med* 75: 815-819
- Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**: 791-815

- Atanassova R., Favet N., Martz F., Chabbert B., Tollier M-T, Monties B., Fritig B. et Legrand M. (1995). Altered lignin composition in transgenic tobacco expressing O-methyltransferase sequences in sense and antisense orientation. *Plant J* 8: 465-77
- Avila J., Nieto C., Canas L., Benito M.J. et Paz-Ares J. (1993). Petunia hybrida genes related to the maize regulatory C1 gene and to animal myb protooncogenes. *Plant J* 3: 553-562
- Badescu G.O. et Napier R.M. (2006). Receptors for auxin: will it all end in TIRs? *Trends Plant Sci* 11: 217-223
- Ballesteros M.L., Bolle C., Lois L.M., Moore J.M., Vielle-Calzada J.P., Grossniklaus U. et Chua N.H. (2001). LAF1, a MYB transcription activator for phytochrome A signaling. *Genes Dev* 15: 2613-2625
- Barrrière Y. et Argillier O. (1993). Brown-midrib genes of maize : a review. *Agronomie* 13: 865-876
- Baucher M., Chabber B., Pilate G., Van Doorsselaere J., Tollier M-T, Petit-Conil M., Cornu D., Monties B., Van Montagu M., Inzé D., Jouanin L. et Boerjan W. (1996). Red xylem and higher lignin extractability by downregulating cinnamyl alcool dehydrogenase in poplar (*Populus trembula* and *Populus alba*). *Plant Physiol* 112: 1479-90
- Baucher M., Halpin C., Petit-Conil M. et Boerjan W. (2003). Lignin: genetic engineering and impact on pulping. *Crit Rev Biochem Mol Biol* 38: 305-350
- **Baudry A. (2004).** Analyse fonctionnelle des protéines TT2, TT8 et TTG1: étude de leur rôle dans la régulation de la biosynthèse des flavonoïdes chez *Arabidopsis thaliana. Thèse de doctorat en Sciences de la Vie, Université Paris 11, Orsay*
- Baudry A., Heim M., Dubreucq B., Caboche M., Weisshaar B. et Lepiniec L.
 (2004). TT2, TT8, andTTG1 synergistically specify the expression of *BANYULS* and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. *Plant J* 39: 366-380
- Bender J. et Fink G. R. (1998). A Myb homologue, ATR1, activates tryptophan gene expression in Arabidopsis. *Proc Natl Acad Sci USA* 95: 5655-60
- Berger F., Hung C.Y., Dolan L. et Schiefelbein J. (1998). Control of cell division in the root epidermis of Arabidopsis thaliana. *Dev Biol* 194: 235-245
- Bernhardt C., Lee M.M., Gonzalez A., Zhang F., Lloyd A. et Schiefelbein J. (2003). The bHLH genes GLABRA3 (GL3) and ENHANCER OF GLABRA3 (EGL3) specify epidermal cell fate in the *Arabidopsis* root. *Development* 130: 6431-6439
- Besseau S., Hoffmann L., Geoffroy P., Lapierre C., Pollet B. et Legrand M. (2007). Flavonoid accumulation in Arabidopsis repressed in lignin synthesis affects auxin transport and plant growth. *Plant Cell* **19**: 148-162
- Bhalerao R.P et Bennett M.J. (2003). The case for morphogens in plants. *Nat Cell Biol* 5: 939-943
- Biedenkapp H., Borgmeyer U., Sippel A.E. et Klempnauer K.H. (1988). Viral myb oncogene encodes a sequence-specific DNA-binding activity. *Nature* 335: 835-837
- Blackwell T.K. et Weintraub H. (1990). Differences and similarities in DNAbinding preferences of MyoD and E2A protein complexes revealed by binding site selection. *Science* 250: 1104-1110
- Blancaflor E.B. (2002). The cytoskeleton and gravitropism in higher plants. *J Plant Growth Regul* 21: 120-136

- Bochar D.A., Friesen J.A., Stauffacher C.V. et Rodwell W. (1999). Biosynthesis of mevalonic acid from acetyl-CoA. In: Cane DE, ed. *Isoprenoids, including carotenoids and steroids*, Vol. 2. *Comprehensive natural products chemistry*. London, UK: Elsevier, p 15-44
- Bodeau J.P. et Walbot V. (1992). Regulated transcription of the maize Bronze-2 promoter in electroporated protoplasts requires the *C1* and *R* gene products. *Mol Gen Genet* 233: 379-387
- Boerjan W., Ralph J. et Baucher M. (2003). Lignin biosynthesis. Annu Rev Plant Biol 54: 519–546
- Bonke M., Thitamadee S., Mähonen A. P., Hauser M. T. et Hedariutta Y. (2003). APL regulates vascular tissue identity in Arabidopsis. *Nature* 426: 181-186
- Borevitz J. O., Xia Y., Blount J., Dixon R. A. et Lamb C. (2000). Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *Plant Cell* 12: 2383–2394
- Boudet A.M. et Grima-Pettenati J. (1996). Lignification genes: structure, function and regulation. Proc. Keystone Symp. *Extracell. Matrix Plants: Mol. Cell. Dev. Biol., Tamarron, Colo.*, p.7 *New York: Wiley-Liss*
- Boudet A.M, Kajita S., Grima-Pettenati J. et Goffner D. (2003). Lignins and lignocellulosics: a better control of synthesis for new and improved uses. *Trends Plant Sci* 8: 576-581
- Bourquin S., Nishikubo N., Abe H., Brumer H., Denman S., Eklund M. Christiernin M., Teeri T.T., Sundberg B. et Mellerowicz E. J. (2002). Xyloglucan endotransglycosylases have a function during the formation of secondary cell walls of vascular tissues. *Plant Cell* 14 : 3073-3088
- Bozhkov P.V., Filonova L.H. et Suarez M.V. (2005). Programmed cell death in plant embryogenesis. *Curr Top Dev Biol* 67: 135-179
- Braun E.L. et Grotewold E. (1999). Newly discovered plant c-myb-like genes rewrite the evolution of the plant *myb* gene family. *Plant Physiol* 121: 21-24
- Brett C.T. (1996). Cellulose microfibrils in plants: biosynthesis, deposition, and integration into the cell wall. *Int Rev Cytol* 199: 161-99
- **Broun P. (2005).** Transcriptional control of flavonoid biosynthesis: a complex network of conserved regulators involved in multiple aspects of differentiation in Arabidopsis. *Cur Opin Plant Biol* **8:** 272-279
- Byrne M.E., Barley R., Curtis M., Arroyo J.M., Dunham M., Hudson A. et Martienssen R. A. (2000). Assymetric leaves 1 mediates leaf patterning and cell function in Arabidopsis. *Nature* 408: 967-971
- Cane D. E. (1999). Sesquiterpene biosynthesis : cyclization mechanisms. In: Cane DE, ed. Isoprenoids, including carotenoids and steroids, Vol. 2. Comprehensive natural products chemistry. London, UK: Elsevier, p 15-44
- Cano-Delgado A., Penfield S., Smith C., Catley M. et Bevan M. (2003). Reduced cellulose synthesis invokes lignification and defense responses in Arabidopsis thaliana. *Plant J* 34: 351-362
- Carey C.C., Strahle J.T., Selinger D.A. et Chandler V.L. (2004). Mutations in the *pale aleurone color1* regulatory gene of the *Zea mays* anthocyanin pathway have distinct phenotypes relative to the functionally similar *TRANSPARENT TESTAGLABRA1* gene in *Arabidopsis thaliana*. *Plant Cell* 16: 450-464
- Carlsbecker A. et HelariuttaY. (2005). Phloem and xylem specification: pieces of the puzzle emerge. *Curr Opin Plant Biol* 8: 512-517

- Cedroni M.L., Cronn R.C., Adams K.L., Wilkins T.A. et Wendel J.F. (2003). Evolution and expression of MYB genes in diploid and polyploid cotton. *Plant Mol Biol* 51: 313-325
- Celenza J.L., Quiel J.A., Smolen G.A., Merrikh H., Silvestro A.R., Normanly J. et Bender J. (2005). The Arabidopsis ATR1 Myb transcription factor controls indolic glucosinolate homeostasis. *Plant Physiol* 137: 253-262
- Chaffey N. (2002). Why is there so little research into the cell biology of the secondary vascular system of trees? *New Phytol* 153: 213-223
- Chappell J. (1995). The biochemistry and molecular biology of isoprenoid metabolism. *Plant Physiol* 107 : 1-6
- Chapple C.C., Vogt T., Ellis B. E. et Sommerville C.R. (1992). An Arabidopsis mutant defective in the general phenylpropanoid pathway. *Plant Cell* 4: 1413-24
- Chen Y., Yang X., He K., Liu M., Li J., Gao Z., Lin Z., Zhang Y., Wang X., Qiu X., Shen Y., Zhang L., Deng X., Luo J., Deng X. W., Chen Z., Gu H. et Qu L. J. (2006). The MYB transcription factor superfamily of *Arabidopsis*: expression analysis and phylogenetic comparison with the rice MYB family. *Plant Mol Biol* 60: 107–124
- Christiansen E., Waring R.H. et Berryman A.A. (1987). Resistance of conifers to bark beetle attack: Searching for general relationships. *Forest Ecol Managt* 22: 89-106
- Citovsky V. et Zambryski P. (2000). Systemic transport of RNA in plants. *Trends Plant Sci* **5**: 52-54
- Colonna P. (2006). La chimie verte. Ed. TEC & DOC, Lavoisier, p. 103
- Cominelli E., Galbiati M., Vavasseur A., Conti L., Sala T., Vuylsteke M., Leonhardt N., Dellaporta S.L. et Tonelli C. (2005). A guard-cell-specific MYB transcription factor regulates stomatal movements and plant drought tolerance. *Curr Biol* 15: 1196-1200
- Cone K.C., Burr F.A. et Burr B. (1986). Molecular analysis of the maize anthocyanin regulatory locus C1. *Proc Natl Acad Sci USA* 83:9631-9635
- Cone K.C., Cocciolone S.M., Burr F.A. et Burr B. (1993). Maize anthocyanin regulatory gene *pl* is a duplicate of *c1* that functions in the plant. *Plant Cell* 5: 1795-1805
- Cosgrove D.J. (1997). Assembly and enlargement of the primary cell wall in plants. *Annu Rev Cell Dev Biol* 13 : 171-201
- **Cosgrove D.J., Li L.C., Cho H.T., Hoffman-Benning S., Moore R.C. et Blecker D. (2002).** The growing world of expansins. *Plant Cell Physiol* **43** : 1436-1444
- Darley C.P., Forrester A.M. et McQueen-Mason S.J. (2001). The molecular basis of plant cell wall extension. *Plant Mol Biol* 47 : 179-195
- **Davis E.M. et Croteau R. (2000).** Cyclization enzymes in the biosynthesis of monoterpenes, sesquiterpenes, and diterpenes. In : Leeper FJ, Vederas JC, eds. *Biosynthesis : aromatic polyketides, isoprenoids, alkaloids*, Vol 209
- Davuluri R.V., Sun H., Palaniswamy S.K., Mathews N., Molina C., Kurts M. et Grotewold E. (2003). AGRIS: Arabidopsis gene regulatory information server, an information resource of Arabidopsis *cis*-regulatory elements and transcription factors. *BMC Bioinformatics* 4: 25
- Demura T., Tashiro G., Horiguchi G., Kishimoto N., Kubo M., Matsuoka N., Ninami A., Nagata-Hiwatashi M., Nakamura K., Okamura Y., Suzuki S., Yazaki J., Kikuchi S. et Fukuda H. (2002). Visualization by

comprehensive microarray analysis of gene expression programs during transdifferentiation of mesophyll cells into xylem cells. *Proc Natl Acad Sci USA* **99:** 15794-15803

- Demura T. et Fukuda H. (2007). Transcriptional regulation in wood formation. Trends Plant Sci 12: 64-70
- **Denekamp M. et Smeekens S.C. (2003).** Integration of wounding and osmotic stress signals determines the expression of the AtMYB102 transcription factor gene. *Plant Physiol* **132**: 1415-1423
- de Vetten N., Quattrocchio F., Mol J. et Koes R. (1997). The *an11* locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals. *Genes Dev* **11**: 1422-1434
- Dewick P.M. (2002). The biosynthesis of C5-C25 terpenoid compounds. *Natural Product Reports* 19 : 181-222
- Diaz I., Vicente-Carbajosa J., Abraham Z., Martinez M., Isabel-LaMoneda I. et Carbonero P. (2002). The GAMYB protein from barley interacts with the DOF transcription factor BPBF and activates endosperm-specific genes during seed development. *Plant J* 29: 453-464
- Dickison W.C. (2000). Integrative Plant Anatomy. Harcourt, Academic Press
- Dixon R.A. et Paiva N. L. (1995). Stress-induced phenylpropanoid metabolism. *Plant Cell* 7: 1085-1097
- **Doblin M.S., Kurek I., Jacob-Wilk D. et Delmer D.P. (2002).** Cellulose biosynthesis in plants : from genes to rosettes. *Plant Cell Physiol* **43** : 1407-1420
- **Dubendorff J.W., Whittaker L.J., Eltman J.T. et Lipsick J.S. (1992).** Carboxyterminal elements of c-Myb negatively regulate transcriptional activation in *cis* and in *trans. Genes Dev* **6:** 2524-2535
- Emery J.F., Floyd S.K., Alvarez J., Eshed Y., Hawker N.P., Izhaki A., Baum S.F. et Bowman J.L. (2003). Radial patterning of Arabidopsis shoots by class III HD-ZIP and KANADI genes. *Curr Biol* 13: 1768-1774
- Esau K. (1965). Plant Anatomy. Wiley and Sons, New York
- Esau K. (1977). Anatomy Of Seed Plants. 2nd Ed, John Wiley & Sons, New York
- Eshed Y., Baum S.F., Perea J.V. et Bowman J.L. (2001). Establishment of polarity in lateral organs of plants. *Curr Biol* 11: 1251-1260
- Fäldt J., Martin D., Miller B., Rawat S. et Bohlmann J. (2003). Traumatic resin defense in Norway spruce (*Picea abies*) : methyl jasmonate-induced terpene synthase gene expression, and cDNA cloning and functionnal characterization of (+)-3–carene synthase. *Plant Mol Biol* **51** : 119-133
- Feng C., Andreasson E., Maslak A., Mock H.P., Mattson O., et Mundy J. (2004). Arabidopsis MYB68 in development and responses to environmental cues. Plant Sci 167: 1099–1107.
- Foos G., Grimm S. et Klempnauer K.H. (1994). The chicken A-myb protein is a transcriptional activator. *Oncogene* 9: 2481-2489
- Fornalé S., Sonbol F.M., Maes T., Capellades M., Puigdomènech P., Rigau J. et Caparrós-Ruiz D. (2006). Down-regulation of the maize and *Arabidopsis thaliana caffeic acid O-methyl-transferase* genes by two new maize R2R3-MYB transcription factors. *Plant Mol Biol* 62: 809-823
- Franceschi V.R., Krekling T., Berryman A.A. et Christiansen E. (1998). Specialized phloem parenchyma cells in Norway spruce (*Pinaceae*) bark are an important site of defense reactions. *Am J Bot* 85 : 601-615

- Franceschi V.R., Krokene P., Krekling T., Berryman A.A. et Christiansen E. (2000). Phloem parenchyma cells are involved in local and distant defense responses to fungal inoculation or bark-beetle attack in Norway spruce (*Pinaceae*). *Am J Bot* 87 : 314-326
- Franke R.,Hemm M.R., Denault J.W., Ruegger M.O., Humphreys J.M. et Chapple C. (2002). Changes in secondary metabolism and deposition of an unusual lignin in the ref8 mutant of Arabidopsis. *Plant J* 30: 47-59
- Fukuda H. (1996). Xylogenesis: initiation, progression, and cell death. Annu Rev Plant Physiol Plant Mol Biol 47: 299-325
- Funk V., Kositsup B., Zhao C. et Beers E.P. (2002). The Arabidopsis xylem peptidase XCP1 is a tracheary element vacuolar protein that may be a papain ortholog. *Plant Physiol* 128: 84-94
- Gabaldon C., Gomez Ros L.V., Pedreno M.A. et Ros Barcelo A. (2005). Nitric oxide production by the differentiating xylem of *Zinnia elegans*. *New Phytol* 165 : 121-130
- Galway M.E., Masucci J.D., Lloyd A.M., Walbot V., Davis R.W. et Schiefelbein J.W. (1994). The TTG gene is required to specify epidermal cell fate and cell patterning in the Arabidopsis root. *Dev Biol* 166: 740-754
- Goff S.A., Cone K.C. et Chandler V.L. (1992). Functional analysis of the transcriptional activator encoded by the maize B gene: evidence for a direct functional interaction between two classes of regulatory proteins. *Genes Dev* 6: 864–875
- Goicoechea M., Lacombe E., Legay S., Milhaevic S., Rech P., Jauneau A., Lapierre C., Pollet B., Verhaegen D., Chaubet-Gigot N. et Grima-Pettenati J. (2005). EgMYB2, a new transcriptional activator from *Eucalyptus* xylem, regulates secondary cell wall formation and lignin biosynthesis. *Plant J* 43: 553–567
- Gonda T.J., Sheiness D.K. et Bishop J.M. (1982). Transcripts from the cellular homologs of retroviral oncogenes: distribution among chicken tissues. *Mol Cell Biol* 2: 617-624
- Groover A. et Jones A.M. (1999). Tracheary element differentiation uses a novel mechanism coordinating programmed cell death and secondary cell wall synthesis. *Plant Physiol* 119 : 375-384
- Grotewold E., Sainz M.B., Tagliani L., Hermandez J.M., Bowen B. et Chandler V.L. (2000). Identification of the residues in the MYB domain of maize C1 that specify the interaction with the bHLH cofactor R. *Proc Natl Acad Sci USA* 97: 13579-13584
- Guimil S. et Dunand C. (2006). Patterning of Arabidopsis epidermal cells: epigenetic factors regulate the complex epidermal cell fate pathway. *Trends Plant Sci* 11: 601-609
- Haberer G. et Kieber J.J. (2002). Cytokinins. New insights into a classic phytohormone. *Plant Phys* 128: 354-362
- Halpin C., Holt K., Chojecki J., Oliver D., Chabbert B., Monties B., Edwards K., Barakate A. et Foxon G. A. (1998). Brown-midrib maize (bm1), a mutation affecting the cinnamyl alcohol dehydrogenase gene. *Plant J* 14: 545-53
- Halpin C., Knight M. E., Foxon G. A., Campbell M. M., Boudet A. M., Boon J. J., Chabbert B., Tollier M-T et Schuch W. (1994). Manipulation of lignin quality by downregulation of cinnamyl alcohol dehydrogenase. *Plant J* 6: 339-350

- Hammond-Kosack K.E et Jones J.D. (1996). Resistance gene-dependent plant defense responses. *Plant Cell* 8: 1773-1791
- Hartley R.D. et Jones E.C. (1977). Phenolics components and degradability of cells walls of grass and legume species. *Phytochem* 16: 1531-1534
- Hartung W., Sauter A. et Hose E. (2002). Abscisic acid in the xylem: where does it come from, where does it go to? *J Exp Bot* 53: 27-32
- Hauffe K.D., Lee S.P., Subramaniam R. et Douglas C.J. (1993). Combinatorial interactions between positive and negative *cis*-acting elements control spatial patterns of 4CL-1 expression in transgenic tobacco. *Plant J* 4: 235-253
- Hazen S.P., Schultz T.F., Pruneda-Paz J.L., Borevitz J.O., Ecker J.R.et Kay S.A. (2005). LUX ARRHYTHMO encodes a Myb domain protein essential for circadian rhythms. *Proc Natl Acad Sci USA* 102 : 10387-10392
- Henderson S.J. et Friend J. (1978). Increases in phenylalanine ammonia-lyase activity and lignin-like compounds as race-specific-resistance responses of potato tubers to Phytophthora infestans. *Biochem Soc Trans* 6: 393
- Hernandez J.M., Heine G.F., Irani N.G., Feller A., Kim M.G., Matulnik T., Chandler V.L. et Grotewold E. (2004). Differents mechanisms participate in the R-dependent activity of the R2R3 MYB transcription factor C1. J Bioch Chem 279: 48205-48213
- Higginson T., Li S.F. et Parish R.W. (2003). *AtMYB103* regulates tapetum and trichome development in *Arabidopsis thaliana*. *Plant J* 35: 177-192
- Hoffmann L., Besseau S., Geoffroy P., Ritzenthaler C., Meyer D., Lapierre C., Pollet B. et Legrand M. (2004). Silencing of hydroxycinnamoyl-coenzyme A shikimate/quinate hydroxycinnamoyltransferase affects phenylpropanoid biosynthesis. *Plant Cell* 16: 1446-1465
- Hovring I., Bostad A., Ording E., Myrset A.H. et Gabrielsen O.S. (1994). DNAbinding domain and recognition sequence of the yeast BAS1 protein, a divergent member of the Myb family of transcription factors. *J Biol Chem* 269: 17663-17672
- Hudgins J.W. et Franceschi V.R. (2004). Methyl jasmonate-induced ethylene production is responsible for conifer phloem defense response and reprogramming of stem cambial zone for traumatic resin duct formation. *Plant Physiol.* **135** : 2134-2149
- Hudgins J.W., Christiansen E. et Franceschi V.R. (2003). Methyl jasmonate induces changes mimicking anatomical defenses in diverse members of the *Pinaceae. Tree Physiol* 23 : 361-371
- Hudgins J.W., Christiansen E. et Franceschi V.R. (2004). Induction of anatomically based defense responses in stems of diverse conifers by methyl jasmonate : a phylogenetic treatment. *Tree Physiol* 24 :251-264
- Igarashi M., Demura T. et Fukuda H. (1998). Expression of the Zinnia TED3 promoter in developing tracheary elements of transgenic Arabidopsis. *Plant Mol Biol* 36: 917-927
- Introna M., Luchetti M., Castellano M., Arsura M. et Goday J. (1994). The myb oncogene family of transcription factor: potent regulators of hematopoietic cell proliferation and differenciation. *Semin Cancer Biol* 5: 113-124
- **Israelsson M., Sundberg B. et Moritz T. (2005).** Tissue-specific localization of gibberellins and expression of gibberellin-biosynthetic and signaling genes in wood-forming tissues in aspen. *Plant J* **44:** 494-504

- Ito J. et Fukuda H. (2002). ZEN1 is a key enzyme in degradation of nuclear DNA during programmed cell death of tracheary elements. *Plant Cell* 14 : 3201-3211
- Jiang C., Gu X. et Peterson T. (2004a). Identification of conserved gene structures and carboxy-terminal motifs in the Myb gene family of Arabidopsis and Oryza sativa L. ssp. indica. *Genome Biol* **5**: R46
- Jiang C., Gu J., Chopra S., Gu X. et Peterson T. (2004b). Ordered origin of the typical two- and three-repeat Myb genes. *Gene* **326**: 13-22
- Jin H.L., Cominelli E., Bailey P., Parr A., Mehrtens F., Jones J., Tonelli C., Weisshaar B. et Martin C. (2000). Transcriptional repression by AtMYB4 controls production of UV- protecting sunscreens in Arabidopsis. *EMBO J* 19: 6150–6161
- Jin H. et Martin C. (1999). Multifunctionality and diversity within the plant MYBgene family. *Plant Mol Biol* 41: 577-585
- Jones L., Ennos A.R. et Turner S.R. (2001). Cloning and characterization of irregular xylem4 (irx4): a severely lignin-deficient mutant of Arabidopsis. *Plant J* 26: 205-216
- Kajita S., Katayama Y. et Omoi S. (1996). Alteration in the biosynthesis of lignin in transgenic plants with chimeric genes for 4-coumarate: CoA ligase. *Plant Cell Physiol* **37**: 957-65
- Kanei-Ishii C., Nomura T., Ogata K., Sarai A., Yasukawa T., Tashiro S., Takahashi Y., Tanaka Y. et Ishii S. (1996). Structure and function of the protein encoded by the myb gene family. *Curr Top Microbiol Immunol* 211: 89-98
- Karpinska B., Karlsson M., Srivastava M., Stenberg A., Schrader J., Sterky F., Bhalerao R. et Wingsle G. (2004). MYB transcription factors are differentially expressed and regulated during secondary vascular development in hybrid aspen. *Plant Mol Biol* 56: 255–270
- Kasahara R.D., Portereiko M.F., Sandaklie-Nikolova I., Rabiger D.S. et Drews G.N. (2005). MYB98 is required for pollen tube guidance and synergid cell differentiation in Arabidopsis. *Plant Cell* 17: 2981-2992
- Kawaoka A. et Ebinuma H. (2001). Transcriptional control of lignin biosynthesis by tobacco LIM protein. *Phytochemistry* 57: 1149-57
- Kawaoka A., Kaothien P., Yoshida K., Endo S., Yamada K. et Ebinuma H. (2000). Functional analysis of tobacco LIM protein Ntlim1 involved in lignin biosynthesis. *Plant J* 22: 289-301.
- Kazama H., Dan H., Imaseki H. et Wastemeys G.O. (2004). Transient exposure to ethylene stimulates cell division and alters the fate and polarity of hypocotyl epidermal cells. *Plant Physiol* **134**: 1614-1623
- Keeling C.I. et Bohlmann J. (2006). Genes, enzymes and chemicals of terpenoid diversity in the constitutive and induced defence of conifers against insects and pathogens. *New Phytol* 170 : 657-675
- Keller T., Abbott J., Moritz T. et Doerner P. (2006). Arabidopsis REGULATOR OF AXILLARY MERISTEMS1 controls a leaf axil stem cell niche and modulates vegetative development. *Plant Cell* 18: 598-611
- Kirik V., Kolle K., Wohlfarth T., Misera S. et Baumlein H. (1998). Ectopic expression of a novel MYB gene modifies the architecture of the *Arabidopsis* inflorescence. *Plant J* 13:729-742
- Kirik V., Schnittger A., Radchuk V., Adler K., Hulskamp M. et Baumlein H. (2001). Ectopic expression of the Arabidopsis AtMYB23 gene induces differenciation of trichomes cells. *Dev Biol* 235: 366-377
- Kirik V., Simon M., Wester K., Schiefelbein J. et Hulskamp M. (2004). ENHANCER of TRY and CPC 2 (ETC2) reveals redundancy in the regionspecific control of trichome development of Arabidopsis. *Plant Mol Biol* 55: 389-398
- Klempnauer K.H., Gonda T. J. et Bishop J.M. (1982). Nucleotide sequence of the retroviral leukemia gene v-myb and its cellular progenitor c-myb : the architecture of a transduced oncogène. *Cell* **31**: 453-463
- Klempnauer K.H. et Sippel A.E. (1987). The highly conserved amino-terminal region of the protein encoded by the v-myb oncogene functions as a DNA-binding domain. *EMBO J* 6: 2719-2725
- Ko J.H., Beers E.P. et Han K.H. (2006). Global comparative transcriptome analysis identifies gene network regulating secondary xylem development in *Arabidopsis thaliana*. *Mol Gen Genomics* 276: 517–531
- Koslowski T.T. (1971). Growth and development of trees. Vol II. Cambial growth, root growth, and reproductive growth. Academic Press, New York
- Koyama T. et Ogura K. (1999). Isopentenyl diphosphate isomerase and prenyltransferase. In: Cane DE, ed. *Isoprenoids, including carotenoids and steroids*, Vol. 2. *Comprehensive natural products chemistry. London, UK: Elsevier*, p 69-96
- Kranz H.D., Denekamp M., Greco R., Jin H., Leyva A., Meissner R.C., Petroni K., Urzainqui A., Bevan M., Martin C., Smeekens S., Tonelli C., Paz-Ares J. et Weisshaar B. (1998). Towards functional characterization of the members of the R2R3-MYB gene family from *Arabidopsis thaliana*. *Plant J* 16: 263–276
- Kranz H., Scholz K. et Weisshaar B. (2000). c-MYB oncogene-like genes encoding three MYB repeats occur in all major plant lineages. *Plant J* 21: 231-236
- Krekling T., Franceschi V. R., Berryman A. A. et Christiansen E. (2000). The structure and development of polyphenolic parenchyma cells in Norway spruce (*Picea abies*) bark. *Flora* **195** : 354-369
- Krokene P., Solheim H., Krekling T. et Christiansen E. (2003). Inducible anatomical defense response in Norway spruce stems and their possible role in induced resistance. *Tree Physiol* 23: 191-197
- Kubo M., Udagawa M., Nishikubo N., Horiguchi G., Yamaguchi M., Ito J., Mimura T., Fukuda H. et Demura T. (2005). Transcription switches for protoxylem and metaxylem vessel formation. *Genes Dev* 19: 1855-1860
- Kuno N., Moller S.G., Shinomura T., Xu X., Chua N.H. et Furuya M. (2003). The novel MYB protein EARLYPHYTOCHROME-RESPONSIVE1 is a component of a slave circadian oscillatorin *Arabidopsis*. *Plant Cell* 15: 2476-2488
- Kuriyama H. (1999). Loss of tonoplast integrity programmed in tracheary element differentiation. *Plant Physiol* 121: 763-774
- Kusumi J., Tsumura Y., Yoshimaru H. et Tachida H. (2002). Molecular evolution of nuclear genes in cupressacea, a group of conifer trees. *Mol Biol Evol* 19: 736-747
- Lachaud S., Catesson A. M. et Bonnemain J. L. (1999). Structure and functions of the vascular cambium. *CR Acad Sci Ser III* 322 : 633-650

- Lacombe E., Van Doorsselaere J., Boerjan W., Boudet A.M. et Grima-Pettenati J. (2000). Characterization of *cis*-elements required for vascular expression of *the Cinnamoyl CoA Reductase* gene and for protein-DNA complex formation. *Plant J* 23: 663-676
- Lai L.B., Nadeau J.A, Lucas J., Lee E.K., Nakagawa T., Zhao L., Geisler M. et Sack F.D. (2005). The Arabidopsis R2R3 MYB proteins FOUR LIPS and MYB88 restrict divisions late in the stomatal cell lineage. *Plant Cell* 17: 2754-2767
- Lange B.M., Lapierre C. et Sandermann H.Jr. (1995). Elicitor-induced spruce stress lignin (structural similarity to early developmental lignins). *Plant Physiol* 108: 1277–1287
- Langenheim J.H. (2003). Plant resins : Chemistry, Evolution, Ecology, and Ethnobotany. *Timber Press, Portland, OR*.
- Lapierre C., Pollet B., MacKay J.J. et Sederoff R.R. (2000). Lignin structure in a mutant pine deficient in cinnamyl alcohol dehydrogenase. J Agric Food Chem 48: 2326-2331
- Lauvergeat V., Rech P., Jauneau A., Guez C., Coutos-Thevenot P. et Grima-Pettenati J. (2002). The vascular expression pattern directed by the *Eucalyptus gunnii* cinnamyl alcohol dehydrogenase EgCAD2 promoter is conserved among woody and herbaceous plant species. *Plant Mol Biol* 50: 497-509
- Lawrence C.J., Malmberg R.L., Muzynski M.G. et Dawe R.K. (2002). Maximum likelihood methods reveal conservation of function among closely related kinesin families. *J Mol Evol* 54: 42-53
- Lawton M.A. et Lamb C.J. (1987). Transcriptional activation of plant defense genes by fungal elicitor, wounding, and infection. *Mol Cel Biol* 7: 335-341
- Lee D., Meyer K., Chapple C. et Douglas C. J. (1997). Down-regulation of 4coumarate:CoA ligase (4CL) in *Arabidopsis*: effect on lignin composition and implications for the control of monolignol biosynthesis. *Plant Cell.* 9: 1985-98
- Lee M. M. et Schiefelbein J. (1999). WEREWOLF, a MYB-related protein in *Arabidopsis*, is a position-dependent regulator of epidermal cell patterning. *Cell* 99: 473-483
- Lee M.M. et Schiefelbein J. (2001). Developmentally distinct MYB genes encode functionally equivalent proteins in *Arabidopsis*. *Development* 128: 1539-1546
- Lee M.M. et Schiefelbein J. (2002). Cell pattern in the Arabidopsis root epidermis determined by lateral inhibition with feedback. *Plant Cell* 14: 611-619
- Legay S., Lacombe E., Goicoechea M., Brière C., Séguin A., MacKay J. et Grima-Pettenati J. (2007). Molecular characterization of EgMYB1, a putative transcriptional repressor of the lignin biosynthetic pathway. *Plant Sci* (sous presse)
- LeProvost G., Paiva J., Pot D., Brach J. et Plomion C. (2003). Seasonal variation in transcript accumulation in wood forming tissues of maritime pine (*Pinus pinaster Ait.*) with emphasis on a cell wall glycine-rich protein. *Planta* 217 : 820-830
- Leyva A, Liang X., Pintor-Toro J.A., Dixon R.A. et Lamb C.J. (1992). cis-Element combinations determine phenylalanine ammonia-lyase gene tissuespecific expression patterns. *Plant Cell* 4: 263-271

- Li L., Osakabe Y., Joshi C.P. et Chiang V.L. (1999). Secondary xylem-specific expression of caffeoyl-coenzyme A 3-O-methyltransferase plays an important role in the methylation pathway associated with lignin biosynthesis in loblolly pine. *Plant Mol Biol* 40: 555-565
- Li S.F., Santini J.M., Nicolaou O. et Parish R.W. (1996). A novel myb-related gene from Arabidopsis thaliana. *FEBS Lett* **379**: 117-121
- Liang Y.K., Dubos C., Dodd I.C., Holroyd G.H., Hetherington A.M. et Campbell M.M. (2005). AtMYB61, an R2R3-MYB transcription factor controlling stomatal aperture in *Arabidopsis thaliana*. *Curr Biol* 15: 1201-1206
- Lichtenthaler H.K. (1999). The 1-deoxy-D-xylulose-5-phosphate pathway of isoprenoid biosynthesis in plants. *Annu Rev Plant Physiol and Plant Mol Biol* 50 : 47-65
- Lipsick J.S. (1996). One billion years of Myb. Oncogene 13: 223-235
- Little C.H.A. et Savidge R.A. (1987). The role of plant growth regulators in forest tree cambial growth. *Plant Growth Regul* 6: 137–169.
- Lloyd A.M., Schena M., Walbot V. et Davis R.W. (1994). Epidermal cell fate determination in *Arabidopsis*: patterns defined by a steroid-inducible regulator. *Science* 266: 436-439
- Lombardero M.J., Ayres M.P., Lorio P.L.J. et Ruel J.J. (2000). Environmental effects on constitutive and inducible resin defences of *Pinus taeda*. *Ecol Lett* **3**: 329–339
- Luchi N., Ma R., Capretti P. et Bonello P. (2005). Systemic induction of traumatic resin ducts and resin flow in Austrian pine by wounding and inoculation with Sphaeropsis sapinea and Diplodia scrobiculata. *Planta* 221: 75-84
- Ludwig S.R., Habera L.F., Dellaporta S.L. et Wessler S.R. (1989). *Lc*, a member of the maize *R* gene family responsible for tissue-specific anthocyanin production, encodes a protein similar to transcription activators and contains the *myc*-homology region. *Proc Natl Acad Sci USA* 86: 7092–7096
- Lynch M et Conery J.S. (2000). The evolutionary fate and consequences of duplicate genes. *Science* 290 : 1151-1155
- Macheix J-J, Fleuriet A. et Jay-Allemand C. (2005). Les composés phénoliques des végétaux. Un exemple de métabolites secondaires d'importance économique. *Presses polytechniques et universitaire romandes*, Lausanne, Chap 1, *Nature et diversité des composés phénoliques des végétaux*, 1-33
- MacKay J.J., O'Malley D. M., Presnell T., Booker FL., Campbell M.M., Whetten R.W. et Sederoff R.R. (1997). Inheritance, gene expression and lignin characterization in a mutant pine deficient in cinnamyl alcohol dehydrogenase. *Proc Natl Acad Sci USA* 94: 8255-8260
- Mandaokar A., Thines B., Shin B., Lange B.M., Choi G., Koo Y.J., Choi Y.D., Choi G., et Browse J. (2006). Transcriptional regulators of stamen development in *Arabidopsis* identified by transcriptional profiling. *Plant J* 46 : 984-1008
- Martin D. M., Fäldt J. et Bohlmann J. (2004) Functional characterization of nine Norway Spruce TPS genes and evolution of gymnosperm terpene synthase of the TPS-d subfamily. *Plant Physiol* 135 : 1908-1927
- Martin D.M., Tholl D., Gershenzon J. et Bohlmann J. (2002). Methyl jasmonate induces traumatic resin ducts, terpenoid resin biosynthesis, and terpenoid

accumulation in developing xylem of Norway spruce stems. *Plant Physiol* **129** : 1003-1018

- Massari M.E. et Murre C. (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol Cell Bio* 20: 429-440
- Maston G.A., Evans S.K. et Green M. (2006). Transcriptional regulatory elements in the Human genome. *Annu Rev Genomics Hum Genet* 7: 29-59
- McKay S.A.B., Hunter W.L., Godard K-A, Wang S.W., Martin D.M., Bohlmann J. et Plant A.L. (2003). Insect attack and wounding induce traumatic resin duct development and gene expression of (–) -pinene synthase in Sitka spruce. *Plant Physiol* 133 : 368-378
- Mehrtens F., Kranz H., Bednarek P. et Weisshaar B. (2005). The Arabidopsis transcription factor MYB12 is a flavonol-specific regulator of phenylpropanoid biosynthesis. *Plant Physiol* **138** : 1083-1096
- Memelink J. (2005). The use of genetics to dissect plant secondary pathways. *Curr* Opin Plant Biol 8: 230-235
- Mengiste T., Chen X., Salmeron J. et Dietrich R. (2003). The BOTRYTIS SUSCEPTIBLE1 gene encodes an R2R3MYB transcription factor protein that is required for biotic and abiotic stress responses in Arabidopsis. *Plant Cell* 15: 2551-2565
- Meyer K., Shirley A.M., Cusumano J.C., Bell-Lelong D.A. et Chapple C. (1998). Lignin monomer composition is determined by the expression of a cytochrome P450-dependent monooxygenase in *Arabidopsis*. *Proc Natl Acad Sci USA* **95**: 6619-6623
- Millar A.A. et Gubler F. (2005). The Arabidopsis GAMYB-like genes, MYB33 and MYB65, are microRNA-regulated genes that redundantly facilitate anther development. *Plant Cell* 17: 705-721
- Milioni D., Sado P-E, Stacey N.J., Domingo C., Roberts K. et McCann M. C. (2001). Differential expression of cell-wall-related genes during the formation of tracheary elements in the *Zinnia* mesophyll cell system. *Plant Mol Biol* 47 : 221-238
- Mitsuda N., Seki M., Shinozaki K. et Ohme-Takagi M. (2005). The NAC transcription factors NST1 and NST2 of *Arabidopsis* regulates secondary wall thickening and are required for anther dehiscence. *Plant Cell* 17 : 2993-3006
- Mitsuda N., Iwase A., Yamamoto H., Yoshida M., Seki M., Shinozaki K. et Ohme-Takagi M. (2007). NAC transcription factors, NST1 and NST3, are key regulators of the formation of secondary walls in woody tissues of *Arabidopsis*. *Plant Cell* (sous presse)
- Morrow B.E., Ju Q. et Warner J.R. (1993). A bipartite DNA-binding domain in yeast Reb1p. *Mol Cell Biol* 13: 1173-1182
- Moyle R., Schrader J., Stenberg A., Olsson O., Saxena S., Sandberg G. et Bhalerao R.P. (2002). Environmental and auxin regulation of wood formation involves members of the Aux/IAA gene family in hybrid aspen. *Plant J* 31: 675-685
- Müller D., Schmitz G. et Theres K. (2006). Blind homologous R2R3 Myb genes control the pattern of lateral meristem initiation in Arabidopsis. Plant Cell 18: 586-597
- Murray F., Kalla R., Jacobsen J. et Gubler F. (2003). A role for HvGAMYB in anther development. *Plant J* 33: 481-491

- Murre C., Bain G., van Dijk M.A., Engle I., Furnari B.A., Massari M.E., Matthews J.R., Quong M.W., Rivera R.R. et Stuiver M.H. (1994). Structure and function of helix-loop-helix proteins. *Biochim Biophys Acta* 1218: 129–135
- Murre C., McCaw P.S. et Baltimore D. (1989). A new DNA binding and dimerization motif in immunoglobulin enhancer binding, daughterless, MyoD, and myc proteins. *Cell* 56: 777-783
- Nagy N.E., Franceschi V.R., Solheim H., Krekling T. et Christiansen E. (2000). Wound-induced traumatic resin duct development in stems of Norway spruce (*Pinaceae*) : anatomy and cytochemical traits. *Am J Bot* 87 : 303-313
- Nakaba S., Sano Y., Kubo T. et Funada R. (2006). The positional distribution of cell death of ray parenchyma in a conifer, *Abies sachalinensis*. *Plant Cell Rep* 25: 1143-1148
- Nei M. (1969). Gene duplication and nucleotide substitution in evolution. *Nature* 221 : 5175-5177
- Nelson M. et Dengler N. (1997). Leaf vascular pattern formation. *Plant Cell* 9: 1121-1135
- Nesi N., Jond C., Debeaujon I., Caboche M. et Lepiniec L. (2001). The *Arabidopsis TT2* gene encodes an R2R3-MYB domain protein that acts as akey determinant for proanthocyanidin accumulation in developing seed. *Plant Cell* **13** : 2099-2114
- Newman L.J., Perazza D.E., Juda L. et Campbell M.M. (2004). Involvement of the R2R3-MYB, At MYB61, in the ectopic lignication and dark-photomorphogenic components of the det3 mutant phenotphype. *Plant J* 37: 239–250
- Nikolov D.B. et Burley S.K. (1997). RNA polymerase II transcription initiation: A structural view. *Proc Natl Acad Sci USA* 94: 15-22
- Nishitani K. (1997). The role of endoxyloglucan transferase in the organization of plant cell wall. *Int Rev Cytol* 173 : 157-206
- Noda K., Glover B.J., Linstead P. et Martin C. (1994). Flower colour intensity depends on specialized cell shape controlled by a Myb-related transcription factor. *Nature* **369**: 661-664
- Ogata K., Hojo H., Aimoto S., Nakai T., Nakamura H., Sarai A., Ishii S. et Nishimura Y. (1992). Solution Structure of a DNA-Binding Unit of Myb: A Helix-Turn-Helix-Related Motif With Conserved Tryptophans Forming a Hydrophobic Core. *Proc Natl Acad Sci USA* **89**: 6428-6432
- Ogata K., Morikawa S., Nakamura H., Sekikawa A., Inoue T., Kanai H., Sarai A., Ishii S. et Nishimura Y. (1994). Solution structure of a specific DNA complex of the Myb DNA-binding domain with cooperative recognition helices. *Cell* **79** : 639-648
- Ohno S. (1970). Evolution by gene duplication. Springer, New York
- **Okamoto-Nakazato A. (2002).** A brief note on the study of yieldin, a wall-bound protein that regulates the yield threshold of the cell wall. *J Plant Res* **115** : 309-313
- **Oparka K.J. et Cruz S.S. (2000).** THE GREAT ESCAPE: Phloem Transport and Unloading of Macromolecules1. *Annu Rev Plant Physiol Plant Mol Biol* **51**: 323-347
- Oppenheimer D.G., Herman P.L., Sivakumaran S., Esch J. et Marks M.D. (1991). A myb gene required for leaf trichome differentiation in *Arabidopsis* is expressed in stipules. *Cell* 67: 483-493

- Ording E., Bergholtz S., Brendeford E.M., Jamin N. et Gabrielsen O.S. (1996). Flexibility in the second half-site sequence recognised by the c-Myb R2 domain in vitro and in vivo analysis. *Oncogene* 13: 1043-1051
- Palaniswamy S.K., James S., Sun H., Lamb R.S., Davuluri R.V. et Grotewold E. (2006). AGRIS and AtRegNet. a platform to link cis-regulatory elements and transcription factors into regulatory networks. *Plant Physiol* 140: 818-829
- Patikoglou G. et Burley S.K. (1997). Eukaryotic transcription factor-DNA complexes. *Annu Rev Biophys Biomol Struct* 26 : 289-325
- Patzlaff A., Newman L.J., Dubos C., Whetten R.W., Smith C., McInnis S., Bevan M.W., Sederoff R. R. et Campbell M.M. (2003a). Characterisation of PtMYB1, an R2R3-MYB from pine xylem. *Plant Mol Biol* 53: 597–608
- Patzlaff A., McInnis S., Courtenay A., Surman C., Newman L.J., Smith C., Bevan M.W., Mansfield S. D., Whetten R.W., Sederoff R.R. et Campbell M.M. (2003b). Characterisation of a pine MYB that regulates lignification. *Plant J* 36: 743–754
- Paz-Ares J., Ghosal D., Wienand U., Peterson P.A. et Saedler H. (1987). The regulatory c1 locus of Zea mays encodes a protein with homology to myb proto-oncogene products and with structural similarities to transcriptional activators. *Embo J* 6: 3553-3558
- Penfield S., Meissner R.C., Shoue D.A., Carpita N.C. et Bevan M.W. (2001). MYB61 is required for mucilage deposition and extrusion in the Arabidopsis seed coat. *Plant Cell* 13: 2777-2791
- Peters C.W., Sippel A.E., Vingron M. et Klempnauer K.H. (1987). Drosophila and vertebrate myb proteins share two conserved regions, one of which functions as a DNA-binding domain. *EMBO J* 6: 3085-3090
- Pilate G., Guiney E., Holt K., Petit-Conil M., Lapierre C., Leple J. C., Pollet B., Mila I., Webster E. A., Marstorp H. G., Hopkins D. W., Jouanin L., Boerjan W., Schuch W., Cornu D. et Halpin C. (2002). Field and pulping performances of transgenic trees with altered lignification. *Nature Biotech* 20: 607-12
- Piquemal J., Lapierre C., Myton K., O'Connell A., Schuch W., Grima-Pettenati J. et Boudet A. M. (1998). Down regulation of cinnamoyl CoA reductase induces significant changes of lignin profiles in transgenic tobacco plants. *Plant J* 13: 71-83
- Plomion C., Leprovost G. et Stokes A. (2001). Wood formation in trees. *Plant Physiol* 127: 1513-1523
- Preston J., Wheeler J., Heazlewood J., Li S.F. et Parish R.W. (2004). AtMYB32 is required for normal pollen development in *Arabidopsis thaliana*. *Plant J* 40: 979-995
- Prince V.E. et Pickett F.B. (2002). Splitting pairs: the diverging fates of duplicated genes. *Nat Rev Genet* **3**: 827-837
- Ptashne M. (1988). How eukaryotic transcriptional activators work. *Nature* 335: 683-689
- Quatrochio F., Wing J.F., van der Woude K., Mol J.N.M. et Koes R. (1998). Analysis of bHLH and MYB domain proteins: species-specific regulatory differences are caused by divergent evolution of target anthocyanin genes. *Plant J* 13: 475-488
- Quatrochio F., Wing J.F., van der Woude K., Souer E., de Vetten N., Mol J.N.M. et Koes R. (1999). Molecular analysis of the *anthocyanin2* gene of

petunia and its role in the evolution of flower color. *Plant Cell* 11: 1433-1444

- Rabinowicz P.D., Braun E.L., Wolfe A.D., Bowen B. et Grotewold E. (1999). Maize R2R3 Myb Genes: Sequence Analysis Reveals Amplification in the Higher Plants. *Genetics* 153: 427-444
- Raes J., Rohde A., Christensen J.H., Van de Peer Y. et Boerjan W. (2003). Genome-wide characterization of the lignification toolbox in Arabidopsis. *Plant Physiol* 133: 1051-1071
- Raffaele S., Rivas S. et Roby D. (2006). An essential role for salicylic acid in AtMYB30-mediated control of the hypersensitive cell death program in Arabidopsis. *FEBS Lett* 580: 3498-3504
- Ralph J., MacKay J.J., Hatfield R.D., O'Malley D.M., Whetten R.W. et Sederoff R.R. (1997). Abnormal lignin in a loblolly pine mutant. *Science* 277: 235-239
- Ramsay N.A. et Glover B.J. (2005). MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. *Trends Plant Sci* 10: 63-70
- Reina J.J., Dominguez E. et Heredia A. (2001). Water sorption-desorption in conifer cuticles: The role of lignin. *Physiol Plant* 112: 372-378
- **Reyes J.L. et Chua N.H. (2007).** ABA induction of miR159 controls transcript levels of two MYB factors during Arabidopsis seed germination. *Plant J* **49**: 592-606
- **Riaño-Pachón D.M., Ruzicic S., Dreyer I. etMueller-Roeber B. (2007).** PlnTFDB: an integrative plant transcription factor database. *BMC Bioinformatics* 8: 42
- Riechmann J.L., Heard J., Martin G., Reuber L., Jiang C., Keddie J., Adam L., Pineda O., Ratcliffe O.J., Samaha R.R., Creelman R., PilgrimM., Broun P., Zhang J.Z.,Ghandehari D., Sherman B.K. et Yu G. (2000). *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. *Science* 290: 2105-2110
- Riechmann J.L. et Ratcliffe O.J. (2000). A genomic perspective on plant transcription factors. *Curr Opin Plant Biol* **3**: 423-434
- Rogers L.A. et Campbell M.M. (2004). The genetic control of lignin deposition during plant growth and development. *New Phytol* 164: 17–30
- Romero I., Fuertes A., Benito M.J., Malpica J.M., Leyva A. et Paz-Ares J. (1998). More than 80 R2R3 MYB regulatory genes in the genome of *Arabidopsis thaliana*. *Plant J* 14: 273-284
- Rosinski J.A. et Atchley W.R. (1998). Molecular evolution of the Myb family of transcription factors: evidence for polyphyletic origin. *J Mol Evol* 46: 74-83
- Rubio V., Linhares F., Solano R., Martín A.C., Iglesias J., Leyva A. et Paz-Ares J. (2001). A conserved MYB transcription factor involved in phosphate starvation signaling both in vascular plants and in unicellular algae. *Genes Dev* 15: 2122-2133
- Ruiz-Medrano R., Xoconostle-Cazares B. et Lucas W. J. (2001). The phloem as a conduit for inter-organ communication. *Curr Opin Plant Biol* 4: 202-209
- Sablowski R.W., Moyano E., Culianez-Macia F.A., Schuch W., Martin C. et Bevan M. (1994). A flower-specific Myb protein activates transcription of phenylpropanoid biosynthetic genes. *Embo J* 13: 128-137.
- Sachs T. (1991). Cell polarity and tissue patterning in plants. *Development* S1: 83-93

- Samuels A.L., Rensing K.H., Douglas C.J., Mansfield S.D., Dharmawardhana D.P. et Ellis B.E. (2002). Cellular machinery of wood production: differentiation of secondary xylem in *Pinus contorta var. latifolia. Planta* 216: 72-82
- Sankoff D. (2001). Gene and genome duplication. Curr Opin Genet Dev 11: 681-684
- Sarkanen K.V. et Hergert H.L. (1971). Classification and distribution. In: Lignins. Occurrence, Formation, Structure and Reactions. Eds. K.V. Sarkanen, C.H. Ludwig. Wiley Interscience, New York. p 43–94
- Savidge R.A. (1983). The role of plant hormones in higher plant cellular differentiation. II. Experiments with the vascular cambium, and sclereid and tracheid differentiation in the pine, *Pinus contorta*. *Histochem J* 15: 447-466
- Savidge R.A. (2001). Intrinsic regulation of cambial growth. *J. Plant Growth Regul* 20: 52-77
- Savidge R.A. et Wareing P.F. (1981). Plant growth regulators and the differentiation of vascular elements. In: Barnett JR, editor. *Xylem cell development. Tunbridge Wells, England: Castle House Publ.* p 192–235
- Saville M.K et Watson R.J. (1998). The cell-cycle regulated transcription factor B-Myb is phosphorylated by cyclin A/Cdk2 at sites that enhance its transactivation properties. Oncogene 17: 2679-2689
- Sawa S. (2002). Overexpression of the *AtmybL2* gene represses trichome development in *Arabidopsis*. *DNA Res* 9: 31-34.
- Scarpella E. et Meijer A. H. (2004). Pattern formation in the vascular system of monocot and dicot plant species. *New Phytol* 164: 209-242.
- Schaffer R., Ramsay N., Samach A., Corden S., Putterill J., Carre I.A. et Coupland G. (1998). The *late elongated hypocotyl* mutation of *Arabidopsis* disrupts circadian rhythms and the photoperiodic control of flowering. *Cell* 93: 1219-1229
- Schellmann S., Schnittger A., Kirik V., Wada T., Okada K., Beermann A., Thumfahrt J., Jurgens G. et Hulskamp M. (2002). TRIPTYCHON and CAPRICE mediate lateral inhibition during trichome and root hair patterning in *Arabidopsis*. *Embo J* 21: 5036-5046
- Schrader J., Baba K., May S.T., Palme K., Bennett M., Bhalerao R.P. et Sandberg G. (2003). Polar auxin transport in the wood-forming tissues of hybrid aspen is under simultaneous control of developmental and environmental signals. *Proc Natl Acad Sci USA* 100: 10096-10101
- Séguin A., Laible G., Leyva A., Dixon R.A. et Lamb C.J. (1997). Characterization of a gene encoding a DNA-binding domain that interact in vitro with vascular specific *cis* element of the *phenylalanine ammonia-lyase* promoter. *Plant Mol Biol* **35**: 281-291
- Seo H.S., Yang J.Y., Ishikawa M., Bolle C., Ballesteros M.L. et Chua N.H. (2003). LAF1 ubiquitination by COP1 controls photomorphogenesis and is stimulated by SPA1. *Nature* **423**: 995-999
- Serna L. et Martin C. (2006) Trichomes: different regulatory networks lead to convergent structures. *Trends Plant Sci* 11: 274-280
- Sewalt V., Ni W., Blount J.W., Jung H.G., Massoud S.A., Howles P.A., Lamb C. et Dixon RA. (1997). 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

- Sharman BC. (1943). Tannic acid and Iron alum with safranin and Orange G in studies of the shoot apex. *Stain Technol* 18: 105–111
- Shin B., Choi G., Yi H., Yang S., Cho I., Kim J., Lee S., Paek N.C., KimJ.H. et Song P.S. (2002). AtMYB21, agene encoding a flower-specific transcription factor, is regulated by COP1. Plant J 30: 23-32
- Schmid G. et Grisebach H. (1982). Enzymatic synthesis of lignin precursors. Purification and properties of UDP glucose: coniferyl-alcohol glucosyltransferase from cambial sap of spruce (*Picea abies L.*). *Eur J Biochem* 123: 363-370
- Smith T.F., Gaitatzes C., Saxena K. et Neer E.J. (1999). The WD repeat: a common architecture for diverse functions. *Trends Biochem Sci* 24: 181-185
- Somerville C. (2006). Cellulose synthesis in higher plants. *Annu Rev Cell Dev Biol* 22: 53-78
- Sompornpailin K., Makita Y., Yamazaki M. et Saito K. (2002). AWD-repeatcontaining putative regulatory protein in anthocyanin biosynthesis in *Perilla frutescens*. *Plant Mol Biol* **50**: 485-495
- Spelt C., Quattrochio F., Mol J.N. et Koes R. (2000). Anthocyanin 1 of petunia encodes a basic helix-loop-helix protein that directly activates transcription of structural anthocyanin genes. *Plant Cell* **12**: 1619-1632
- Srivastava L.M. (2002). Plant Growth and Development. *Academic Press, London,* p 329–339
- Steiner-Lange S., Unte U.S., Eckstein L., Yang C., Wilson Z.A., Schmelzer E., Dekker K. et Saedler H. (2003). Disruption of Arabidopsis thaliana MYB26 results in male sterility due tonon-dehiscent anthers. Plant J 34: 519-528
- Stracke R., Werber M. et Weisshaar B. (2001). The R2R3-MYB gene family in *Arabidopsis thaliana. Curr Opin Plant Biol* 4: 447–456
- Sundberg B., Uggla C. et Tuominen H. (2000). Cambial growth and auxin gradients. Savidge, B R., J., R Napier, eds. *Cell and Molecular Biology of Wood Formation. BIOS Scientific Publishers Ltd, Oxford* : 169-188
- Tamagnone L., Merida A., Parr A., Mackay S., Culianez-Macia F.A., Roberts K. et Martin C. (1998). The AmMYB308 and AmMYB330 transcription factors from *Antirrhinum* regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. *Plant Cell* 10: 135–154
- Thomas T.H., Wareing P.F. et Robinson P.M. (1965). Action of the sycamore "dormin" as a gibberellin antagonist. *Nature* 205: 1270-1272
- Timell T. E. (1969). The chemical composition of tension wood. *Swen Papperstidn* 72: 173-181
- Timell T. E. (1986). Compression wood in gymnosperms. Springer, Berlin Heidelberg New York
- Toledo-Ortiz G., Huq E. et Quail P.H. (2003). The *Arabidopsis* basic/helix-loophelix transcription factor family. *Plant Cell* 15: 1749-1770
- Trapp S. et Croteau R. (2001). Defensive resin biosynthesis in conifers. *Annu Rev Plant Physiol Plant Mol Biol* 52 : 689-724
- Turner S., Gallois P. et Brown D. (2007). Tracheary element differentiation. *Annu Rev Plant Biol* 58 : 407-433
- Turner S.R et Hall M. (2000). The gapped xylem mutant identifies a common regulatory step in secondary cell wall deposition. *Plant J* 24 : 477-488

- Tuskan G.A. et al. (109 co-auteurs). (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313: 1596-1604
- Uggla C., Mellerowicz E. J. et Sundberg B. (1998). Indole-3 acetic acid controls cambial growth in Scots pine by positional signaling. *Plant Physiol* 117: 113-121
- Uggla C., Moritz T., Sandberg G. et Sundberg B. (1996). Auxin as a positional signal in pattern formation in plants. *Proc Natl Acad Sci USA* 93: 9282-9286
- Vailleau F., Daniel X., Tronchet M., Montillet J.L., Triantaphylides C. et Roby D. (2002). A R2R3-MYB gene, AtMYB30, acts as a positive regulator of the hypersensitive cell death program in plants in response to pathogen attack. *Proc Natl Acad Sci USA* 99: 10179-10184
- van Bel A.J., Ehlers K. et Knoblauch M. (2002). Sieve elements caught in the act. *Trends Plant Sci* 7: 126-32
- Vance C.P., Kirk T.K. et Sherwood R.T. (1980). Lignification as a mechanism of disease resistance. *Annu Rev Phytopathol* 15 : 259-288
- Van Doorsselaere J., Baucher M., Chognot E., Chabbert B. et Tollier M-T.(1995). A novel lignin in poplar trees with reduced caffeic acid/5hydroxyferulic acid O-methyltransferase activity. *Plant J.* 8: 855-64
- Vignols F., Rigau J., Torres M.A., Capellades M. et Puigdomenech P. (1995). The brown midrib (bm3) mutation in maize occurs in the gene encoding caffeic acid *O*-methyltransferase. *Plant Cell* **7**: 407-416.
- Voronova A. et Baltimore D. (1990). Mutations that disrupt DNA binding and dimmer formation in the E47 helix-loop-helix protein map to distinct domains. *Proc Natl Acad Sci USA* 87 : 4722-4726
- Wada T., Tachibana T., Shimura Y. et Okada K. (1997). Epidermal cell differentiation in *Arabidopsis* determined by a Myb homolog, CPC. *Science* 277:1113-1116
- Wagner A. (2001). Birth and death of duplicated genes in completely sequenced eukaryotes. *Trends Genet* 17: 237-239
- Walker A.R., Davison P.A., Bolognesi-Winfield A.C., James C.M., Srinivasan, N. Blundell T.L., Esch J.J., Marks M.D. et Gray J.C. (1999). The *TRANSPARENT TESTAGLABRA1* locus, which regulates trichome differentiation and anthocyanin biosynthesis in *Arabidopsis*, encodes aWD40 repeat protein. *Plant Cell* 11:1337-1350
- Wang D.M., Dubendorff J.W., Woo C.H. et Lipsick J.S. (1999) Functional analysis of carboxy-terminal deletion mutants of c-Myb. *J Virol* 73: 5875-5886
- Wang S., Wang J.W., Yu N., Li C.H., Luo B., Gou J.Y., Wang L.J. et Chen X.Y. (2004). Control of plant trichome development by a cotton fiber *MYB* gene. *Plant Cell* 16 : 2323-2334
- Wang Z.Y., Kenigsbuch D., Sun L., Harel E., Ong M.S. et Tobin E.M. (1997). A Myb-related transcription factor is involved in the phytochrome regulation of an *Arabidopsis Lhcb* gene. *Plant Cell* **9**: 491-507
- Wareing P.F. (1965). Dormancy in plants. Science Progress 53: 529-537
- Watson R.J., Robinson C. et Lam E.W. (1993). Transcription regulation by murine B-myb is distinct from that by c-myb. *Nucleic Acid Res* 21: 267-272
- Werner T. (1999). Models for prediction and recognition of eukaryotic promoters. *Mamm Genome* 10 : 168-175

- Weston K. (1998). MYB proteins in life, death and differentiation. *Curr Opin Genet Dev* 8 : 76-81
- Weston K. et Bishop J.M. (1989). Transcriptional activation by the v-myb oncogene and its cellular progenitor, c-myb. *Cell* 58: 85-93
- Wise M.L. et Croteau R. (1999). Monoterpene biosynthesis. In: Cane DE, ed. Isoprenoids, including carotenoids and steroids, Vol. 2. Comprehensive natural products chemistry. London, UK: Elsevier, p 97-153
- Xue B., Charest P.J., Devantier Y. et Rutledge R.G. (2003). Characterization of a *MYBR2R3* gene from black spruce (*Picea mariana*) that shares functional conservation with maize C1. *Mol Genet Genomics* 270: 78-86
- Yang M. et Sack F.D. (1995). The *too many mouths* and *four lips* mutations affect stomatal production in Arabidopsis. *Plant Cell* 7: 2227–2239
- Yang X.Y., Li J.G., Pei M., Gu H., Chen Z.L et Qu L-J. (2007). Over-expression of a flower-specific transcription factor gene AtMYB24 causes aberrant anther development. *Plant Cell Rep* 26: 219-228
- Yanhui C., Xiaoyuan Y., Kun H., Meihua L., Jigang L., Zhaofeng G., Zhiqiang L., Yunfei Z., Xiaoxiao W., Xiaoming Q., Yunping S., Li Z., Xiaohui D., Jingchu L., Xing-Wang D., Zhangliang C., Hongya G. et Li-Jia Q. (2006). The MYB transcription factor superfamily of Arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family. *Plant Mol Biol* 60: 107-123
- Ye Z-H. (2002). Vascular tissue differentiation and pattern formation in plants. Annu Rev Plant Biol 53: 183-2002
- Ye Z.H et Varner J.E. (1996). Induction of cysteine and serine proteases during xylogenesis in *Zinnia elegans*. *Plant Mol Biol* **30** : 1233-1246
- Yoo J.H., Park C.Y., Kim J.C., Do Heo W., Cheong M.S., Park H.C., Kim M.C., Moon B.C., Choi M.S., Kang Y.H., Lee J.H., Kim H.S., Lee S.M., Yoon H.W., Lim C.O., Yun D-J., Lee S.Y., Chung W.S. et Cho M.J. (2005). Direct Interaction of a Divergent CaM Isoform and the Transcription Factor, MYB2, Enhances Salt Tolerance in Arabidopsis. J Biol Chem 280: 3697-3706
- Zhang F., Gonzalez A., Zhao M., Payne C.T. et Lloyd A. (2003). A network of redundant bHLH proteins functions in all TTG1-dependent pathways of *Arabidopsis*. *Development* 130 : 4859-4869
- Zhang J., Rosenberg H.F. et Nei M. (1998). Positive Darwinian selection after gene duplication in primate ribonuclease genes. *Proc Natl Acad Sci USA* 95 : 3708-3713
- Zhong R., Demura T. et Ye Z-H. (2006). SND1, a NAC domain transcription factor, is a key regulator of secondary wall synthesis in fibers of *Arabidopsis*. *Plant Cell* 18 : 3158-3170
- Zhong R., Taylor J.J. et Ye Z.H. (1997). Disruption of interfascicular fiber differentiation in an Arabidopsis mutant. *Plant Cell* 9: 2159-2170
- Zhu J., Verslues P.E., Zheng X., Lee B.H., Zhan X., Manabe Y., Sokolchik I., Zhu Y., Dong C.H., Zhu J.K., Hasegawa P.M. et Bressan R.A. (2005). HOS10 encodes an R2R3-type MYB transcription factor essential for cold acclimation in plants. *Proc Natl Acad Sci USA* 102: 9966-9971
- Zimmermann I.M., Heim M.A., Weisshaar B. et Uhrig J.F. (2004). Comprehensive identification of Arabidopsis thaliana MYB transcription factors interacting with R/B-like BHLH proteins. *Plant J* 40: 22-34

CHAPITRE II

Conifer R2R3-MYB transcription factors: sequence analyses and gene expression in wood-forming tissues of white spruce (*Picea glauca*)

Frank Bedon^{1, 2}, Jacqueline Grima-Pettenati² and John Mackay¹

¹Centre d'étude de la Forêt, Université Laval, Pavillon Charles-Eugène Marchand, Sainte Foy G1K7P4, Québec, Canada ²UMR CNRS/UPS 5546 Surfaces Cellulaires et Signalisation chez les Végétaux, Pôle de Biotechnologie Végétale, BP426 17 - Auzeville 31226, Castanet Tolosan, France

II-1-Avant-propos

Les résultats de ce chapitre ont été publiés le 30 mars 2007 dans la revue "*BMC Plant Biology*", volume 7, numéros 17. Mon rôle a été de participer à l'élaboration de cette étude, faire l'ensemble du travail expérimental (i.e. isoler et cloner les *MYB* d'épinette blanche, effectuer l'étude par PCR quantitative, réaliser l'expérience de bois de compression), analyser les données et produire toutes les figures, rédiger le manuscrit. Les séquences des MYB de Pin m'ont été fournies par Dr. J. MacKay. Le protocole d'extraction des ARN adapté de Chang et *al.* (1993) présenté en annexe II-1 a été mis au point au laboratoire avec l'aide de Dr. J. Cooke et Francis Boileau. Mes co-directeurs Dr. J. Grima-Pettenati et Dr. J. MacKay ont participé à l'élaboration de ce projet et à la préparation du manuscrit.

II-2-Résumé

Plusieurs membres de la famille des facteurs de transcription de type MYB-R2R3 agissent comme des régulateurs du métabolisme des phénylpropanoïdes et des lignines pendant la formation du bois chez les angiospermes et les gymnospermes. La plante angiosperme Arabidopsis possède plus d'une centaine de gènes MYB-R2R3; cependant, seulement quelques membres de cette famille ont été découverts chez les gymnospermes. Nous avons isolé et caractérisé des gènes MYB-R2R3 codant pour des ADNc pleines longueurs chez les gymnospermes suivant: l'épinette blanche, Picea glauca (13 séquences), et le pin loblolly, Pinus taeda L. (5 séquences). Les similarités de séquence et les analyses phylogénétiques positionnent les séquences d'épinette et de pin dans différents sous-groupes de la grande famille des MYB-R2R3, bien que plusieurs de ces séquences se rapprochent ensemble. Nous avons recherché des séquences d'acides aminés conservés dans les régions C-terminales hautement variable des MYBs de plantes et identifié 20 motifs dans les MYBs d'épinette, parmi lesquels neuf n'ont jamais été rapporté et trois sont spécifique des conifères. Le nombre et la longueur des introns des gènes MYB d'épinette varient significativement mais leurs positions sont bien conservées comparativement aux gènes MYB d'angiospermes. L'analyse des niveaux de transcrits des gènes MYB par PCR quantitative dans les tissus des racines et des tiges ont révélé différent profils d'expression; trois gènes MYB sont préférentiellement exprimés dans le xylème secondaire alors que d'autres sont préférentiellement exprimés dans le phloème ou sont ubiquitaires. Les gènes MYB exprimés dans le xylème, et trois autre, ont leurs niveaux d'expression augmentés dans le bois de compression des arbres penchés au sein des 76 heures d'induction. Notre étude portant sur 18 gènes MYB-R2R3 de conifère montre clairement que cette famille possède une structure génique similaire à celle d'Arabidopsis. Trois de ces séquences semblent avoir un rôle dans le métabolisme des lignines et/ou la formation du bois chez les arbres de type gymnospermes, incluant un homologue très proche de PtMYB4 chez le pin loblolly, impliqué dans la régulation de la lignine dans des tabacs transgéniques.

II-3-Abstract

Several members of the R2R3-MYB family of transcription factors act as regulators of lignin and phenylpropanoid metabolism during wood formation in angiosperm and gymnosperm plants. The angiosperm Arabidopsis has over one hundred R2R3-MYBs genes; however, only a few members of this family have been discovered in gymnosperms. We isolated and characterised full-length cDNAs encoding R2R3-MYB genes from the gymnosperms white spruce, Picea glauca (13 sequences), and loblolly pine, Pinus taeda L. (five sequences). Sequence similarities and phylogenetic analyses placed the spruce and pine sequences in diverse subgroups of the large R2R3-MYB family, although several of the sequences clustered closely together. We searched the highly variable Cterminal region of diverse plant MYBs for conserved amino acid sequences and identified 20 motifs in the spruce MYBs, nine of which have not previously been reported and three of which are specific to conifers. The number and length of the introns in spruce MYB genes varied significantly, but their positions were well conserved relative to angiosperm MYB genes. Quantitative RTPCR of MYB genes transcript abundance in root and stem tissues revealed diverse expression patterns; three MYB genes were preferentially expressed in secondary xylem, whereas others were preferentially expressed in phloem or were ubiquitous. The MYB genes expressed in xylem, and three others, were up-regulated in the compression wood of leaning trees within 76 hours of induction. Our survey of 18 conifer R2R3-MYB genes clearly showed a gene family structure similar to that of Arabidopsis. Three of the sequences are likely to play a role in lignin metabolism and/or wood formation in gymnosperm trees, including a close homolog of the loblolly pine PtMYB4, shown to regulate lignin biosynthesis in transgenic tobacco.

II-4-Introduction

Insights into the regulation of lignin biosynthesis during vascular development of plants are being derived from angiosperm model plants like Arabidopsis thaliana (reviewed by Rogers and Campbell, 2004) and from investigations unravelling the molecular basis of wood formation in trees like Populus (reviewed by Chaffey, 2002). Members of the R2R3-MYB transcription factor family have been implicated as regulators of phenylpropanoid and lignin metabolism (Rogers and Campbell, 2004) as well as pattern formation and differentiation of primary and secondary vascular tissues, (reviewed by Scarpella and Meijer, 2004). The MYB proteins comprise one of the largest families of plant transcription factors, which is represented by over one hundred members in the model plant Arabidopsis (Riechmann et al., 2000). The biological roles of MYBs have been deduced primarily from flowering plants (angiosperms) including snapdragon (Tamagnone et al., 1998), maize (Grotewold et al., 1994), Arabidopsis (Jin et al., 2000; Newman et al., 2004) and eucalyptus (Goicoechea et al., 2005). By contrast, the biological roles of only a few R2R3-MYBs has been examined in non-flowering plants (gymnosperms) and relatively little is known about their gene family structure (Patzlaff et al., 2003a, 2003b; Xue et al., 2003). The loblolly pine genes PtMYB1 and PtMYB4 were shown to be transcriptional activators which have the ability to regulate lignin synthesis enzymes (Patzlaff et al., 2003a, 2003b). They are expressed in xylem tissues, bind AC elements and activate transcription in transient assays in yeast or plant cells (Patzlaff et al., 2003a, 2003b; Gomez-Maldonaldo et al., 2004). Overexpression of pine MYBs resulted in ectopic lignification in tobacco (Patzlaff et al., 2003a) and in Arabidopsis (Newman et al., 2004) The reports are strong evidence supporting a role for MYBs in the lignifying process in gymnosperm trees. Lignins play an important role in trees because they confer rigidity and impermeability to wood by accumulating in thickened secondary vascular tissues (Rogers and Campbell, 2004; Boerjan et al., 2003), therefore its regulation is of interest for understanding the genetic basis of wood properties. However, the number of MYB transcription factors that may participate in regulating lignification in gymnosperms, and their potential roles remain an open question.

Gymnosperms, especially conifers of the *Pinaceae* family are ecologically and economically important due to their abundance in forests in many parts of the world (North America, Europe, Asia) and because of their use in diverse wood products (pulp and paper, solid wood and engineered lumber). Despite recent largescale gene discovery initiatives for conifer trees like pine and spruce (e.g. Kirst et al., 2003; Pavy et al., 2005), only a few regulatory gene families have been characterised systematically in any conifer species. In one such study, it was recently shown that the structure of the *knox-I* gene family appears to be monophyletic in the *Pinaceae*, whereas angiosperms have several distinct clades (four in dicots and three in monocots) (Guillet-Claude et al., 2004). The R2R3-MYB family is very large with over 120 members in angiosperms (Riechmann et al., 2000) and has been divided into several subgroups (Kranz et al., 1998; Romero et al., 1998). One may predict that several MYBs are likely to regulate lignin metabolism and other aspects of wood formation in conifer trees; however no data have been available from which to infer the size or the structure of the family in gymnosperms. Therefore, a broader survey of MYB genes expressed in the vascular and other tissues of gymnosperms seems essential for developing a better understanding of their roles in gymnosperm lignin biosynthesis and wood formation.

MYB proteins have two structural regions, an N-terminal DNA-binding domain (DBD or MYB domain) and a C-terminal modulator region that is responsible for the regulatory activity of the protein. The MYB domain is well conserved among plants, yeast and animals (Lipsick, 1996). Its consensus sequence contains around 50 amino acid residues with regularly spaced tryptophans giving rise to a helix-turn-helix structure (Ogata et *al.*, 1992). There are usually one to three imperfect repeats of the MYB domain. Proteins with two repeats (R2R3-MYBs) are specific to plants and yeast (Jin and Martin, 1999) and are the most abundant type in plants (Riechmann et *al.*, 2000). Plant R2R3-MYBs take part in many biological processes including seed development and germination (Penfield et *al.*, 2001), the stress response (Vailleau et *al.*, 2002) and epidermal cell fate in addition to their involvement in phenylpropanoid and lignin biosynthesis (Tamagnone et *al.*, 1998; Newman et *al.*, 2004; Goicoechea et *al.*, 2004) (for a review, see Schiefelbein, 2003).

The genetic selection and breeding activities of a few commercial conifer species are being expanded to include genetic mapping and marker development. Candidate gene approaches are being adopted to identify robust genetic markers derived from genes that have a physiological role in the traits that are targeted by breeders (Neale and Savolainen, 2004). Our goal is to characterize several members of a gene family proposed to play a role in controlling lignin synthesis and wood properties in conifer trees, in order to support candidate gene approaches for marker discovery. In this report, we characterize 13 different R2R3-MYB gene sequences from the white spruce, Picea glauca, (designated PgMYB) and five from loblolly pine, Pinus taeda L (designated PtMYB). The fulllength coding sequences we obtained enabled us to explore their phylogenetic relationships to other plant MYB genes and to search for novel amino acid motifs within this large protein family. We also compared the gene structures, i.e. number, size, position and splice sequences of introns, to gain further insights into their evolution. The steady-state levels of MYB and cell wall-related gene mRNAs were examined by Q-RTPCR in various spruce tissues and organs with an emphasis on wood-forming tissues and compression wood formation. We identify three MYBs that are preferentially expressed in secondary xylem and are also upregulated during the formation of compression wood.

II-5-Materials and Methods

II-1-1-Plant material and RNA isolation

Several tissues were isolated from two 33-year-old *P. glauca* trees felled in July 2003, from a progeny trial established near Quebec City (Canada). All tissues were frozen in liquid nitrogen immediately upon removal from the tree and stored at -80 °C until further use. We collected newly formed needles from the upper crown. Differentiating secondary xylem and phloem, as well as bark tissues were collected from three 30-40 cm bolts taken from the lower third of the main stem. These vascular tissues were scraped with a scalpel immediately after peeling the bark. Tissues scrapped from the exposed inner side of the bark and from surface of the exposed wood were labelled as differentiating secondary phloem and xylem, respectively. Similarly, differentiating xylem and bark (including phloem) were collected from large roots located in a one-meter radius from the base of the

stem. Samples from each tree and each tissue were kept separate for RNA extraction and gene expression studies.

A gravitropic treatment to induce compression wood formation was performed on 3-year-old spruce seedling stock. The seedlings were transferred to 3 L pots one month before the experiment, grown in a greenhouse with 16 hours light per day, and fertilised weekly with 20 g/L N-P-K. A randomised design of 24 young trees was established in which 12 trees were maintained at 45° angle by leaning the pots and tying the plants to stakes (also at 45°); 12 seedlings were grown in the normal vertical position. Destructive tissue samplings were carried out 4, 28 and 76 hours after the beginning of the treatment. The average diameter of the plants near the base was 7.2 ± 0.61 mm, their average height was 60.63+/- 7.25 cm, and the terminal leader was 19.83 +/- 2.91 cm. For each time point, four vertical and four leaning trees were harvested mid-morning in a randomised order, and three randomly selected trees were used for gene expression analyses. Secondary xylem was collected as described above from the two sides of the main stem of the seedlings; the lower and upper sides representing compression wood and opposite side wood, respectively, for the leaning trees, or left and right side for vertical tree. The whole terminal leader was also collected from each plant. Total RNAs were isolated from tissues described above and ground in liquid nitrogen with a pestle and mortar, except for the gravitropic treatment where a Mixer Mill MM300 engine (Retsch) was used to grind in Microtubes (Eppendorf). The RNAs were extracted from each tissue sample and each tree or seedling separately, following the procedure of Chang et al. (1993), in Oakridge tubes or in Microtubes. RNA concentration and quality were determined with a bioAnalyser (model 2100, Agilent Technologies; RNA 6000 Nano Assay kit).

II-1-2-DNA cloning and accession numbers

Previous reports described two complete coding sequences of *R2R3-MYB* genes from pine (*PtMYB1* and *PtMYB4*) (Patzlaff et *al.*, 2003a, 2003b) and one from spruce (Xue et *al.*, 2003) as well as several partial sequences (MacKay et *al.*, 2004) expressed in xylem tissues of pine. We isolated partial spruce cDNA clones representing putative orthologues of the pine *MYB* genes by PCR amplification,

and identified several partial and putative full-length sequences among the spruce EST sequences data of the ARBOREA project derived from 17 different cDNA libraries (Pavy et *al.*, 2007). For each of the partial spruce and pine gene sequences, we obtained complete coding sequence and UTRs by using 3' RACE, 5' RACE or both cloning methods on spruce or pine mRNA from needles or xylem (SMART RACE cDNA Amplification Kit, Invitrogen, Carlsbad, CA). DNA was cloned in pCR2.1 with the TA cloning Kit (Invitrogen, Carlsbad, CA) and sequenced. The sequence analyses presented hereafter are based upon cDNA clones containing the complete coding sequences of the *MYB*, which were isolated as a single fragment from reverse transcribed RNA, with gene specific primer pairs for each of the 13 sequences from spruce and *PtMYB14*; no *PtMYB5* and *6* reported) is in accordance with Patzlaff *et al.* (2003a, 2003b); the numbering of spruce genes was established on the basis of putative orthology with the pine sequences.

In addition, we isolated the 13 corresponding sequences from spruce genomic DNA (gDNA). Genomic DNA was extracted from needles of white spruce using the Genomic-Tip Kit (Qiagen, Mississauga, Ontario). The entire coding region with introns was isolated by PCR amplification with gene specific primer pairs spanning each gene's coding region (Appendix II-2) and cloned in pCR2.1 with the TA cloning Kit (Invitrogen, Carlsbad, CA). The gDNA was from *Picea glauca* genotype Pg653 and so did most of the cDNA clones (although a few came from wild *Picea glauca* genotypes). Each clone was sequenced at least through the MYB DBD in order to determine the number of introns present in this region. Some nucleotide differences were observed between cDNA and gDNA sequences due to the genotypic variation, but no non-sense mutations were detected. The genomic sequences of *PgMYB* 3, 6, 7 and 11 showed 1 to 3 non synonymous substitutions giving no less than 99.2% amino acids identity; however, we did not find nucleotide mismatches in spruce *MYB* 2, 4, 8, 9 and 12.

The 13 *MYB* genes from spruce and five *MYB* genes from pine have the following accession numbers: *PgMYB1* [GenBank: DQ399073], *PgMYB2* [GenBank: DQ399072], *PgMYB3* [GenBank: DQ399071], *PgMYB4* [GenBank:

DQ399070], *PgMYB5* [GenBank: DQ399069], *PgMYB6* [GenBank: DQ399068], *PgMYB7* [GenBank: DQ399067], *PgMYB8* [GenBank: DQ399066], *PgMYB9* [GenBank: DQ399065], *PgMYB10* [GenBank: DQ399064], *PgMYB11* [GenBank: DQ399063], *PgMYB12* [GenBank: DQ399062], *PgMYB13* [GenBank: DQ399061] and *PtMYB2* [GenBank: DQ399060], *PtMYB3* [GenBank: DQ399059], *PtMYB7* [GenBank: DQ399058], *PtMYB8* [GenBank: DQ399057], *PtMYB14* [GenBank: DQ399056].

The nucleotides sequences of candidate genes involved in wood formation come from the spruce EST assembly directory number 8 (dir8) of the ARBOREA project (Pavy et *al.*, 2007) Their percentage amino acid sequence similarity with other species is given in brackets. They are:

phenylalanine ammonia lyase (PAL) [dir8: contig10199] partial coding sequence (cds), 85% to Pinus taeda [GenBank: U39792]; 4-coumarate:CoA ligase (4CL) [dir8: contig10433] partial cds, 86% to Pinus taeda [GenBank: U39405]; caffeoyl-CoA 3-O-methyltransferase (CCOaOMT) [dir8: contig5884] complete cds, 92 % to Pinus taeda [GenBank: AF036095]; arabinogalactan protein (AGP) [dir8: contig10745] complete cds, 75 % to Pinus taeda [GenBank: U09556]; cinnamyl alcohol dehydrogenase (CAD) [dir8: contig9065] complete cds, 95 % to Pinus radiata [GenBank: AF060491], and the housekeeping gene elongation factor alpha (EF1- α) [dir8: contig10829] complete cds, 99% to Picea abies [GenBank: AJ132534].

II-1-3-Sequence analyses and phylogenetic studies

Nucleotide and amino acid sequence alignments were obtained with Clustal W (Thompson et *al.*, 1994) using BioEdit software version 6.0.7 and BoxShade 3.21 to highlight sequences. Delineation of introns was achieved by aligning the cDNA and genomic nucleotide sequences of the *PgMYB* from the start codon to stop codon on the basis that introns begin with GT and end with AG dinucleotides. Phylogenetic studies were performed with the 13 predicted MYB protein sequences from white spruce, one sequence from black spruce (PmMBF1, Xue et *al.*, 2003), and seven loblolly pine sequences including previously reported PtMYB1 and 4 (Patzlaff et *al.*, 2003a, 2003b). We also included 11 diverse

Arabidopsis MYB sequences and two *R1R2R3-MYB* genes from human and mouse as landmarks to classify the MYBs according to previous reports (Kranz et *al.*, 1998; Romero et *al.*, 1998; Jiang et *al.*, 2004). We constructed a neighbourjoining tree based on a Clustal W amino acid alignment generated with the Mega 2.0 method (Kumar et *al.*, 2001) and using 1000 bootstraps to estimate the node strength (parameters are Poisson correction and pair-wise deletion as described in (Jiang et *al.*, 2004).

The accession numbers of the Arabidopsis genes analysed are: AtMYB13 [GenBank: At1g06180], PtMYB1[GenBank: AY356372], AtMYB4 [GenBank: At4g38620], AtMYB103 [GenBank: At1g63910], AtMYB46 [GenBank: At5g12870], PtMYB4 [GenBank: AY356371], [GenBank: AtMYB61 At1g09540], AtMYB52 [GenBank: At1g17950], AtMYB44 [GenBank: At1g66230], AtMYB101 At5g67300], AtMYB20 [GenBank: [GenBank: At2g32460], AtMYB106 [GenBank: At3g01140], AtMYB33 [GenBank: At5g06100], PmMBF1 [GenBank: U39448]. The accession numbers of the conifer MYB genes are as above.

MEME analysis software (Bailey and Elkan, 1994) was used to identify amino acid regions conserved between several members of a subgroup of sequences (according to Kranz *et al.*, 1998) containing one or more spruce MYBs. The parameters setting was the number of motifs to find: 5; minimum width of motif: 5 and maximum: 15. We used complete protein sequences of MYBs from 12 conifer species and from 14 angiosperm species (Appendix II-3). We also included 10 partial conifer sequences that encompassed the C-terminal region (Kusumi et *al.*, 2002) and were closely related to PgMYB6, 7 and 9 (Appendix II-4).

II-1-4-Analysis of transcript accumulation by Q-RTPCR

To analyse the transcript abundance of *PgMYBs*, first-strand cDNA was synthesised starting from one microgram of RNA treated with amplification grade DNAse I (Invitrogen) and purified on an RNeasy column (Qiagen), using oligo(dT) primers and SuperScript II RT (Invitrogen) according to the manufacturer's instructions. The resulting cDNA was diluted 1:5 with sterile RNAse-free water and stored at -20°C; the same cDNA was used for all Q-RTPCR (quantitative reverse transcription PCR) assays for any given tissue sample. Steady-state mRNA levels were determined on cDNA by quantitative RT-PCR using Opticon Monitor 2 fluorescence detection system (MJ research) and DyNAmo SYBR Green QPCR kit (Finnzymes Oy, Espoo, Finland). Gene primer pairs were designed using the Primer 3 software (Rozen and Skaletsky, 2000) to anneal near the 3' end of each transcript (usually in 3' UTR) to ensure primer specificity. The forward and reverse primers were as follows (amplicon length indicated in brackets):

 $EF1-\alpha$ (elongation factor alpha): 5'-aactggagaaggaacccaag-3' and 5'aacgacccaatggaggatac-3', (114bp). Forward and reverse primer pairs used for the Q-RTPCR of the wood formation-related genes are:

PAL: 5'-tggatttgcatcctactg-3' and 5'-tccatcttcaactataggac-3' (103bp);

4CL: 5'-cattectcaaaagcatgaagag-3' and 5'-ategcatecaaaagtacac-3' (150bp);

CCOaOMT: 5'-attgagatcagccaaatcc-3' and 5'-gcgctctccctataatcag-3' (124bp);

AGP: 5'-gcgtccattgttttaatgtag-3' and 5'-tgtatttatccctctgtctgc-3' (181bp) and

CAD: 5'-ctggactacatcatactgc-3' and 5'-gatttactcattctgcacg-3' (141bp).

The PCR reaction mixture (20 µL) consisted of 10 µL DyNAmo SYBR Green qPCR mix, 1 µL primers (0.25 µM forward and 0.25 µM reverse), 1 µL of cDNA and 8 µL of RNase-free water. The cycling conditions were 95°C for 2 min, then 35 cycles of 95°C for 10 sec, 55°C for 10 sec, 72°C for 8 sec, followed by a plate reading. The melting curve readings were carried out every 0.2°C from 65 to 95°C, holding 1 sec at each temperature. Standard curves were established for each spruce MYB and cell-wall-related cDNAs and were used to determine RNA abundance in each sample. The standards consisted of serial dilutions of PCR amplicons prepared from each cDNA, cloned in pCR2.1, with M13 reverse and forward primers. Amplicon standards were gel purified (Qiagen), and product length and concentration were verified using a bioAnalyser (model 2100 Agilent Technologies, DNA 1000 LabChip kit). The standard curves were determined from duplicate reactions from the dilutions series of each amplicon. Raw data were converted with the following parameters: no blank subtraction, subtract baseline and average over cycle range according each case. We calculated the number of transcript molecules per ng of total RNA [(DNA quantity quantified in g * DNA base pair mass per gram of DNA) / M13 Reverse-Forward amplified fragment length in bp]. For within tissue comparisons that were carried on differentiation xylem (including compression wood induction) and apical shoots young trees, the RNA abundance was normalised using the abundance of RNA of the elongation factor alpha (*EF1-\alpha*), as an endogenous control (calculated as the following ratio: *target gene* (ng) / *EF1-* α (ng)).

II-6-Results

II-6-1-Isolation and sequence analysis of 18 *R2R3-MYB* genes from spruce and pine

We isolated and sequenced 18 full-length cDNAs encoding *R2R3-MYB* genes from conifer trees: 13 from spruce (*PgMYB1–13*) and five from pine (*PtMYB2, 3, 7, 8* and *14*). Each of the full length cDNA sequences were obtained starting from partial or full length clones identified by EST database mining (all of the pine sequences and most of the spruce sequences), or starting from the pine sequence and using RT-PCR amplification with conserved primers to amplify a spruce fragment (Table II-1). For partial clones, we used RACE cloning to

identify flanking sequences and, full length PCR amplification to generate a single full length cDNA. Their predicted amino acid sequences were aligned together with the three available full-length MYB sequences from gymnosperms (Patzlaff et *al.*, 2003a, 2003b; Xue et *al.*, 2003). The DNA-binding domains of these 21 gymnosperm sequences showed a high level of amino acid conservation particularly in the R3 helix-turn-helix repeat, consistent with its involvement in DNA binding (Figure II-1). Most of the variations among the spruce *PgMYB* sequences were located in the turn of each R repeat.

The conifer sequences were consistent with the consensus DBD sequence identified by Avila *et al.* (1993), which was largely based on angiosperm sequences. Only a few amino acid residues differed from this consensus; these were mainly in *PgMYB 3, 6, 7* and *9* and *PtMYB3* (black arrows in Figure II-1). We found a motif similar to that involved in the interaction with basic helix-loop-helix (bHLH) proteins in *Arabidopsis* ([DE]Lx2[RK]x3Lx6Lx3R; Zimmermann et *al.*, 2004) in the R3 repeat of three spruce MYBs (*PgMYB5, 10* and *13*) as well as in *PmMBF1* (Xue et *al.*, 2003; Figure II-1). *PtMYB14* had a similar motif but with two differences: an R instead of an L, and a gap before the last R residue. In addition, several conifer *MYBs*, including those with the bHLH motif (except *PmMBF1*), encoded an Rx5Rx3RR motif similar to the calmodulin-interaction site previously described in the DBD of *Arabidopsis MYB2* (Yoo et *al.*, 2005). The highest level of conservation with the calmodulin-binding motif was observed in *PgMYB2* and *PtMYB2*, but most of the conifer *MYB* genes shown in Figure II-1 had a similar motif.

		Full	DNA	C-						
		length	Binding	terminal			Angio-	MEME	Start	
_Sg ¹		cDNA ²	Domain ³	Domain ³	Motifs	Consensus sequences	Gymno⁴	E-value	motif	⁹ Ref.
4	Pg MYB5	-b-	115	142	F	LlsrGiDP(at)tHrp(li)n	13/13-5/5	6.00 ^{e-14}	1	a)
					G	e(re)cpdLNLel(cr)ispp	13/13-4/5	3.31 ^{e-16}	67	a), b)
	Pg MYB10	-a-	115	95	F	LlsrGiDP(at)tHrp(li)n	13/13-5/5	6.52^{e-14}	1	a)
					G	e(re)cpdLNLel(cr)ispp	13/13-4/5	4.32 ^{e-15}	67	a), b)
	Pg MYB13	-b-	116	80	F	LlsrGiDP(at)tHrp(li)n	13/13-5/5	1.18^{e-14}	1	a)
					G	e(re)cpdLNLel(cr)ispp	13/13-4/5	5.43 ^{e-14}	65	a), b)
22	Pg MYB6	-a-	115	235	Н	(cs)s(sv)DPpT(ls)LsLslPg	7/7-14/14	2.02 ^{e-14}	99	d)
					Ι	YlkaedaismmsaAv	0/7-13/14	1.87^{e-13}	141	d)
					J	vmremvakEVrsYmn	7/7-14/14	1.07 ^{e-17}	188	a), b), c)
	Pg MYB7	-c-	116	257	K	egdyEVesrgLKRln	0/7-13/14	1.34^{e-12}	43	d)
					Н	(cs)s(sv)DPpT(ls)LsLslPg	7/7-14/14	4.28 ^{e-13}	113	d)
					Ι	YlkaedaismmsaAv	0/7-13/14	2.30^{e-10}	161	d)
					J	vmremvakEVrsYmn	7/7-14/14	6.15 ^{e-16}	206	a), b), c)
	Pg MYB9	-a-	119	297	K	egdyEVesrgLKRln	0/7-13/14	6.03 ^{e-16}	60	d)
					Р	hRQSAFksYesqktp	0/7-11/14	1.19^{e-13}	116	d)
					Н	(cs)s(sv)DPpT(ls)LsLslPg	7/7-14/14	2.91 ^{e-13}	144	d)
					Ι	YlkaedaismmsaAv	0/7-13/14	9.50 ^{e-16}	205	d)
					J	vmremvakEVrsYmn	7/7-14/14	1.09 ^{e-16}	256	a), b), c)
8	Pg MYB1	-b-	115	217	Α	lr(kq)mGiDP(lv)THkpl	5/5-2/2	1.79 ^{e-18}	1	<u>a)</u>
21	Pg MYB3	-c-	130	177	C	(fg)Re(rq)S(rs)(is)(rg)(kr)R	4/5-2/2	4.69 ^{e-14}	1	d)
					D	e(en)s(l)(vs)(pt)ffDfl(g)vG(cn)	5/5-2/2	1.26^{e-13}	35	a), b)
					E	(cy)xi(sg)h(in)nh(v)q(sf)(jr)Kef	3/5-2/2	4.76 ^{e-14}	123	d)
13	Pg MYB8	-c-	115	411	L	LrrGIDP(n)THkpl	4/4-2/2	2.54 ^{e-17}	1	a)
					Μ	VC(dv)(yk)(np)SIm(al)nPsm(yn)	2/4-2/2	1.94^{e-18}	199	d)
					N	e(ye)(ae)vKWSEml	2/4-2/2	6.45^{e-14}	317	d)
					0	(pk)D(fl)(hq)R(im)Aa(vs)(lf)(dg)q	2/4-2/2	4.89 ^{e-15}	399	a)
9	Pg MYB11	-a-	115	384	Q	L(lv)kMGIDPvTHkp(k)	6/6-1/1	4.08^{e-16}	1	a), b), c)
					R	h(m)AQWEsARleAear	6/6-1/1	3.10^{e-13}	35	a), b), c)
					S	(yc)eDnknYw(nd)silnlV	4/6-1/1	6.79 ^{e-12}	360	c)
_ 2	Pg MYB12	-a-	115	254	Т	MdfW(fl)(dn)v(fl)(t)	5/5-1/1	2.39 ^{e-09}	237	a)
nd	Pg MYB2	-c-	115	333	В	(c)SylPPL(y)d(v)	2/2-2/2	3.29 ^{e-13}	249	d)
	Pg MYB4	-c-	120	214	none	none	0/3-0/2	none	none	none

Table II-1. Predicted lengths and C-terminal motifs of spruce MYB proteins

Conserved amino acid regions were identified in angiosperm and gymnosperm Cterminal sequences by the use of MEME software (setting described in Methods). Motifs were detected among the sequences belonging to each phylogenetic clade comprised of at least one spruce MYB (Appendix II-3). Sequences from appendix II-4 were used to identify more conifer members of the PgMYB6, 7, 9 clade. Within the consensus sequences, upper-case letters indicate amino acids found in all members of a subgroup, lower-case letters indicate amino acids conserved in more than 50% of the members, pairs of lower-case amino acid in brackets show the two most abundant amino acids present for 50% each and above, x indicates that no amino acid is conserved among the sequences.

(See end of legend next page)

Table II-1. End of legend

¹Sg, MYB subgroups identified by Kranz et al. (1998).

² Source of full length cDNA sequence: -a-, full length cDNA clone identified from EST of *Picea glauca* database; b- partial cDNA clone identified from EST database of *P. glauca*, extended by RACE amplifications and finally amplified as a single clone by PCR with gene specific primers, and –c-, from non degenerates primers based on *Pinus taeda MYB* sequences and used on spruce cDNA followed by RACE amplifications.

³Lengths are expressed in amino acid (aa) residues.

⁴ The number of MYB sequences, separately from angiosperm and gymnosperms, sharing the motif among all those used in each case.

⁵ The position of the motif relative to the beginning of the C-terminal domain (5' end).

Ref: references for previously reported motifs, a) Kranz *et al.* 1998, b) Stracke *et al.* (2001) and c) Jiang *et al.* (2004) and d) new motifs.



Figure II-1. Alignment of predicted MYB domain protein sequences from spruce and pine (see legend next page)

Figure II-1. (legend) Amino acid sequence alignments of the 21 conifer MYB R2R3 domains were obtained with Clustal W (see Methods) and then separated into three groups based on their homologies to the consensus R2R3-MYB DNAbinding domain (MYBR2R3-DBD, top panel), the bHLH protein-binding motif (bHLH motif, middle panel) or the Arabidopsis calmodulin-interaction motif (AtMYB2 CaMBD, bottom panel), as indicated. Black shading indicates identical amino acid residues and grey shading the similar residues that agree with the fraction sequence of 0,4 (BoxShade 3.21) and dashes indicate gaps. The numbers on the left and right indicate the amino acid position relative to the translation start codon. The boxes and dotted line above the sequences show the predicted helix and turn structures in the R2 and R3 regions of the MYB domain. Stars show positions of conserved tryptophan residues and black arrows indicate unusual amino acid residues compared to the consensus amino acid sequence of the MYB DNA-binding domains of several plant R2R3-MYB proteins described by Avila et al. (1993).The bHLH protein-binding motif ([DE]Lx2[RK]x3Lx6Lx3R) identified by Zimmerman et al. (2004) and the calmodulin-interaction motif (Yoo et al., 2005) are shown above the middle and bottom panels, respectively (major amino acids in upper-case, bold). Ia or Ib and II indicate the positions of the first and second introns, respectively (Ib is specific to PgMYB3). Accession numbers of the newly identified spruce and pine MYBs are listed in Methods. Pg, Picea glauca; Pt, Pinus taeda; Pm, Picea mariana; At, Arabidopsis thaliana

II-6-2-Phylogenetic relationships and gene family structure of conifer *R2R3-MYBs*

We used the Mega 2.0 method to construct a phylogenetic tree using full length cDNA sequences (Figure II-2). The result of our analysis is congruent with the three major groups of R2R3-MYBs (A, B and C) defined by Romero *et al.* (1998) on the basis of their binding affinities to MYB recognition elements. On this phylogenetic tree, the predicted spruce MYB proteins sequences fell into several subgroups with bootstrap values ranging from 96–100%, indicating the high grouping robustness. All of the spruce and pine MYB sequences fell into group A (PgMYB3, 6, 7 and 9, and putative pine orthologues) or C (all the other conifer sequences in figure II-2) and none belonged to the B group. The conifer sequences were assigned to 7 of the 22 subgroups previously defined based upon *Arabidopsis* sequences (Kranz et *al.*, 1998). Four of the conifer sequences (PgMYB2 and 4; PtMYB2 and 4) clustered with *Arabidopsis* sequences that do not fit into a defined subgroup. Several pairs of spruce and pine sequences clustered closely together with short branch-lengths indicative of a high degree of homology (Figure II-2). Indeed, pair-wise optimal alignments with the Clustal W algorithm of the pine and spruce pairs 1, 2, 3, 4, 7 and 8 gave amino acid identities from 95% to 100% for the DBD and of 79%–93% for the complete coding sequence (Table II-3), suggesting that they are putative orthologous pairs. By comparison, PtMYB14 was less homologous to its neighbouring spruce sequences PgMYB5, 10 and 13 (60% to 67% homologous for the full CDS).

We also analysed the number, size and sequences of introns in PCRamplified genomic DNA, as a complement to the phylogenetic analysis based on the coding sequences. In angiosperm R2R3-MYBs the introns are located in the Myb DBD, therefore we sequenced this specific region in genomic DNAs of the 13 spruce *R2R3-MYBs*, isolated by PCR amplification with gene specific primer pairs spanning each gene's coding region (Appendix II-2). Most of the gDNA sequences were identical to the cDNAs, ranging from 100% to 99.3% in amino acid identities (data not shown), due to a few variations in the predicted amino acid sequences. The sequences were also verified for the lack of non-sense mutations (stop codons or frameshifts). As observed in angiosperms, we found spruce *MYB* genes with one (I), two (I, II) or no introns (Table II-2).

Similarity was found between the spruce MYBs in terms of intron position, phases and, in some cases, between intron sequences, but the number and length of introns was quite variable. Generally, the second intron (II) was longer than the first (I) except in *PgMYB11* where intron I was five times longer than intron II. The spruce sequences belonging to group A MYBs fell into two subfamilies with distinct gene structures, i.e. with one intron (Sg21) and one without introns (Sg22). The group C sequences all had one or two introns, as in *Arabidopsis* (Jiang et *al.*, 2004). The intron I occurred before the GL amino acid pair in repeat 2 and the intron II occurred after the GN amino acid pair in repeat 3 (Figure II-1), as found in the majority of

Arabidopsis R2R3-MYB genes (Jiang et *al.*, 2004). Only *PgMYB3* had a different intron I site, named Ib, located before the GKS amino acids. Moreover, the phase (1 or 2) of insertion was consistent, and the end sequences (GT in 5' and AG in 3') were conserved among the sequences we analysed (Table II-2). Phylogenetically close sequences, like *PgMYB5*, *10* and *13*, had similar 5' and 3' splice junctions for both introns (Table II-2). The intron of these three genes also showed strong nucleotide sequence conservation, although the first intron of *PgMYB5* was much longer due to a 20-nucleotide triplicated sequence (not shown).

	L	ength (bp)	Intron I	(phase 1)	Intron II (phase 2)		
	Coding							
	sequence	Intron I	Intron II	5'Splice site	3'Splice site	5'Splice site	3'Splice site	
Pg MYB 1	999	83	101	CCG:GTAAAT	TTGCAG:GTC	TAG:GTATAT	CACCAG:GTG	
Pg MYB 2	1347	501	88	CAG:GTACTC	TGACAG:GTC	CAG:GTTTGT	GTGCAG:GTG	
Pg MYB 3	924	1427	none	CAG:GTAAAG	ATGCAG:GGA	none	none	
Pg MYB 4	1005	194	267	CTG:GTAAGC	GTACAG:GTC	CAG:GT TTTT	GCGCAG:GTG	
Pg MYB 5	774	191	186	CAG:GTTGAA	TTGCAG:GGC	<u>CAA: GTATGT</u>	<u>GCGC AG:GTG</u>	
Pg MYB 6	1053	none	none	none	none	none	none	
Pg MYB 7	1122	none	none	none	none	none	none	
Pg MYB 8	1581	97	90	CTG:GT AAAG	TCGCAG:GCC	CAG:GTAATG	ACACAG:GTG	
Pg MYB 9	1251	none	none	none	none	none	none	
Pg MYB 10	633	94	187	<u>CAG: GTTTCT</u>	ATGC AG:GGC	<u>CAA: GTATGT</u>	<u>GTGC AG:GTG</u>	
Pg MYB 11	1500	644	132	CAG:GTATTT	ATGCAG:GAC	CAA:GT AAGG	TTACAG:ATG	
Pg MYB 12	1110	94	294	CAG:GTCACT	TTGCAG:GGC	CAG:GTGAGT	ATGTAG:ATG	
Pg MYB 13	591	94	139	<u>CAG: GTTTCT</u>	ATGC AG:GGC	<u>CAA: GTATGT</u>	<u>GTGC AG:GTG</u>	

Table II-2. Length of spruce MYB coding sequences and introns with their predicted splice junctions

Coding sequences indicate the length in nucleotides from the translation start codon to the stop codon. Introns I and II represent the first and second introns, respectively. Intron phase refers to the position in a codon where the intron is inserted: after the first nucleotide (phase 1) or after the second nucleotide (phase

2) of a codon. Italic nucleotide pairs GT and AG represent the beginning and the end of the introns, respectively; and underlined nucleotides are conserved splice-site sequences.



Figure II-2. Phylogenetic tree of gymnosperm and angiosperm R2R3-MYB proteins

This neighbour-joining (1000 Bootstraps) tree was based on the Clustal W alignment of the complete coding sequences of 13 spruce and five pine MYB proteins identified in this study (represented by filled and empty lozenges, respectively). The bar indicates an evolutionary distance of 0.2%. *Arabidopsis* proteins were chosen as landmarks representing the three main groups (circles A, B and C) and subgroups (Sg next to bracket; nd, not determined) defined by Romero *et al.* (1998) and Kranz *et al.* (1998). Human c-MYB [GenBank: P10242] and *Mus musculus* MmMYBA [GenBank: X82327] were not used as out groups but as landmarks. The accession numbers of the *Arabidopsis* genes are given in Methods. Other abbreviations are in Figure II-1.

	DNA Bii	nding Domains	Full coding sequences		
Amino acids percentage	Identity	Similarity	Identity	Similarity	
PgMYB1 / PtMYB1	99,1	100	87,1	91,3	
PgMYB2 / PtMYB2	100	100	88	91,6	
PgMYB3 / PtMYB3	94,6	95,4	79,2	82,3	
PgMYB4 / PtMYB4	96,6	97,5	84,1	88,6	
PgMYB7 / PtMYB7	94,8	96,5	90,4	93,6	
PgMYB8 / PtMYB8	98,3	100	93,1	95,5	
PgMYB5 / PtMYB14	88	94,8	67,3	77,4	
PgMYB10 / PtMYB14	87,8	95,6	64,3	73	
PgMYB13 / PtMYB14	82,7	92,2	60	70	

Table II-3. Pair-wise sequence amino acids identities of the DBD and fullCDS of closest spruce and pine homologs

Percent similarity calculations were performed using pairwise optimal alignment with Bioedit software (Clustal W, matrix blosum62).

II-6-3-Sequence analysis of conserved regions in the C-terminal of *P. glauca* MYBs

The coding regions of the spruce *PgMYB* sequences ranged widely in length, encoding between 196-526 amino acid residues depending on the length of the Cterminal region (Table II-1). We used the predicted C-terminal coding regions of the spruce MYB proteins to search for conserved sequences, reasoning that such motifs might be important for the function or post-translational regulation of MYB. We used the MEME motif-detection software to analyse the C-terminal region of spruce MYBs using a set of protein sequences selected for their high degree of similarity to each of the spruce MYBs (Table II-1, Appendix II-3 et II-4). Our approach incorporated a large diversity of sequences; it identified a total of 20 different motifs (A-T) in the spruce MYBs, including nine new unpublished motifs (Table II-1) and 11 that were reported previously (Kranz et al., 1998; Romero et al., 1998; Jiang et al., 2004). The probability scores for each of the motifs identified in this study ranged from 2.39^{e-09} to 1.79^{e-18} . The lowest previously published score for such motif was 6.79^{e-12} (motif S in PgMYB11) (Jiang et al., 2004). We detected between zero (in PgMYB4) and five (in PgMYB9) motifs per protein in the predicated spruce MYB sequences. The large number of conifer sequences enabled us to detect three amino acids regions, I, K and P, that appeared to be specific to gymnosperms (in PgMYB6, 7 and 9). Other motifs, such as F and G, were shared between gymnosperm and angiosperm sequences. Four of the

conserved amino acid sequences (A, F, L and Q) shared the central core residues GIDPxTH but displayed differences in neighbouring amino acids between the consensus sequences of subgroups 4, 8, 9 and 13 defined by Kranz *et al.* (1998).

II-6-4-Expression of *P. glauca MYB* genes in tissues of young and mature trees

We surveyed the abundance of each of the 13 *PgMYB* gene transcripts by Q-RTPCR, in mature (33-year-old trees) and young (3-year-old) green-house-grown trees to determine their tissue distribution during normal development. Six different organs and differentiating tissues (the young needles; the periderm, phloem and xylem from the stems; and the periderm with phloem, or bark, and xylem from the roots) were collected from two different mature trees (Figure II-3). For tissue comparisons, we calculated the number initial *MYB* RNA molecules per ng of total RNA. Spruce *PgMYBs 2, 4* and *8* were expressed preferentially in differentiating xylem from stem and root. Other *MYBs* were abundant in the needles along with one to two other tissues from the stem or the roots or both, Some *MYB* mRNAs also appeared to have rather ubiquitous profiles or low abundance. The RNA abundance of lignin biosynthesis enzymes PAL, 4CL, CCoAOMT and CAD were also determined in the same tissue samples. The lignin enzymes RNAs all gave very similar profiles, and they were most abundant in differentiating xylem (only *4CL* is shown; Figure II-3).

We also compared the abundance of the different *MYB* transcripts in the differentiating secondary xylem and in the elongating apical leader of young spruce trees (Figure II-4). The cell wall-related genes *PAL*, *4CL*, *CAD*, *CCoAOMT* and an *arabinogalactan protein* (*AGP*) were included in this analysis. For these within tissue comparisons, the data were normalized against the *EF1-* α transcript levels. Again, the spruce *MYB* transcripts *2*, *4* and *8* were clearly the most abundant among the MYBs detected in the secondary xylem, consistent with the data from the mature trees. In the apical leader, the relative abundance of the *MYB* transcripts was quite different than in the secondary xylem, except that *PgMYB4* transcripts remained very abundant. Some *MYB* genes that were weakly expressed or not detectable in secondary xylem were among the most highly expressed in apical stem (*PgMYB6*, 7 and *11*; Figure II-4a).



Figure II-3. Transcript abundance for 13 spruce MYB genes and 4CL in various organs and tissues

Transcript abundance was determined by Q-RTPCR of six tissues from two different 33-year-old trees (number of molecules per ng of total RNA, see methods). The transcript level of an *elongation factor* (*EF1-* α) gene was used as an RNA control. N, needles; Stem tissues: P, periderm; Ph, differentiating phloem; X, differentiating xylem; Root tissues: PPh, root periderm with differentiating phloem; X, root differentiating xylem. Data are based on three technical repetitions per tree, i.e. six measurements per data point. Vertical bars represent the standard error. *4CL: 4-coumarate: CoA ligase*. NS, no PCR product detected.





Transcript abundance was determined as in figure II-3 for, a) 13 spruce *MYB* genes, and b) five cell-wall-related genes in differentiating secondary xylem from stem and in the elongating terminal leader (apical stem) from 3-year-old spruce seedlings. The standard error (bars) was calculated from three biological replicates and two independent technical repetitions (i.e. six independent measurements). *PAL*, *phenylalanine ammonia lyase*; *4CL*, *4-coumarate*: *CoA ligase*; *CCOaOMT*, *caffeoyl-CoA 3-O-methyltransferase*; *AGP*, *arabinogalactan protein*; *CAD*, *cinnamyl alcohol dehydrogenase*. NS, no PCR product detected.

II-6-5-Spruce *MYB* genes are differentially expressed in compression wood

We followed the expression of the 13 spruce MYB genes and five cellwall-related genes during the early phases of compression wood formation, in order to explore further the potential involvement of MYBs in wood formation and lignin biosynthesis. Gymnosperm trees form a type of reaction wood (known as compression wood) on the lower side of a bent or leaning stem, or in branches. Compression wood is enriched in lignin and contains lignins that are more condensed. Therefore compression wood formation requires the modulation of lignin biosynthesis, which we hypothesized to involve such gene sequences as R2R3-MYBs. We induced the formation of compression wood in actively growing 3-year-old spruces by maintaining at a 45° angle (relative to vertical) (Figure II-5). After 21 days of growth in this leaning position, characteristic compression wood was well developed on the lower side of the stems (Figure II-5a). We chose to monitor transcript abundance over a 76-hour period immediately after induction, and found that several transcripts accumulated between 28 and 76 hours (Figure II-5c). The transcripts of PgMYB2, 4 and 8 clearly increased in the xylem forming compression wood compared to the opposite wood and compared to the vertical trees (0 hour time point). The transcripts of PgMYB9, 11 and 13 RNA were slightly increased and the seven others did not fluctuate significantly. By contrast, no significant variation in spruce MYB RNA abundance was observed in the opposite wood, which is found on the upper side of the stem (opposite to the the compression wood). Transcripts for PAL, 4CL, CCoAOMT and CAD lignin biosynthesis enzymes as well as the AGP also increased within the same timeframe as the MYB transcripts (Figure II-5b). In the opposite wood, only CCoAOMT RNA transcripts decreased. No significant variation in transcript abundance was observed for the spruce *MYB* or lignin genes in the terminal shoots of the same seedlings (data not shown).


Figure II-5. Transcript accumulation for *MYB* **genes and secondary cell-wallrelated genes in differentiating compression wood and opposite wood.** (see legend next page)

Figure II-5. (legend) a) Compression wood and opposite wood formed in a leaning spruce seedling after 21 days of treatment, compared to the control from vertical seedling. Exposed wood (compression wood is light brown) and wood cross-sections (10 μ m thick) were stained by the safranin-orange procedure (Sharman, 1943) (magnification, x40). Steady-state mRNA levels were determined as in Figures 4 and 5 for cell-wall-related genes (b) and for several *PgMYB* genes (c) in the compression wood (left panels) and opposite side wood (right panels) of spruce seedlings leaning at a 45° angle from vertical. Continuous lines indicate genes with significant variation, and standard error bars are shown three trees (biological replicates) with two independent technical repetitions). Discontinuous lines indicate examples of gene transcripts that do not fluctuate in abundance. The zero time point represents vertical control trees only. *PgMYB4* (1/15) means that mRNA level is divided by 15.

II-7-Discussion

In this paper, we report the complete coding sequences of 18 conifer gene sequences that share the characteristic features of the *R2R3-MYB* gene family. Thirteen sequences were from *P. glauca* (white spruce; *PgMYBs*) and five from *P. taeda* L. (loblolly pine; *PtMYBs*). We characterised the full-length cDNA sequences, as well as the spruce exon-intron structure. We assigned the conifer sequences to several phylogenetic clades of the R2R3-MYB family and identified conserved motifs within them based on predicted amino acid sequences. The steady-state mRNA levels of spruce MYBs were surveyed in several tissues to identify those genes that are preferentially expressed in wood-forming tissues. Furthermore, we identified *PgMYBs* whose transcript levels are upregulated, along with those of an *AGP* and enzymes of lignin biosynthesis, during the induction of compression wood in young spruce trees.

II-7-1-Sequence conservation and identification of amino acid motifs in spruce R2R3-MYBs

Our data show that the DBDs of conifer MYBs are highly conserved, whereas the C-terminal region are highly variable, as shown in prior studies of other plant MYBs. The predicted amino acid sequences of some of the spruce MYB DBDs

contain a motif for interaction with bHLH proteins and/or with calmodulin. We identified twenty amino acid motifs in the variable C-terminal region, of which nine were previously unreported. The amino acid motifs in the DBD and in the C-terminal region are useful to better characterise the spruce R2R3-MYB sequences belonging to each phylogenetic clade.

The R2R3-MYBs are specific to plants and are subdivided into three major groups according to their binding affinities (Romero et al., 1998). The more than 120 Arabidopsis sequences were placed into 22 subgroups based on their overall amino acid sequences and C-terminal motifs (Kranz et al., 1998). Amino acid motifs are conserved among members of several of the 22 phylogenetic clades or subgroups (Kranz et al., 1998; Jiang et al., 2004; Stracke et al., 2001), including the bHLH-interaction and calmodulin-binding motifs in the DNAbinding domain (Zimmermann et al., 2004; Yoo et al., 2005), and the repression domain pdLNLD/ELxiG/S in the C-terminus (Jin et al., 2000). In our study, the spruce group C sequences were dispersed among seven phylogenetics clades, five of which were previously defined as distinct subgroups by Kranz et al. (1998). All the spruce members of subgroup 4 harboured the bHLH-interaction motif as well as the C-terminal motifs F and G, except for PmMBF1 (Xue et al., 2003), which lacked the motif G (pdLNLD/ELxiG/S) described by Kranz et al. (1998). The bHLH-interaction motif identified by Zimmermann et al. (2004) is required for MYB proteins to transactivate some of the phenylpropanoid and anthocyanin genes through protein-protein interactions (Goff et al., 1992). The G motif in the C-terminus has been linked to transcriptional repression of the cinnamate 4hydroxylase (C4H) gene by AtMYB4 (Jin et al., 2000). Several genes in group C also encoded a conserved GIDP sequence located after the end of the DBD, suggesting a conserved molecular function for this motif. The DBDs of PgMYB2, 4 and 8, which were upregulated during compression wood formation, harboured a motif similar to the calmodulin-interaction site of AtMYB2 (Yoo et al., 2005), suggesting a potential link with the calcium signalling pathway implicated in the regulation of secondary wall formation (Kobayashi and Fukuda, 1994). No conserved regions were detected in the C-terminal region of PgMYB4 and its closest homolog PtMYB4, even though experimental evidences indicate that PtMYB4 is a regulator of lignin synthesis enzymes (Patzlaff et al., 2003a), as is

the case for the closely related EgMYB2 (Goicoechea et *al.*, 2005). The presence of a regulator motif in PgMYB4 may have escaped our analysis because the parameters were set to detect motifs ranging from 5–15 amino acids in length; motifs of less than five amino acids or scattered in several small modules may thus remain undetected.

Spruce MYBs were relatively under-represented in group A, where they fell into subgroups 21 and 22. In our analysis, spruce group A MYBs contained six of the nine newly identified C-terminal consensus amino acid sequences. Three of these motifs were specific to conifers assigned to subgroup 22: motifs I, K and P found in PgMYB6, 7 and 9. The motifs might be involved in protein or DNA interactions; however, it remains to be seen whether they play a role in protein structure or function.

II-7-2-Spruce MYB phylogeny and evolution

There are very few reports from which to estimate the number of R2R3-MYB genes in gymnosperms or to gain insights into the molecular evolution of this protein family (Patzlaff *et al.*, 2003a, 2003b; Xue *et al.*, 2003; Kusumi *et al.*, 2002; MacKay *et al.*, 2004). According to the phylogenetic relationship with other MYB genes in angiosperms and gymnosperms, the spruce MYB sequences described here belong to nine different MYB clades distributed between group A and group C described by Romero *et al.* (1998). None of the conifer sequences identified in this study and none of the reported gymnosperm R2R3-MYBs were assigned to the B group (Romero *et al.*, 1998). We may hypothesize that group B sequences are present only in angiosperms, however, more gene discovery work is needed to draw conclusions since only four of the 125 Arabidopsis MYB genes belong to this group B (Romero *et al.*, 1998; Stracke *et al.*, 2001).

Despite recent large-scale gene discovery initiatives for conifers like pine and spruce (e.g. Kirst *et al.*, 2003; Pavy *et al.*, 2005), only a few regulatory gene families have been characterised in any conifer species. The R2R3-MYBs family has evolved and expanded very rapidly through numerous gene duplications in

Angiosperms (Rabinowicz et al., 1999). Given the very distant separation of gymnosperms and angiosperms (approx. 300 million years), we were interested in assessing whether a similar gene family evolution would be present in both taxonomic groups. In other words, is the R2R3-MYB gene family structure similar in these two groups? In the knox-I gene family of conifer trees, the structure and number of genes was shown to be very different from that of angiosperms, in a recent study investigating evolution of the family in great detail (Guillet-Claude et al., 2004). Several of the angiosperm clades are missing in conifers which appear to have undergone several recent gene duplications with relatively low sequence divergence levels. Our work provides a clear indication that the conifer MYB family structure is not all that divergent from that of the angiosperms, in contrast to the Knox-I report, suggesting that the basic family structure predates the gymnosperm - angiosperm split. In maize, several subgroups of R2R3-MYB genes have expanded within the past 50 million years (Rabinowicz et al., 1999; Braun and Grotewold, 1999). Consistent with this, our analysis of coding sequences and introns in spruce MYB genes also suggests more recent gene duplications in at least in some of the clades. For example, PgMYB5, 10 and 13 have high levels of nucleotide sequence similarity in coding sequence as well as introns I and II.

Further investigation is needed to discover the full complement of conifer MYB sequences. By comparison to the angiosperms, we predict that the set of sequences described here represents a fraction of the conifer R2R3-MYB family. Identification of new sequences would complete the evolutionary picture of this conspicuous family of regulators and help to determine its position in the evolution of plant lineages.

II-7-3-Potential involvement of the spruce R2R3-MYBs in the lignification of woody tissue

The spruce and pine sequences we analysed represent diverse subgroups of the R2R3-MYB family. Thus, we hypothesized that they could play diverse roles in metabolism and development. The involvement of specific R2R3-MYB gene products in lignin biosynthesis and/or wood formation is suggested by their expression profiles and by their sequence homology with genes from pine (Patzlaff et *al.*, 2003a, 2003b) and in other species whose functions have been previously tested. The AC cis-regulatory elements, for example, which are found in many promoters of phenylpropanoid and lignin biosynthesis genes, play an important role in gene regulation in lignifying xylem cells, thus linking R2R3-MYB genes with lignin biosynthesis. AC elements have been implicated in the transcriptional regulation of PAL in bean (Leyva et *al.*, 1992), 4CL in parsley (Hauffe et *al.*, 1993), CCR and CAD in Eucalyptus (Lauvergeat et *al.*, 2002; Lacombe et *al.*, 2000).

We compared the abundance of the 13 different spruce MYB mRNAs in selected tissues and organs that develop a secondary vasculature in mature spruce, and in the primary stems and differentiating secondary xylem in young trees. PgMYB2, 4 and 8, all of which belong to the same phylogenetic clade, were expressed preferentially in the secondary differentiating xylem of both juvenile plants and mature trees. Interestingly, all three genes were also expressed preferentially in xylem tissues isolated from large roots. By comparison, 4CL had a very similar transcript profile in prospected tissues. The other MYB genes had various patterns of expression including phloem-preferential and ubiquitous patterns.

We also compared the RNA levels in differentiating secondary xylem during the induction of compression wood in spruce seedlings. Compression wood development in conifers that are leaning or bent is characterised by the formation of thicker cell walls, increased lignin content and the deposition of more condensed lignin polymers, among other features (Timell, 1986). The plasticity of lignin biosynthesis and cell wall architecture observed in compression wood have been linked to the fluctuation in abundance of several gene transcripts and proteins (Plomion et *al.*, 2001; Gion et *al.*, 2005). Although the transcriptional regulators that orchestrate this plasticity are unknown, they might include MYB transcription factors due to their implication in xylem differentiation and in lignin biosynthesis. The three PgMYBs (2, 4 and 8) that were preferentially expressed in xylem are likely candidates because they were also upregulated on the compression wood-forming side (downward side) of the stem but remained relatively constant on the opposite side. The time-course and relative magnitude of the changes in transcript levels of the PgMYBs 2, 4 and 8 were quite similar to those seen for genes encoding lignin biosynthesis enzymes and an AGP surveyed in the same samples. Three other PgMYBs (9, 11 and 13) also showed an increase in transcript abundance in secondary xylem upon induction of compression wood. In the control trees, the mRNA for PgMYB11 was one of the highest we examined in the apical portion of the stem but it was low in secondary xylem. These observations might imply a role for PgMYB11 in processes that are common to primary stem growth and compression wood formation. The MIXTA gene from *Antirrhinum majus*, which belongs to the same phylogenetic subgroup, is involved in cellular development (Noda et *al.*, 1994) and may provide clues to the role of PgMYB11 in conifers.

The expression profiles of a few of the spruce MYB genes are consistent with previous reports describing the putative function of homologous genes. For example, spruce PgMYB4 is a close homolog of PtMYB4 (Patzlaff et al., 2003a), which induced ectopic lignification when overexpressed in transgenic tobacco. A putative role for PgMYB4 in lignification is consistent with our data showing a higher mRNA level in compression wood, characterised by increased lignin deposition. The gene PgMYB8 (subgroup 13) showed strong similarity with AtMYB61, which is expressed in xylem tissues of Arabidopsis and was shown to play an important role in regulating lignification (Newman et al., 2004). AtMYB61 is also expressed in developing seeds, where it regulates the extrusion of seed coat-derived rhamnogalacturonan mucilage (Penfield et al., 2001). The accumulation of PgMYB8 transcripts in compression wood is consistent with the ectopic lignification resulting from the constitutive overexpression of AtMYB61 in Arabidopsis (Newman et al., 2004). By contrast, AtMYB103, the most similar Arabidopsis sequence to PgMYB2, is not expressed in the stem but is involved in trichome and tapetum development (Higginson et al., 2003), suggesting a putative role in xylem differentiation other than lignin biosynthesis.

The spruce sequence *PgMYB1* has a close pine homolog, *PtMYB1*, that has been linked to lignin biosynthesis (Patzlaff et *al.*, 2003b), however the spruce

sequence was not expressed preferentially in secondary xylem (it was also expressed in needles, phloem and the shoot apex) nor was it induced during compression wood formation. It was demonstrated that *PtMYB1* is able to bind the AC-I and AC-II elements (PAL-Box) (Patzlaff et *al.*, 2003b). Recently, Gomez-Maldonaldo *et al.* (2004) showed that pine MYB1 and MYB4 bind to glutamine synthetase AC elements and that MYBs are linked to several metabolic pathways by shared *cis*-acting elements. Based on these observations, it appears that the *MYB1* genes of pine and spruce may regulate phenylpropanoid metabolism as well as nitrogen assimilation in various plant tissues.

II-8-Conclusion

Through a systematic survey of EST sequence data followed by full length sequencing, we characterised 18 conifer R2R3-MYB gene sequences (13 from P. *glauca*, white spruce; 5 from P. *taeda*, loblolly pine). Three R2R3-MYBs from spruce, namely MYB 2, 4 and 8 were shown to be expressed preferentially in secondary xylem. We also found that transcript levels of six PgMYB genes (including the MYB 2, 4 and 8 genes), were upregulated in differentiating secondary xylem from young trees during the induction of compression wood along with cell-wall-related genes. Our study highlights a small set of spruce MYB transcription factors that could be good candidate genes for marker development studies. Gain-of-function / loss-of-function studies using transgenic plants are also needed to delineate the roles of these different MYBs. Such studies are expected to lead to greatly lacking insights into the regulation of wood formation in conifers.

II-9-Acknowledgments

We are grateful to R.R Sederoff (North Carolina State University, Raleigh, NC) for kindly providing clones of *Pinus taeda*. We acknowledge the technical assistance of S. Blais for cloning of pine sequences and F. Boileau for tissue and RNA isolation. The authors thank C. Bomal for critical reading of the manuscript. Funding from Genome Canada and Génome Québec to JM for the ARBOREA project supported this research.

II-10-References

- Avila J., Nieto C., Canas L., Benito M.J. and Paz-Ares J. (1993). Petunia hybrida genes related to the maize regulatory C1 gene and to animal myb proto-oncogenes. *Plant J* **3**: 553-562
- Bailey T.L. and Elkan C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol 2: 28-36
- Boerjan W., Ralph J. and Baucher M. (2003). Lignin biosynthesis. Annu Rev Plant Biol 54: 519–546
- **Braun E.L. and Grotewold E. (1999).** Newly discovered plant c-myb-like genes rewrite the evolution of the plant *myb* gene Family. *Plant Physiol* **121:** 21-24
- Chaffey N: (2002). Why is there so little research into the cell biology of the secondary vascular system of trees? *New Phytol* 153: 213-223
- Chang S., Puryear J., and Cairney J. (1993). A Simple and Efficient Method for Isolating RNA from Pine Trees. *Plant Mol Biol* Reporter 11: 113-116
- Gion J.M., Lalanne C., Le Provost G., Ferry-Dumazet H., Paiva J., Chaumeil P., Frigerio J.M., Brach J., Barré A., de Daruvar A., Claverol S., Bonneu M., Sommerer N., Negroni L. and Plomion C. (2005). The proteome of maritime pine wood forming tissue. *Proteomics* 5: 3731-3751
- Goff S.A., Cone K.C. and Chandler V.L. (1992). Functional analysis of the transcriptional activator encoded by the maize B gene: evidence for a direct functional interaction between two classes of regulatory proteins. *Genes Dev* 6: 864–875
- Goicoechea M., Lacombe E., Legay S., Milhaevic S., Rech P., Jauneau A., Lapierre C., Pollet B., Verhaegen D., Chaubet-Gigot N. and Grima-Pettenati J. (2005). EgMYB2, a new transcriptional activator from *Eucalyptus* xylem, regulates secondary cell wall formation and lignin biosynthesis. *Plant J* 43: 553–567
- Gomez-Maldonaldo J., Avila C., de la Torre F., Canas R., Canovas F. M. and Campbell M. M. (2004). Functional interactions between a glutamine synthetase promoter and MYB proteins. *Plant J* 39: 513–526
- Grotewold E., Drummond B. J., Bowen B. And Peterson T. (1994). The mybhomologous P gene controls phlobaphene pigmentation in maize floral organs by directly activating a flavonoid biosynthetic gene subset. *Cell* 76: 543-553
- Guillet-Claude C., Isabel N., Pelgas B. and Bousquet J. (2004). The evolutionary implications of knox-I gene duplications in conifers: correlated evidence from phylogeny, gene mapping, and analysis of functional divergence. *Mol Biol Evol* 21: 2232-45
- Hauffe K.D., Lee S.P., Subramaniam R. and Douglas C.J. (1993). Combinatorial interactions between positive and negative *cis*-acting elements control spatial patterns of 4CL-1 expression in transgenic tobacco. *Plant J* 4: 235-253
- Higginson T., Li S.F. and Parish R.W.(2003). *AtMYB103* regulates tapetum andtrichome development in *Arabidopsisthaliana*. *Plant J* 35: 177-192
- Jiang C., Gu X. and Peterson T. (2004). Identification of conserved gene structures and carboxy-terminal motifs in the Myb gene family of Arabidopsis and Oryza sativa L. ssp. indica. *Genome Biol* **5**: R46

- Jin H. L., Cominelli E., Bailey P., Parr A., Mehrtens F., Jones J., Tonelli C., Weisshaar B. and Martin C. (2000). Transcriptional repression by AtMYB4 controls production of UV- protecting sunscreens in Arabidopsis. *EMBO J* 19: 6150–6161
- Jin H. and Martin C. (1999). Multifunctionality and diversity within the plant MYB-gene family. *Plant Mol Biol* 41: 577-585
- Kirst M., Johnson A.F., Baucom C., Ulrich E., Hubbard K., Staggs R., Paule C., Retzel E., Whetten R. and Sederoff R. (2003). Apparent homology of expressed genes from wood-forming tissues of loblolly pine (*Pinus taeda* L.) with Arabidopsis thaliana. *Proc Natl Acad Sci USA* 100: 7383-7388
- Kobayashi H. and Fukuda H. (1994). Involvment of calmodulin and calmodulin binding proteins in the differentiation of tracheray elements in *Zinnia* cells. *Planta* 194: 388-394
- Kranz H. D., Denekamp M., Greco R., Jin H., Leyva A., Meissner R. C., Petroni K., Urzainqui A., Bevan M., Martin C., Smeekens S., Tonelli C., Paz-Ares J. and Weisshaar B. (1998). Towards functional characterization of the members of the R2R3-MYB gene family from *Arabidopsis thaliana*. *Plant J* 16: 263–276
- Kumar S., Tamura K., Jakobsen I.B. and Nei M. (2001). MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* 17: 1244-1245
- Kusumi J., Tsumura Y., Yoshimaru H. and Tachida H. (2002). Molecular evolution of nuclear genes in cupressacea, a group of conifer trees. *Mol Biol Evol* 19: 736-747
- Lacombe E., Van Doorsselaere J., Boerjan W., Boudet A.M. and Grima-Pettenati J.(2000). Characterization of *cis*-elements required for vascular expression of *the Cinnamoyl CoA Reductase* gene and for protein-DNA complex formation. *Plant J* 23: 663-676
- Lauvergeat V., Rech P., Jauneau A., Guez C., Coutos-Thevenot P. and Grima-Pettenati J. (2002). The vascular expression pattern directed by the Eucalyptus gunnii cinnamyl alcohol dehydrogenase EgCAD2 promoter is conserved among woody and herbaceous plant species. *Plant Mol Biol* 50: 497-509
- Leyva A,. Liang X., Pintor-Toro J.A., Dixon R.A. and Lamb C.J. (1992). cis-Element combinations determine phenylalanine ammonia-lyase gene tissue-specific expression patterns. *Plant Cell* **4**:263-271
- Lipsick, J. S. (1996). One billion years of Myb. Oncogene 13: 223-235
- Mackay J., Bérubé H., Regan S. and Séguin A. (2004). Functional genomics in forest tress: Application to the investigation of defense mechanisms and wood formation. *Plantation Forest Biotechnology for the 21st century: Research Signpost*: 446p
- Neale D. B. and Savolainen O. (2004). Association genetics of complex traits in conifers. *Trends Plant Sci* 9: 325-30
- Newman L.J., Perazza D.E., Juda L. and Campbell M.M. (2004). Involvement of the R2R3-MYB, At MYB61, in the ectopic lignication and darkphotomorphogenic components of the det3 mutant phenotphype. *Plant J* 37: 239–250
- Noda K., Glover B.J., Linstead P. and Martin C. (1994). Flower colour intensity depends on specialized cell shape controlled by a Myb-related transcription factor. *Nature* 369: 661-664

- Ogata K., Hojo H., Aimoto S., Nakai T., Nakamura H., Sarai A., Ishii S. and Nishimura Y. (1992). Solution Structure of a DNA-Binding Unit of Myb: A Helix-Turn-Helix-Related Motif With Conserved Tryptophans Forming a Hydrophobic Core. *Proc Natl Acad Sci USA* **89**: 6428-6432
- Patzlaff A., McInnis S., Courtenay A., Surman C., Newman L. J., Smith C., Bevan M. W., Mansfield S. D., Whetten R. W., Sederoff R. R. and Campbell M. M. (2003a). Characterisation of a pine MYB that regulates lignification. *Plant J* 36: 743–754
- Patzlaff A., Newman L. J., Dubos C., Whetten R. W., Smith C., McInnis S., Bevan M. W., Sederoff R. R. and Campbell M. M. (2003b). Characterisation of PtMYB1, an R2R3-MYB from pine xylem. *Plant Mol Biol* 53: 597–608
- Pavy N., LarocheJ., Bousquet J. and MacKay J. (2005). Large-scale statistical analysis of secondary xylem ESTs in pine. *Plant Mol Biol* 57: 203-224
- Penfield S., Meissner R.C., Shoue D.A., Carpita N.C. and Bevan M.W. (2001). MYB61 is required for mucilage deposition and extrusion in the Arabidopsis seed coat. *Plant Cell* 13: 2777-2791
- Plomion C., Leprovost G. and Stokes A. (2001). Wood formation in trees. *Plant Physiol* **127**: 1513-1523
- Rabinowicz P.D., Braun E.L., Wolfe A.D., Bowen B. and Grotewold E. (1999). Maize R2R3 Myb Genes: Sequence Analysis Reveals Amplification in the Higher Plants. *Genetics* 153: 427-444
- Riechmann J.L., Heard J., Martin G., Reuber L., Jiang C., Keddie J., Adam L., Pineda O., Ratcliffe O.J., Samaha R.R., Creelman R., PilgrimM., Broun P., Zhang J.Z., Ghandehari D., Sherman B.K. and Yu G. (2000). Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. Science 290: 2105-2110
- Rogers L. A. and Campbell M. M. (2004). The genetic control of lignin deposition during plant growth and development. *New Phytol* 164: 17–30
- Romero I., Fuertes A., Benito M. J., Malpica J. M., Leyva A. and Paz-Ares J. (1998). More than 80R2R3 MYB regulatory genes in the genome of Arabidopsis thaliana. *Plant J* 14: 273-284
- Rozen S. and Skaletsky H. (2000). Primer3 on the WWW for general users and for biologist programmers. In: S. *Krawetz and S. Misener* (Eds.), *Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa*, N.J.: 365-386
- Scarpella E. and A. H. Meijer (2004). Pattern formation in the vascular system of monocot and dicot plant species. *New Phytol* 164: 209-242
- Schiefelbein J. (2003). Cell-fate specification in the epidermis: a common patterning mechanism in the root and shoot. *Curr Op Plant Biol* 6: 74-78
- Sharman B. C. (1943). Tannic acid and iron alum with safranin and orange G in studies of the shoot apex. *Stain Techn* 18: 105-111
- Stracke R., Werber M. and Weisshaar B. (2001). The R2R3-MYB gene family in Arabidopsis thaliana. Curr Opin Plant Biol 4: 447–456
- Tamagnone L., Merida A., Parr A., Mackay S., Culianez-Macia F.A., Roberts K. and Martin C. (1998). The AmMYB308 and AmMYB330 transcription factors from *Antirrhinum* regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. *Plant Cell* 10: 135–154
- Thompson J. D., Higgins D. G. and Gibson T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment

through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673-4680

- Timell T. E. (1986). Compression wood in gymnosperms. Springer, Berlin Heidelberg New York
- Vailleau F., Daniel X., Tronchet M., Montillet J.L., Triantaphylides C. and Roby D. (2002). A R2R3-MYBgene, AtMYB30, acts as a positiveregulator of the hypersensitive cell deathprogram in plants in response topathogen attack. *Proc Natl Acad SciUSA* 99: 10179-10184
- Xue B., Charest P.J., Devantier Y. and Rutledge R.G. (2003). Characterization of a *MYBR2R3* gene from black spruce (*Picea mariana*) that shares functional conservation with maize C1. *Mol Genet Genomics* 270: 78-86
- Yoo J.H., Park C.Y., Kim J.C., Do Heo W., Cheong M.S., Park H. C., Kim M.C., Moon B.C., Choi M.S., Kang Y.H., Lee J.H., Kim H.S., Lee S.M., Yoon H.W., Lim C.O., Yun D-J., Lee S.Y., Chung W.S. and Cho M.J. (2005). Direct Interaction of a Divergent CaM Isoform and the Transcription Factor, MYB2, Enhances Salt Tolerance in Arabidopsis. J Biol Chem 280: 3697-3706
- Zimmermann I.M., Heim M.A., Weisshaar B. and Uhrig J.F. (2004). Comprehensive identification of Arabidopsis thaliana MYB transcription factors interacting with R/B-like BHLH proteins. *Plant J* 40: 22-34

CHAPITRE III

Distinct and overlapping effects of *Pinus taeda* MYB1 and MYB8 overexpression on lignification and related metabolism in conifers

Claude Bomal¹, Frank Bedon^{1,2}, Sébastien Caron¹, Shawn Mansfield⁴, Caroline Levasseur³, Janice E. K. Cooke⁶, Sylvie Blais¹, Laurence Tremblay⁵, Marie-Josée Morency,³ Nathalie Pavy¹, Jacqueline Grima-Petenatti², Armand Séguin³ and John MacKay¹

¹Centre d'Étude de la Forêt, Université Laval, Québec (QC), G1K 7P4, Canada ²UMR UPS,CNRS 5546, Pôle de Biotechnologies Végétales, 24 chemin de Borde Rouge, BP42617, Auzeville Tolosane, 31326 Castanet Tolosan, France ³Natural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre, Québec (QC), G1V 4C7, Canada ⁴Canada Research Chair in Wood and Fibre Quality, Department of Wood

Science, University of British Columbia, 4030-2424 Main Mall, Vancouver (BC), V6T 1Z4, Canada

⁵DPSP, Ministère des Ressources naturelles et de la Faune, 2700, rue Einstein, Québec (QC), G1P 3W8, Canada

⁶Department of Biological Sciences, CW405 Biological Sciences Building, University of Alberta, Edmonton (AB), T6G 2E9, Canada.

III-1-Avant-propos

Ce chapitre III a été soumis pour publication à la revue "*Plant Physiology*". Il est présentement en évaluation. En tant que deuxième auteur, mon rôle à été d'analyser l'expression des gènes MYB1 et MYB8 chez le pin (Figure III-1), et de participer au design et à la réalisation expérimentale de l'étude des épinettes trangéniques avec Dr. C. Bomal. Plus précisément, j'ai participé à toutes les collectes d'échantillons, j'ai fait les extractions des ARN utilisés pour les biopuces à ADN, et participé à l'obtention des résultats de PCR quantitative (Figures III-5 et III-6) *via* l'identification des gènes et la conception des amorces. J'ai aussi participé à l'interprétation des résultats. Dr. C. Bomal a effectué les analyses de PCR quantitative sur les transgéniques, mis en forme les résultats microarrays. J'ai collaboré à la rédaction du manuscrit réalisée principalement par Dr. C. Bomal et Dr. J. MacKay. Dr. J. Cooke a assuré le développement des biopuces et l'optimisation des méthodes d'analyses. S. Caron a effectué les hybridations microarrays, et leur analyse statistique avec Dr. N. Pavy. S. Blais et M.-J Morency ont préparé les constructions moléculaires pour la surexpression MYB. Dr. S. Mansfield a dosé les lignines des transgéniques. La transgénèse s'est effectué au Centre de Foresterie des Laurentides à Québec par C. Levasseur et L. Tremblay, au sein de l'équipe du Dr. A. Séguin.

III-2-Résumé

Nous avons examiné l'implication de deux MYB-R2R3 issus de Pinus taeda L., PtMYB1 et PtMYB8, dans le métabolisme des phénylpropanoïdes et la lignification. Des expériences antérieures ont montré que PtMYB1 était capable de fixer des éléments AC du promoteur PAL et d'en activer la transcription (Patzlaff et al., 2003), alors que PtMYB8 n'a pas été fonctionnellement caractérisé. Nous avons surexprimé ces deux MYB de pin dans un système conifère hétérologue par transformation de Picea glauca (Moench) Voss. Après germination des embryons somatiques, la croissance racinaire était réduite dans les transgéniques surexprimant chacun des gènes; cependant, PtMYB8 exerce un effet plus fort et plus consistent, empêchant complètement le développement de la racine dans plusieurs lignées. La lignification ectopique était un caractère commun aux deux transgéniques mais avec une distribution très distincte. Les analyses par biopuces à ADN ont montré que la surexpression des MYB conduisait à la dérégulation de l'expression des gènes associés à la paroi cellulaire, augmentant clairement l'expression des gènes liés à la synthèse des monomères de lignine, mais conduisait aussi à la diminution de l'expression des gènes des métabolismes des flavonoïdes et des terpènes. Toutefois, chaque MYB confère des impacts distincts sur les profils d'expression de ces gènes. La nature et la force des phénotypes observés sur les transgéniques fournissent des évidences que les deux gènes MYB jouent des rôles distincts mais chevauchants en lien avec la lignification. Nos résultats suggèrent aussi que le contrôle des gènes liés aux monolignols et à la paroi cellulaire peut impliquer un réseau d'interactions ou une hiérarchie régulatrice parmi les gènes MYB.

III-3-Abstract

We investigated the implication of two R2R3-MYB genes from Pinus taeda L., PtMYB1 and PtMYB8, in phenylpropanoid metabolism and lignification. PtMYB1 was previously shown to specifically bind AC elements and activate transcription from PAL promoter (Patzlaff et al., 2003), whereas PtMYB8 had not been functionally characterized. We overexpressed these pine MYBs in a heterologous conifer system by transformation of Picea glauca (Moench) Voss. Upon germination of somatic embryos, root growth was decreased in transgenics overexpressing either one of the genes; however, PtMYB8 exerted a stronger and more consistent effect, completely preventing root development in some lines. Ectopic lignification was a common feature of both transgenics but followed very distinct distributions. Microarray analysis showed that overexpression of MYBs led to a misregulation of cell wall genes, clearly upregulating genes related to lignin monomer synthesis, but also led to the downregulation of flavonoid and terpenoid metabolism genes; however, each MYB conferred distinct impacts on these gene expression profiles. The nature and strength of the phenotypes observed upon overexpression provide evidence that the two MYB genes play distinct but overlapping roles with regard to lignification. Our results also suggest that the control of monolignol-, and cell wall-related genes could involve networking or regulatory hierarchy among the MYB genes.

III-4-Introduction

The plant cell wall is a composite assembly of complex polymers that confer structure and provide protection to the cell and to the whole plant. Plant cell growth and differentiation depend on the proportion of the different cell wall polymers, which are temporally and spatially accumulated in response to developmental and environmental cues. The major polymers of the primary cell wall include cellulose, hemicelluloses, and pectins (Cosgrove, 1999). The secondary cell wall that is formed mostly in vascular cells and fibres also contains a large proportion of lignin. In trees, lignin synthesis takes on a greater importance because of the production of large amounts of wood which is typically made up of 20 to 30% lignin on a dry weight basis. Therefore, the formation of fixed carbon resources into the synthesis of lignin building blocks through the phenylpropanoid pathway (Amthor, 2003; Boerjan *et al.*, 2003).

The phenylpropanoid biosynthesis pathway provides precursors for the biosynthesis of natural products such as flavonoids, hydroxycinnamate conjugates, lignins and lignans (Humphreys and Chapple, 2002). Lignin is built up from p-hydroxycinnamyl alcohol precursors, also called monolignols. The resulting H, G, S lignin units vary in proportion according to plant species, tissue type and response to environmental stress among others (Campbell and Sederoff, 1996). Although the enzymes and genes of phenylpropanoid and monolignol biosynthetic pathways have been extensively studied (Humphreys and Chapple, 2002; Boerjan *et al.*, 2003), the mechanisms governing their regulation are still a matter of debate. In addition, molecular mechanisms regulating metabolic flux through the phenylpropanoid biosynthetic pathway are very complex and, to date, remain unclear.

Recent experimental evidence indicates that monolignol synthesis genes are under transcriptional control (Rogers and Campbell, 2004). Among the different classes of transcription factors (TF) directly or indirectly implicated in lignification, the R2R3-*MYB*s are strong candidates for regulation of phenylpropanoid enzymes and monolignol biosynthesis (Rogers and Campbell, 2004; Groover and Robischon, 2006). The R2R3-*MYB*s form one of the largest families of TFs in plants. Numerous studies characterizing their molecular function have assigned diverse roles to these transcription factors. In some cases, they appear to control cell fate or identity in different plant organs (Rogers and Campbell, 2004; Paux *et al.*, 2005; Chen *et al.*, 2006; Groover and Robischon, 2006). They also govern many aspects of secondary metabolism in *Arabidopsis* (Vom Endt *et al.*, 2002) as well as in angiosperm and gymnosperm trees (Rogers and Campbell, 2004; Groover and Robischon, 2006).

The implication of R2R3-MYBs in the transcriptional control of phenylpropanoid and flavonoid metabolic pathways has been well documented in plant model systems. The promoters of most genes in these pathways contain AC elements (Rogers and Campbell, 2004) to which different R2R3-MYBs may bind either alone (Patzlaff et al., 2003a, 2003b; Gomez-Maldonado et al., 2004; Goicoechea et al., 2005), or in heteroduplex with bHLH proteins (Goff et al., 1992) or with WD-40 proteins (Koes et al., 2005; Morita et al., 2006). Activation or repression of transcription has been attributed to R2R3-MYBs belonging to different subgroups of this very large plant-specific branch of the MYB gene family (Jin and Martin, 1999; Vom Endt et al., 2002). While the DNA binding domains of MYBs are highly conserved, MYB regulatory domains (activation or repression) are more variable and have only been characterized for a small subset (Kranz et al., 1998; Stracke et al., 2001; Chen et al., 2006; Li et al., 2006;). In woody species, only a few functional studies have addressed this question for MYBs. For example, the *PttMYB21a* gene of poplar was proposed to be a negative regulator of CCoAOMT expressed in vascular tissues (Karpinska et al., 2004). On the other hand, the genes PtMYB1 and PtMYB4 from loblolly pine (Patzlaff et al., 2003a, 2003b; Gomez-Maldonado et al., 2004), and EgMYB2 from eucalyptus (Goicoechea et al., 2005) appear to be transcriptional activators of lignin synthesis. These three MYBs are preferentially expressed in developing xylem tissues, bind AC elements and activate transcription from lignin biosynthetic gene promoters in transient assays in yeast or plant cells (Patzlaff et al., 2003a, 2003b; Gomez-Maldonado et al., 2004; Goicoechea et al., 2005). Overexpression of pine MYBs resulted in ectopic lignification in tobacco (Patzlaff et al., 2003b) and in Arabidopsis (Newman et al., 2004) while overexpression of EgMYB2 in tobacco

lead to altered lignin structure, to thicker secondary cell walls and to the upregulation of lignin-related genes (Goicoechea *et al.*, 2005). Although convergent data from these studies and others support a role for *MYB* in the lignifying process in trees, knowledge is still limited about the number and the respective implication of *MYB* TFs in wood formation.

Recently, Bedon et al. (2007) characterized 18 conifer R2R3 MYB gene sequences, including 13 genes from spruce (Picea glauca [Moench] Voss) and 5 genes from loblolly pine (Pinus taeda L.). To gain new insights into the role of MYBs in regulating lignification in gymnosperm trees, we focussed our work on characterizing two of these Pinus taeda R2R3 MYB genes, namely PtMYB1 and *PtMYB8*, through gain-of-function experiments. As mentioned above, PtMYB1 is a strong activator of phenylalanine ammonia-lyase (PAL) and is able to bind AC elements (Patzlaff et al., 2003a). Whereas PtMYB1 and its close spruce homolog, *PgMYB1*, are expressed in several tissues including primary and secondary xylem, a close homolog of *PtMYB8* from spruce has strong preferentially expression in secondary xylem (Bedon et al., 2007). Overexpression experiments were carried out in spruce because it is a member of the Pinaceae, and thus, provides an expression system that is much closer to pine than Arabidopsis or tobacco which have been employed in previous reports. The overexpression strategy was preferred since it has been particularly effective in revealing TF function and offers an alternative strategy that is less affected by TF functional redundancy than knockout/knockdown analysis (Zhang, 2003). This functional redundancy problem can be especially acute for TFs from large gene families such as MYBs that often include closely related genes (Schwechheimer et al., 1998; Zhang, 2003). We produced transgenic spruce plantlets overexpressing these MYBs by Agrobacterium-mediated transformation of embryogenic spruce tissue, from which mature somatic embryos were recovered and germinated. The nature of the ectopic lignification phenotypes that were observed, together with the suites of differentially expressed genes that were revealed by microarray analyses provide evidence that PtMYB1 and PtMYB8 play distinct but perhaps overlapping roles in regulating lignification or developmental processes that lead to lignification. Our results also point to potential networking or regulatory hierarchy among MYB genes for transcriptional regulation of genes important in xylogenesis.

III-5-Materials and Methods

III-5-1-Pine tissues collection for *PtMYB1* and *PtMYB8* expression profile

Several tissues were collected from two-year old *Pinus taeda* L. plants to survey preferential tissue expression of *PtMYB1* and *PtMYB8*. Selected tissues, obtained from three biological replicates, were Needle (N), Primary Shoot (PS), phloem and periderm (Bark), Stem Secondary Xylem (X), Root Tip (RT) and Whole fine Root (WR). All tissue were frozen in liquid nitrogen immediately upon removal and stored at -80°C until required for RNA extraction.

III-5-2-Vector construction and spruce transformation

The *PtMYB1* (AY356372, Patzlaff *et al.*, 2003a) and *PtMYB8* (DQ399057, Bedon *et al.*, 2007) cDNA from *Pinus taeda* L. (Loblolly pine) were used to conduct gain-of-function experiments in *Picea glauca* [Moench] Voss (white spruce). Gain of function was obtained by inserting the full length cDNA in front of the maize ubiquitin promoter (Christensen *et al.*, 1992), followed at the 3' end by the 35S terminator. This constitutive expression vector, named pMJM, was obtained by modifying the pRT106 (Topfer *et al.*, 1993) plant expression vector. For *Agrobacterium tumefaciens*-mediated stable transformation, cassettes were digested by *Hin*dIII and cloned into pCAMBIA1305.2 (www.cambia.org). The resulting plasmid was then transferred into the *A. tumefaciens* strain C58 pMP90 (Koncz and Schell, 1986).

The white spruce embryogenic line Pg653 was used in the present study, and was initiated and maintained as described by Klimaszewska *et al.* (2001). Genetic transformation was carried out also as described in Klimaszewska *et al.* (2004). Once co-cultivated, explants were decontaminated from *A. tumefaciens* with cefotaxim and transferred onto fresh medium containing cefotaxim alone. Kanamycin resistant embryonal tissues were screened according to positive X-gluc tests (Klimaszewska *et al.*, 2004) and assayed for *PtMYB1* and *PtMYB8* mRNAs accumulation by QPCR (see below). Lines exhibiting a range of *PtMYB* mRNAs levels were selected for somatic embryo maturation and somatic seedling production.

III-5-3-Transgenic lines and culture conditions used for somatic seedling production

For *in vitro* monitoring, somatic embryos produced from selected transgenic lines overexpressing PtMYB1 (lines 4, -12, -14, -24, and -26) and PtMYB8 (lines 1 to -13) as well as wild-type and pCAMBIA control lines were germinated according to Klimaszewska *et al.* (2004). When allowed by the phenotype, ten to fourteen week old transgenic plantlets were transplanted in a mix of moss, vermiculite and turface (ratio 4:2:1, v/v/v) and grown in a mist environment for 15 days before being transferred in the greenhouse under a 16h day, 8h night photoperiod at 24°C day / 20°C night.

For the microarray experiments, somatic embryos were germinated for 3 weeks on MLVG medium supplemented with 58 mM sucrose (Klimaszewska *et al.*, 2004). Two transgenic lines per transgene (lines 4 and 14 for PtMYB1, and lines 1 and 2 for PtMYB8) as well as four control lines (Pg653) were used in a randomized complete block design, with four biological replicates per line. Each experimental unit, comprising 50 whole germinants, was divided into two 25-germinant subsamples at the time of collection.

III-5-4-Seedling growth quantification and tissue sampling

Image capturing was used for seedling growth quantification. The length of hypocotyl and root was determined by image analysis using ImageJ 1.32 Software (W. Rasband, National Institute of Health, USA, <u>http://rsb.info.nih.gov/ij/</u>). For each sampling date and when required, roots and hypocotyls were rapidly separated on the medium by using scalpel and were immediately fixed as described below when used for histology, or frozen into liquid nitrogen and stored at -80°C for further molecular analysis.

III-5-5-Histology

Tissue samples were fixed for 24 hours under low vacuum in 2% (v/v) paraformaldehyde, 3% (v/v) glutaraldehyde, 0.1 M cacodylate buffer (pH 7.2) supplemented with 1 mM CaCl₂, and 1% (w/v) sucrose. For the preparation of thin sections (5 μ m), samples were dehydrated in graded ethanol series and toluene

before infiltration with paraffin for 7 days. Paraffin-free sections, prepared by using a microtome were transferred to water and, after incubation in the mordant 2% (w/v) ZnCl₂, were stained in Sharman's safranin O-orange G-tannic acid (Sharman, 1943). All sections were observed under a Zeiss Axioskop microscope (Jena, Germany) fitted with a digital camera. Lignin gold staining was obtained after Sharman's staining under UV excitation (Ex 365 nm, Em 420 nm, FT 395 nm) coupled to a Zeiss 09 filter cube (450–490, FT510, LP520).

III-5-6-RNA extraction, cDNA synthesis and real-time RT-PCR analysis

RNA extraction, cDNA synthesis and real-time RT-PCR (QPCR) analysis were essentially performed as described in Bedon et al. (2007) with minor changes. Briefly, total RNA samples were obtained by using a Mixer Mill MM300 engine (Retsch, Germany) to grind tissue (from spruce and pine) in 1.5 mL Ependorf microtubes, and were extracted following the procedure of Chang et al. (1993). Total RNA samples were treated with amplification grade DNAse I (Invitrogen) and purified using RNeasy columns (Qiagen). cDNAs were synthesized from one microgram of purified RNA using SuperScript II (Invitrogen, Carlsbad CA) according to the manufacturer's instructions. For QPCR quantification of each target RNA, a five-fold dilution of cDNA mixture was used as template. Thermocycling, conducted using an Opticon2 DNA Engine (MJ Research Inc.), was initiated by a 2 min incubation at 95°C followed by 35 cycles (95°C for 10 s; 55°C for 40 s; 72°C for 6 s) with a single fluorescent reading taken at the end of each cycle. A standard dilution series covering five orders of magnitude were prepared for the target and reference genes from 1 ng μ L⁻¹ PCR amplicon of each cDNA to produce solutions covering 10⁻¹ to 10⁻⁶ ng μ L⁻¹. All mRNA levels were calculated from threshold cycle values and as relative to controls and normalized with respect to the transcript level of ELONGATION FACTOR 1-alpha (EF1-alpha, 100% homolog to Picea abies AJ132534). The specific primer pairs of target and reference genes, designed using Primer3 software (Rozen and Skaletsky, 2000), are given in Appendix III-1.

III-5-7-Microarray experiment

All the information regarding microarray manufacture and quality control are detailed in Appendix III-2. The microarray experiment was designed in total compliance with MIAME guidelines. Briefly, two transgenic lines per transgene (lines 1 and 4 for PtMYB1, and lines 1 and 2 for PtMYB8) as well as four control lines (Pg653) were used to produce 3-week old somatic seedlings as described above. Total RNA was extracted from four biological replicates per line as described above. For each sample, 1.0 µg of total RNA was used for indirect RNA amplification that was performed with SuperscriptTM Indirect RNA Amplification System kit (Invitrogen, Carlsbad CA) according to manufacturer's instructions. Dye coupling was performed on 5 µg of amplified RNA (aRNA) with either Alexa Fluor® 555 or Alexa Fluor® 647 reactive dyes (Invitrogen, Carlsbad CA).

Slides were pre-hybridized at 65°C for 2 h in 5x SSC, 0.1% (w/v) SDS, 0.2 mg mL⁻¹ BSA, 0.1 mg mL⁻¹ herring sperm DNA solution. They were subsequently washed twice in 0.1x SSC, and water for 30 s each before boiling into water for 3 min. Slides were then dried by centrifugation for 2 min at 480 g. Prior to hybridization, both samples of labeled aRNA to be compared were combined, concentrated, heated at 95°C for 2 min and added to 52 µL hybridization solution (50% formamide, 5x SSC, 0.1% (w/v) SDS, 0.1 mg mL⁻¹ herring sperm DNA) pre-warmed to 45°C. Slides were then hybridized overnight at 45°C and finally washed once in 2x SSC plus 0.5% (w/v) SDS for 15 min at 45°C, twice in 0.5x SSC plus 0.1% (w/v) SDS for 15 min at 45°C, twice in 0.1x SSC for 1 min at room temperature and twice in water for 1 min at room temperature. Slides were dried by centrifugation at 480 g for 2 min and kept in a slide box in the dark until ready to scan. Complete pre-hybridization and hybridization protocols are available at http://atarrays.tigr.org/PDF/Probehyb.pdf. Eight hybridizations (including four dye swaps) were carried out for each of the PtMYB1 and PtMYB8 versus wild type comparisons, and 16 hybridizations (eight dve swaps) for PtMYB1 versus PtMYB8 comparisons. Array images were scanned using ScanArray Express (PerkinElmer) and image analysis was performed with Quantarray software (Packard BioSciences, version 3.0.0.0, 2001).

III-5-8-Statistical analysis of microarray data

Data were analyzed in R (Ihaka and Gentleman, 1996), mainly using the BioConductor suite of packages (Gentleman *et al.*, 2004). Spot intensities were analyzed with the LIMMA package from Bioconductor (Smyth, 2005). Data normalization was performed using the composite method based on Lowess curves (Yang *et al.*, 2002). Normalized data were then statistically analyzed using the linear model and empirical Bayes analyses in LIMMA. The results were corrected using the Benjamini and Hochberg (1995) method of false discovery rate (FDR, 1%). In this study, we focused on differentially expressed genes that gave a P value < 0.01 (from the LIMMA), and met log₂ ratio threshold of 0.8 (1.75-fold change) between wild type and PtMYB transgenics. For each transgenic construct, data from the two different lines were analyzed separately and only the differentially-expressed gene identified in both lines for a given transgene were considered to reduce the impact of positional effect of transgene.

III-5-9-Soluble phenolic metabolite analysis

Spruce hypocotyl tissue (~20 mg) was suspended in 1.5 mL of methanol:water:HCl (48.5:48.5:1), pulverized in a cell disrupter (FastPrep) at high power for 20 s, and then extracted for 4 h at 50°C in heating block. The homogenate was centrifuged for 10 min at 13,000 rpm, and the supernatant equally divided into two 600-µL aliquots. One aliquot was saponified with 1 M NaOH for 18 h at room temperature to facilitate the release of esterified phenolics, and then acidified with 120 µL of glacial acetic acid. Subsequently, one milliliter of distilled water was added to both methanolic and saponified methanolic aliquots, followed by an equal volume of ethyl acetate. After mixing thoroughly, the ether phase was retained. This extraction was performed twice and the ether phases were pooled. The ether phase was concentrated to dryness in a speedvac, resuspended in 100 µL methanol before injection in a Summit HPLC system (Dionex) fitted with a 0.2×150 mm Pursuit column (Waters, 5 µm particle size), an autosample and a photodiode array detector. The elution at a rate of 0.2 mL min⁻¹ was performed at 45°C using a linear gradient from 100% eluant A (CH₃COOH, 5%) to 80% eluant B (CH₃COOH (20%): CH₃CN, 75:25) over 60 min, followed by 10 min wash with 100% B, and 10 min reacclimatizing with 100% A. Compounds were identified based on their retention time and UV spectra as compared to their respective standards. Peak identity was confirmed by LC-MS (data not shown).

III-5-10-Acetyl bromide lignin determination

Oven dried 2-year old spruce stems (Pg653) were ground in a Wiley mill to pass through a 40-mesh screen, and then soxhlet-extracted with acetone for 24 h. Holocellulose was purified from the extractive free wood by reacting 200 mg of ground wood with 1 mL of NaClO₂ solution (400 mg 80% sodium chlorite, 4 mL distilled water, 0.4 mL glacial acetic acid) in a 25 mL round bottom flask maintained at 90°C in an oil bath. An additional milliliter of NaClO₂ solution was added every half hour and the samples removed to a cold water bath after 2 h. Samples were then filtered through a coarse crucible and dried overnight. Holocellulose composition was determined gravimetrically and characterized for residual lignin content by Klason lignin determination. For lignin isolation, ground, extract-free 2-year old spruce wood was ball milled as per Stewart *et al.* (2006) and lignin samples were then recombined in varying ratios, ranging from 100:0 to 0:100 and validated by Klason analysis, to produce a standard curve for lignin content determination of unknown samples via the acetyl bromide technique.

Seven-week old hypocotyls from transgenic and wild-type spruce were preextracted overnight in hot acetone to remove free phenolics prior to lignin determination using the acetyl bromide method. Acetone-extracted samples as well as the standard curve mixtures (~5 mg) were accurately weighed into Tefloncapped glass vials, in which 0.5 mL of acetyl bromide solution was added (25% (v/v) acetyl bromide in glacial acetic acid). Samples were incubated in a heating block at 50°C for 2 h with occasional mixing. After two hours, the solution was transferred to a 10 mL volumetric flask with ~5 mL of glacial acetic acid and 2 mL of 2 M NaOH. Then, 0.35 mL of 0.5 M hydroxylamine was added. The volume was made up to exactly 10 mL with glacial acetic acid and the solution thoroughly mixed. The hypocotyls and standard curve solution were then read at 280 nm and lignin content interpreted.

III-6-Results

III-6-1-Transcript abundance of *PtMYB1* and *PtMYB8* in various tissues of *Pinus taeda* trees

Transcript abundance of *PtMYB1* and *PtMYB8* was surveyed in several organs or tissues (needle, primary shoot, bark, stem secondary xylem, root tip, and whole root) from 2-year-old *Pinus taeda* trees by using real time RT-PCR (QPCR) analysis. *PtMYB8* transcripts appeared to be strongly preferential to differentiating secondary xylem in stem (X), and to a lesser extend in whole roots (WR) and in primary shoots (PS) (Figure III-1). *PtMYB1* transcripts were also preferential to stem differentiating secondary xylem (X) with lowest expression level in primary stem and whole roots, but were also abundant in needles. Interestingly, the expression profiles of these two *MYB* genes were similar to those reported for their closest homologs from spruce, *PgMYB1* and *PgMYB8* (Bedon *et al.*, 2007).



Figure III-1. Expression pattern of *PtMYB1* and *PtMYB8* in *Pinus taeda*.

Total RNA was isolated from various tissue of two year-old *Pinus taeda* tree. Transcript levels were determined by real time RT-PCR from three biological replicates and were normalized relative to *EF1-alpha* expression level. Error bars represent standard deviation. N, needle; PS, primary shoot; Bark, phloem and periderm; X, stem differentiating xylem; RT, root tip; WR, whole roots.

III-6-2-PtMYB1- and PtMYB8-OE in spruce leads to reduced root growth and to ectopic lignification

To gain insight into the role of *PtMYB1* and *PtMYB8* as potential regulators of lignification, we generated transgenic spruces overexpressing (OE) these *MYB* genes under the control of maize ubiquitin (*UB1*) promoter. We used an efficient *Agrobacterium*-mediated protocol for transformation of somatic embryogenic spruce cells (Klimaszewska *et al.*, 2004), and obtained several GUS positive transgenic lines, which were used to produce mature somatic embryos. We selected five PtMYB1-OE lines and thirteen PtMYB8-OE lines with strong transgene expression (determined by QPCR analysis of transgene RNA levels, not shown) for *in vitro* growth monitoring and phenotypic analyses. Upon germination of the somatic embryos, both transgenes conferred morphological phenotypes within the first weeks of development, and clear histological differences were observed when the transgenics were compared to control, which included the untransformed (WT) and empty vector transformed (pCAMBIA) lines (Figure III-2 and III-3).

The PtMYB1 transgenic spruce plantlets displayed reduced root growth (Figure III-2 B) compared to controls (Figure III-2 A). The roots were 30 to 80% shorter (mean decrease of 40%) depending on the transgenic line compared to WT and pCAMBIA controls (Figure III-2 C). PtMYB1-OE plantlets survived subsequent transfer to soil but their growth was moderately to severely delayed, particularly for line 4 (Figure III-2 D). Histological observations of PtMYB1 transgenic plantlets provided evidence of ectopic lignification in a few cells displaying the pink color that is diagnostic of lignified cell walls following Sharman's staining (Sharman, 1943). Significant cell wall thickening characteristic of sclerenchyma-like elements was observed among the parenchyma cells, located close to the shoot apical meristem (Figure III-2 H) and peripheral to vascular tissues (Figure III-2 J). This phenomenon was never observed in control plantlets (Figure III-2, G and I). The hypocotyls of PtMYB1 transgenic plantlets also had altered vascular radial patterning as the phloem zone was expanded and contained longer files of phloem cells (Figure III-2 F) compared to pCAMBIA controls (Figure III-2 E).



Figure III-2. Phenotypes induced by PtMYB1 overexpression in spruce.

A to C, In vitro phenotype of 7-week old plantlets in wild type(A) and PtMYB1 transgenic (B) spruce. A *reduced-root-growth* phenotype was observed in plantlets produced from 5 independent transgenic lines (C). D, Plantlet morphology in wild type (WT) and two independent lines (L4 and L14) of PtMYB1 transgenic spruce after transfer to soil. E to J, 5-µm section in paraffin embedded hypocotyl of wild type (E, G, I) and PtMYB1 (F, H, J) in vitro plantlets. Sections were stained in safranin O-orange G-tannic acid after mordanting in 2% ZnCl₂ (Sharman, 1943). (E, F) Cross sections in hypocotyl of 7-week old plantlets. Supernumerary phloem cells can be observed in PtMYB1 section (F). Longitudinal (G, H) and cross (I, J) section in hypocotyl of 16-day old in vitro plantlets. The presence of sclerenchyma-like elements around the vascular cylindar was revealed in PtMYB1 transgenic (arrows) by the pink staining specific to lignified cell wall (Sharman, 1943). Bars correspond to 5 mm in (A) and (B), 20 µm in (E) and (F), 50 µm in (G) and (H), and 25 µm in (I) and (J).

The root growth of PtMYB8 transgenic plantlets was even more strongly affected than PtMYB1 transgenics (Figure III-3, A and B), with a decrease in length ranging from 55 to more than 80% reduction, and a mean of 68% (Figure III-3 C). The two PtMYB8 transgenic lines that displayed the strongest root phenotype also developed shorter hypocotyls (Figure III-3 C, lines 1 and 2). In vitro development of PtMYB8-OE plantlets was rapidly arrested; plantlets showed clear signs of senescence and none of them survived subsequent transfer to soil. Histological analyses using Sharman's staining coupled with UV light microscopy provided evidence of extensive ectopic cell wall thickening and altered lignin deposition in PtMYB8 transgenics. While pink and gold staining was limited to xylem tissue in control plants (Figure III-3, D, H, and L), it extended to the pith, parenchyma and phloem cells in the hypocotyls in PtMYB8 transgenics (Figure III-3, E and G). The presence of sclerenchyma-like elements was also observed around the vascular tissue in the hypocotyls (Figure III-3 E) as well as in the cotyledons (Figure III-3 M). Thickened cell walls were also observed ectopically in the peripheral cell layers and in the cortical area of PtMYB8-OE roots (Figure III-3, I and K). Compared to controls (Figure III-3 D), the PtMYB8 transgenics had fewer tracheids (Figure III-3 E) indicating that development of xylem tissue was also perturbed.

III-6-3-Lignin and phenolics in PtMYB8 overexpressing spruce

Due to the extensive staining suggestive of ectopic lignification observed in PtMYB8 transgenic spruces, we aimed to gather further information on their lignin content and phenolic compound profiles. The samples were first preextracted overnight in hot acetone to remove free phenolics prior analyses. The lignin determinations were carried out with the acetyl bromide method and indicated that PtMYB8 overexpression led to a drastic increase in total lignin content in the acetone-extracted cell wall residues of spruce plantlets (Figure III-4 A). Compared to the controls, the total cell wall lignin content was increased by 66.3% and 54.2% in transgenic lines 1 and 2, respectively. We also examined the amount and nature of low molecular weight phenolic compounds that could be recovered by methanol extraction from transgenic and wild type plantlet tissues. Visual inspection of the HPLC chromatograms shows that all the phenolic compounds detected in the methanol extracts of PtMYB8-OE were significantly decreased, but the profiles appear quite similar to the controls (Figure III-4 B). These observations would indicate that there was an overall decrease in the amount of major soluble phenolic compounds, and no clear selective decrease in any particular metabolite.





A to C, In vitro phenotype of 7-week old plantlets in wild type (A) and PtMYB8 transgenic (B) spruce. (C) A *reduced-root-growth* phenotype was observed in plantlets produced from 13 independent transgenic lines but no plantlet survived transfer to soil. D to M, 5-µm sections in paraffin embedded hypocotyl (D to G), root (H to K), and cotyledons (L, M) in wild type (D, F, H, J, L) and PtMYB8 (E, G, I, K, M) in vitro plantlets. Sections were stained in safranin O-orange G-tannic acid after mordanting in 2% ZnCl₂ (Sharman, 1943) and observed under UV for gold specific lignin staining (F, G, I, K). Ectopic lignification was revealed in PtMYB8 by staining that was specific to lignified cell wall (pink and gold). The presence of sclerenchyma-like elements around the vascular cylindar was revealed in PtMYB8 transgenic (arrows) in hypocotyl (E) and cotyledons (M). Positive lignified cell wall staining was also observed in parenchyma (E) and cortical (J) cells. Bars correspond to 5 mm (A, B), 50 µm (D to K), and 10 µm (L, M).



Figure III-4. Lignin content and free phenolic profiles by induced PtMYB8 overexpression in spruce.

A, Lignin determination in wild-type (WT) and transgenic spruce overexpressing PtMYB8 (lines 1, 2, and 8). Lignin content was determined with the AcBr method (due to the small sample size), expressed as percent of dry weight, acetone extracted whole hypocotyls of 7-week old plantlets, and was calculated from 3 biological replicates per line and 10 plantlets per replicate. Error bars represent standard deviation. Numbers above bars indicate the percent increase in lignin content compared to control. B, HPLC profile of saponified phenols in wild-type (WT) and PtMYB8 overexpressing spruce (lines 1, 2, and 8). Peaks were identified by retention time matches with phenol standards and confirmed by LC/MS (data not shown) as follows: peak 1, 3,4 dihydroxy-cinnamic acid (caffeic acid); peaks 2 and 3, *cis* and *trans* 4-hydroxy, 3-methoxy cinnamic acid (ferulic acid); peak 6, flavonol (absorption maxima at 250 and 350); and peak 7, 4-methoxycinnamic acid.

III-6-4-PtMYB1- and PtMYB8-OE have distinct but also overlapping impact on spruce transcriptome

As described above, some redundancy was observed in the phenotypes associated with PtMYB1 and PtMYB8 overexpression, such as reduced root growth and ectopic lignification. However, the phenotypes conferred by the two transgenes are clearly distinct, implying that the functions of these two genes may not be completely redundant. Microarray transcript profiling was carried out in order to further characterize similarities and differences in the function of the two MYB genes. We analysed four biological replicates in two independent lines for each of PtMYB1-OE and PtMYB8-OE with custom cDNA microarrays comprised of 9,000 non redundant spruce amplicons, representing most of the cDNA described in Pavy et al. (2005). Each transgenic line was analysed separately, but we only considered those genes that had differential transcript levels in both lines for each MYB transgenic, so as to reduce the potential influence of position effects of the transgene. A large number of genes on the array showed statistically significant differences between the transgenics and the controls; namely 640 genes were mis-regulated for PtMYB1 and 1919 for PtMYB8 (P value < 0.01; FDR, 1%). Therefore, we focussed on statistically significant genes that also gave a 1.75-fold difference or greater (Log₂ ratio of 0.8) between the transgenic and controls. In the end, the number of misregulated genes that we considered decreased to 46 and 217 for MYB1 and MYB8, respectively. Among the differentially expressed genes, a large proportion of them had a predicted function that could be linked to cell wall biogenesis, secondary metabolism (phenylpropanoid, flavonoid, and terpenoid) and related pathways (shikimate and S-adenosyl methionine) as well as stress and detoxification (Table III-1). Many genes with unknown function were also misregulated in PtMYB1 (19%) and PtMYB8 (31%) (complete gene list not shown). Distinct and overlapping responses in both transgenics are described below.

The overexpression of PtMYB1 had a moderate impact on the spruce transcriptome with a total of 34 upregulated and 12 downregulated genes, based on the criteria outlined above. Several of the upregulated genes were related to the phenylpropanoid pathway such as *phenylalanine ammonia-lyase (PAL)*, *trans*-

cinnamate 4-hydrolase (C4H), 4 coumarate coA ligase (4CL) and two *pinoresinol-lariciresinol reductase (PLR)* as well as an *hydroxycinnamoyl transferase (HCT)* and an *O-methyl transferase (OMT)* (Table III-I). A cell-wall related alpha-expansin was also upregulated two-fold in PtMYB1-OE. In turn, several genes involved in stress response and detoxification as well as lignin polymerization and the terpenoid pathway were downregulated. All the genes upregulated in PtMYB1-OE were preferential to secondary xylem whereas downregulated genes were assigned to needles, phloem-bark or did not show preference to any particular tissue (Table III-I).

The effect of PtMYB8 overexpression on the spruce transcriptome was stronger than PtMYB1-OE, with a total of 79 upregulated and 138 downregulated genes. Upregulated transcripts represented genes from the shikimate pathway (DAHP and chorismate mutase), and from the phenylpropanoid pathway, such as PAL, 4CL, PLR1, HCT, and OMT, as well as genes linked to S-adenosyl-Lmethionine, a potential methyl group donor for methoxylation reactions in the monolignol biosynthetic pathway (Table III-I). They also included genes associated with cell wall organization and biogenesis. Indeed, several genes related to cellulose (CesA), hemicellulose (CLS, GT2) and pectin (GT8) were upregulated, together with an epimerase (NSI) and an arabinogalactan protein (AGP). Interestingly, nearly all the genes that were upregulated following PtMYB8-OE have been assigned as preferential to differentiating secondary xylem (X). Among the genes associated to stress and detoxification, the few that were upregulated in PtMYB8-OE, were also preferentially associated to secondary xylem (Table III-I). In turn, a large number of genes differentially expressed in PtMYB8-OE were downregulated. They include 4CL-, CCoAOMT-, *COMT*-, and *CCR*-like genes that may be related to phenylpropanoid pathway, as well as a gene for an N-hydroxycinnamoyl benzoyltransferase. Several genes related to cell wall expansion and reassembly were also downregulated in PtMYB8-OE such as GH17, GH28, UDPGT, XTH, and PAE, in addition to a CesA and a CSL-like transcript. Transcripts corresponding to enzymes involved in the early steps of flavonoid pathway also showed clear downregulation in both PtMYB8 transgenic lines, such as chalcone synthase, chalcone isomerase, dihydroflavonol 4-reductase and flavonol synthase (Table III-I). Upstream of the

phenylpropanoid pathway, two *prephenate dehydratase* genes linked to the shikimate pathway were downregulated following PtMYB8-OE. A clear decrease in transcript abundance was observed for several peroxidase genes, some of which could be associated with lignin polymerization. Genes associated with stress and detoxification were generally downregulated, as shown for *glutathione S-transferase*, as well as those related to the terpenoid pathway. In PtMYB8-OE, downregulated genes were generally identified as preferential to needles (N), phloem-bark (P) or not assigned to any particular tissue (Table III-I).

Several genes had similar patterns of differential expression in both PtMYB1 and PtMYB8 transgenics (Table III-I, shaded genes). Transcripts encoding 3-deoxy-D-arabino-heptulosonate 7-phosphate synthetase (DAHP) showed strong upregulation in both PtMYB1-OE and PtMYB8-OE. DAHP is an important enzyme of the shikimate pathway which provides precursors for secondary metabolites leading to lignins and flavonoids (Hermann, 1995; Amthor, 2003). Interestingly, five genes that were associated with the phenylpropanoid pathway were overexpressed in both transgenics, i.e. PAL, HCT, OMT, and 4CL, as well as *PLR1* which is involved in lignan synthesis primarily derived from the monolignols, E-coniferyl, and E-p-coumaryl alcohols (Kwon et al., 2001). Both transgenics also displayed a strong upregulation for an uclacyanin, a plant-specific blue copper protein that might participate in redox processes occurring during the primary defense response and/or lignin formation in plants (Nersissian et al., 1998). All the upregulated genes that overlapped in both transgenics were also preferentially expressed in differentiating secondary xylem (Table III-I). Conversely, the transcripts that had reduced expression overlapping in both transgenics were associated to phloem-bark or needle tissue or did not shown any tissue preference. Two genes that could be associated with cell wall biogenesis and reassembly, a xyloglucan endo-transglycosylase as well as a canonical class III peroxidase, demonstrated reduced transcript levels in both transgenics. Three other genes, that could be associated with stress and detoxification, such as glyoxalase, LEA EMB7 (MN5251934) and r4g30 protein coding gene, were strongly downregulated, particularly in PtMYB8-OE. A beta-primeverosidase (MN5233446), which could be linked to geraniol hydrolysis, was also strongly down-regulated in both transgenics, as well as two genes with unknown function.

Only a single transcript, a carbonic anhydrase (nacrein), was strongly upregulated in PtMYB1 transgenics and downregulated in PtMYB8 transgenics. Carbonic anhydrases have been shown to facilitate the transport of inorganic carbon and to catalyze carboxylation-decarboxylation reactions essential to photosynthesis and/or respiration in plants (Shiraiwa and Miyachi, 1985).

Table III-1. Selected genes differentially expressed in PtMYB1 and PtMYB8transgenic spruce. (Table on next page)

Microarray experiment has been designed with four control (WT) lines, two independent line per transgenic (lines 4 and 14 for PtMYB1, and lines 1 and 2 for PtMYB8), and four biological replicates.

^aPT, Preferential tissue: independent microarray experiments performed on a tissue panel of spruce tree (Pavy *et al.*, unpublished results) showed that some ESTs were preferentially expressed in either differentiating secondary xylem (X), phloem-bark (P), or needles (N), or did not show any preference to particular tissue (-). The first and second letters correspond to the tissue preference in a "differentiating secondary xylem versus needle", and a "differentiating secondary xylem versus phloem-bark" comparison, respectively.

^bFold, Fold change: the minus and plus signs indicate that, compared to wild type, the gene is under- or over-represented, respectively, in the PtMYB transgenic plantlets.

^cStudent's t test level of confidence.

^dGenes that are either up-regulated in both PtMYB1 and PtMYB8, or downregulated in both PtMYB1 and PtMYB8, or up-regulated in PtMYB1 and downregulated in PtMYB8, are shaded in grey.

^eAccording to Cell Wall Navigator annotations (<u>http://bioinfo.ucr.edu/projects/Cellwall/index.pl</u>).

			PtMYB1-4 F		PtMYB1-14		PtMYB8-1		PtMYB8-2	
Putative Gene Function	PT^{a}	EST ID	Fold ^b	Pc	Fold	Р	Fold	Ρ	Fold	Ρ
Shikimate related										
3-deoxy-DAHP synthase ^d	X,X	MN5251829	+2.9 2	2E-05	+2.6	7E-05	+3.6	4E-06	+4.1 2	2E-06
Chorismate mutase	-,X	MN5236775					+2.1	2E-05	+1.8 2	2E-04
Prephenate dehydratase	N,X	MN5171113					-2.2	1E-05	-1.8 ´	IE-04
Prephenate dehydratase	N,-	MN5240131					-2.3	8E-06	-1.8 2	2E-04
Phenylpropanoid,Monolignol related										
Phenylalanine ammonia-lyase (PAL)	X,X	MN5159896	+1.9 8	8E-05	+1.9	1E-04	+2.2	1E-04	+2.0 4	1E-04
Trans-cinnamate 4-hydrolase (C4H)	-,X	MN5158256	+2.1 3	8E-06	+2.0	1E-05				
4-coumarate-CoA ligase (4CL)	X,X	MN5255003	+2.3 9	9E-06	+2.0	5E-05	+3.1	9E-02	+2.6 8	3E-02
Pinoresinol-lariciresinol reductase (PLR1)	X,X	MN5177685	+3.2 6	8E-07	+2.7	4E-06	+2.2	4E-05	+2.3 2	2E-05
Pinoresinol-lariciresinol reductase (PLR2)	X,X	MN5194637	+2.4 4	E-03	+2.6	4E-03				
Hydroxycinnamoyl transferase (HCT)	X,X	MN5171597	+2.0 9)E-06	+1.8	6E-05	+2.0	1E-04	+2.1	IE-04
O-methyltransferase (OMT)	X,X	MN5194345	+2.3 4	E-05	+2.1	2E-04	+2.3	2E-04	+3.4 2	2E-05
4-coumarate-CoA ligase-like	-,-	MN5255491					-2.3	3E-06	-2.0 ²	1E-05
Caffeic acid O-methyltransferase-like	N,-	MN5255308					-3.3	1E-05	-3.7 8	3E-06
Caffeoyl-CoA 3-O-methyltransferase-like	N,-	MN5164253					-6.3	6E-07	-6.8 4	4E-07
Cinnamoyl CoA reductase-like	N,P	MN5235414					-2.4	5E-06	-2.3 ´	IE-07
Cytochrome P450	N,P	MN5181371					-2.1	1E-06	-2.1 ´	1E-06
N-hydroxycinnamoyl,benzoyltransferase	-,P	MN5238027					-2.4	3E-06	-2.1 ´	IE-05
Peroxidases										
Peroxidase (class III)	-,-	MN5232900	-1.8 2	2E-04	-1.8	3E-04	-4.0	4E-05	-3.6 9	9E-05
Peroxidase	-,P	MN5173202					-2.9	5E-05	-2.3 4	1E-04
Peroxidase precursor	-,-	MN5181972					-2.2	6E-05	-2.0 2	2E-04
Peroxidase ATP2a	-,-	MN5233396					-2.3	3E-06	-1.8 4	4E-05
Cationic peroxidase 1 precursor	N,P	MN5235950					-3.2	9E-06	-2.5 8	3E-05
Cell Wall related ^e										
Cellulose synthase catalytic subunit (CesA)	X,X	MN5175268					+2.2	2E-05	+1.9 2	2E-04
Cellulose synthase catalytic subunit (CesA)	X,X	MN5176643					+2.3	2E-05	+1.9 2	2E-04
Cellulose synthase (CesA)	X,X	MN5195335					+3.6	4E-06	+3.5 6	6E-06
Cellulose synthase (CesA)	N,P	MN5234677					-2.1	5E-05	-2.4 2	2E-05
Cellulose synthase-like (CSL)	N,-	MN5236143					-3.3	1E-06	-3.1 2	2E-06
Glycosyl transferase 2 (CSL)	X,X	MN5194187					+2.2	1E-05	+2.2 8	3E-06
Glycosyl transferase 2 (CSL)	X,X	MN5159543					+2.1	8E-05	+2.3 4	4E-05
Epimerase (NSI)	X,X	MN5177923					+2.3	5E-06	+2.6 2	2E-06
Expansin (EXP)	X,P	MN5194355	+2.1 9	9E-06	+2.1	2E-05				
Glycohydrolase (GH17)	-,-	MN5251191					-3.9	4E-06	-2.6	7E-05
Glycohydrolase (GH17)	-,X	MN5256029					-2.9	4E-06	-2.2 6	6E-05
Glycosyl transferase 8 (GT8)	X,X	MN5170784					+3.8	8E-07	+3.0 3	3E-06
Hydroxyproline-rich glycoprotein-like (HRGP	Y,P	MN5173049					-1.9	2E-04	-1.8 6	6E-04
Arabinogalactan-like protein (AGP)	X,X	MN5191579					+3.8	8E-07	+4.1 4	4E-07
Pectin Acetylesterase (PAE)	-,P	MN5240041					-2.1	3E-05	-1.9 ´	1E-04
Xyloglucan endo-transglycosylase (XTH)	-,-	MN5255573	-3.6 1	E-05	-2.9	1E-04	-8.2	5E-07	-8.3 4	4E-07
Glycoside Hydrolase (AGAL)	-,-	MN5236822					-2.3	1E-06	-2.0 5	5E-06

Chitinase	-,-	MN5181335					-3.2	3E-05	-2.2 6E-04
Chitinase	-,-	MN5182703					-2.6	4E-06	-2.5 7E-06
Transferase	-,P	MN5238027					-2.4	3E-06	-2.1 1E-05
Xyloglucanase inhibitor	-,X	MN5163020					-2.5	2E-05	-1.8 7E-04
Polygalacturonase (GH28)	-,X	MN5233983					-2.9	7E-06	-1.9 5E-04
UDP-glycosyltransferase 85A8 (UDPGT)	-,X	MN5181672					-2.4	6E-05	-2.4 9E-05
Glucosyltransferase-like protein (UDPGT)	N,-	MN5232972					-2.1	1E-06	-1.8 8E-04
SAM,SAH metabolism									
Adenosylhomocysteinase	X,X	MN5251787	+1.9 11	E-05	+1.9	2E-05	+1.8	3E-04	+1.8 4E-04
Nicotianamine synthase	N,-	MN5182064					+2.3	2E-04	+2.7 8E-05
CBS-domain-containing protein	X,X	MN5174213					+2.4	5E-05	+2.0 3E-04
Homocysteine methyltransferase	X,X	MN5252997	+1.9 41	E-06	+1.9	1E-05			
Flavonoid related									
Flavonol synthase	-,-	MN5172805					-2.1	1E-04	-2.1 1E-04
Chalcone-flavonone isomerase	N,P	MN5163731					-2.1	5E-06	-2.3 2E-06
Flavanone 3-hydroxylase	N,-	MN5192796					-2.0	5E-06	-1.8 2E-05
Dihydroflavonol 4-reductase	N,-	MN5235944					-3.0	5E-07	-2.8 7E-07
Chalcone synthase	N,-	MN5252787					-3.3	2E-06	-2.4 2E-05
Anthocyanidin synthase	N,-	MN5255315					-3.7	1E-06	-2.8 2E-06
Flavonoid 3'.5'-hydroxylase	N,-	MN5255946					-2.7	2E-04	-3.5 3E-05
Terpenoid related									
Alpha-bisabolene synthase	-,-	MN5251845	-2.6 91	E-04	-2.1	8E-03			
Geraniol 10-hydroxylase	-,-	MN5233989					-4.3	3E-05	-4.5 3E-05
Beta-primeverosidase	-,-	MN5233446	-3.4 21	E-06	-2.1	1E-04	-8.3	1E-06	-9.2 9E-07
Beta-primeverosidase	-,-	MN5251221					-2.9	8E-07	-3.1 4E-07
Beta-primeverosidase	-,-	MN5182631					-6.0	1E-06	-6.0 1E-06
Hydroxymethylglutaryl-CoA synthase	-,-	MN5236845					-2.2	2E-06	-2.2 2E-06
Acetyl Co-A acetyltransferase	-,-	MN5260082					-1.8	8E-06	-2.0 4E-06
Gamma-tocopherol methyltransferase	-,P	MN5254253					-7.3	1E-06	-7.1 1E-06
CPRD12 protein	N,P	MN5255892					-2.2	2E-06	-2.3 1E-06
Serine carboxypeptidase II	X,X	MN5246267					-2.1	1E-05	-2.0 3E-05
Stress and Detoxification									
Beta-cyanoalanine synthase	N,X	MN5258469					+3.2	7E-06	+3.3 7E-06
Vignain precursor	Х,-	MN5160240	+2.1 1	E-03	+1.8	1E-02			
Vignain precursor	Х,-	MN5170536	+2.3 51	E-04	+1.9	5E-03			
Gamma-tocopherol methyltransferase	-,P	MN5254253					-7.3	1E-06	-7.1 1E-06
Glutathione S-transferase	-,-	MN5238832					-2.3	8E-05	-2.3 7E-05
Glutathione S-transferase	-,-	MN5255937					-2.3	2E-05	-2.5 9E-06
Glutathione S-transferase	N,-	MN5256468					-1.9	1E-04	-2.1 5E-05
Glyoxalase	-,-	MN5257610	-2.1 1	E-03	-2.4	6E-04	-4.3	2E-06	-4.9 1E-06
Phytochelatin synthetase	Х,-	MN5207547					+2.8	4E-05	+2.3 2E-04
Phytochelatin synthetase-like	Х,-	MN5259467					+2.7	5E-05	+2.1 5E-04
LEA EMB35	-,-	MN5239233	-2.0 6l	E-04	-2.2	3E-04			
LEA-EMB7	-,-	MN5181285					-3.2	7E-05	-2.7 2E-04
LEA EMB7	-,-	MN5251934	-2.6 1	E-03	-2.6	2E-03	-5.9	3E-06	-7.6 1E-06
Dehydrin	-,-	MN5161489					-1.8	4E-06	-1.8 3E-06
Dehydrin	-,-	MN5163790					-4.1	7E-06	-3.8 1E-05
r40g3 protein	N,P	MN5256759	-2.2 1	E-04	-2.3	2E-04	-10.1	3E-06	-9.8 3E-06
Lipase, Thioesterase	X,X	MN5176766	+3.0 2E-06	+2.4 1E-05	+2.3	8E-05	+2.4 9E-05		
--	-------	-----------	------------	------------	------	-------	------------		
BURP domain-containing protein	-,P	MN5163000			-4.8	5E-05	-4.3 1E-04		
Arm repeat-containing protein	-,-	MN5195907			-2.0	2E-06	-1.9 4E-06		
Lanatoside 15'-O-acetylesterase	-,P	MN5256118			-2.8	5E-05	-2.1 6E-04		
EF-hand Ca2+-binding protein	N,-	MN5244381			-2.4	1E-04	-2.4 1E-04		
Delta-1-pyrroline-5-carboxylate dehydrogenas	ε X,X	MN5259351			+2.0	3E-06	+1.8 8E-06		
Disease resistance gene	-,X	MN5192231			+2.0	2E-05	+2.0 1E-05		
Rapid alkalinization factor 3	-,P	MN5243656			-2.3	3E-05	-2.0 2E-04		
Fatty acid hydroxylase	-,-	MN5196748			-1.8	2E-05	-1.8 3E-05		
Allene oxide synthase precursor	N,-	MN5241220			-2.9	2E-05	-2.7 3E-05		
Metallothionein-like protein	N,-	MN5182650			-3.9	8E-06	-3.4 2E-05		
Metallothionein-like protein	-,P	MN5241564			-2.9	7E-06	-2.3 5E-05		
Lipid transfer protein	X,X	MN5239177			-2.6	7E-06	-5.7 3E-07		
Thioredoxin H-type	-,-	MN5254070			-2.1	9E-05	-1.8 1E-04		
Miscellanous									
Glycine decarboxylase	X,X	MN5241924	+2.0 9E-06	+1.8 1E-04	+1.8	2E-04	+2.2 2E-05		
Uclacyanin	N,X	MN5177006	+3.6 2E-06	+3.2 6E-06	+4.6	1E-05	+5.2 8E-06		
Carbonic anhydrase (nacrein)	-,-	MN5182742	+3.3 1E-07	+2.6 2E-06	-3.4	5E-07	-3.5 3E-07		
Unknown	N,P	MN5161611	–1.9 1E-03	-1.8 4E-03	-2.6	2E-06	-2.4 3E-06		
Unknown	-,P	MN5250897	-4.3 8E-06	-3.1 1E-04	-6.0	1E-06	-9.0 4E-07		
Transgene	-	MN5173128	+3.1 2E-04	+2.7 7E-04	+5.7	3E-06	+4.7 7E-06		

Table III-1. Selected genes differentially expressed in PtMYB1 and PtMYB8

transgenic spruce.

The legend is presented before the table.

QPCR analyses were performed on a subset of genes that were related to cell wall biogenesis and to flavonoid and shikimate pathways in order to validate the results from our microarray analyses (Figure III-5). We extended this analysis to all the genes involved in phenylpropanoid pathway to complement microarray transcript profiles. The misregulation of cell wall-, flavonoid-, and shikimate-related genes was confirmed in both transgenics as shown in figure III-5 B, C, and D, respectively. We also confirmed the upregulation of genes involved in phenylpropanoid, monolignol biosynthesis compared to controls (Figure III-5 A). In PtMYB1-OE, transcript abundance was increased for all of the genes tested in this pathway, going from PAL to cinnamyl alcohol dehydrogenase (CAD). In PtMYB8 transgenics, several genes of the pathway also had increased transcript levels, except for caffeate O-methyltransferase (COMT) and coumaroyl CoA Omethyltransferase (CCoAOMT) for which transcript abundance was similar to control, and, PAL and C4H gave smaller and less consistent increases. The upregulation of the two *pinoresinol-lariciresinol reductase* transcripts (*PLR1* and 2) was confirmed in both transgenics compared to control (Figure III-5 A).

Figure III-5. (Legend, see figure on next page). A, Monolignol related genes: Phenylalanine ammonia-lyase (PAL), Cinnamic acid 4-hydroxylase (C4H), 4-Coumarate-CoA ligase (4CL), Coumarate 3-hydroxylase (C3H), Caffeate Omethyltransferase (COMT), Caffeovl-CoA 3-O-methyltransferase (CCoAOMT), Cinnamoyl-CoA reductase (CCR), Cinnamyl-alcohol dehydrogenase (CAD), Pinoresinol-lariciresinol reductase 1 and 2 (PLR1 and PLR2). B, Cell wall related Cellulose synthase (CesA), Glycosyltransferase family 8 (GT8), genes: Arabinogalactan (AGP1), Xyloglucan endotransglycosylase, hydrolase (XTH). C, Flavonoid related genes: Chalcone synthase (CHS), Dihydroflavonol 4-reductase (DFR). D, Shikimate related gene: 3-deoxy-D-arabino-heptulosonate 7-phosphate synthetase (DAHP). For both transgenics, transcript accumulation was assessed by QPCR on 3-week old plantlets produced from two independent transgenic lines for each gene construct (lines 4 and 14 for PtMYB1, and lines 1 and 2 for PtMYB8). Transcript levels were determined from three biological replicates (25 plantlets per replicate) and were normalized relative to EF1-alpha expression level. Error bars represent standard deviation. nd, not detected.



Figure III-5. Real time RT-PCR analysis of transcript accumulation in wild type, PtMYB1 and PtMYB8 transgenic plantlets: validation of microarray and expression data from genes related to secondary metabolism and cell wall assembly.

The legend is presented before the figure.

III-6-5-PtMYB1- and PtMYB8-OE in spruce affected the expression levels of spruce endogenous MYBs

We also examined the effect of PtMYB1 and PtMYB8 overexpression on the RNA level of their respective putative spruce orthologs, PgMYB1 and PgMYB8. Pairwise optimal global alignments previously identified putative orthologous pairs of spruce and pine for MYB1 and MYB8 (Bedon et al., 2007); these pairs had a sequence similarity of 91.3% and 95.5%, respectively, based on complete amino acids sequence alignment, and of 100% based on DNA binding domain (DBD). In addition, transcript levels of three other spruce MYBs, PgMYB2, PgMYB3, and PgMYB4, that were shown to be preferentially expressed in differentiating xylem from 33-year old white spruce trees (Bedon et al., 2007), were also evaluated. First, we determined the transcript levels of the two transgene by QPCR and found that they were similar when normalized relative to EF1-alpha (Figure III-6) but represented a 30-fold and 130-fold increase in transcript abundance compared to their endogenous counterparts PgMYB1 and *PgMYB8*, respectively. In PtMYB8-OE plantlets, transcript levels of *PgMYB1* and *PgMYB3* were upregulated compared to control (Figure III-6). A clear decrease in transcript abundance was also observed for PgMYB4, as well as for PgMYB8. A more variable result was obtained for PgMYB2, for which an increase in transcript level was only recorded in line 1 (Figure III-6). In contrast, PtMYB1 overexpression led to a decrease in steady state RNA of its putative homolog PgMYB1. It also gave decreased levels for PgMYB4 and, to a lesser extent, for *PgMYB2* but *PgMYB3* and *PgMYB8* were unaffected (Figure III-6).





For both transgenics, transcript accumulation was assessed by real time RT-PCR on 3-week old plantlets produced from two independent transgenic lines (line 4 and 14 for PtMYB1, and lines 1 and 2 for PtMYB8). Transcript levels were normalized relative to *EF1-alpha* expression level and were obtained from 3 biological replicates (15 plantlets per replicate). Error bars represent standard deviation. nd, not detected.

III-7-Discussion

Previous studies have inferred that lignin biosynthesis and deposition in wood are regulated by MYBs in gymnosperm and angiosperm trees, based on gain-of-function experiments carried out in heterologous systems like tobacco (Patzlaff et al., 2003b; Goicoechea et al., 2005) or Arabidopsis (Newman et al., 2004). In the present study, functional analyses of two pine MYBs, PtMYB1 and *PtMYB8*, were carried out to examine the role that these genes play in a conifer. Both gene sequences were overexpressed with a maize ubiquitin promoter by Agrobacterium transformation of somatic embryogenic cells in white spruce (Picea glauca [Moench] Voss). Upon germination of mature somatic embryos, root growth was decreased in transgenics overexpressing either one of the genes, however PtMYB8 had a stronger and more consistent effect, completely preventing root development in some lines. Ectopic lignification was also apparent in both transgenics, but with very different intensities and spatial patterns. Microarray and QPCR analyses of the transgenics identified several genes linked to secondary metabolism and cell wall assembly with altered expression; however, there were notable differences in the suite of differentially expressed genes observed for PtMYB1 and PtMYB8. Taken together, these data suggest that the *PtMYB1* and *PtMYB8* regulons are distinct but also overlapping. The interpretation of these phenotypes is addressed in more detail below.

III-7-1-Involvement of pine MYBs in lignin deposition

The phenotypes resulting from overexpression of PtMYB1 and PtMYB8 in spruce were consistent with the hypothesis that they both regulate lignification or developmental processes that lead to lignification. Both gene sequences lead to ectopic lignification, localized in a few cells in the PtMYB1-OE hypocotyl (Figure III-2) or in several cell types in the hypocotyl, the cotyledons and the root in PtMYB8-OE spruce plantlets (Figure III-3). The large increase in lignin content observed in plantlets overexpressing PtMYB8 (Figure III-4 A) was consistent with the decreased pool of (free phenolic) intermediates in the lignin monomer biosynthesis pathway (Figure III-4 B) and was congruent with our histological observations. The increased lignin content probably reflects the increased proportion of lignifying cells in PtMYB8-OE plantlets compared to wild type, indicating that extensive cell wall thickenings observed in PtMYB8-OE may be linked to altered lignin deposition.

It is apparent from other studies that ectopic lignification is not sufficient to establish that PtMYB1 and PtMYB8 directly regulate lignin biosynthesis or assembly. For example, ectopic lignification and altered tissue organisation have been observed in Arabidopsis as a result of diverse mutations that do not directly implicate lignin enzymes or regulators. For instance, the mutants ectopic lignification of the pith 1 (elp1) (Zhong et al., 2000) and ectopic lignification 1 (eli1) (Cano-Degado et al., 2000) are impaired in genes encoding a chitinase-like protein and a CesA protein, respectively. Ectopic lignin accumulation was also shown to result from a mutation in a gene that is not linked to cell wall structure, as shown by the *det3* mutation in a vacuolar-type ATPase subunit (Schumaker *et* al., 1999). Interestingly, the ectopic lignin phenotype of the det3 mutant is most likely a consequence of the altered expression of the R2R3-MYB AtMYB61 resulting from the det3 mutation. AtMYB61 was sufficient and necessary to account for the lignin phenotype in det3 mutants (Newman et al., 2004). Thus, in some cases at least, ectopic lignification that has no obvious or direct link to the mutation may in fact be mediated by MYB genes.

Despite these observations, restricting the role of PtMYB1 and PtMYB8 to the sole control of lignin biosynthetic pathway may lead to a biased interpretation or to a restricted view of the putative metabolic pathways affected by R2R3-MYB overexpression in spruce. PtMYB1 was shown to regulate transcription from cisacting AC elements present in the *PAL* promoter (Patzlaff *et al.*, 2003a) but also in the *glutamine synthase* (*GS*) promoter (Gomez-Maldonado *et al.*, 2004). In the GS/GOGAT nitrogen-recycling system, GS participates in the regeneration of glutamate which is required for conversion of prephenate to arogenate, which in turn leads to phenylalanine (Amthor, 2003). It thus seems plausible that PtMYB1 could also regulate metabolic pathways other than or linked to lignin biosynthesis (Gomez-Maldonado *et al.*, 2004). R2R3-MYBs have been implicated in the regulation of other aspects of secondary metabolism such as flavonoids (Borevitz *et al.*, 2000; Von Endt *et al.*, 2002) as well as benzenoids through the regulation of the shikimate pathway (Verdonk *et al.*, 2005). Thus, R2R3-MYB overexpression may affect metabolic flux through the phenylpropanoid pathway and indirectly stimulate lignification, by acting upon closely related pathways. We decided to carry out microarray experiments to develop further insights regarding the implication of both genes in conifer secondary metabolism and xylogenesis, using an unbiased approach.

III-7-2-Transcriptional regulation of genes linked to lignin biosynthesis and secondary metabolism pathways

The impact of PtMYB overexpression on the spruce transcriptome was assessed by gene expression profiling with spruce 9K microarrays. We conducted side by side analyses of transgenics overexpressing the two genes, in light of their phenotypic similarities suggesting partly redundant functions. The results from theses analyses identified an overlapping set of misregulated genes in the PtMYB1 and PtMYB8 transgenics, including several upregulated genes that are involved in lignin monomer synthesis (Table III-I). These findings were verified and extended to nearly all of the pathway genes by QPCR (Figure III-5). Whereas PtMYB1-OE led to a clear upregulation of all the genes tested in the phenylpropanoid pathway (from PAL to CAD), PtMYB8-OE upregulated most of the genes but had no effect on the transcript levels of COMT and CCoAOMT. As we observed here in spruce, the possibility that different MYB genes may act differentially on the same target genes within the lignin biosynthetic pathway has been highlighted in previous reports. For example, overexpression of the Eucalyptus EgMYB2 in tobacco strongly upregulated transcript levels of some lignin biosynthesis genes involved in the monolignol-specific portion of the pathway, including COMT and CCoAOMT genes, but not PAL and C4H genes (Goicoechea et al., 2005). Similarly, the results obtained upon overexpression of PtMYB4, also in tobacco (Patzlaff et al., 2003b), showed that PAL transcript levels were decreased and that only part of the pathway genes were increased, including COMT and CCoAOMT. Other authors have also proposed that more than one MYB may act on the same target gene to regulate its expression (Moyano et al., 1996; Tamagnone et al., 1998; Borevitz et al., 2000) and thus provide multiple levels of control for lignin biosynthesis in plants (Tamagnone et al., 1998; Borevitz et al., 2000). Our results provide evidence that several lignin

biosynthetic genes can be recruited in spruce through either PtMYB1- or PtMYB8-OE. Furthermore, the degree to which theses genes are recruited (i.e. which targets are preferentially misregulated) may influence the extent of lignification, as shown in PtMYB1 and PtMYB8 overexpressers, respectively. In addition, the strong upregulation of DAHP in both PtMYB1-OE and PtMYB8-OE may indicate an overlapping routing of metabolic flow into the upstream shikimate pathway, which provides precursors required for lignin biosynthesis (Hermann, 1995; Amthor, 2003). The monolignol and shikimate pathways appear to be co-ordinately upregulated through overexpression of PtMYB1 and PtMYB8. The misregulation of genes linked to S-adenosyl-L-methionine (Table III-I), which is a methyl group donor for lignin monomers (Amthor, 2003), may also influence methoxylation in the monolignol biosynthetic pathway, thus potentially contributing to the different lignification responses observed in PtMYB1- and PtMYB8-OE. Based on the results, it would appear that PtMYB1 and PtMYB8 are likely to be involved in the lignification process in conifers, probably by directing metabolic flux into shikimate and monolignol pathways to sustain a differential methoxylated lignification.

There were also large differences in the number and nature of transcripts affected upon overexpression of each of the two pine MYBs. First, PtMYB8-OE led to a strong downregulation of many flavonoid-, terpenoid- and bezenoidrelated genes contrary to PtMYB1-OE (Table III-I). In addition, PtMYB8-OE misregulated transcript levels of many stress and detoxification genes and, surprisingly, downregulated several putative lignin polymerisation genes. The differences in the effect of the two MYBs may entail functional specificity and could involve interaction with others proteins, such as bHLH and WD40 which have been identified as important MYB partners in flavonoid pathways (Von Endt et al., 2002). Due to the highly conserved DNA binding domain observed among conifer MYBs (Bedon et al., 2007), it may also reflected the indirect effects linked to non-specific binding to gene promoters caused by MYB overexpression. Such a dosage-linked effect would influence most strongly the expression of genes with similar cis regulatory motifs, as suggested in Arabidopsis (Jin et al., 2000). Another explanation for our results is that the very strong lignification had pleiotropic effects on development like dwarfing, most evident on root growth,

and that at least some of the shifts in gene expression are reflection of these indirect consequences of MYB overexpression. For example in *Arabidopsis*, dwarfing effects has also been observed with the overexpression of AtMYB61 which was sufficient to cause ectopic lignification (Newman *et al.*, 2004).

III-7-3-Different roles of MYB1 *versus* MYB8 in metabolism and plant development

Taken together, the different phenotypes observed upon overexpression of PtMYB1 and PtMYB8 provide a clear indication that lignification may be the primary target for both of these MYBs. Nevertheless, PtMYB1 and PtMYB8-OE clearly impacted lignification quite differently as assessed by the distribution of the ectopic lignification in the plants. Accordingly, development was compromised to very different degrees by the two genes, PtMYB8-OE plants being unable to survive transfer to soil.

In pine and spruce, MYB1 clearly has a ubiquitous expression profile while displaying a preference to lignifying tissues. Transcripts of PtMYB1 accumulated in several organs of young pine trees, but are most abundant in secondary xylem as shown by QPCR analysis (Figure III-1) and northern hybridizations (Patzlaff et al., 2003a). PtMYB1 proteins have also been localized to epidermal cells and in the primary xylem of cotyledons, hypocotyls and roots of young pine seedlings (Gomez-Maldonado et al., 2004). In mature spruce trees, the transcripts of its closest homolog *PgMYB1* accumulated preferentially in secondary xylem as well as in the primary stem and in needles (Bedon et al., 2007). Constitutive overexpression of PtMYB1 in spruce only led to localized and restricted ectopic lignification (Figure III-2, H and J). It induces a delay in growth that is not incompatible with further plant development (Figure III-2 D), and leads to the coordinated upregulation of all the phenylpropanoid-related genes (Figure III-5 A) which represent a large part of differentially expressed genes in PtMYB1-OE plantlets (Table III-I). Together, these results indicate that *PtMYB1* may be viewed as a general regulator of phenylpropanoid metabolism but suggests, however, that PtMYB1-OE is not sufficient for the deposition of lignin in most cells. Given these observations, it appears that lignification could require other factors in conjunction with MYB1 activity, potentially involving other regulatory proteins that have been linked to lignin such as LIM, MADS, KNOX or others R2R3-MYBs (Rogers and Campbell, 2004). Alternatively, PtMYB1 activity may require specific protein partners which would be limiting in several cell types. For example, bHLH or WD40 proteins have been shown to form heteroduplexes with MYBs in order to bind to *cis*-regulatory elements (Goff *et al.*, 1992; Koes *et al.*, 2005; Morita *et al.*, 2006).

In contrast, PtMYB8 has a more restricted expression profile in pine and spruce. In young pine trees, transcripts of *PtMYB8* preferentially accumulated in differentiating secondary xylem of stem (Figure III-1). In spruce, PgMYB8, the closest homolog of *PtMYB8*, was also clearly associated with secondary xylem in mature trees (Bedon et al., 2007). MYB8 expression appears to be tightly regulated, since transcripts were generally present at lower levels than MYB1 in young pine trees (Figure III-1) as well as spruce plantlets (Figure III-6) and mature trees (Bedon et al., 2007). PtMYB8 overexpression was sufficient to lead to extensive ectopic lignification in several cell type and organs (Figure III-3), suggesting that it could be a strong regulator that is less dependent on other factors for its activity and function. It led to the coordinated upregulation of specific monolignol-related genes, excluding COMT and CCoAoMT, in addition to affect numerous cell wall-related genes (Table III-1, Figure III-5). PtMYB8-OE was also associated with strong decrease in transcript levels of flavonoid biosynthesis enzymes (Table III-1). Therefore, our data support the hypothesis that MYB8 targets lignin biosynthesis as well as other metabolic or developmental pathways linked to secondary cell wall but its expression profiles suggest that its normal activity is confined to secondary xylem, and thus to a narrower subset of tissues and cell types than MYB1. In a recent report (Bedon et al., 2007), the three spruce MYBs, PgMYB2, 4, and 8, that had the highest levels of steady state RNA in secondary xylem also displayed increased transcript accumulation upon the induction of compression wood. In conifers, compression wood is produced in bent or leaning trees, and is charaterized by a higher lignin content and by the deposition of more condensed lignins that are enriched in H units, in spruce (Lange et al., 1995) and pine (Yeh et al., 2006). Interestingly, PgMYB2, PgMYB4 and PgMYB8 displayed altered transcript level in PtMYB8

overexpressers (Figure III-6). PtMYB8-OE also led to the misregulation of cellulose, hemicellulose and pectin related genes, as well as arabinogalactan protein whose upregulation has been associated to compression wood in spruce (Bedon *et al.*, 2007) and pine (Zhang *et al.*, 2000) and tension wood in poplar (Lafarguette *et al.*, 2004). A possible explanation for these results would be that PtMYB8 recruited the spruce genes necessary for compression wood deposition, together with or including PgMYB2 and 4. Consistent with this idea, the overexpression of PtMYB8 appears to lead to the upregulation of genes required to produce the less methoxylated precursors leading to H lignin units (Lange *et al.*, 1995).

III-7-4-Do MYBs form a regulatory network to control lignification?

The misregulation of four spruce MYBs (PgMYB1, 2, 3, 4) upon overexpression of PtMYB8 is an intriguing finding from the present study. If these data are substantiated, they would lead to the hypothesis that the effect of PtMYB8 on the lignification of cell walls may be mediated, in part by the induction of other MYBs, including MYB1 which itself is a strong candidate for direct regulation of lignin biosynthesis. Our data only represent circumstantial evidence that such a mechanism exists, but are consistent with other studies in which the regulation of MYB expression by other MYBs has been proposed (Perez-Rodriguez et al., 2005). The authors reported that the restriction of AmMYBML1 expression in snapdragons might be a consequence of the loss of activity of its upstream regulator, DIVARICATA, which encodes another MYBrelated transcription factor. Consistent with this idea, a recent analysis of genes regulating secondary xylem development in Arabidopsis identified a MYB (AtMYB69) whose promoter contains an AC element, which is a putative binding site for MYB transcription factors (Ko et al., 2006). As such, the idea of a network of *MYB*s acting to integrate developmental, metabolic and environmental cues to regulate lignification is not inconsistent with previous studies, but further experimentation is needed to directly test the hypothesis itself. DNA binding experiments should be considered to test whether MYB8 regulates other MYBs.

By extrapolation from the *Arabidopis* and *Populus* genomes which each contain well over 100 R2R3-*MYB*s, it would appear that many more conifer *MYB*s remain to be characterized. Furthermore, several *MYB*s have been shown to regulate different aspects of vascular tissue differentiation and patterning, which are events that occur upstream of lignification (Scarpella and Meijer, 2004). Therefore, other *MYB*s in addition to those discussed here are likely to play direct or indirect roles in lignin biosynthesis or in developmental pathways that lead to lignification of primary and secondary xylem in conifers.

III-9-Acknowledgments

We thank RR Sederoff (North Carolina Sate University, Raleigh, NC) for kindly providing clones of *Pinus taeda*, Frédéric Morency for assistance in tissue culture and Pierre-Olivier Nadeau for RNA extractions, and quality control. Thanks are extended to Dr Brian Boyle for critical comments on the manuscript.

III-10-References

- Amthor J. (2003). Efficiency of lignin biosynthesis: a quantitative analysis. *Ann Bot* 91: 673–695
- Bedon F., Grima-Pettenati J. and Mackay J. (2007). Conifer R2R3-MYB transcription factors: sequence analysis and gene expression in wood-forming tissues of white spruce. (*Picea glauca*). *BMC Plant Biol* 7: 17
- **Benjamini Y. and Hochberg Y. (1995).** Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Statist Soc Ser B* **57:** 289–300
- Boerjan W., Ralph J. and Baucher M. (2003). Lignin biosynthesis. Annu Rev Plant Biol 54: 519–546
- Borevitz J.O., Xia Y., Blount J., Dixon R.A. and Lamb C. (2000). Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *Plant Cell* 12: 2383–2394
- Björkmann A. (1956). Studies on finely divided wood. Part 1. Extraction of lignin with neutral solvents. *Suen Papperstidn* 59: 477
- Campbell M.M. and Sederoff R.R. (1996). Variation in lignin content and composition. *Plant Physiol* 110: 3–13
- Cano-Delgado A.I., Metzlaff K. and Bevan M.W. (2000). The eli1 mutation reveals a link between cell expansion and secondary cell wall formation in *Arabidopsis thaliana*. *Development* 127: 3395–3405
- Chang S., Puryear J. and Cairney J. (1993). A simple and effcient method for isolating RNA from pine trees. *Plant Mol Biol Rep* 11: 113–116
- Chen Y., Yang X., He K., Liu M., Li J., Gao Z., Lin Z., Zhang Y., Wang X., Qiu X., Shen Y., Zhang L., Deng X., Luo J., Deng X.W., Chen Z., Gu H. and Qu L.J. (2006). The MYB transcription factor superfamily of

Arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family. *Plant Mol Biol* **60**: 107–124

- Christensen A.H., Sharrock R.A. and Quail P.H. (1992). Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Mol Biol* **18:** 675–689
- Cosgrove D.J. (1999). Enzymes and other agents that enhance cell wall extensibility. *Annu Rev Plant Physiol Plant Mol Biol* 50: 391–417
- Gentleman R.C., Carey V.J., Bates D.M., Bolstad B., Dettling M., Dudoit S., Ellis B., Gautier L., Ge Y., Gentry J., Hornik K., Hothorn T., Huber W., Iacus S., Irizarry R., Leisch F., Li C., Maechler M., Rossini A.J., Sawitzki G., Smith C., Smyth G., Tierney L., Yang J.Y.H. and Zhang J. (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 5: R80
- Goff S.A., Cone K.C. and Chandler V.L. (1992). Functional analysis of the transcriptional activator encoded by the maize B gene: evidence for a direct functional interaction between two classes of regulatory proteins. *Genes Dev* 6: 864–875
- Goicoechea M., Lacombe E., Legay S., Milhaevic S., Rech P., Jauneau A., Lapierre C., Pollet B., Verhaegen D., Chaubet-Gigot N. and Grima-Pettenati J. (2005). EgMYB2, a new transcriptional activator from *Eucalyptus* xylem, regulates secondary cell wall formation and lignin biosynthesis. *Plant J* 43: 553–567
- Gomez-Maldonaldo J., Avila C., de la Torre F., Canas R., Canovas F.M. and Campbell M.M. (2004). Functional interactions between a glutamine synthetase promoter and MYB proteins. *Plant J* 39: 513–526
- Groover A. and Robischon M. (2006). Developmental mechanisms regulating secondary growth in woody plants. *Curr Opin Plant Biol* 9: 55-58
- Herrmann K.M. (1995). The shikimate pathway: Early steps in the biosynthesis of aromatic compounds. *Plant Cell* 7: 907–919
- Humphreys J.M. and Chapple C. (2002). Rewriting the lignin roadmap. *Curr* Opin Plant Biol 5: 224–229
- Ihaka R. and Gentleman R.. (1996). R: A language for data analysis and graphics. *J Comp Graph Stat* 5: 299–314
- Jin H.L., Cominelli E., Bailey P., Parr A., Mehrtens F., Jones J., Tonelli C., Weisshaar B. and Martin C. (2000). Transcriptional repression by AtMYB4 controls production of UV- protecting sunscreens in Arabidopsis. *EMBO J* 19: 6150–6161
- Jin H. and Martin C. (1999). Multifunctionality and diversity within the plant MYB-gene family. *Plant Mol Biol* 41: 577–585
- Karpinska B., Karlsson M., Srivastava M., Stenberg A., Schrader J., Sterky
 F., Bhalerao R. and Wingsle G. (2004). MYB transcription factors are differentially expressed and regulated during secondary vascular development in hybrid aspen. *Plant Mol Biol* 56: 255–270
- Klimaszewska K., Rutledge R.G. and Séguin A. (2004). Genetic transformation of conifers utilizing somatic embryogenesis. In L Peña, ed, *Methods in Molecular Biology, Humana Press, Totowa*, pp 151–164
- Klimaszewska K., Lachance D., Pelletier G., Lelu M.A. and Séguin A. (2001). Regeneration of transgenic *Picea glauca*, *P. mariana*, and *P. abies* after

cocultivation of embryogenic tissue with *Agrobacterium tumefaciens*. In Vitro Cell Dev Biol-Plant **37**: 748–755

- Ko J.H., Beers E.P. and Han K.H. (2006). Global comparative transcriptome analysis identifies gene network regulating secondary xylem development in *Arabidopsis thaliana*. *Mol Gen Genomics* **276**: 517–531
- Koes R., Verweij W. and Quattrocchio F. (2005). Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. *Trends Plant Sci* 10: 236–242
- Koncz C.and Schell J. (1986). The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. *Mol Gen Genet* 204: 383–396
- Kranz H.D., Denekamp M., Greco R., Jin H., Leyva A., Meissner R.C., Petroni K., Urzainqui A., Bevan M., Martin C., Smeekens S., Tonelli C. Paz-Ares J. and Weisshaar B. (1998). Towards functional characterization of the members of the R2R3-MYB gene family from *Arabidopsis thaliana*. *Plant J* 16: 263–276
- Kwon M., Davin L.B. and Lewis N.G. (2001). In situ hybridization and immunolocalization of lignan reductases in woody tissues: implications for heartwood formation and other forms of vascular tissue preservation. *Phytochem* 57: 899–914
- Lange B.M., Lapierre C. and Sandermann H.Jr. (1995). Elicitor-induced spruce stress lignin. (structural similarity to early developmental lignins). *Plant Physiol* 108: 1277–1287
- Lafarguette F., Leplé J.C., Déjardin A., Laurans F., Costa G., Lesage-Descauses M.C. and Pilate G. (2004). Poplar genes encoding fasciclinlike arabinogalactan proteins are highly expressed in tension wood. *New Phytol* 164: 107–121
- Li J., Yang X., Wang Y., Li X., Gao Z., Pei M., Chen Z., Qu L.J. and Gu H. (2006). Two groups of MYB transcription factors share a motif which enhances trans-activation activity. *BBRC* 341: 1155–1163
- Morita Y., Saitoh M., Hoshino A., Nitasaka E. and Iida S. (2006). Isolation of cDNAs for R2R3-MYB, bHLH and WDR Transcriptional regulators and identification of c and ca mutations conferring white flowers in the Japanese Morning Glory. *Plant Cell Physiol* **47**: 457–470
- Moyano E., Martinez-Garcia J.F. and Martin C. (1996). Apparent redundancy in myb gene function provides gearing for the control of flavonoid biosynthesis in Antirrhinum flowers. *Plant Cell* 8: 1519–1532
- Nersissian A.M., Immoos C., Hill M.G., Hart P.J., Williams G., Herrmann R.G. and Valentine J.S. (1998). Uclacyanins, stellacyanins, and plantacyanins are distinct subfamilies of phytocyanins: Plant-specific mononuclear blue copper proteins. *Protein Sci* 7: 1915–1929
- Newman L.J., Perazza D.E., Juda L. and Campbell M.M. (2004) Involvement of the R2R3-MYB, At MYB61, in the ectopic lignication and darkphotomorphogenic components of the det3 mutant phenotphype. *Plant J* 37: 239–250
- Patzlaff A., Newman L.J., Dubos C., Whetten R.W., Smith C., McInnis S., Bevan M.W., Sederoff R.R. and Campbell M.M. (2003a). Characterisation of PtMYB1, an R2R3-MYB from pine xylem. *Plant Mol Biol* 53: 597–608

- Patzlaff A., McInnis S., Courtenay A., Surman C., Newman L.J., Smith C., Bevan M.W., Mansfield S.D., Whetten R.W., Sederoff R.R. and Campbell M.M. (2003b) Characterisation of a pine MYB that regulates lignification. *Plant J* 36: 743–754
- Paux E., Carocha V., Marques C., Mendes de Sousa A., Borralho N., Sivadon
 P. and Grima-Pettenati J. (2005). Transcript profiling of Eucalyptus xylem genes during tension wood formation. New Phytol 167: 89–100
- Pavy N., Paule C., Parsons L., Crow J.A., Morency M.J., Cooke J.K., Johnson J.E., Noumen E., Guillet-Claude C., Butterfield Y., Barber S., Yang G., Liu J., Stott J., Kirkpatrick R., Siddiqui A., Holt R., Marra M., Séguin A., Retzel E., Bousquet J. and MacKay J. (2005) Generation, annotation, analysis and database integration of 16,500 white spruce EST clusters. *BMC Genomics* 6: 144
- Perez-Rodriguez M., Jaffe F.W., Butelli E., Glover B.J. and Martin C. (2005). Development of three different cell types is associated with the activity of a specific MYB transcription factor in the ventral petal of *Antirrhinum majus* flowers. *Development* 132: 359–370
- Qu L.J. and Zhu Y.X. (2006). Transcription factor families in Arabidopsis: major progress and outstanding issues for future research. *Curr Opin Plant Biol* 9: 544–549
- Rogers L.A. and Campbell M.M. (2004). The genetic control of lignin deposition during plant growth and development. *New Phytol* 164: 17–30
- Rozen S. and Skaletsky H. (2000). Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* 132: 365–386.
- Scarpella E. and Meijer A.H. (2004). Pattern formation in the vascular system of monocot and dicot plant species. *New Phytol* 164: 209–242
- Schumacher K., Vafeados D., McCarthy M., Sze H., Wilkins T. and Chory J. (1999). The Arabidopsis det3 mutant reveals a central role for the vacuolar H. (+)-ATPase in plant growth and development. *Genes Dev* 13: 3259–3270
- Schwechheimer C., Zourelidou M. and Bevan M.W. (1998). Plant transcription factor studies. *Annu Rev Plant Physiol Plant Mol Biol* 49:127–150
- Sharman B.C. (1943). Tannic acid and Iron alum with safranin and Orange G in studies of the shoot apex. *Stain Technol* 18: 105–111
- Shiraiwa Y. and Miyachi S. (1985). Role of carbonic anhydrase in photosynthesis of blue-green alga. (*Cyanobacterium*). Anabaena variabilis ATCC 29413. Plant Cell Physiol 26: 109–116
- Smyth G.K. (2005). Limma: linear models for microarray data. In R Gentleman, V Carey, S Dudoit, R Irizarry, W Huber, eds, *Bioinformatics and Computational Biology Solutions using R and Bioconductor, Springer, New York*, p 397–420
- Stracke R., Werber M. and Weisshaar B. (2001). The R2R3-MYB gene family in *Arabidopsis thaliana*. *Curr Opin Plant Biol* 4: 447–456
- Stewart J.J., Kadla J.F. and Mansfield SD. (2006). The Influence of lignin chemistry and ultrastructure on the pulping efficiency of clonal aspen. (*Populus tremuloides* Michx.). *Holzforschung* 60: 111-122
- Tamagnone L., Merida A., Parr A., Mackay S., Culianez-Macia F.A., Roberts K. and Martin C. (1998). The AmMYB308 and AmMYB330 transcription factors from *Antirrhinum* regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. *Plant Cell* 10: 135–154

- Topfer R., Maas C., Horicke-Grandpierre C., Schell J. and Steinbiss H.H. (1993). Expression vectors for high-level gene expression in dicotyledonous and monocotyledonous plants. *Methods Enzymol* 217: 67– 78
- **Vom Endt D., Kijn, J.W. and Memelink J. (2002).** Transcription factors controlling plant secondary metabolism: what regulates the regulators? *Phytochemistry* **61:** 107–114
- Yang Y.H., Dudoit S., Luu P., Lin D.M., Peng V., Ngai J. and Speed T.P. (2002). Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* 30: e15
- Yeh T.F., Braun J.L., Goldfarb B., Chang H.M. and Kadla J.F. (2006). Morphological and chemical variations between juvenile wood, mature wood, and compression wood of loblolly pine (*Pinus taeda* L.). *Holzforschung* 60: 1–8
- Zhang J.Z. (2003). Overexpression analysis of plant transcription factors. *Curr* Opin Plant Biol 6: 430–440
- Zhang Y., Sederoff R.R. and Allona I. (2000). Differential expression of genes encoding cell wall proteins in vascular tissues from vertical and bent loblolly pine trees. *Tree Physiol* 20: 457–466
- Zhong R.Q., Ripperger A. and Ye Z.H. (2000). Ectopic deposition of lignin in the pith of stems of two Arabidopsis mutants. *Plant Physiol* 123: 59–69

CHAPITRE IV

Résultats complémentaires sur la caractérisation phénotypique et fonctionnelle des jeunes épinettes surexprimant le gène *PtMYB1*

IV-1-Avant-propos

Ce chapitre n'a pas fait l'objet de publication ou de soumission dans des revues scientifiques. Les résultats présentés sont complémentaires à ceux du chapitre III réalisés sur des plantules transgéniques *in vitro* et découlent de la caractérisation d'épinettes transgéniques surexprimant le gène *PtMYB1* (*Pro_{UBI}PtMYB1*) après leur transfert en sol et culture en serre. Les analyses décrites dans ce chapitre IV sont exclusivement réalisées sur les pousses terminales d'arbres transgéniques en fin de troisième cycle de croissance en serre et exprimant la construction *Pro_{UBI}PtMYB1*. Il s'agit aussi d'un travail préliminaire à l'analyse de la dérégulation de l'expression des gènes dans le xylème secondaire de jeunes épinettes « Pro_{UBI}PtMYB1 ». J'ai réalisé toutes les étapes de l'analyse et de la rédaction. La transgénèse et l'entretien des plants ont été effectués au Centre de Foresterie des Laurentides à Québec, au sein de l'équipe du Dr. Armand Séguin.

IV-2-Résumé

Le rôle du gène PtMYB1 a été analysé par surexpression constitutive dans de jeunes épinettes blanches en deuxième et troisième cycle de croissance en serre. Des études antérieures avaient montré la capacité de la protéine PtMYB1 à fixer les éléments AC présents dans les promoteurs des gènes codant la phénylalanine amonia lyase (PAL) de haricot et la glutamine synthetase (PsGS1b) de pin sylvestre, impliqués respectivement dans la synthèse des phénylpropanoïdes et le recyclage de l'azote, pour en activer la transcription (Patzlaff et al., 2003; Gomez-Maldonaldo et al., 2004). Dans ce chapitre, le suivi du développement de jeunes arbres transgéniques surexprimant PtMYB1 indique une réduction de la croissance en hauteur et en largeur qui semble être liée au niveau d'expression du transgène. Malgré l'absence de différences histologiques dans les pousses terminales conparativement aux arbres témoins, l'analyse par PCR quantitative des gènes CAD et CCR liés à la synthèse des lignines montre clairement une augmentation de leurs niveaux d'expression. Les gènes codant des glutamine synthétases voient aussi leurs niveaux de transcrits significativement augmentés. En revanche, nous n'avons pas mis en évidence la dérégulation de l'expression des facteurs de transcription PgMYB2, 4 et 8.

IV-3-Introduction et mise en contexte

Les plantules « Pro_{UBI}PtMYB8 » décrites dans le chapitre III présentent des phénotypes hyperlignifiés, non-viables. Nous n'avons donc pas pu étudier leur développement passé le stade de la germination, ce qui aurait permis de mieux caractériser le rôle de PtMYB8 dans la différenciation du xylème secondaire et la lignification du bois d'épinette. En revanche, les plantules « Pro_{UBI}PtMYB1 » ont pu être étudiées au cours de leur développement.

Les plantules « Pro_{UBI}PtMYB1 » ont été transférées en sol et cultivées en serre tel que décrit dans le chapitre III. Après un premier cycle de croissance en serre (environ trois à quatre mois), suivi d'environ 6 à 8 semaines de dormance en chambre froide (à +4°C), elles ont été transférées dans des pots de 3 litres (entretien cf. chapitre II). Les observations de sections transversales prélevées à la base des arbres « Pro_{UBI}PtMYB1 » à la fin du premier cycle de croissance n'ont pas mis en évidence de différences histologiques (e.g. épaississement des parois cellulaire du xylème) par comparaison aux témoins (résultats non montrés). Les arbres ont ainsi été maintenus pour un deuxième et troisième cycle de croissance (environ 3 mois chacun), pendant lesquels nous avons fait des observations phénotypiques générales (hauteurs et diamètres des jeunes arbres). A la fin du troisième cycle de croissance, nous avons prélevé les jeunes pousses terminales de cinq arbres dans les lignées 10-Témoin (non transformées) et dans les lignées transgéniques 10-11 et 10-14 qui présentaient un stade de croissance identique. Ces échantillons ont servi à réaliser des analyses préliminaires en histologie et en PCR quantitative afin de voir si l'expression du transgène était en mesure d'induire des différences comparativement aux témoins.

IV-4-Matériels et méthodes

Les ARN totaux ont été extraits suivant le protocole décrit au Chapitre II (Annexe II-1). La concentration et la qualité des ARN ont été determinées avec un bioAnalyseur (model 2100, Agilent Technologies; RNA 6000 Nano AssayKit). L'analyse de transcrits d'ARN par RT-PCR quantitative a été effectuée comme décrit dans le chapitre V avec la trousse Qiagen et l'appareil Roche LC480. Les paires d'amorces ont été conçues à l'aide du logiciel Primer 3 (Rozen et Skaletsky, 2000) pour se fixer dans la région UTR-3' des gènes transcrits, dans le but de favoriser une plus grande spécificité. Les paires d'amorces sont indiquées ci-après (le numéro d'identification du clone est indiqué entre parenthèses ainsi que la longueur de l'amplicon), et les acronymes des gènes sont expliqués dans la légende de la figure I-6:

PtMYB1 (les amorces s'hybridant dans la région du terminateur 35S de la construction ProUBIPtMYB1 ont été utilisé, pas d'identifiant de clone, 101 pb): 5'-CAAAAATCACCAGTCTCTCTC-3' 5'et ACCCTATAAGAACCCTAATTCC-3'; PgMYB14a (pas d'identifiant de clone, 153 5'-AGGGATGAAGGGAACCATCG-3' 5'pb): et TGAATTATAGCGGTGGACTC-3'; PgMYB14b (pas d'identifiant de clone, 120 5'-AGGGATGAAAGGAACGCG-3' 5'et pb): ATAAGCAATCCATTAATCTTCC-3'; C4H (GQ0021b H10, 91 pb): 5'-TCTCACTTAAGTCGGTCCAG-3' et 5'-CTTCCTGGTAATGCAAAAAG-3'; COMT (GQ0071.TB A01, 142 pb): 5'-TCATTGGGGGCTTGTAATG-3' et 5'-ACATGACTTCCAGAGATAAAGC-3'; CCR (GQ00411 J17, 134 pb): 5'-CCCTCAGTATCCAATTCCTACC-3' et 5'-CTGTTTCATACAGGCATTGTTTC-3'; GS1a (GQ026M07 E03, 136 pb): 5'-TCGAGGAGCTTCAGTTAGAG-3' et 5'-GCTTCCATAGAATTGTCGTC-3'; GS1b (GQ0198.B3 G17, 119 pb): 5'-CTCAGTTGAATTTCTCTGTGTG-3' et 5'-ACACAAATCTGAATGACAAACTAC-3'; CHS8 (GQ02011 A08, 85 pb): 5'-TTCCAGCAGTTCGGAATCTC-3' et 5'-CCTCCACCTGATCCAGAATG-3'; SUSY (GQ0045 D20, 128 pb): 5'-ATGAGCAAAGATAGAATCAGG-3' et 5'-GATTGGGCTCTTCTCTTTTAC-3'. Les séquences des paires d'amorces des gènes CAD, 4CL, EF1-a, PgMYB1 à PgMYB4 et PgMYB8 sont précisées dans le chapitre II.

IV-5-Résultats

IV-5-1-Phénotypes des arbres surexprimant PtMYB1

Les jeunes arbres d'épinette blanche en deuxième et troisième cycle de croissance surexprimant *PtMYB1* montrent clairement une réduction de la taille (hauteur totale et apex) et du diamètre, pour plusieurs évènements de transformation indépendants (Tableau IV-1 et Figure IV-1). De plus, on observe un dessèchement prématuré des aiguilles sur les branches des cycles de croissance antérieurs des arbres Pro_{UBI}PtMYB1, par comparaison aux arbres témoins (Figure IV-1 A et B). Enfin, les observations histologiques faites sur des sections transversales à mi-hauteur des jeunes pousses terminales des lignées 10-témoin, 10-10 et 10-14 n'ont pas permis de mettre en évidence de phénotype affectant le développement ou la structure du tissu vasculaire, ou encore de la paroi cellulaire secondaire du xylème (Figure IV-2).

	Second cycle	e de croissance	Troisième cycle de croissance			
Lignées	Hauteur sans pousse terminale (cm)	Longueur de la pousse terminale (cm)	Hauteur sans pousse terminale (cm)	Longueur de la pousse terminale (cm)	Diamètre au dessus du collet (mm)	
10-Témoin	19,225 +/- 3,4	3,79 +/- 0,79	23,89 +/- 3,79	2,18 +/- 0,82	7,81 +/- 0,6	
10-4	3,47 +/- 2,65	1,76 +/- 0,74	6,87 +/- 1,85	1,66 +/- 0,71	2,54 +/- 0,64	
10-10	10,93 +/- 1,9	2,19 +/- 0,9	13,55 +/- 2,64	2,38 +/- 1,06	5,02 +/- 0,61	
10-11	10,17 +/- 1,35	2,11 +/- 1,31	12,3 +/- 1,65	2,13 +/- 1,07	5,07 +/- 0,6	
10-14	6,95 +/- 3,11	3,15 +/- 0,87	9,46 +/- 3,28	1,37 +/- 1,07	3,79 +/- 0,57	

Tableau IV-1. Hauteurs et diamètres des épinettes transgéniques exprimantla construction ProUBIPtMYB1.

Les plants ont été transplantés en pot de 3 litres le 27 mai 2005 et maintenus en serre, ou en chambre froide pour la dormance. Chaque moyenne et écart-type a été calculé à partir de 20 arbres pour le deuxième cycle de croissance (le 21 juin 2006) et des mêmes arbres pour le troisième cycle de croissance (26 février 2007) exception faite des lignées 10-témoin, 10-11 et 10-14 comprenant 15 arbres au troisième cycle.





A. Arbres en troisième cycle de croissance « Pro_{UBI}PtMYB1 », lignées 10-Témoin (non transformées), 10-4, 10-10, 10-11 et 10-14; B. Gros plan de deux arbres de la lignée 10-4 montrant le dessèchement graduel des aiguilles apparues sur les deux cycles de croissance antérieurs.



Figure IV-2. Sections transversales des pousses terminales des épinettes « Pro_{UBI}PtMYB1 ».

Arbres de deuxième cycle de croissance « $Pro_{UBI}PtMYB1$ », lignées 10-T (témoin, non-transformé), 10-11 et 10-14. Les sections transversales ont une épaisseur de 15 µm et sont traitées à la safranine orangée (Sharman, 1943), qui colore en rose les parois lignifiées (grossissement 10x).

IV-5-2-Nuveaux de transcrits de gènes reliés à la formation de la paroi dans les pousses terminales

L'analyse de l'accumulation des transcrits par RT-PCR quantitative montre que le transgène *PtMYB1* est plus fortement exprimé que le gène endogène présumé *PgMYB1*, de 1,5 à 3 fois respectivement dans les lignées 10-10 et 10-14 (Figure IV-3). Nous avons étudié les niveaux d'expression de gènes du métabolisme des phénylpropanoïdes (*4CL*, *COMT*, *C4H*) et de la synthèse des monolignols (*CAD* et *CCR*), du métabolisme de l'azote (*GS1a* et *GS1b*), du métabolisme du sucrose (*SUSY*) et du métabolisme des flavonoïdes (*CHS*). De plus, nous avons analysé des facteurs de transcription *MYB-R2R3* précédement décrits dans les chapitres II et III. Les résultats montrent que les niveaux d'expression des gènes *4CL*, *CCR* et *CAD*, *GS1a* et *GS1b* sont significativement augmentés dans les deux lignées transgéniques 10-10 et 10-14 comparativement aux témoins (Figure IV-4). Par contre, les niveaux d'expression des gènes *PgMYB14a* et *14b* sont augmentés (*MYB-R2R3* décrit dans les chapitres V et VI), et l'expression du gène *COMT* est diminuée, seulement pour la lignée 10-14 qui exprime le plus fortement le transgène (Figure IV-4).



Figure IV-3. Niveaux d'expression du transgène *PtMYB1* et de l'endogène présumé *PgMYB1* chez les arbres transgéniques « Pro_{UBI}PtMYB1 ».

L'accumulation des ARN messagers dans les pousses terminales a été déterminée par PCR quantitative sur des arbres en deuxième cycle de croissance de la lignée témoin (10-T) (non transformée, 4 arbres), et sur deux lignées surexprimant *PtMYB1* avec le promoteur *ubiquitine* de maïs (10-11, 5 arbres; 10-14, 4 arbres).





Le matériel végétal utilisé dans cette figure est le même que dans la figure IV-3. Les étoiles indiquent le niveau de signification (*P*-value) du test de Student entre le témoin (10-T) et le transgénique, * :P < 0,05, **: P < 0,01, ***: P < 0,001. La signification des acronymes et noms de gènes 4*CL*, *COMT*, *C*4*H*, *CCR*, *CAD* et *CHS* est expliquée dans la figure I-6 de l'introduction de la thèse. GS1a et *GS1b*: *Glutamine synthase a et b*, *SUSY*: *Sucrose synthase*.

IV-6-Conclusions préliminaires et perspectives

La surexpression de *PtMYB1* a induit une réduction de croissance en hauteur et en diamètre qui se maintient pendant les deuxième et troisième cycles de croissance de jeunes épinettes. Cette réduction de croissance semble être corrélée avec le niveau d'expression du transgène. Les observations histologiques ne permettent pas d'associer PtMYB1 à un rôle dans la lignification et la différenciation du xylème primaire dans les pousses terminales. Toutefois l'expression légèrement augmentée des gènes CCR et CAD suggère une lignification potentiellement accrue. De plus, contrairement à ce que nous avions observé chez les plantules in vitro « Pro_{UBI}PtMYB1 » du chapitre III, nous n'avons observé aucun épaississement de la paroi cellulaire, ni aucune altération dans l'organisation des tissus vasculaires dans les nouvelles pousses des plants. Cette différence pourrait être explicable par un niveau d'expression de PtMYB1 qui n'est pas largement supérieur à l'endogène présumé PgMYB1, comparativement aux plantules in vitro où la différence est très grande (voir chapitre III). De plus, les conditions de culture (e.g. plus de sucrose dans le milieu in vitro que dans la terre des plants en serre), le stade de développement et l'âge des plants peuvent aussi être des paramètres à l'origine de la disparition du phénotype observé dans les plantules in vitro..

Des études récentes ont montré que PtMYB1 est capable d'interagir avec des éléments *cis* régulateurs en amont du gène *glutamine synthase GS1b* pour en réguler l'expression (Gomez-Maldonaldo et *al.*, 2004). La glutamine synthase permet la condensation de l'ammonium et du glutamate en glutamine dans une réaction ATP-dépendante. Chez le Pin sylvestre, deux gènes différents codent pour GS1, *PsGS1a*, exprimé exclusivement dans les tissus parenchymateux à vocation photosynthétique, et *PsGS1b*, exprimé principalement dans les tissus vasculaires de la plante (Avila et *al.*, 2001). Dans notre étude, l'augmentation des niveaux de transcrits des gènes *GS1a* et *GS1b* appuie le rôle de PtMYB1 (ou d'une séquence similaire) dans l'assimilation de l'ammonium et potentiellement dans le recyclage de l'azote dans les tissus en lignification, si on tient compte de son expression tissulaire préférentielle dans le xylème (chapitre II et III).

Les analyses préliminaires de ce chapitre associées aux précédentes observations du chapitre III au stade plantule dressent les perspectives des études futures qui pourraient être entreprises sur les arbres transgéniques « Pro_{UBI}PtMYB1 » encore disponibles :

- Récolter le xylème secondaire dans le but d'analyser si le niveau d'expression du transgène est supérieur au niveau d'expression de l'endogène des individus témoins. Si tel est le cas, une étude du transcriptome du xylème secondaire de ces arbres pourrait être envisagée par microarray. L'objectif étant de connaître les gènes dérégulé par PtMYB1 dans des tissus ou se déroule la formation de la paroi secondaire du xylème.
- Analyser le contenu en lignine du xylème secondaire.
- Analyser le rapport carbone/azote des aiguilles et de l'écorce sous l'hypothèse que PtMYB1 est impliqué dans le système de recyclage de l'azote en se fixant aux éléments AC en amont des gènes *PAL* et *glutamine synthase* (Gomez-maldonaldo et *al.*, 2004; Patzlaff et *al.*, 2003a).

IV-7-Références

- Avila C., Suárez M.F., Gómez-Maldonado J. et Cánovas F.M. (2001) Spatial and temporal expression of two cytosolic glutamine synthetase genes in Scots pine: functional implications on nitrogen metabolism during early stages of conifer development. *Plant J* 25: 93–102
- Gomez-Maldonaldo J., Avila C., de la Torre F., Canas R., Canovas F.M. et Campbell M.M. (2004). Functional interactions between a glutamine synthetase promoter and MYB proteins. *Plant J* 39: 513–526
- Patzlaff A., Newman L.J., Dubos C., Whetten R.W., Smith C., McInnis S., Bevan M.W., Sederoff R.R. et Campbell M.M. (2003a). Characterisation of PtMYB1, an R2R3-MYB from pine xylem. *Plant Mol Biol* 53: 597–608
- Rozen S. et Skaletsky H. (2000). Primer3 on the WWW for general users and for biologist programmers. In: S Krawetz and S Misener (Eds), Bioinformatics Methods and Protocols: Methods in Molecular Biology Humana Press, Totowa, NJ:365-386
- Sharman B. C. (1943). Tannic acid and iron alum with safranin and orange G in studies of the shoot apex. *Stain Techn* 18:105-111

CHAPITRE V

Pinus taeda MYB14, a member of a new phylogenetic clade of R2R3 *MYB* genes, is identified as a putative regulator in defence-response in pine and spruce

Frank Bedon^{a,b}, Claude Bomal^a, Sébastien Caron^a, Shawn Mansfield^d, Axel Schmidt^e, Caroline Levasseur^c, Brian Boyle^a, Jacqueline Grima-Pettenati^b, Armand Séguin^c and John MacKay^a

^aCentre d'Étude de la Forêt, Université Laval, Québec (QC), G1K 7P4, Canada. ^bUMR UPS/CNRS 5546, Pôle de Biotechnologies Végétales, 24 chemin de Borde Rouge, BP42617, Auzeville Tolosane, 31326 Castanet Tolosan, France ^cNatural Resources Canada, Canadian Forest Service, Laurentian Forestry Centre, Québec (QC), G1V 4C7, Canada

^dCanada Research Chair in Wood and Fibre Quality, Department of Wood Science, University of British Columbia, 4030-2424 Main Mall, Vancouver (BC), V6T 1Z4, Canada

^eMax Planck Institute- for Chemical Ecology, Hans-Knoell-Str.8, Beutenberg-Campus, D-07745 Jena, Germany

V-1-Avant-propos

Ce chapitre sera prochainement soumis pour publication à la revue « Plant Cell ». Les résultats ont fait l'objet d'une présentation par affiche lors du XV ème congrés de la «Federation of European Societies of Plant Biology» du 17 au 21 juillet 2006 à Lyon. J'ai cloné le promoteur de l'alcool cinnamylique deshydrogénase d'épinette blanche et testé son activité. J'ai isolé et analysé les séquences du sousgroupe 4 de conifère, j'ai aussi réalisé le travail d'identification, de clonage et de PCR quantitative sur tous les gènes présentés dans ce chapitre. J'ai participé à la mise en place et la réalisation des dispositifs expérimentaux avec l'aide de Dr. C. Bomal. J'ai joué un rôle majeur dans l'interprétation des résultats et la rédaction de cet article. Dr. C. Bomal est intervenu dans le choix des gènes à étudier en PCR quantitative, le traitement et la mise en forme des données de microarray ainsi que leurs interprétations. Les hybridations microarrays et les analyses statistiques ont été effectuées par S. Caron. Dr. S. Mansfield et Dr. A. Schmidt sont intervenu dans le dosage, respectivement, des amidons et anthocyanes, et des terpènes. Les étapes de transgénèse ont été effectuées dans le laboratoire de Dr. A, Séguin par C. Levasseur. Dr. B. Boyle a collaboré aux analyses de RT-PCR quantitative.

V-2-Résumé

Nous avons réalisé la caractérisation fonctionnelle d'un facteur de transcription MYB-R2R3 (PtMYB14) de Pinus taeda qui appartient au sousgroupe 4, dont plusieurs membres sont impliqués dans la régulation négative de la biosynthèse des phénylpropanoïdes chez les angiospermes. Les membres de ce sous-groupe isolés chez le pin et l'épinette (*Picea glauca*) montrent des niveaux élevés de conservation de séquence et définissent un nouvel ensemble nommé le sous-groupe 4 de conifères (Sg4-C). Pour caractériser fonctionnellement PtMYB14, nous avons donc choisi une stratégie de surexpression dans le système quasi-homologue que représente l'épinette blanche. La surexpression constitutive de PtMYB14 conduit à une augmentation de l'accumulation des anthocyanines, de l'amidon et des terpènes (mono-et sesquiterpènes) chez les plantules d'épinettes transgéniques comparativement au témoins. Nous avons utilisé en parallèle le promoteur du gène codant l'alcool cinnamylique deshydrogénase (CAD) de l'épinette blanche de façon à diriger plus spécifiquement l'expression de PtMYB14 dans les cellules du xylème en différenciation. Les plantules transgéniques obtenues n'ont pas de phénotypes aussi forts que celles qui expriment constitutivement PtMYB14 mais elles accumulent aussi davantage de sesquiterpènes que les témoins. Des expériences microarrays couplées à des analyses de RT-PCR quantitative sur les deux types de plantules transgéniques révèlent que, les gènes dérégulés en commun sont associés avec l'activation des mécanismes de défense et plus particulièrement avec les gènes de la biosynthèse des isoprénoïdes et des terpènes. Chez les conifères, les composés de défense comme les terpènes (mono-, di- et sesquiterpènes) s'accumulent dans certains tissus en réponse à des blessures mécaniques et à l'application d'acide jasmonique, stress qui provoquent également des changements anatomiques. Nous avons montré ici que ce type de stress appliqué à des plantules sauvages de pin et d'épinette conduit à l'accumulation de plusieurs des transcrits impliqués dans la synthèse des isoprénoïdes et des terpènes, ainsi que de membres du sous groupe MYBs Sg4-C, et en particulier les gènes MYB14.

V-3-Abstract

We investigated the function of the Pinus taeda MYB14 (PtMYB14) a transcription factor that belongs to a new subclade of the subgroup 4 of the R2R3-MYB family. In angiosperms, several members of subgroup 4 are involved in the negative regulation of phenylpropanoid biosynthesis. Members of subgroup 4 isolated from pine and spruce (Picea glauca) showed a high level of sequence conservation, and define a new clade specific to conifers, named subgroup 4 of conifers (Sg4-C). To functionnaly characterize PtMYB14 we thus chose an overexpression strategy in the white spruce homologous-like system. Constitutive strong overexpression of PtMYB14 lead to transgenic spruce plantlets with increased accumulation of anthocyanins, starch and terpenes (mono- and sesquiterpenoids) as compared to controls. In parallel we used the *cinnamvl* alcohol dehydrogenase (CAD) gene promoter of white spruce to drive preferential expression of *PtMYB14* in differentiating xylem cells. The resulting transgenic plantlets did not show the strong phenotype observed in the transgenics expressing *PtMYB14* constitutively but they also accumulated more sequiterpenoids than the controls. Microarrays experiments coupled to quantitative PCR analyses on both types of transgenics plantlets revealed that overlapping misregulated genes were associated with the activation of the defence mechanisms and more particularly with isoprenoid and terpene biosynthesis genes. In conifers, defense compounds like terpenes (mono-, di- and sesquiterpenoids) are accumulated in response to insects attack or mechanical wounding and methyl jasmonate application, resulting also in anatomical changes. We also showed that mechanical wounding and methyl-jasmonate application in wild-type pine and spruce plantlets led to the accumulation of many of the same transcripts involved in isoprenoids and terpenoids synthesis, in addition to members of the MYBs Sg4-C, most notably the MYB14 genes.

V-4-Introduction

Secondary metabolism in plants is the target of several regulatory mechanisms, including most notably transcriptional regulators such as the MYB transcription factor. MYBs belong to one of the largest classes of regulatory proteins able to control biological processes through coordinate gene transcription activation or repression (Riechmann *et al.*, 2000). Indeed, the regulation of many developmental and biosynthetic pathways have been demonstrated to involve one or several MYBs. These include seed development and germination (Penfield *et al.*, 2001), epidermal cell fate (Lee and Schiefelbein, 1999), and several aspects of vascular structure (Scarpella and Meijer, 2004), as well as the production of phenylpropanoids (Rogers and Campbell, 2004), monolignols (Goiecoechea *et al.*, 2005), flavonoids (Stracke *et al.*, 2007) and benzenoids (Verdonk *et al.*, 2005). Several studies indicate that members of the MYB family are under the control of environmental conditions and internal physiological signals (Kranz *et al.*, 1998, Vom Endt *et al.*, 2002).

Members of the MYB protein family are characterized by the highly conserved DNA binding domain, or MYB domain, which is able to bind to cognate cis-regulatory DNA elements usually located in the promoter region of target genes (Romero et al., 1998). According to the number of MYB repeats, ranging from one to three, they are organized into three subfamilies (Jin et al., 2000). Proteins with two MYB repeats (R2R3-MYB) are by far the most abundant type in plants and are estimated to be represented by 126 distinct genes in Arabidopsis thaliana (Yanhui et al., 2006). Plant R2R3 MYB genes can further be divided into 22 subgroups based on aminoacids sequences similarities due to conserved motifs located in their carboxy-terminus (Kranz et al., 1998; Stracke et al., 2001; Jiang et al., 2004). Modulation of gene transcription by MYBs involves the carboxy-terminal region of the proteins in which diverse aminoacids motifs may be directly implicated. Nevertheless relatively few aminoacids regions in the MYB domain or in the carboxy-terminal have been tested to establish activation (Li et al., 2006), repression (Jin et al., 2000) or interaction activities (Zimmermann et al., 2004; Yoo et al., 2005).

Multigenic families like MYB transcription factors present a well-known challenge when it comes to characterize their function due to the genetic redundancy and to the high level of sequence conservation among members of a particular subgroup or clade (Bouché and Bouchez, 2001; Briggs et al., 2006; Pickett and Meeks-Wagner, 1995). This phenomenon is believed to be responsible for the lack of phenotypes in several single mutants. In Arabidopsis, this problem has been circumvented through the use of double or triple mutants to overcome functional redundancy (Briggs et al., 2006). For example, the Arabidopsis single mutants of AtMYB33 and AtMYB65 both belonging to subgroup 18, have no discernable phenotype but the double mutant Atmyb33 myb65 exhibited abnormal anther development (Millar and Gubler, 2005). This drawback of loss of function experiments was also shown in the investigations of transcription factors belonging to such families as the MADS domain (Lehti-Shiu et al., 2005), bHLH (Wang et al., 2007) and Aux/IAA (Overvoorde et al., 2005). Therefore, various gain of function experiments with tissue specific or inducible promoters and molecular dominant gene constructs have been used to gain valuable insights into gene function (Moore et al., 2006).

In a previous report, we described R2R3-MYBs from pine (Pinus taeda PtMYB14) as well as three from white spruce (*Picea glauca* PgMYB5, 10 and 13) that are phylogenetically related to the Arabidopsis MYB4 involved in phenylpropanoid gene regulation, and belong to the plant R2R3-MYB subgroup 4 (Bedon et al., 2007). In this study, we used degenerated primers based on previously identified conifers R2R3-MYB genes from subgroup 4 to isolate several additional conifer R2R3-MYBs from *Pinus taeda* and *Picea glauca*. In a single species (P. glauca), we identified as many as nine distinct sequences which form a monophyletic clade within subgroup 4. We named this clade subgroup 4 of conifers (Sg4-C). The pine MYB14 gene was used for overexpression studies in stably transformed P. glauca, as we anticipated that the members of this clade would be functionally redundant and thus would not lend themselves well to lossof-function studies using methods like RNAi. Two kinds of promoters were used : the constitutive Ubiquitin maize gene promoter (Pro_{UBI}, Christensen et al., 1992, 1996) and a promoter from the P. glauca gene encoding cinnamyl alcohol deshydrogenase (Pro_{CAD}) gene which we showed to confer preferential expression

in vascular tissue (i.e. differentiating xylem). The $Pro_{UBI}PtMYB14$ and $Pro_{CAD}PtMYB14$ transgenic plantlets accumulated terpenes, one of the main compounds of conifer defence mechanisms. They also accumulated starch and developed more polyphenolic parenchyma cells. Microarray transcript profiling and quantitative RT-PCR on the Pro_{UBI} and $Pro_{CAD}PtMYB14$ plantlets strongly suggest that PtMYB14 participates in the regulation of defence genes pathways in conifers. To support this hypothesis we also examined the expression of *MYB14* genes and other Sg4-C MYBs from pine and spruce in response to mechanical wounding and JA application.

V-5-Materials and Methods

V-5-1-DNA cloning and accession number

In a previous studies we identified the pine R2R3 PtMYB14 gene (AC: DQ399056) and three other spruce related genes : PgMYB5 (AC: DQ399069), 10 (AC: DQ399064) and 13 (AC: DQ399061) with elevated amino acid similarity levels (Bedon et al., 2007). Based on these genes and other conifer nucleotide gene sequences (PtMYB1 and PtMYB4, Patzlaff et al., 2003a and b; PmMBF1, Xue et al., 2003) we performed 3' Rapid Amplification of cDNA Ends (3'RACE) with a degenerated primer in the more specific region of the PtMYB14 related genes, including PgMYB5, 10 and 13: 5'- TGGCGYTCSCTTCCCAAGGCC-3' (Y = C, T or U, S = C or G) corresponding to the amino acid sequence WRSLPKAA in the second helix of the R2 MYB repeat (Bedon et al., 2007). 3'RACE cDNA libraries were produced from total RNA extracted from organs and tissues of pine and spruce separately, ground in liquid nitrogen with pestle and mortar following the procedure of Chang et al. (1993) as described in Bedon et al. (2007). 3'-RACE cDNA synthesis and PCR was carried out using the SMART RACE cDNA Amplification Kit (BD Biosciences Clontech, CA, USA) with 1 ug of total RNA from (1) wounded terminal leader (new flush of the year), needles and whole stem of one year old white spruce, and (2) non wounded terminal leader and bark of one year old pine. Touchdown PCR for the nested 3'-RACE reaction was performed using a DNA engine PTC-225 thermal cycler (Biorad, Hercules, CA) with the following parameters: five cycles of two steps at 94°C for 30 sec and 72°C for 3 min, five cycles of tree steps at 94°C for 30 sec,

70°C for 30 sec and 72°C for 3 min and twenty five cycles of tree steps at 94°C for 30 sec, 68°C for 30 sec and 72°C for 3 min. Amplification products were run on 1% agar gels, several gel blocks containing amplification products were removed and gel extracted (Gel Extraction Kit, QIAquick, Qiagen, Mississauga, CA). The eluted PCR products were concentrated under air dry vacuum, ligated to pCR2.1, and transformed into TOPOF' competent cells (TA cloning Kit, Invitrogen, Carlsbad, CA) prior to sequencing.

The numbering of newly identified *MYB* genes is in accordance with previous studies and was established on the basis of putative orthology between pine and spruce sequences as well as similar sequences in both species (Patzlaff *et al.*, 2003a and b; Bedon *et al.*, 2007). We isolated six *MYB* genes in pine (*PtMYB5*, 10a, 10b, 13, 15 and 17) and six in spruce (*PgMYB*14a, 14b, 15, 16a, 16b and 16c). The PgMYB5b cDNA sequence was identified by database mining of the Arborea spruce EST project (GQ0045.B3_G24) and full length sequencing. These sequences will be submitted to GenBank database.

Cloning of the pine genes of the mevalonate (MEV) pathway (cytosolic mevalonate pathway named MVA, plastidic non-mevalonate pathway named MEP or DOX) was necessary to prepare standard curves for Q-RTPCR. We used of the TIGR contig sequences pine index gene (http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=pine) to design primer pairs with the Primer 3 software (Rozen and Skaletsky, 2000). Primers for all of the spruce genes that were analysed by RT-PCR were determined from the sequences of the cDNA clones that were spotted on the microarray (see further). Appendix V-1 lists the nucleotide sequence of the primers used to amplify the ten pine genes from the MEV pathway. PCR was performed in a 50 µL reaction mixture with 0.2 µM of each forward/reverse primers, 2 mM MgSO₄, 0.2 mM dNTPs at 10 mM each, 1 µL of pine cDNA, and 1 unit of Platinum ® Taq DNA polymerase High Fidelity (Invitrogen, CAT#11304-011). The cycling conditions were 94°C for 15 min followed by 35 cycles of 94°C for 30 sec, 60°C for 30 sec, 68°C for 2 min, and 68°C for 5 min.

V-5-2-Sequences analysis of the PtMYB14 and related sequences in pine and spruce

Amino acid sequence alignments used for phylogenetic studies were obtained with Clustal W (Thompson et al., 1994) with Bioedit software version 6.0.7 and BoxShade 3.21. Phylogenetic analyses of the partial PtMYB14 related sequences and partial DBD from pine and spruce were carried out with aminoacid sequences using the Mega 2.0 method (Kumar et al., 2001) and using 1000 bootstraps to estimate the node strength (parameters are Poisson correction and pair-wise deletion). Conifers amino acid sequences which are not from this study are described in Bedon et al. (2007). They are PtMYB14 (ABD60279), PgMYB10 (ABQ51226), PgMYB5 (ABQ51221) and PgMYB13 (ABQ51229). Angiosperm gene sequences of the subgroup 4 were found as the most homologous to PtMYB14 after Blast search on the NCBI database (e.g. Fornalé et al., 2006). GenBank accession numbers are TfMYB2 (AAS19476), TfMYB6 (AAS19480), TfMYB1 (AAS19475), GhMYB1 (AAN28270), AmMYB330 (P81395), AtMYB3 (NP564176), AtMYB8 (NP849749), AtMYB6 (AL161515), HvMYB5 (CAA50221), OsAAV59423 (AAV59423), OsNP 915716 (NP915716), OsT02984 (T02984), DspMYB8 (AAO49417) , DspMYB10 (AAO49419), HvMYB1 (CAA50224), TaMYB1 (AAT37167), ZmMYB31 (CAJ42202), ZmMYB42 (CAJ42204), Zm38 (P20025). OsPTYPE1 (AAL84628), OsPTYPE2 (XP483665), EgMYB1 (CAE09058), AtMYB32 (NM119665), AtMYB7 (AAA98762), AtMYB4 (AAS10085), LeMYB27 (CAA64614), AmMYB308 PttMYB4a (AJ567346), GhMYB9 (AAK19619), VvMYB4 (P81393), (ABL61515), ZmMYB8 (CAJ42201). The other sequences used as landmark of distant subgroups are: AtMYB13 (AT1G06180), AtMYB106 (AT3G011340), FaMYB1 (AAK84064), AtMYB20 (AT1G66230), AtMYB61 (AT1G09540), PmMBF1 (U39448), AtMYB123 (AT5G35550), and ZmMYBC1 (M37153).

Pair-wise percentage identity calculations between pine and spruce MYB14 were performed with the Smith-Waterman algorithm from the Emboss package (Matrix: EBLOSUM62, Gap penalty: 3.0, Extend penalty: 0.1). The MEME analysis software (Bailey and Elkan, 1994) was used to identify six regions ranging from 4 to 6 amino acids between the 56 carboxy terminal
sequences of the same sequences used for the phylogenetics analysis. The parameter setting was to find zero or one motif per sequence. The MEME resulting motifs were used as given by the software to build the diagram's presented in figure V-2.

V-5-3-Vectors construction and genetic transformation of white spruce

The Pt*MYB14* cDNA (DQ399056) from *Pinus taeda* L. were used to conduct gain-of-function experiments in white spruce (*Picea glauca* [Moench] Voss). Gain of function was obtained by inserting the full length cDNA in front of the maize ubiquitin promoter (Christensen *et al.*, 1992; 1996), followed at the 3' end by the nopaline synthase terminator. This constitutive expression vector, named pMJM, was obtained by modifying the pRT106 (Topfer *et al.*, 1993) plant expression vector, refered to as the Pro_{UBI}PtMYB14 construct. For *Agrobacterium tumefaciens*-mediated stable transformation, cassettes were digested by *Hin*dIII and cloned into pCambia1305.2 (www.cambia.org). The resulting plasmid was then transferred into the *A. tumefaciens* strain C58 pMP90 (Koncz and Schell, 1986).

The Picea glauca (white spruce) cinnamyl alcohol deshydrohenase promoter (Pro_{CAD}) is 1.12 Kb in length (see chapter VI) and was isolated by a PCR based genome walking procedure using the GenomeWalker Universal Kit (BD Biosciences, USA). The CAD promoter sequence was amplified with primers 5'containing restrictions sites (proCAD-Forward (PstI): CTGCAGTACTATAGGGCACGCGTGGTCGACGGC-3', proCAD-Reverse (EcoRI): 5'-GAATTCTTTTTTAAAACCACTTCACAGGATTCG-3') and was also cloned into the pMJM expression vector after removal of the the ubiquitin promoter sequence, to drive the expression of the MYB14 gene (Pro_{CAD}PtMYB14 construct). In addition, the UBI and CAD promoter sequences were cloned into the pCAMBIA 1391Z in phase with the *uid-A* gene to drive the overexpression of β glucoronidase (GUS) reporter gene.

The white spruce embryogenic line Pg653 was used in the present study, and was initiated and maintained as described by Klimaszewska *et al.* (2001). Genetic transformation was carried out also as described in Klimaszewska *et al.* (2004). After co-cultivation, explants were decontaminated from *A. tumefaciens* with cefotaxim and transferred onto fresh medium containing cefotaxim alone. Kanamycin resistant embryonal tissues were screened according to positive X-gluc tests (Klimaszewska *et al.*, 2004) and assayed for transgene (*PtMYB14* or *GUS*) mRNA accumulation by real-time PCR (see below). Lines exhibiting a range of *PtMYB14* or *GUS* mRNA levels were selected for somatic embryo maturation and somatic seedling production.

V-5-4-Histological analysis, X-gluc histochemical staining, and MUG fluorometric enzymatic assay

Tissue samples (hypocotyl and root) were fixed, and paraffin-embedded as described in Bomal *et al.* (under review, chapter III). For tissue structure observations, paraffin-free sections were prepared by using a microtome and were stained in Sharman's safranin O/orange G/tannic acid (Sharman, 1943). For X-gluc (beta-glucoronidase) histochemical staining assay, samples were prepared as essentially described in Hawkins *et al.* (1997). Briefly, samples were pre-treated for 30 min in cold 90% acetone to prevent any potential wound-induced induction of spruce *CAD* promoter and to facilitate substrate penetration. Samples were then rinsed twice in 100 mM potassium phosphate buffer (pH 7.0) and incubated for 4 h in 5-bromo-4-chloro-3-indolyl-P-D-glucuronic acid reaction medium in darkness at 37°C until sufficient blue color had developed. All sections were observed under a Zeiss Axioskop microscope (Jena, Germany) fitted with a digital camera.

MUG (4-methylumbelliferyl glucuronide) fluorometric enzymatic assay was performed as described in Côté and Rutledge (2003) on hypocotyls, and needles of wild-type (Pg653), Pro_{UBI}, and Pro_{CAD}GUS transgenic spruce plantlets. MUG assay data were obtained from four biological replicates per line, with 25 hypocotyls (including cotyledons) per replicate.

V-5-5-Starch content determination

Soluble metabolites were extracted from the spruce hypocotyl tissue by overnight incubation at -20°C with methanol:chloroform:water (12:5:3). The sample was centrifuged, the supernatant removed, and the remaining pellet washed twice with fresh methanol:chloroform:water (12:5:3). The dried residual pellet was weighed, and then hydrolyzed using 4% sulfuric acid at 121°C for 4 min. The liberation of glucose, representing starch content, was quantified directly by Anion exchange HPLC (Dionex, Sunnyvale, CA) on a DX-600 equipped with a Carbopac PA1 column and an electrochemical detector, as in Coleman *et al.* (2006).

V-5-6-Anthocyanin extraction and analysis

Spruce hypocotyls tissue (~20 mg) was suspended in 1.5 mL of methanol:HCl (95:5) and pulverized at high power for 20 sec in a cell disrupter (FastPrep) and then extracted for 4 h at 40°C in a constant temperature heating block. Following extraction, the homogenate was centrifuged for 10 min at 13,000 rpm, and the supernatant removed from the pellet, dried on a speedvac and resuspended in 200 μ l HPLC grade methanol. Detection and identification of anthocyanins was achieved by LC/MS on an HP 1100 LC-MSD-Trap XCT plus. The methanolic extracts were separated on a SB C-18 Zorbax Rapid Resolution 4.6x150mm 3.5 μ m column at 40°C at a flow rate of 1 mL min⁻¹. Separation was achieved by using a linear gradient from 95% solvent A and 5% solvent B to 75% solvent A and 30% solvent B over 37 min. Solvent A was water with 0.2% formic acid and solvent B was acetonitrile with 0.2 % formic acid. Detection was monitored with a HP 1100 photodiode array detector with the reference wavelength set to 520 nm, while mass determination was achieved by ESI in negative ion polarity.

V-5-7-Extraction of resin terpenes

Terpene extractions were based on the procedures of Martin et al. (2002). All steps were carried out in 2 mL vials with a teflon-coated screw cap (Hewlett-Packard, Palo Alto, CA). Sample pieces (200 mg) were submerged into 1.5 mL of tert-butyl methyl ether containing 150 µg/mL isobutylbenzene and 200 µg/mL pimaric acid as internal standards, and extracted 14 h overnight with constant shaking at room temperature. To purify extracted terpenes from other organic acids, ethereal extract was transferred to a fresh vial and washed with 0.3 mL of $0.1 \text{ M} (\text{NH}_4)_2 \text{CO}_3$ (pH 8.0). The extract was then split into two equal portions. For the analysis of diterpene resin acids, one aliquot was methylated by adding 50 μ L of 0.2 M N-trimethylsulfonium hydroxide in methanol (Macherey-Nagel GmbH, Germany) to 0.4 mL of the washed ethereal extract in a separate vial. After 2-h incubation at room temperature, the solvent was evaporated under nitrogen to leave 100 μ L of sample, which was stored at -20°C until analysis by GC-MS. For the analysis of monoterpenes and sesquiterpenes, the remainder of the original sample was prepared for GC-MS analysis by filtering through a Pasteur pipette column filled with 0.3 g of silica gel (Sigma 60Å) overlaid with 0.2 g of anhydrous MgSO₄. The column was washed with 1 mL of diethyl ether, and the combined eluant collected in a fresh vial and evaporated to an approximate volume of 100 µL and then stored at -20°C until GC-MS analysis.

V-5-8-Analysis of monoterpenes, sesquiterpenes and diterpenes

GC-MS analysis of monoterpenes and sesquiterpenes was carried out with a Hewlett-Packard 6890 GC-MSD system, using a DB-WAX column (0.25 mm × 0.25 μ m × 30 m, J W Scientific, Folsom, CA). Split injections (1 μ L ethereal extract) were made at a ratio of 1:5 with an injector temperature of 220°C. The instrument was programmed to hold an initial temperature of 40°C for 3 min and then increased at a rate of 3°C min⁻¹ to 80°C. The temperature was then increased at a rate of 5°C min⁻¹ to 180°C with a 5 min hold, followed by a final ramp of 15°C min⁻¹ up to 240°C. Helium was used at a constant flow of 1 mL min⁻¹. For compound identification, the MS detector was operated using a mass range of 40-350 for monoterpenes and sesquiterpenes, and of 40-550 for diterpenes. Spectra were collected under standard conditions (electron impact ionization at 70 eV). Identification of terpenes was based on retention time comparison and mass spectra with authentic standards or to mass spectra in the Wiley or National Institute of Standards and Technology libraries.

Analysis of diterpene constituents was performed on the same GC-MS instrument fitted with an HP-5 column (0.25 mm \times 0.25 μ m \times 30 m, Hewlett-Packard). Injections were made with 1 μ L of the concentrated, derivatized ethereal extract. GC-MS split ratios were 1:10 with an injector temperature of 220°C. The instrument was programmed for an initial temperature of 120°C and increased at a rate of 1°C min⁻¹ to 150°C, followed by 5°C min⁻¹ up to 280°C (6 min hold). Helium was used as carrier gas at a constant flow of 1 mL min⁻¹.

GC-MS generated peaks were quantified using Hewlett-Packard Chemstation software. Isobutylbenzene (IBB) was used as the internal standard for quantification of both monoterpenes and sesquiterpenes while the methylated pimaric acid was employed as an internal standard to calculate diterpene concentrations. For quantitative analysis of monoterpenes and sesquiterpenes, the MS detector was operated in the single ion mode monitoring ions at 91, 93 and 161 m/z, corresponding to the internal standard, monoterpenes, and sesquiterpenes, respectively. For diterpene quantification, the following selected ions were monitored: 121 135 m/z, 135 m/z, 239 m/z, 241m/z. Compounds were identified by comparison of retention time to those of authentic standards (Aldrich Chemicals).

V-5-9-RNA Extraction, cDNA Synthesis and real-time PCR analysis

Fifteen transgenic germinants (Pro_{UBI} or Pro_{CAD}PtMYB14) or five wild type (spruce or pine) seedlings from the jasmonate/wound experiment were used per sample to extract total RNA. Samples were ground using a Mixer Mill MM300 engine (Retsch, Germany) in 1.5 mL microtubes. Total RNA was extracted using the procedure of Chang *et al.* (1993) (Appendix II-1), treated with amplification grade DNAse I (Invitrogen, Carlsbad, USA) and quality checked for integrity with a bioAnalyser (model 2100, RNA 6000 Nano Assay kit, Agilent Technologies, Santa Clara, USA). cDNAs were synthesized from one microgram of purified RNA using SuperScript II First-Strand Synthesis System for RT-PCR (Invitrogen) according to

the manufacturer's instructions. For real-time PCR quantification, a five-fold dilution of cDNA mixture was used as template. The PCR reaction mixture (15 µL) consisted on 7.5 µL of 2X Quantitect SYBR Green I mixture (Qiagen, Germantown, USA), 0.9 µL of primers (0.3 µM forward and 0.3 µM reverse), 2 µL of cDNA, and 4.6 µL of RNase-free water. Reactions were prepared in LightCycler® 480 Multiwell plate 384 (Roche, Basel, Switzerland) using a pipetting robot (EpMotion 5075, Eppendorf, Hamburg, Germany). The cycling conditions, conducted with a LightCycler® 480 (Roche), were: 15 min activation period at 95°C followed by 40 cycles (95°C for 10 s; 55°C for 60 s; 72°C for 30 s) with a single fluorescent read taken at the end of each cycle. A melting curve was performed at the end of cycling to ensure that there was single amplification. Crossing point (Cp) values were determined with the LC480 software supplied with the instrument. Standard curves were used to transform Cp values obtained from total RNA samples into transcript number. Standard dilution series covering five orders of magnitude (10^{-1} to 10^{-6} ng/µL) were prepared for target and reference genes from 1 ng/µL of each cDNA cloned in pCR2.1, linearized by EcoRI or BamHI, purified on Qiaquick columns (Qiagen) and product length verified on bioAnalyser (model 2100, DNA 1000 LabChip kit., Agilent Technologies). Transcript levels were normalized with respect to the transcript level of ELONGATION FACTOR 1-alpha (EF1-alpha) or CELL DIVISION CYCLE 2 (CDC2). RNA and DNA concentrations were determined optically with a Multiskan Spectrum spectrophotometer (Thermo Scientific, Waltham, USA). The specific primer pairs of target and reference genes (Appendix V-2) used in pine and spruce were designed using Primer3 software (Rozen and Skaletsky, 2000).

V-5-10-Microarray experiment

All the information regarding microarray manufacture and quality control are detailed elsewhere (Bomal *et al.*, under review, chapter III). The microarray experiment was designed in total compliance with MIAME guidelines. Two transgenic embryogenic lines per construct (lines 16 and 18 for $Pro_{UBI}PtMYB14$, and lines 10 and 14 for $Pro_{CAD}PtMYB14$) as well as an untransformed control line (WT) were used to produce 4-week old somatic seedlings. Total RNA was extracted as described above from four biological replicates per line, with 25 hypocotyls (including cotyledons) per replicate. For each sample, 1 µg of total RNA was used for indirect RNA amplification that was performed with SuperscriptTM Indirect RNA Amplification System kit (Invitrogen, Carlsbad, CA) according to manufacturer's instructions. Dye coupling was performed on 5 μ g of amplified RNA (aRNA) with either Alexa Fluor® 555 or Alexa Fluor® 647 reactive dyes (Invitrogen, Carlsbad CA). Expression profiling was obtained by using a 9K custom cDNA spruce microarray. All RNA processing, and microarray hybridization procedures were described in Bomal *et al.* (under review, chapter III).

V-5-11-Statistical analysis of microarray data

Data were analyzed in R (Ihaka and Gentleman, 1996), mainly using the BioConductor suite of packages (Gentleman *et al.*, 2004). Spot intensities were analyzed with the LIMMA package from Bioconductor (Smyth, 2005). Data normalization was performed using the composite method based on lowess curves (Yang *et al.*, 2002). Normalized data were then statistically analyzed using the linear model and empirical Bayes analyses in LIMMA. The results were corrected using the Benjamini and Hochberg (1995) method of false discovery rate (FDR, 1%). In this study, we focused on differentially expressed genes that gave a P value < 0.01 (from the LIMMA), and met log₂ ratio threshold of 0.5 (1.41-fold change) between wild type and PtMYB14 transgenics. For each transgenic construct, data from the two different lines were analyzed separately and only the differentially-expressed gene identified in both lines for a given transgene were considered to reduce the impact of positional effect of transgene.

V-5-12-Wounding and jasmonic acid treatments in pine and spruce in vitro plantlets

Pine seeds were first sterilized in ethanol (95%) for 5 min, then in sodium hypohlorite (2%) solution containing 0.1% Tween 80 for 10 min, H_2O_2 (10%) for 5 min, followed by ethanol (70%) for 5 min, and finally rinsed three times in sterile distilled water for 10 min each. Seeds were then stratified by soaking for 24 hours in sterile MilliQ water at room temperature in a drawer and were subsequently scarified by cutting the radicular tegument. Seeds were placed in a Magenta box containing half-strength MS medium supplemented with 0.5% activated charcoal

and incubated at 25°C under continuous light (40 μ mol m⁻² s⁻¹). Spruce plantlets were obtained from mature somatic embryos (Pg653) that were produced and germinated for 4 weeks according to Klimaszewska *et al.* (2001).

Wounding treatment was inflicted by pinching hypocotyl and cotyledons with forceps and by cuting one-third of cotyledons and first needles with scissors. Samples were harvested 0, 1.5, 6, and 24 hours after wounding. Jasmonic acid (JA) treatment was obtained by spraying a 100 μ M jasmonic acid (Sigma-Aldrich, Buchs, Switzerland) solution (approx. 5 mL per Magenta) on the entire plantlets. The control treatment consisted of spraying a solution containing only JA solvent which is methanol. JA experiment samples were collected at 0 h and 24 h after treatment. In each experiment, 3 and 10 plantlets were used for pine and spruce, respectively in a randomized complete block design, with three biological replicates per treatment.

V-6-Results

V-6-1-Isolation of several R2R3-MYB sequences similar to PtMYB14 in pine and spruce

In a previous study we characterized several conifer R2R3-MYBs belonging to the plant *MYB* gene subgroup 4 (Sg4) described by Kranz *et al.* (1998) which includes Arabidopsis MYB4 (Jin *et al.*, 2000). We described PtMYB14 from *Pinus taeda* and the closely related sequences PgMYB5, 10 and 13 from *Picea glauca* (Bedon *et al.*, 2007) each sharing the two conserved aminoacids motifs previously identified in plants as the signature of the R2R3-MYB subgroup4 (Krantz *et al.*, 1998, Jiang *et al.*, 2004). Here we describe several new sequences that are also closely related to *PtMYB14* and were isolated from spruce and pine, by using 3' RACE with a degenerate primer. Based on several *R2R3-MYB* nucleotides sequences including available conifer sequences, we targeted the more specific region of the *PtMYB14* related genes to design a degenerate primer corresponding to the sequence WRSLPKAA, located in the second helix of the second MYB repeat (Figure V-1). We analysed more than 20 3'RACE products in pine and spruce putatively representing the entire coding sequence, except for a 36 amino acids segment in the N-terminal.

Through this analysis, we found six different partial cDNAs in spruce ranging from 189 to 230 amino acids to the stop codon (PgMYB- 14a, 14b, 15, 16a, 16b, 16c) and six partial cDNAs in pine from 159 to 219 amino acids (PtMYB- 5, 10a, 10b, 13, 15, and PtMYB17) (Figure V-1). We also identified one full length cDNA from spruce (PgMYB5b, 236 amino acids) by EST data mining. The newly identified sequences are numbered in continuation with previously published sequences in pine and spruce but new sequences closest to previous annotated sequences are annotated alphabetically (Patzlaff *et al.*, 2003a and b; Bedon *et al.*, 2007). We found that spruce had two close homologues to PtMYB14: PgMYB14a, which shares 81.1% identity with the pine MYB14 (similarity is 85.4%) and PgMYB14b, which shares 82.4% of identity (similarity is 87.3%) based on pair-wise comparisons of the overlapping partial amino acid identity and 95% identity at the nucleotide level.



Figure V-1. CLUSTAL W alignments of the predicted partial amino acids sequences of pine and spruce close to PtMYB14.

Residues highlighted in black or grey are respectively identical or homologous in at least thirteen sequences. MYB repeat R2 and R3 are indicated by a black and grey bar respectively. C1 and C2 are two conserved motifs in carboxy-terminal identified in other members of subgroup 4 MYB proteins (Kranz *et al.*, 1998). The dashed rectangle show a highly conserved region between the conifers sequences where was found a six aminoacids motifs ADFFQ[YH] unique to conifers sequences, except PtMYB17, present in subgroup 4. Numbers on the right of the sequences indicates the predicted aminoacid lenght. The Primer indicates the position of the degenerate primer used to isolate pine and spruce sequences in this study (underlined names). Filled and empty lozenges symbols indicate spruce (Pg) and pine (Pt) sequences respectively.

V-6-2-PtMYB14 belongs to a new conifer specific clade in the subgroup 4 of plant R2R3-MYBs

Compared to plant species like Arabidopsis, rice or poplar, the conifers we analysed have many more Sg4 R2R3-MYBs (Figure V-1 and V-2). A phylogenetical analysis was carried out with newly identified pine and spruce sequences, based on partial amino acid sequence, along several other members from subgroup 4 and representatives of several other R2R3-MYB subgroups. The results indicated that the conifer sequences group within the subgroup 4, except for PtMYB17, and that they form a single conifer-specific clade without angiosperm members (Figure V-2 B). In constrast, the phylogenetic tree built only with amino acid sequences of the DNA binding domain (without the 36 N-terminal residues) has a different topology whichs groups all the members of subgroup 4 into one large clade, including PtMYB17 (Figure V-2 A). Furthermore, the Clustal W alignment shows that three conserved amino acids region are present in the C-terminal of most of the new conifer sequences, whereas only the C1 and C2 motifs (Krantz al al., 1998), are found in the angiosperm sequences (Figure V-1). The MEME motif-detection software also identified the motif ADFFQ[YH] as common to the conifer sequences that clustered together in subgroup 4 and as absent from the angiosperm members (Figure V-2). We named this sub-clade and its members as Sg4-Conifers (Sg4-C) because of the close similarities between DBD of conifers and angiosperms from subgroup 4, and the divergence in their C-terminal sequences particularly due to the ADFFQ[YH] signature motif (Figure V-1 and V-2).

V-6-3-RNA transcript profiles of MYB14 sequences in pine and spruce

Transcript abundance of *MYB14* sequences was surveyed in young pine and spruce green-house-grown trees to determine their tissues distribution. The different organs and tissues included young needles, apical stem or elongating apical leader, whole stem xylem, periderm stem with phloem, root tips or whole roots (Figure V-3). In pine, *PtMYB14* transcripts were detected in all the investigated tissues and did not show strong preferential expression; however whole fine roots and bark had the highest levels (Figure V-3). In spruce, *PgMYB14a* and *14b* have a similar transcript profile and level trough the tissues analysed, higher level in bark and lower level in xylem and needles (Figure V-3).





Neighbour-joining trees (1000 bootstraps) was based on the ClustalW alignment of (A) the DNA-binding domain with partial repeat R2 and entire repeat R3 both covering 79 aminoacids, (B) the MYB coding sequences starting from WRSLPKAAG region of the R2 repeat to stop codon. These two figures highlight the specific conifers clade, namely Sg4-C, and the other sequences regrouping principaly angiosperms sequences, of the subgroup 4. Only boostrap values up to 50% are presented. The bar indicates an evolutionary distance of 0,05 % in (A) and 0,1 % in (B). On the right part of the figure (B) are presented the diagrams of the aminoacid motifs identified by MEME software in

the carboxy terminal region of the R2R3-MYB proteins (see material and methods for details). Double vertical line show that all members of the subgroup 4 of conifers contains the motif [4] in bold. General disposition of the six identified motifs and their aminoacids compositions are given below the phylogenetic trees. R2R3-MYB protein name in bold have been previously characterized as repressors of lignin biosynthesis. Seven spruce and six pine MYB proteins identified in this study are mentioned by an asterisk, the others (PtMYB14, PgMYB10, PgMYB13 and PgMYB5 renamed PgMYB5a) comes from Bedon et al., 2007. Spruce and pine MYBs are represented by filled and empty lozenges symbols, respectively. Proteins which do not belong to subgroup 4 are used as landmarks representing other closest subgroups defined by Kranz et al. (1998). Pg: Picea glauca; Pt: Pinus taeda; Pm: Picea marianna; At: Arabidopsis thaliana; Tf: Tradescantia fluminensis; Os: Oriza sativa; Eg: Eucalyptus gunii; Zm: Zea mais; Ptt: Populus tremula x P. tremuloides; Gh: Gossypium hirsutum; Hv: Hordeum vulgare; Dsp: Dendrobium sp; Ta: Triticum aestivum; Le: Lycopersicum esculentum; Am: Antirrhinum majus; Vv: Vitis vinifera; Fa: Fragaria ananassa. The accession numbers of the used genes are given in Methods.



Figure V-3. Tissue survey of RNA transcripts of *PtMYB14* sequences in *Pinus* taeda and *PgMYB14a* and *14b* in *Picea glauca*.

Average transcript abundance was determined by quantitative RT-PCR in total RNA from various tissues of 2 year-old *Pinus taeda* and *Picea glauca* trees, from three biological replicates. Transcript levels were normalized relative to *ef1-a* expression level in pine and *cdc2* in spruce. Organs and tissues that were tested (varying slightly between the two species) are N: needle; AS: apical stem (young terminal leader); PPh: periderm and differenciating phloem (i. e. Bark); RT: root tip, R: whole fine roots; X: differenciating xylem, and; WX: whole stem xylem (see methods for details).

V-6-4-Constitutive and xylem preferential overexpression of PtMYB14 in transgenic spruce

We aimed to investigate the biological role of PtMYB14 in a conifer, therefore we stably transformed two different overexpression constructs into Picea glauca (Klimaszewska et al., 2004). Picea has been well developed for transformation and tissue culture, thus affording heterologous expression for Pinus genes in a closely related species. Because of the high level of functional redundancy among plant R2R3 MYB genes (e.g. Millar and Gubler, 2005; Müller et al., 2006) and, particularly, in light of the multiple gene sequences in the new and uncharacterized conifer subgroup 4 sequences, our analysis was focused on the overexpression experiments. Two overexpression constructs were prepared with the *PtMYB14* gene, driven by either a strong constitutive promoter (*Pro_{UBI}*), *ubiquitin* gene from maize (Christensen et al., 1992, 1996) or by a tissue preferential promoter (Pro_{CAD}), resulting in strong transgene expression difference in transgenic plantlets (about 100 times) normalized to endogenous CDC2 reference gene (Appendix V-3 and V-4). We studied the cellular localisation of *Pro_{UBI}* and *Pro_{CAD}* by the use of the uid-A gene (GUS) to visualize each promoter activities as well as to perform GUS enzymatic essays. Consequently, we validated the tissue constitutive expression of Pro_{UBI} and the xylem preferential expression of Pro_{CAD} in roots and hypocotyls of spruce plantlets (Figure V-4).

Constitutive overexpression of PtMYB14 (Pro_{UBI}PtMYB14-OE) produced a strong phenotype in spruce, greatly altering plantlet morphology (Figure V-5 B) compared to control (Figure V-5 A). After a few weeks in germination, transgenic plantlets displayed hypertrophic hypocotyls and cotyledons with obvious anthocyanin pigmentation. Compared to the control (Figure V-5 D, G), root development was not affected by constitutive PtMYB14-OE (Figure V-5 E, H). Nevertheless, plantlets rapidly withered and were unsuitable for transfer to soil. As shown in figure V-5 K, histological analyses of Pro_{UBI}PtMYB14 plantlets revealed an expanded paremchymatous tissue. Most cells of the inner parenchyma accumulated a large amount of starch, where as the outer parenchyma accumulated polyphenolic compounds. Cross sections of the transgenic hypocotyls also showed the vascular system was disorganized, developing relatively few tracheids which were variable in

diameter (Figure V-5 K). Polyphenolic compounds also accumulated in cells within the rays of the transgenic plantlets (Figure V-5 K) but were absent from the controls (Figure V-5 J).

We also produced transgenic spruce with tissue preferential overexpression of PtMYB14, aiming to circumvent or reduce the potential deleterious effects resulting from constitutive overexpression of PtMYB14 in spruce. In agreement with our initial hypothesis linking PtMYB14 to the transcriptional regulation of the phenylpropanoid pathway, the *CAD* gene promoter was selected due to its obvious involvement in differentiating xylem cells for monolignol biosynthesis (Boerjan *et al.*, 2003). The 1.2 Kb fragment of *CAD* gene promoter was isolated from a digested *Picea glauca* genomic DNA library by genome walking (see chapter VI).

Pro_{CAD}PtMYB14-OE plantlets were morphologically very similar to the control plantlets (Figure V-5). The hypocotyls had no sign of hypertrophy (Figure V-5 C) compared to the $Pro_{UBI}PtMYB14$ line (Figure V-5 B) and developed normally although a light increase in anthocyanin pigmentation was noted. However, in contrast to the wild-type and to the $Pro_{UBI}PtMYB14$ plantlets (Figure V-5 D, E, G, H), the roots of the $Pro_{CAD}PtMYB14-OE$ plantlets produced significantly more and longer root hairs (Figure V-5 F, I). Histological observations also revealed altered cellular organization and development compared to control and $Pro_{UBI}PtMYB14$ plantlets (Figure V-5 J and L). Inner parenchyma cells accumulated starch (Figure V-5 L) but to a lesser extent than those observed in $Pro_{UBI}PtMYB14-OE$ (Figure V-5 K). Within the vasculature, tracheids of $Pro_{CAD}PtMYB14$ were smaller in diameter and presented a decreased thickness compared to wild type. Polyphenolic cells were also observed within pith and ray cells.



Figure V-4. Histochemical analysis of Pro_{UBI-} , and Pro_{CAD-} driven GUS expression in root and hypocotyl of spruce transgenic plantlets.

(A) to (E) 5- μ m cross sections in paraffin-embedded root (A to C) and hypocotyl (D,E) of 10-week old plantlets. Compared to root of wild-type untransformed control (A), blue GUS staining is spread to the whole eustele area in Pro_{UBI}GUS plantlets (B) whereas it is restricted to differentiating primary xylem and cambial cells (arrowheads) in Pro_{CAD}GUS plantlets (C). In hypocotyl, GUS staining was observed in parenchyma, differentiating xylem and ray cells in Pro_{UBI}GUS plantlets (D) while it was only limited to differentiating xylem and ray cells in Pro_{CAD}GUS transgenic plantlets (E).

(F) Enzymatic MUG assay in hypocotyls (dark bars) and needles (grey bars) of spruce plantlets produced from different PtMYB14-OE transgenic lines.





(A) to (C) 6-week old plantlets. Compared to wild-type (A) and Pro_{CAD}PtMYB14 (C) plantlets, Pro_{UBI}PtMYB14 (B) plantlets exhibit hypertrophic hypocotyl and cotyledons with a pronounced anthocyanic pigmentation.

(D) to (I) root morphology and root-hair density in a wild type (D, G), Pro_{UBI}PtMYB14 (E, H), and Pro_{CAD}PtMYB14 (F, I) transgenic plantlets.

(J) to (L) 10- μ m cross section in paraffin-embedded hypocotyl of (I) wild-type, (K) Pro_{UBI}PtMYB14 and (L) Pro_{CAD}PtMYB14 transgenic plantlets. A disorganized vascular system with a few but large tracheids (white arrowhead) as well as an accumulation of starch grains in parenchyma cells (black arrowhead) and phloem polyphenolics cells (grey arrowhead) can be seen in Pro_{UBI}PtMYB14 transgenics.



Figure V-6. Accumulation of anthocyanins and starch in wild type and PtMYB14-OE transgenic spruce.

(A) HPLC chromatograms of anthocyanins produced from wild type (WT), *ProUBIPtMYB14* (line 18), and *ProCADPtMYB14* (line 10) spruce plantlets. Compared to wild type (WT), increase in peak intensity in transgenic line is indicated by closed arrowheads. (1) Delphinidin 3-O-Glucoside; (2) Cyanidin 3-O-Glucoside; (3) Petunidin 3-O-Glucoside; (4) Peonidin 3-O-Glucoside; (5) Malvidin 3-O-Glucoside.

(B) Starch accumulation in controls (WT, and pCAMBIA), ProUBIPtMYB14 (lines 3, 16, and 18), and ProCADPtMYB14 (lines 5,10, and 14) spruce plantlets. Each bar is a mean of 3 biological replicates, with 15-20 plantlets per replicate.

V-6-5-Determination of anthocyanin and starch contents in transgenic PtMYB14-OE plantlets

Anthocyanins and starch determinations were performed to confirm histological observations in spruce plantlets overexpressing PtMYB14. $Pro_{UBI}PtMYB14$ plantlets accumulated large amounts of anthocyanin Oglucosides, as indicated by HPLC chromatograms (Figure V-6 A). Compared to wild type (WT), large increase in peak intensity was observed for delphinidin 3-O-glucoside, cyanidin 3-O-glucoside, petunidin 3-O-glucoside, peonidin 3-Oglucoside, and malvidin 3-O-glucoside in Pro_{UBI} line 18. The $Pro_{CAD}PtMYB14$ overexpressers only had increased accumulation of cyanidin 3-O-glucoside (Figure V-6 A). Similar HPLC profiles were obtained from other transgenic lines with Pro_{UBI} (line 16), and Pro_{CAD} (line 14) (data not shown). Starch content determinations in $Pro_{UBI}PtMYB14$ plantlets gave a two-fold increase compared to control (Figure V-6 B) but no increase was recorded in any of the $Pro_{CAD}PtMYB14$ lines which were similar to the control (Figure V-6 B).

V-6-6-Constitutive and tissue-preferential overexpression of PtMYB14 has an overlapping impact on spruce transcriptome

As described above, partly overlapping phenotypes affecting hypocotyl development were observed with constitutive and tissue-preferential overexpression of PtMYB14 in spruce. Thus, we carried out large-scale transcriptome analysis of Pro_{UBI} and $Pro_{CAD}PtMYB14$ plantlets and compared the outcomes in order to assess the impact of both types of overexpression upon gene expression. Total RNAs were isolated from hypocotyls (including cotyledons) and profiled with a 9K custom cDNA spruce microarray as described in Bomal *et al.* (under review, chapter III). The experiment was carried out with two independent lines for each constructs (two lines for Pro_{UBI} and two lines for Pro_{CAD}), and one untransformed control line (WT)). We used four biological replicates per line, each replicate being comprised of a pool of 25 hypocotyls. We only considered those genes that were determined to be differentially expressed in both lines of each construct.

The two constructs gave quite different transcript profiles, as reflected by the numbers of genes with statistically significant transcript level variation compared to the controls (Figure V-7 A). A total of 1045 genes were misregulated for Pro_{UBI} whereas 164 genes were mis-regulated for Pro_{CAD}PtMYB14 (P value < 0.01; FDR, 1%). Interestingly, the proportion of up- and downregulated genes also varied significantly depending on the construct (Figure V-7 A). We focused our analysis on statistically significant genes having differences in expression of 1.4 fold or greater (Log₂ ratio $\geq |0.5|$), which led to a decrease of the number of misregulated genes to 517 and 78 for Pro_{UBI} and Pro_{CAD} , respectively; however, it did not change the up-down proportion which stayed around the same range, i.e. 59%–41% for the Pro_{UBI} constructs and 94%–6% for the Pro_{CAD} construct (Figure V-7 A). These misregulated genes were assigned to functional classes (Figure V-7 B) based on BlastX results against Uniref100, AGI10.0, PGI5.0 and SGI1.0 (the complete list of differentially expressed genes is not shown in this thesis because of its lenght). According to KEEG and GO categories ascribed to those genes, 8.9% and 13.5% of misregulated genes were assigned to "biotic/abiotic stress" category for ProUBI and ProCAD construct, respectively (Figure V-7 B). However, based on the literature, many other genes that were assigned to different other categories may also be related to stress and defense responses, increasing these proportions to 30.1% and 32.3% for Prouble and Pro_{CAD} construct, respectively. Therefore, despite the very significant divergence in the number of misregulated genes, the gene annotations suggested similarity with regard to the nature of the misregulated genes.

A comparison of the transcription profiles from Pro_{UBI-} and $Pro_{CAD}PtMYB14$ overexpression (PtMYB14-OE) indeed showed that there was significant overlap and identified many genes in common (Table V-1.). There were 50 differentially expressed genes common to both constructs (P value < 0.01, and Log₂ ratio $\geq |0.5|$). These include 48 genes with an increase in transcript abundance with both constructs compared to untransformed control, and only two genes, a heat shock protein (MN5181585) and an expansin that were downregulated in both constructs. Approximetaly two thirds of the genes that were misregulated with the $Pro_{CAD}PtMYB14$ were misregulated in a similar manner with the $Pro_{UBI}PtMYB14$ construct, i.e. the genes were up in both or down

in both. Only one gene gave opposite results (Chaperone DNAJ). The fold change ratios (relative to the control) were generally higher in Pro_{UBI} - than in Pro_{CAD} -overexpressers. The genes in common and their putative role in metabolism or in stress and defense response are described below.

The overlapping transcript profile data included several upregulated genes associated with the cytosolic mevalonate (MVA) pathway leading to isoprenoid formation, as well as the downstream pathway leading to sesquiterpenes (Table V-1). Indeed, increased transcript abundance was observed for acetoacetyl-CoA thiolase (AACT), 3-hydroxy-3-methylglutaryl-CoA-synthase (HMGS), mevalonate diphosphate decarboxylase (MVD1) as well as a terpene synthase (AFS) (Table V-1). Also within secondary metabolism, PtMYB14-OE led to the upregulation of flavonoid and phenylpropanoid related genes such as *chalcone synthase* (CHS), and cinnamoyl-CoA reductase-like (CCR-like) and 4 coumarate CoA ligase-like (4CL-like), respectively. Congruent with secondary metabolism modification in PtMYB14-OE spruces, several transcripts associated to electron transport were also misregulated as observed for a cytochrome p450 mono-oxygenase, a flavonoid 3',5'-hydroxylase and a FAD binding domain containing protein. Both constructs also lead to increased transcripts from ATP-binding cassette (ABC) transporter family proteins (PDR-like), defined as important partners in processes such as pigment accumulation, detoxification, and oxidation damage (Theodoulou, 2000).

As with the gene lists derived from the analysis of each construct seperately, many of the differentially expressed genes in the common gene list could be directly or indirectly linked to defense and stress responses. These included genes coding for CC-NBS-LRR resistance-like and disease resistance proteins which were upregulated in both construct (Table V-1). Genes defined by functional domains such as *leucine-rich repeat (LRR)*, *Toll-interleukin-1 receptor-like (TIR)*, *coil coil (CC)*, *receptor-like kinase*, and *NBD (nucleotide binding)* have been associated to defense response and plant resistance (Fluhr, 2001). The transcript profiling also identified several genes associated to stress response; they included *dehydrin* as well as *glycine-rich (GRP)*, and *heat-shock (HSP)* proteins in addition to *glutathione S-transferases (GST)*. Furthermore, transcripts that could be associated

with jasmonic acid (JA), and salicylic acid (SA) signaling pathways were also misregulated. For example, the upregulation was observed for SA response-associated genes such as *EDS1-like* and *phytoalexin deficient 4 (PAD4) proteins*. Also upregulated were transcript that could be linked to JA response, including *lipoxygenase (LOX), acyl-CoA synthase* and *acyl-CoA thioesterase* genes as well as *proteasome-associated proteins (F-Box associated protein)* and *chitinases*. In addition, hydrolytic enzyme coding genes linked to regulated proteolysis (serine carboxypeptidases) showed increased transcript abundance. Consistent with the metabolic and defense responses, a number of transcripts associated with transcriptional regulation and signal transduction were upregulated; these included a *zinc finger CCC H-type* as well as a *MAP kinase* and a *receptor-like protein kinase*.

As already indicated, significant differences were noted between the differentially expressed genes of the Pro_{UBI} and the Pro_{CAD}PtMYB14 overexpressers, particularly with regard to downregulated genes. In terms of metabolism modification, constitutive PtMYB14-OE (Pro_{UBI}PtMYB14) lead to the misregulation of genes associated to aromatic amino acid biosynthesis that serve as starting material for many secondary metabolites such as lignin, pigments, phytoalexins (data not shown). As such, transcripts of genes from shikimate (DAHP, chorismate mutase), pathway and from pyruvate and glycine/serine/threonine metabolisms were downregulated. Downstream of this pathway, several genes that could be linked to secondary metabolism including the phenylpropanoid, flavonoid, terpenoid, and benzenoid metabolic pathways were misregulated. Briefly, we identified the downregulation of transcripts from monolignol related enzymes such as PAL, 4CL, COMT, CCR, and CAD, as well as lignin polymerization (laccase, diphenol oxidase, multicopper oxidase) and lignan biosynthesis (HCBT, PCBER, and PLR). Numerous transcripts associated to Sadenosyl-L-methionine cycle, an important methyl group donor in monolignol pathway, were also downregulated. Two prenyl transferases were also misregulated (FPP synthase and GGPP synthase). A few genes were associated to cell wall biogenesis (CesA6 and expansin), but more prominently, several genes were related to cell wall relaxation including several pectin acetylesterases, xyloglucan endotransglucosylase/hydrolases, as well as extensins and structural proteins such as arabinogalactan and hydroxyproline rich proteins, with potential implication in

defense response. Numerous genes linked to carbohydrate metabolism were also misregulated, particularly genes coding for hydrolictic enzymes such as alkaline alpha-galactosidase, beta 1,3-glucanase as well as chitinases, in addition to several enzymes linked to sucrose and starch metabolism, like cell wall apoplastic invertase, glucan endo-1,3 beta glucosidase, beta amylase, starch phosphorylase. In contrast, genes associated to gibberellin biosynthesis and response stimuli were upregulated as observed for GASA family proteins and a *gibberellin 3-hydroxylase*. Gibberellins are diterpenoid derived from isoprenoid units that have been implicated in amylase induction. Also linked to stress and defense response, numerous genes associated to regulated proteolysis such as proteases and peptidases were misregulated. The list also includ²es several differentially expressed transcription factors (TFs) including *NAC (ATAF1-like)* and *MYBs*, TFs as well as numerous partners of the regulation of transcription.

In Pro_{CAD}PtMYB14-OE spruce, there were several upregulated transcripts associated with transcriptional regulation and signal transduction, including TFs like *MBF1*, a spruce *R2R3-MYB* and putative regulator of the anthocyanin pathway (Xue *et al.*, 2003), *GRAS2* which has been associated to stress and the JA signalling cascade (Mayrose *et al.*, 2006) as well as *no-apical-meristem* (*nam-like protein 6*), putatively implicated in meristem maintenance (Table V-1 and data not shown).

RT-QPCR analyses were performed on a subset of genes that were related to cell wall biogenesis as well as phenylpropanoid, flavonoid, and shikimate pathways to validate the results obtained with our microarray analyses. The misregulation was tested and confirmed for a total of 24 genes for the Pro_{UBI}PtMYB14 and 9 genes for the Pro_{CAD}PtMYB14 transgenics plantlets (Table V-1 and data not shown); i.e. no false positives were detected. Some of the genes surveyed by RT-QPCR both in hypocotyls and in roots transgenic plantlets are presented in appendix V-3 and V-4 respectively, including transgene *PtMYB14* and other *PgMYB* from Sg4-C.

EST ID ^a Accession Number Annotation E- Value FC ^b P- Value ^c GO:0008152 Metabolic process MN5234789 P37221 NAD-dependent malic enzyme 5E-101 1.26 5.8E-03 MN5159719 P55217 Cystathionine gamma-synthase 9E-69 1.33 5.8E-04 MN5252080 Q9C524 Fructokinase, putative 4E-41 1.48 1.1E-05 GO:0006508 Proteolysis MN5234749 Q9SFB5 Serine carboxypeptidase, putative 3E-117 1.51 2.0E-05 MN5191997 Q8GYW7 Acyl-CoA thioesterase 3E-17 1.53 1.2E-03 GO:0006629 Lipid metabolic process; GO:0016042 Lipid catabolic process MN5191997 Q8GYW7 Acyl-CoA thioesterase 3E-17 1.53 1.2E-03 MN5236677 Q2TNK3 Phytoalexin-deficient 4-1 protein (PAD4) 1E-13 1.91 3.6E-07 MN5161001 Q8L12 EDS1 3E-20 1.54 2.5E-06 MN5236944 Q1SPM3 EDS1-like, lipase, active site 7E-20 2.22 4.1E-03	FC	P- Valuo
GO:0008152 Metabolic process MN5234789 P37221 NAD-dependent malic enzyme 5E-101 1.26 5.8E-03 MN5159719 P55217 Cystathionine gamma-synthase 9E-69 1.33 5.8E-04 MN5252080 Q9C524 Fructokinase, putative 4E-41 1.48 1.1E-05 GO:0006508 Proteolysis MN5195534 Q1EP77 Serine carboxypeptidase, putative 3E-117 1.51 2.0E-05 MN5234749 Q9SFB5 Serine carboxypeptidase II, putative 3E-47 1.29 6.2E-03 GO:0006629 Lipid metabolic process; GO:0016042 Lipid catabolic process MN5191997 Q8GYW7 Acyl-CoA thioesterase 3E-17 1.53 1.2E-03 MN5236657 Q2TNK3 Phytoalexin-deficient 4-1 protein (PAD4) 1E-13 1.91 3.6E-07 MN5161001 Q8LL12 EDS1 3E-20 1.54 2.5E-06 MN5236944 Q1S9M3 EDS1-like, lipase, active site 7E-20 2.22 4.1E-08 MN5250118 Q9AXR6 ATP:citrate lyase 3E-78 1.57 1.1E-03 MN5250118 Q9AXR6 <th></th> <th>value</th>		value
GO:0008152 Metabolic process MN5234789 P37221 NAD-dependent malic enzyme 5E-101 1.26 5.8E-03 MN5159719 P55217 Cystathionine gamma-synthase 9E-69 1.33 5.8E-04 MN522080 Q9C524 Fructokinase, putative 4E-41 1.48 1.1E-05 GO:0006508 Proteolysis		
MN5234789 P37221 NAD-dependent malic enzyme 5E-101 1.26 5.8E-03 MN5159719 P55217 Cystathionine gamma-synthase 9E-69 1.33 5.8E-04 MN5252080 Q9C524 Fructokinase, putative 4E-41 1.48 1.1E-05 GO:0006508 Proteolysis MN5195534 Q1EP77 Serine carboxypeptidase, putative 3E-117 1.51 2.0E-05 MN5234749 Q9SFB5 Serine carboxypeptidase II, putative 3E-47 1.29 6.2E-03 GO:0006629 Lipid metabolic process; GO:0016042 Lipid catabolic process MN5191997 MN5251754 Q94622 Acyl-CoA thioesterase 3E-17 1.53 1.2E-03 MN5236657 Q2TNK3 Phytoalexin-deficient 4-1 protein (PAD4) 1E-13 1.91 3.6E-07 MN5161001 Q8L12 EDS1 3E-20 1.54 2.5E-06 MN5236944 Q1S9M3 EDS1-like, lipase, active site 7E-20 2.22 4.1E-08 MN5246065 Q93YH3 ATP:citrate lyase 3E-78 1.57 1.1E-03 GO:0006810 Transport <td< td=""><td>4.04</td><td></td></td<>	4.04	
MNS159719 PS217 Cystatritonine gamma-synthase 9E-69 1.33 5.8E-04 MN5252080 Q9C524 Fructokinase, putative 4E-41 1.48 1.1E-05 GO:0006508 Proteolysis MN5195534 Q1EP77 Serine carboxypeptidase, putative 3E-117 1.51 2.0E-05 MN5234749 Q9SFB5 Serine carboxypeptidase II, putative 3E-47 1.29 6.2E-03 GO:0006629 Lipid metabolic process; GO:0016042 Lipid catabolic process MN5191997 Q8GYW7 Acyl-CoA thioesterase 3E-17 1.53 1.2E-03 MN5251754 Q946Z2 Acyl-CoA synthetase 5E-46 1.80 6.9E-05 MN5236657 Q2TNK3 Phytoalexin-deficient 4-1 protein (PAD4) 1E-13 1.91 3.6E-07 MN5161001 Q8L12 EDS1 3E-20 1.54 2.5E-06 MN5236944 Q1S9M3 EDS1-like, lipase, active site 7E-20 2.22 4.1E-08 MN5246065 Q93YH3 ATP:citrate lyase 3E-78 1.57 1.1E-03 GO:0006810 Transport MN5161865 Q8GU87<	1.61	5.5E-05
GO:0006508 Proteolysis 3E-41 1.46 1.1E-05 MN5195534 Q1EP77 Serine carboxypeptidase, putative 3E-117 1.51 2.0E-05 MN5234749 Q9SFB5 Serine carboxypeptidase II, putative 3E-47 1.29 6.2E-03 GO:0006629 Lipid metabolic process; GO:0016042 Lipid catabolic process MN5191997 Q8GYW7 Acyl-CoA thioesterase 3E-17 1.53 1.2E-03 MN5251754 Q946Z2 Acyl-CoA synthetase 5E-46 1.80 6.9E-05 MN5236657 Q2TNK3 Phytoalexin-deficient 4-1 protein (PAD4) 1E-13 1.91 3.6E-07 MN5161001 Q8LL12 EDS1 3E-20 1.54 2.5E-06 MN5236055 Q93YH3 ATP:citrate lyase b-subunit 4E-121 1.38 3.8E-03 MN5250118 Q9AXR6 ATP:citrate lyase 3E-78 1.57 1.1E-03 GO:0006810 Transport MN5161865 Q8GU87 ABC transporter, PRD-like 1E-35 2.20 1.2E-05	1.40	9.7E-04
GO:0006508 Proteolysis MN5195534 Q1EP77 Serine carboxypeptidase, putative 3E-117 1.51 2.0E-05 MN5234749 Q9SFB5 Serine carboxypeptidase II, putative 3E-47 1.29 6.2E-03 GO:0006629 Lipid metabolic process; GO:0016042 Lipid catabolic process MN5191997 Q8GYW7 Acyl-CoA thioesterase 3E-17 1.53 1.2E-03 MN5251754 Q946Z2 Acyl-CoA synthetase 5E-46 1.80 6.9E-05 MN5236657 Q2TNK3 Phytoalexin-deficient 4-1 protein (PAD4) 1E-13 1.91 3.6E-07 MN5161001 Q8LL12 EDS1 3E-20 1.54 2.5E-06 MN52360544 Q1S9M3 EDS1-like, lipase, active site 7E-20 2.22 4.1E-08 MN5246065 Q93YH3 ATP:citrate lyase b-subunit 4E-121 1.38 3.8E-03 MN5250118 Q9AXR6 ATP:citrate lyase 3E-78 1.57 1.1E-03 GO:0006810 Transport MN5161865 Q8GU87 ABC transporter, PRD-like 1E-35 2.20 1.2E-05	1.39	1.0E-03
MNS195534 Q1EP77 Serine carboxypeptidase, putative 3E-117 1.51 2.0E-05 MN5234749 Q9SFB5 Serine carboxypeptidase II, putative 3E-47 1.29 6.2E-03 GO:0006629 Lipid metabolic process; GO:0016042 Lipid catabolic process 3E-17 1.53 1.2E-03 MN5191997 Q8GYW7 Acyl-CoA thioesterase 3E-17 1.53 1.2E-03 MN5251754 Q946Z2 Acyl-CoA synthetase 5E-46 1.80 6.9E-05 MN5236657 Q2TNK3 Phytoalexin-deficient 4-1 protein (PAD4) 1E-13 1.91 3.6E-07 MN5161001 Q8L12 EDS1 3E-20 1.54 2.5E-06 MN5236944 Q1S9M3 EDS1-like, lipase, active site 7E-20 2.22 4.1E-08 MN5246065 Q93YH3 ATP:citrate lyase 3E-78 1.57 1.1E-03 GO:0006810 Transport MN5161865 Q8GU87 ABC transporter, PRD-like 1E-35 2.20 1.2E-05		
MN5234749 Q9SFB5 Serine carboxypeptidase II, putative 3E-47 1.29 6.2E-03 GO:0006629 Lipid metabolic process; GO:0016042 Lipid catabolic process; MN5191997 Q8GYW7 Acyl-CoA thioesterase 3E-17 1.53 1.2E-03 MN5251754 Q946Z2 Acyl-CoA thioesterase 5E-46 1.80 6.9E-05 MN5236657 Q2TNK3 Phytoalexin-deficient 4-1 protein (PAD4) 1E-13 1.91 3.6E-07 MN5161001 Q8LL12 EDS1 3E-20 1.54 2.5E-06 MN5236944 Q1S9M3 EDS1-like, lipase, active site 7E-20 2.22 4.1E-08 MN5246065 Q93YH3 ATP:citrate lyase b-subunit 4E-121 1.38 3.8E-03 MN5250118 Q9AXR6 ATP:citrate lyase 3E-78 1.57 1.1E-03 GO:0006810 Transport ABC transporter, PRD-like 1E-35 2.20 1.2E-05	1.49	3.3E-04
GO:0006629 Lipid metabolic process; GO:0016042 Lipid catabolic process MN5191997 Q8GYW7 Acyl-CoA thioesterase 3E-17 1.53 1.2E-03 MN5251754 Q946Z2 Acyl-CoA synthetase 5E-46 1.80 6.9E-05 MN5236657 Q2TNK3 Phytoalexin-deficient 4-1 protein (PAD4) 1E-13 1.91 3.6E-07 MN5161001 Q8LL12 EDS1 3E-20 1.54 2.5E-06 MN5236044 Q1S9M3 EDS1-like, lipase, active site 7E-20 2.22 4.1E-08 MN5246065 Q93YH3 ATP:citrate lyase b-subunit 4E-121 1.38 3.8E-03 MN5250118 Q9AXR6 ATP:citrate lyase 3E-78 1.57 1.1E-03 GO:0006810 Transport MN5161865 Q8GU87 ABC transporter, PRD-like 1E-35 2.20 1.2E-05	1.47	1.2E-03
MN5191997 Q8GYW7 Acyl-CoA thioesterase 3E-17 1.53 1.2E-03 MN5251754 Q946Z2 Acyl-CoA synthetase 5E-46 1.80 6.9E-05 MN5236657 Q2TNK3 Phytoalexin-deficient 4-1 protein (PAD4) 1E-13 1.91 3.6E-07 MN5161001 Q8LL12 EDS1 3E-20 1.54 2.5E-06 MN5236944 Q1S9M3 EDS1-like, lipase, active site 7E-20 2.22 4.1E-08 MN5246065 Q93YH3 ATP:citrate lyase b-subunit 4E-121 1.38 3.8E-03 MN5250118 Q9AXR6 ATP:citrate lyase 3E-78 1.57 1.1E-03 GO:0006810 Transport MN5161865 Q8GU87 ABC transporter, PRD-like 1E-35 2.20 1.2E-05		
MN5251754 Q946Z2 Acyl-CoA synthetase 5E-46 1.80 6.9E-05 MN5236657 Q2TNK3 Phytoalexin-deficient 4-1 protein (PAD4) 1E-13 1.91 3.6E-07 MN5161001 Q8LL12 EDS1 3E-20 1.54 2.5E-06 MN5236944 Q1S9M3 EDS1-like, lipase, active site 7E-20 2.22 4.1E-08 MN5246065 Q93YH3 ATP:citrate lyase b-subunit 4E-121 1.38 3.8E-03 MN5250118 Q9AXR6 ATP:citrate lyase 3E-78 1.57 1.1E-03 GO:0006810 Transport MN5161865 Q8GU87 ABC transporter, PRD-like 1E-35 2.20 1.2E-05	1.48	9.7E-03
MN5236657 Q2TNK3 Phytoalexin-deficient 4-1 protein (PAD4) 1E-13 1.91 3.6E-07 MN5161001 Q8LL12 EDS1 3E-20 1.54 2.5E-06 MN5236944 Q1S9M3 EDS1-like, lipase, active site 7E-20 2.22 4.1E-08 MN5246065 Q93YH3 ATP:citrate lyase b-subunit 4E-121 1.38 3.8E-03 MN5250118 Q9AXR6 ATP:citrate lyase 3E-78 1.57 1.1E-03 GO:0006810 Transport MN5161865 Q8GU87 ABC transporter, PRD-like 1E-35 2.20 1.2E-05	1.53	9.0E-03
MN5161001 Q8LL12 EDS1 3E-20 1.54 2.5E-06 MN5236944 Q1S9M3 EDS1-like, lipase, active site 7E-20 2.22 4.1E-08 MN5246065 Q93YH3 ATP:citrate lyase b-subunit 4E-121 1.38 3.8E-03 MN5250118 Q9AXR6 ATP:citrate lyase 3E-78 1.57 1.1E-03 GO:0006810 Transport MN5161865 Q8GU87 ABC transporter, PRD-like 1E-35 2.20 1.2E-05	1.49	1.0E-03
MN5236944 Q1S9M3 EDS1-like, lipase, active site 7E-20 2.22 4.1E-08 MN5246065 Q93YH3 ATP:citrate lyase b-subunit 4E-121 1.38 3.8E-03 MN5250118 Q9AXR6 ATP:citrate lyase 3E-78 1.57 1.1E-03 GO:0006810 Transport MN5161865 Q8GU87 ABC transporter, PRD-like 1E-35 2.20 1.2E-05	1.31	2.8E-03
MN5246065 Q93YH3 ATP:citrate lyase b-subunit 4E-121 1.38 3.8E-03 MN5250118 Q9AXR6 ATP:citrate lyase 3E-78 1.57 1.1E-03 GO:0006810 Transport MN5161865 Q8GU87 ABC transporter, PRD-like 1E-35 2.20 1.2E-05	1.39	6.2E-03
MN5250118 Q9AXR6 ATP:citrate lyase 3E-78 1.57 1.1E-03 GO:0006810 Transport MN5161865 Q8GU87 ABC transporter, PRD-like 1E-35 2.20 1.2E-05	1.41	9.2E-03
GO:0006810 Transport MN5161865 Q8GU87 ABC transporter, PRD-like 1E-35 2.20 1.2E-05	1.68	2.0E-03
MN5161865 Q8GU87 ABC transporter, PRD-like 1E-35 2.20 1.2E-05		
•	1.91	1.0E-03
MN5182341 Q0IRX8 ABC transporter, PDR-like 1E-30 2.30 3.4E-06	1.78	1.7E-03
MN5232791 Q9FKS8 Amino acid permease 1E-16 2.16 2.5E-07	1.61	8.0E-04
MN5254357 Q9AT12 Sulfate transporter ST1 5E-58 1.73 3.7E-05	1.51	5.4E-03
GO:0006118 Electron transport		
MN5197010 Q6RK07 UDP-glucose dehydrogenase 0 146 12E-05	1 29	5 1E-03
MN5236981 Q10RL8 FAD binding domain containing protein 1E-50 1 60 6 7E-06	1 33	7 7E-03
MN5250508 Q50EK4 Cytochrome P450 CYPC 7E-31 1 88 1 7E-04	2 16	2 6F-04
MN5252599 O04773 Flavonoid 3',5'-hydroxylase 7E-40 1.87 2.5E-06	1.96	2.3E-05
GO:0006950 Response to stress		
MN5177396 Q03683 Heat shock protein (HSP70) 5E-08 1 42 3 3E-04	1.38	5 3E-03
MN5181585 097S70 Heat shock protein (HSP17 2IA) 3E-03 -1.48 1.4E-03	_1.00	6.6E-03
MN5232825 08H6B6 Clucing_rich protein (Arn94) 8E-96 1.47 7.4E-04	1.40	1 3E-03
MN5232823 Q01000 Clycine-nch protein (Crps+) 0L-50 1.47 7.4L-04	1.47	1.3⊑_03
MN5256468 O65056 Clutathione Stransferase probable 3E 73 2.14 9.6E 07	1.74	2 3 = 03
MNS230400 000000 Glutatilione 5-transierase, probable 5E-75 2.14 0.0E-07 MNS242202 0114D2 Chaptering protein Dag L 6E 09 1.20 2.6E 02	1.50	2.3L-03
MN5250987 Q3ZDL1 Dehydrin 0 1.42 1.4E-04	1.47	4.1E-03 3.8E-04
CO.0006050 Defense meneros CO.0000611 Despense to wounding		
GO:0006952 Defense response; GO:0009611 Response to wounding	4 50	
MN5190556 Q1L6F3 CC-NBS-LRR resistance-like protein 3E-77 1.40 2.2E-03	1.59	9.7E-04
MN5192026 Q76122 Lipoxygenase 1 (LOX) 6E-77 2.48 1.1E-03	3.05	1.2E-03
MN5235055 Q2VWL9 Disease resistance protein ADR1 9E-18 2.69 1.3E-06	1.72	6.2E-03
MN5250426 Q9M0X9 4-coumarate-CoA ligase-like 5E-49 2.00 1.9E-07	1.65	1.8E-04
GO:0007165 Signal transduction; GO:0009737 Response to abscisic acid stimulus		
MN5159144 Q39026 MAP kinase 6 (<i>AtMPK6</i>) 1E-109 1.51 1.4E-04	1.80	4.6E-05
MN5173638 Q8LP72 Receptor-like protein kinase 1E-27 2.89 9.5E-08	2.30	4.6E-05
GO:0045449 Regulation of transcription		
MN5182156 Q10EL1 Zinc finger CCCH type family protein 7E-05 1.48 5.7E-03		
MN5196607 Q6ZFI5 Parathymosin-like 2E-21 1.47 1.3E-04	1.59	8.0E-03

Flavonoid biosynthetic process										
MN5260082	Q2L8A7	Acetoacetyl-CoA thiolase	3E-46	1.49	1.6E-05	1.58	5.5E-05			
MN5236845	P93773	3-hydroxy-3-methylglutaryl-CoA-synthase	0E+00	1.35	2.6E-03	1.59	3.4E-04			
MN5234648	Q675K8	E,E-alpha-farnesene synthase	2E-13	2.62	2.7E-05	2.60	3.4E-04			
MN5232972	Q7Y232	Zeatin O-glucosyltransferase 1	8E-35	2.13	2.2E-05	1.92	1.0E-03			
MN5255975	Q4EVY6	Thiohydroximate S- glucosyltransferase	1E-28	1.33	1.8E-03	1.52	3.8E-04			
MN5252787	Q2ENB1	Chalcone synthase	0	1.44	2.4E-03	1.49	6.6E-03			
MN5235414	Q9M0B3	Cinnamoyl-CoA reductase-like	2E-29	1.39	2.9E-03	1.41	9.0E-03			
GO:0009664 Cell wall organization and biogenesis										
MN5194355	Q84UT0	Expansin	1E-69	-2.85	3.4E-06	-1.79	9.0E-03			
MN5234630	Q6E6M9	Class II chitinase	3E-94	3.86	1.4E-08	2.20	2.2E-04			
MN5254248	Q6WSR9	Class IV chitinase (Chia4-Pa1)	2E-66	1.87	1.2E-07	1.39	2.0E-03			
Miscellanous										
MN5174632	Q1ECE0	Vesicle associated membrane protein	1E-34	1.45	7.4E-04	1.62	5.0E-04			
MN5182122	Q84V46	Alternative oxidase 1b	0	1.32	6.2E-04	1.45	3.4E-04			
MN5232933	Q65GA1	DnaB (Membrane attachment protein)	0	2.31	4.8E-07	2.30	1.4E-05			
MN5182536	O82812	Expressed protein	1E-19	1.52	4.8E-04	1.41	1.0E-02			
MN5253353	No hit	No hit	N/A	1.67	5.5E-06	1.39	4.3E-03			
MN5173128	ABD60279	Transgene (PtMYB14)	0	9.06	7.3E-08	2.01	3.2E-02			

GO:0008299 : Isoprenoid biosynthetic process; GO:0008202 Steroid metabolic process; GO:0009813 Flavonoid biosynthetic process

Table V-1. Overlapping set of genes differentially expressed (P value < 0.01, and Log₂ ratio $\geq |0.5|$) following both *Pro_{UBI}*- and *Pro_{CAD}*-driven overexpression of *PtMYB14* in spruce.

^aEST ID, Expressed Sequence Tag Identifier; Bolded EST ID indicates that transcript abundance and fold change was validated through RT-qPCR,

^bFC, Fold change. The minus and plus signs indicate that, compared to wild type, the gene is under- or over-represented, respectively;

^cStudent's t test level of confidence.



Figure V-7. Number and functional grouping of genes differentially expressed in *Pro_{UBI}PtMYB14* and *Pro_{CAD}PtMYB14* overexpressing spruce plantlets.

(A) Number of differentially expressed genes in Pro_{UBI} and Pro_{CAD} construct.

(B) Pie diagram showing the repartition (%) of differentially expressed genes based on molecular function. Categories ascribed to genes were determined from KEEG and Gene Ontology. For each construct, the numbers of genes associated with a specific function are in brackets.

V-6-7-PtMYB14 overexpression affects terpene production in spruce plantlets

In light of the overlapping effects of the *Pro_{UBI}*- and *Pro_{CAD}*-driven *PtMYB14*-OE on isoprenoid and terpenoid-related gene transcript accumulation, we aimed to gather further information on terpene content of transgenic plantlets. We focused our analyses on mono-, di- and sesqui- terpenoids known as defence compounds in plants. We were able to identify and quantify several compounds belonging to the monoterpenoids (MT: α -pinene and camphene) and sesquiterpenoids (ST: α farnesene and β -farnesene) in transgenic spruce plantlets but no diterpenoids were found (Figure V-8). A major accumulation of terpenoids (MT and ST) was observed in the Pro_{UBI}PtMYB14 plantlets and there was a weak accumulation in Pro_{CAD}PtMYB14 plantlets, both compared to the wild type control which had been used for the microarray experiments (Figure V-8). Only the monoterpenoids α -pinene and camphene significantly accumulated in Pro_{CAD}PtMYB14 plantlets compared to control but to a lesser extent than in Pro_{UBI}PtMYB14 plantlets. No diterpenoids accumulation was found in transgenic compare to control plantlets.

V-6-8-Wounding and JA treatments induce the expression of *PtMYB14* related genes and isoprenoid/defense pathway genes in pine and spruce seedlings

Microarray data and chemical determinations from both the constitutive and xylem preferential PtMYB14-OE most strongly point to the misregulation of genes of the isoprenoid/terpenoids metabolism involved in defence pathways. This finding led us to investigate the putative link between *PtMYB14* and the defence-induced gene expression mechanisms in the cytosolic (MVA) and the plastidial mevalonate (MEP) pathways (also named DOX pathway for 1-deoxy-D-xylulose). We carried out induction experiments with pine and spruce *in vitro* seedlings and followed gene transcript accumulation of highly similar sequences, in order to develop a complementary view in these two closely related systems. Rapid transcript accumulation of MYB14 sequences was observed in pine (*PtMYB14*) and spruce (*PgMYB14a* and *14b*), suggesting that their expression is induced early after wounding and upregulated by jasmonic acid (Figures V-9 and V-10). Hypocotyls and cotyledons of pine and spruce seedling were wounded with a pliers and analysed 1.5,

6 and 24 hours after induction. The JA was sprayed onto in vitro seedlings and hypocotyls were harvested after 24 hours. The upregulation of transcripts encoding allene oxide cyclase (AOC, jasmonic acid precussor) and defensine (DEF, Pervieux et al., 2004) under wound and JA treatments confirmed that the plantlets are in stress/defence situation. Several transcripts of the MVA pathway accumulate after wounding with upregulation, including AACT, HMGS, MVAK and MVD1, but only the DXPS of the DOX pathway in spruce was clearly up-regulated. Jasmonate application also led to increased mRNA abundance of MVA pathway genes in pine (AACT, HMGS, MVAK and MCT) but not in spruce. In contrast to the wounding treatment, JA induced DOX pathway gene transcript accumulation including DXPS, DXR, MCT in both systems and of MECPS in pine only (Figures V-9 and V-10). Differences in the JA-induction of MVA and DOX pathway genes between pine and spruce may be related to a less efficient elicitation of defence responses as suggested by the lower increase of AOC and DEF transcripts in spruce compare to pine. As a follow up on our microarray results, we also analysed the expression pattern of two key enzymes of the isoprenoids pathway (FPPS and GGPPS) upstream of terpenoid derivatives biosynthesis. Both FPPS and GGPPS were clearly upregulated after JA treatment in pine, while FPPS was upregulated after wounding in spruce. Among the sesquiterpenoids genes investigated we found AFS to be upregulated both in pine and spruce by wounding but not by AJ.

Moreover we investigated the gene expression level of all the *R2R3-MYB* genes identified in Sg4-C in pine and spruce. Wounding in both systems led to rapid (and usually transient) down-regulation of several MYB transcripts except the *MYB14* sequences (pine and spruce) and *PgMYB15* and *16a* (spruce) which were up-regulated (Figures V-9 and V-10). In contrast, JA application led to increased transcript levels of *PtMYB5, PtMYB13, PtMYB17* (of pine) in addition to the PtMYB14 sequences in both pine and spruce.



Figure V-8. Accumulation of terpenoid compounds in wild type and PtMYB14-OE trasngenic spruce.

(A) Monoterpenes and (B) sesquiterpenes identification and accumulation in controls (WT, and pCAMBIA), $Pro_{UBI}PtMYB14$ -OE (lines 3, 16 and 18), and $Pro_{CAD}PtMYB14$ -OE (lines 5, 10 and 14) spruce plantlets. RT: retention time. Three biological samples per lines with two technical replicates each were analyzed. Every amount of 'detector response' was calculated against the internal standard and calibrated against the freshweight of the plantlets sample. Mono- and sesquiterpenes were are analyzed together in one run and IBB standard was used for MT and ST normalization.





Wounding treatment was inflicted to 4-week old pine seedlings by pinching hypocotyl with forceps and by cuting one-third of cotyledons and first needles with scissors. Samples were harvested 0, 1.5, 6, and 24 hours after wounding. Each point is a mean (\pm SE) of three biological replicates, with 3 hypocotyls per replicate. Jasmonic acid (AJ) solution (100 µM) was spread on 4-week old pine seedlings (5 mL per replicate). Samples were harvested 0, and 24 hours after treatment. Control treatment consisted of spraying the JA solvent solution (0.08% ethanol, MeOH). Each point is a mean (\pm SE) of three biological replicates, with 3, and 10 hypocotyls per replicate in pine and spruce respectively. In each graph, asterids indicate a significant difference between control point (open bars) and treatment (closed bars) according to Student's test at 0.05 (*), 0.01 (**), or 0.001 (***). Meaning of abbreviations is given in figure V-10.



Figure V-10. Expression of R2R3-MYB, MVA- and DOX pathway related genes in response to wounding and jasmonic treatments in spruce somatic plantlets.

Wounding and jasmonic treatment was perfomed on 4-week old somatic spruce plantlets (Pg653) as described in figure V-9. Each point is a mean (\pm SE) of three biological replicates, with 10 hypocotyls per replicate. In each graph, asterids indicate a significant difference between control point (open symbol) and treatment (closed symbol) according to Student's test at 0.05 (*), 0.01 (**), or 0.001 (***). MVA, mevalonate pathway; AACT, Acetoacetyl-CoA thiolase; HMGS, 3-hydroxy-3-methylglutaryl-CoA synthase; HMGR, 3-hydroxy-3-

methylglutaryl-CoA reductase; MVAK, mevalonate kinase; MVD1, mevalonate-5-pyrophosphate decarboxylase. Prenyltransferases: FPPS, farnesylpyrophosphate synthase; GGPPS, geranyl-geranyl-pyrophosphate synthase. DOX, 1-Deoxy-D-xylulose-5-phosphate pathway; DXPS, 1-deoxy-D-xylulose-5phosphate synthase; DXR, 1-deoxy-D-xylulose-5-phosphate reductoisomerase; MCT, 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase; MECPS, 2-Cmethyl-D-erythritol 2,4-cyclodiphosphate synthase. Sesquiterpene synthases: AFS, E,E-alpha farnesene synthase; AG1, alpha-bisabolene synthase. Reporter, defense reporter genes: AOC, allene oxide cyclase; DEF, defensin. Reference, gene used for transcript normalization (ratio), CDC2: cell division cycle 2

V-7-Discussion

The R2R3MYB sequences described in this report combined with a previously described sequences (Bedon et al., 2007) define a new phylogenetic clade of MYB genes that is specific to conifers (pine and spruce) and which belongs to the previously defined subgroup 4 (Krantz et al., 1998). We named it "Subgroup 4 of Conifers" or Sg4-C. We functionally characterized the Pinus taeda MYB 14 belonging to this group, by stable transformation and plantlet regeneration (Klimaszewska et al., 2004) in white spruce (Picea glauca), which represents a closely related conifer system for heterologous expression. Because of the high level of sequence similarity among the multiple members of subgroup 4-C and the incomplete characterization of the spruce genome, we decided to circumvent potential gene redundancies through the use of over-expression constructs. In one construct, the PtMYB14 gene was driven by a strong constitutive promoter from the maize *ubiquitin* gene (Christensen *et al.*, 1992, 1996). In the other, we used a xylem preferential promoter from the white spruce gene encoding cinnamyl alcohol dehydrogenase (CAD) which catalyzes the last step of the monolignin biosynthesis. Analyses of the resulting transgenic plantlets involving chemical determinations and transcript profiling lead to the conclusion that PtMYB14 is most likely involved in defence mechanisms, more specifically impacting on the JA and mevalonate pathways. Such a role is highly consistent with the rapid induction of MYB14 transcript accumulation upon wounding and application of jasmonate.

V-7-1-Structure and evolution of a new subgroup of R2R3MYB genes specific to conifers pine and spruce

The pine and spruce sequences we analysed belong to Subgroup 4, one of the 22 previously defined subgroups of R2R3-MYBs (Krantz et al., 1998). Compared to the angiosperms, the conifers have significantly more members in Sg4, with up to 10 members (although only six were isolated to date from pine) with highly conserved sequences. To date, extensive investigations of the MYB family at the level of sequence, expression and function have focused on angiosperm species like Arabidopsis, rice and maize, due among other things to the availability of their genome sequences (Stracke et al., 2001; Yanhui et al., 2006; Rabinowicz et al., 1999). Little attention has been given to gymnosperm MYB sequences despite the relevance for plant evolution studies and their potential interest for breeding applications (Kusumi et al. 2002; Xue et al., 2003; Patzlaff et al, 2003a, 2003b; Bedon et al., 2007). The high level of sequence similarity and the conservation of their DNA binding domains as well as a potentially unique amino acid motif within this set of expressed sequences may indicate that several recent gene duplication events have given rise to this clade from an original ancestral sequence.

A common and recent origin of the Sg4-C sequences is suggested by the fact that they share the specific amino acid motif ADFFG[HY] which appears to be unique to conifer sequences. We acknowledge that there are specific challenges for identifying short conserved motifs by sequence analysis alone, due to the possibility of false positives for one or more residues in such a short consensus motif. However, a retrospective analysis of the C2 motif, another conserved amino acids region found in Sg4 sequences (in angiosperms and gymnosperms), suggests that searching short motifs can be informative. The C2 motif, pdLNLD/ElxiG/S was shown to be involved in transcriptional repression of phenylpropanoid genes (Jin *et al.*, 2000). By restricting the MEME software parameters to find short motifs ranging from 4 to 6 residues, we identified two short motifs within the C2 motif (Figure V-2). One of them is likely the same as the EAR-motif-containing transcriptional repressors (core EAR-motif: PDLNL in AtMYB4; Kazan, 2006). Recently several studies reported important roles for

EAR repressor mainly in regulating plant defence and stress responses (Chern *et al.*, 2005; McGrath *et al.*, 2005; Yang *et al.*, 2005; Song *et al.*, 2005). In turn the high level of conservation of the ADFFG[HY] motif we identified as putatively unique to conifer sequences, strongly argues in favour of continued investigations of protein structure and function to determine its biological significance.

The maintenance of this intriguingly large and potentially redundant group of sequences belonging to the subgroup 4-C has also led us to hypothesize that at least some of the sequences may have diverged in terms of their biochemical functions or biological roles. Selective pressures may act together with random genetic variation within newly duplicated copies of a gene to modify their function or expression profile. Indeed, expression studies in Arabidopsis have indicated that highly homologous sequences may have different expression profiles under particular conditions or treatments (Yanhui et al., 2006). Not surprisingly, widespread duplication has been previously invoked as an important mechanism for MYB gene evolution leading to the acquisition of new functions, as well as enabling plant adaptation to new environments. For example, related to flavonoid regulation in maize, duplication of the MYB regulator p1 and p2 genes from the ancestral p gene resulted in two different expression profiles. The p1gene is expressed in the kernel pericarp, and p2 is expressed in the developing anther and silk, probably to provide a selective advantage during the evolution of the maize kernel morphology (Zhang et al., 2000).

The data presented in this report indeed suggest that the Sg4-C genes have diverged functionally to some extent. A survey of transcript profiles across different tissues in spruce (Bedon *et al.*, 2007), as well as their transcript accumulation upon wounding and JA application clearly indicate that the regulation of their expression has differentiated. Such variations would directly imply that different members of the subgroup may act in different organs or tissues, or in different types of stress responses. Variations in the expression profiles of Sg4-C genes may also indirectly lead to variations in their targets. Indeed, the predicted amino acid sequences of all the Sg4-C genes contain a putative bHLH interaction domain suggesting that they participate in protein-protein interactions to regulate gene expression. Thus variations in MYB

expression could lead to variation in protein partners and impact upon the nature of their targets. Finally, we may also consider the significant sequence variation which occurs in the carboxy-terminal region of these genes, knowing that the Cterminus of MYBs has been implicated in activation or repression of transcription.

V-7-2-Constitutive and tissue-preferential overexpression as a robust strategy for the functional analysis of transcription factors

Overexpression (OE) has been effective in revealing transcription factors (TF) function and offers an alternative strategy that is less affected by TF functional redundancy than knockout/knockdown analysis (Zhang, 2003). The problem of functional redundancy can be especially acute for TFs from large gene families that often include closely related gene sequences (Schwechheimer et al., 1998; Zhang, 2003). The advantages of the OE approach appeared well suited to investigate the role of PtMYB14 in conifers due to the large number of closely related sequences in Sg4-C (Figures V-1 and V-2). Nonetheless, in order to decrease the drawbacks which may arise from non-specific or pleiotropic effects of overexpression, two different and complementary OE strategies were compared. A constitutive (Pro_{UBI}) and a tissue-preferential (Pro_{CAD}) promoter were used to drive PtMYB14 OE. Different expression modes in terms of transgene expression level and targeted tissues were thus achieved as confirmed by MUG assay for GUS activity and histochemical GUS staining (Figure V-4). These results were also consistent with the different levels of *PtMYB14* transcript abundance determined in the Pro_{UBI} and Pro_{CAD}PtMYB14 overexpressors (Table 1).

Results indicate that combining both constitutive and tissue-preferential strategies may be a way to discriminate between direct and pleiotropic effects induced by transgene OE. First, a significant subset of genes was misregulated in both types of PtMYB14 overexpressors, based on transcript profile data from microarray and qRT-PCR analyses (Table V-1 and data not shown). We propose that these misregulated sequences more likely represent genes and pathways directly linked to the MYB14 activity, because their transcripts were modulated similarly (Table V-1) despite the very different levels and spatial distributions of

PtMYB14 transgene expression observed with the two OE strategies (Figure V-4). Although the Pro_{CAD}PtMYB14 plants had many fewer misregulated sequences, the majority of them (around two thirds) were also part of the common set. This large proportion of overlapping sequences further supports the idea that the set of common genes are indeed directly linked to PtMYB14 activity, as opposed to purely pleiotropic effects due to unspecific or random effects. Furthermore, the fact that nearly all of this common set was comprised of upregulated sequences suggests that PtMYB14 is a positive regulator. In support of these interpretations, transcript profiling analysis identified common upregulated sequences (flavonoidand MVA-related) which are consistent with overlapping phenotypes revealed through histological and chemical analyses of the Pro_{UBI} and the Pro_{CAD}PtMYB14 overexpressors. Indeed, both types of transgenics accumulated sesquiterpenes like alpha- and beta-farnesene (Figure V-8) as well as cyanindin 3-O-glucoside (Figure V-6). The development of polyphenolic accumulating ray cells was also observed within the altered vasculature (Figure V-5). The overlapping transcript profiles and the phenotypes thus point to candidate genes and metabolic pathways which fall under the positive regulation of PtMYB14.

Additional effects were observed in transcript profiles of the Pro_{UBI}PtMYB14 overexpressors (data not shown), including genes belonging to other pathways of secondary metabolism than those reported in both OE (Table V-1). For instance, a large part of the downregulated sequences in ProuBlPtMYB14 could be linked to shikimate, monolignol, and benzenoid biosynthetic pathways (data not shown) while the accumulation of these transcripts was not affected in Pro_{CAD}PtMYB14. R2R3 MYB transcription factors have been implicated in the control of transcription for genes associated to phenylpropanoid (Rogers and Campbell, 2004), and benzenoid metabolism through the upstream shikimate pathway (Verdonk et al., 2005). Due to the highly conserved DNA binding domain observed among conifer MYBs (Bedon et al., 2007, chapter II), the downregulation of these genes may reflect indirect effects, such as non-specific binding to gene promoters, caused by MYB overexpression. Indeed, the transgene level was about 100-fold higher in Pro_{UBI} than Pro_{CAD}PtMYB14 when normalized to CDC2 gene transcripts. Such a dosagelinked effect would influence most strongly the expression of genes with similar
cis regulatory motifs, as suggested in *Arabidopsis* (Jin *et al.*, 2000). In addition, some of the differences in the effect of the two OE strategies may be due to a spatial expression of the transgene. Variations in the spatial profile may entail functional specificity and could involve interaction with others proteins, such as bHLH and WD40 which have been identified as important MYB partners in flavonoid pathways (Vom Endt *et al.*, 2002). In Pro_{UBI}PtMYB14, most of the anthocyanin accumulating cells were located in the outer part of paremchyma, which is expected to be unaffected by Pro_{CAD} -driven OE (Figure V-5). Congruent with anthocyanin accumulation (Figure V-6), transcripts associated with anthocyanin biosynthetic pathway such as a leucoanthocyanidin reductase and several anthocyanidin reductases were upregulated in $Pro_{UBI}PtMYB14$ (Data not shown).

Finally, as a member of the Pinaceae, spruce provides an expression system that is much closer to pine than Arabidopsis (Brassicaceae) or Nicotiana (Solanaceae) which have been employed in previous reports (Patzlaff et al., 2003a, 2003b; Newman et al., 2004). Performing PtMYB14 OE in the spruce plant system was thus particularly appropriate to study R2R3-MYB from Sg4-C, especially when considering the large number of sequences in this conifer-specific clade without angiosperm members (Figure V-2). In a recent report analysing TF binding in human and mouse experimental systems, it was shown that TF binding may be largely species-specific even for highly conserved transcription factors (Odom et al., 2007). Thus, the choice of the heterologous plant system may be of primary importance to study conserved and closely related TFs, as reported here for R2R3-MYB from Sg4-C. We have shown that the Sg4 sequences have undergone similar evolution in spruce and pine. We also showed that PtMYB14 in pine and its closest homologues in spruce, PgMYB14a and PgMYB14b, are rapidly induced in response to wounding and JA application (Figures V-9 and V-10), suggesting that putative functional orthologs are present in these species. The implication of MYB14-related sequence in defense response in spruce and pine is discussed in more detail below.

V-7-3-PtMYB14 as a putative partner in isoprenoid-oriented response leading to terpene accumulation

Most of the defensive compounds produced by plants are derived from phenylpropanoid, alkaloid, fatty acid and polyketide, and isoprenoid pathways (Dixon, 2001). Isoprenoids are important biomolecules implicated in numerous cellular processes and serving as hormones, defensive agents, membrane constituents, components of signal transduction networks, mating pheromones, and photoprotective agents (Kleinig, 1989; Sacchettini and Poulter, 1997). They are derived from isoprenic C5 unit, the isopentenyl diphosphate (IPP) and from its isomer dimethylallyl diphosphate (DMAPP), that are synthesized through two metabolic routes, the mevalonate (MVA) and the non-mevalonate (DOX) pathways (Lichtenthaler, 1999; Lange et al., 2000). The condensation of IPP and DMAPP through the action of prenyltransferases (FPPS, GPPS, and GGPPS) lead to the formation of terpene precursors such as farnesyl diphosphate, geranyl diphosphate, and geranylgeranyl diphosphate. Finally, cyclization and oxidation steps catalyzed by terpene synthase (TPS) lead to the formation of mono-, di- and sesquiterpenes. Therefore, managing the IPP pool (cytosolic and/or plastidic) may be of primarily importance for plants to produce defensive compounds such as terpenes.

Sesquiterpenes are important components of mechanical and herbivoreinduced wound response in conifers (Keeling and Bohlman, 2006; Trapp and Croteau, 2001). Nevertheless, little is known about the implication of transcription factors in the regulation of terpenoid metabolism as recently underlined (Memelink and Gantet, 2007). WRKY1 from cotton and ORCA3, a jasmonateresponsive APETALA2 (AP2)-domain TF from *Catharanthus*, have been recently reported as downstream and upstream regulators of the terpenoid pathway, respectively. GaWRKY1 was suggested to be a transcriptional activator of the cadinene synthase coding gene participating in cotton sesquiterpene biosynthesis (Xu *et al.*, 2004). ORCA3 OE led to upregulation of gene encoding DXPS and to the accumulation of terpenoid indole alkaloids (van der Fits and Memelink, 2000). To date however, and to our knowledge, no MYBs has been directly associated to the isoprenoid pathway. Here, we showed that MYB14 is likely an important partner in isoprenoid accumulation in conifers by co-ordinately acting on MVArelated genes. More specifically, both PtMYB14-OE led to the upregulation of AACT and HMGS sequences associated to MVA pathway, as well as those coding for downstream terpene synthases (Table V-1). These transcription profiles confirmed by RT-qPCR were congruent with terpene analysis performed on transgenic plantlets and indicating an accumulation of sesquiterpenes (afarnesene, β-farnesene) in both PtMYB14 overexpressers (Figure V-8). The present study also indicates a rapid and transient induction of MYB14 related sequence in spruce and pine following mechanical wounding, together with MVA related sequences (Figures V-9 and V-10). Strong induction was observed for *PtMYB14*, and *PgMYB14a* and *14b* when pine and spruce plantlets were subjected to JA treatment (Figures V-9 and V-10). These results are consistent with upregulation of transcripts that may be linked to JA biosynthetic pathway, such as lipoxygenase observed in PtMYB14 overexpressors (Table V-1) as well as slightly increased of *allene oxide cyclase* gene expression (AOC) in PtMYB14-OE plantlets (Appendix V-3 and V-4). Moreover, a gene coding for the woundinducible pathogenenis related protein (PR4b) is clearly upregulated in both transgenics (Appendix V-3 and V-4). Therefore, our results suggest a possible role of PtMYB14 in the regulation of isoprenoid synthesis and more particularly sesquiterpenoid synthesis throught a JA mediated response which could involve mechanical wounding as starting point.

V-8-Références

- Bedon F., Grima-Pettenati J. and Mackay J. (2007). Conifer R2R3-MYB transcription factors: sequence analysis and gene expression in wood-forming tissues of white spruce. (*Picea glauca*). BMC Plant Biol 7: 17
- Bailey T.L. and Elkan C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol 2: 28-36
- **Benjamini Y. and Hochberg Y. (1995).** Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Statist Soc Ser B* **57:** 289–300
- Bouché N. and Bouchez D. (2001). Arabidopsis gene knockout: phenotypes wanted. *Curr Opin Plant Biol* 4: 111-117
- Briggs G.C., Osmont K.S., Shindo C., Sibout R. and Hardtke C.S. (2006). Unequal genetic redundancies in *Arabidopsis* – a neglected phenomenon? *Trends Plant Sci* 11: 492-498
- Chang S., Puryear J., and Cairney J. (1993). A Simple and Efficient Method for Isolating RNA from Pine Trees. *Plant Mol Biol* Reporter 11: 113-116

- Chern M., Canlas P.E., Fitzgerald H.A. and Ronald P.C. (2005). Rice NRR, a negative regulator of disease resistance, interacts with Arabidopsis NPR1 and rice NH1. *Plant J* 43: 623-635
- Christensen A.H., Sharrock R.A. and Quail P.H. (1992). Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Mol Biol* **18:** 675–689
- Christensen A.H. et Quail P.H. (1996). Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Res* 5: 213-221
- **Coleman H.D., Ellis D.D., Gilbert M. and Mansfield S.D. (2006)** Up-regulation of sucrose synthase and UDP-glucose pyrophosphorylase impacts plant growth and metabolism. *Plant Biotechol J* **4**: 87-101
- Côté C. et Rutledge R.G. (2003). An improved MUG fluorescent assay for the determination of GUS activity within transgenic tissue of woody plants. *Plant Cell Rep* 21: 619-214
- **Dixon R.A. (2001).** Natural products and plant disease resistance. *Nature* **411:** 843-847
- Fluhr R. (2001). Sentinels of disease. Plant resistance genes. *Plant Physiol* 127: 1367-1374
- Gentleman R.C., Carey V.J., Bates D.J., Bolstad B.M., Dettling M., Dudoit S., Ellis B., Gautier L., Ge Y., Gentry J., Hornik K., Hothorn T., Huber W., Iacus S., Irizarry R., Leisch F., Li C., Maechler M., Rossini A.J., Sawitzki G., Smith C., Smyth G.K., Tierney L., Yang Y.H. and Zhang J. (2004). Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biol* 5: R80
- Goicoechea M., Lacombe E., Legay S., Milhaevic S., Rech P., Jauneau A., Lapierre C., Pollet B., Verhaegen D., Chaubet-Gigot N. and Grima-Pettenati J. (2005). EgMYB2, a new transcriptional activator from *Eucalyptus* xylem, regulates secondary cell wall formation and lignin biosynthesis. *Plant J* 43: 553–567
- Hawkins S., Samaj J., Lauvergeat V., Boudet A. and Grima-Pettenati J. (1997). Cinnamyl Alcohol Dehydrogenase: Identification of New Sites of Promoter Activity in Transgenic Poplar. *Plant Physiol* 113: 321-325
- Ihaka R. and Gentleman R. (1996). R: A language for data analysis and graphics. *J Comp Graph Stat* 5: 299-314
- Jiang C., Gu X. and Peterson T. (2004). Identification of conserved gene structures and carboxy-terminal motifs in the Myb gene family of Arabidopsis and Oryza sativa L. ssp. indica. *Genome Biol* **5**: R46
- Jin H.L., Cominelli E., Bailey P., Parr A., Mehrtens F., Jones J., Tonelli C., Weisshaar B. and Martin C. (2000). Transcriptional repression by AtMYB4 controls production of UV- protecting sunscreens in Arabidopsis. *EMBO J* 19: 6150–6161
- Kazan K. (2006). Negative regulation of defence and stress genes by EAR-motifcontaining repressors. *Trends Plant Sci* 11: 109-112
- Keeling C. I. and Bohlmann J. (2006). Genes, enzymes and chemicals of terpenoid diversity in the constitutive and induced defence of conifers against insects and pathogens. *New Phytol* 170 : 657-675
- Kleinig H. (1989). The role of plastids in isoprenoid biosynthesis. *Annu Rev Plant Physiol Plant Mol Biol* 40:39–59

- Klimaszewska K., Rutledge R.G. and Séguin A. (2004). Genetic transformation of conifers utilizing somatic embryogenesis. In L Peña, ed, *Methods in Molecular Biology, Humana Press, Totowa*, pp 151–164
- Klimaszewska K., Lachance D., Pelletier G., Lelu M.A. and Séguin A. (2001). Regeneration of transgenic *Picea glauca*, *P. mariana*, and *P. abies* after cocultivation of embryogenic tissue with *Agrobacterium tumefaciens*. *In Vitro Cell Dev Biol-Plant* 37: 748–755
- Koncz C.and Schell J. (1986). The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. *Mol Gen Genet* 204: 383–396
- Kranz H.D., Denekamp M., Greco R., Jin H., Leyva A., Meissner R.C., Petroni K., Urzainqui A., Bevan M., Martin C., Smeekens S., Tonelli C. Paz-Ares J. and Weisshaar B. (1998). Towards functional characterization of the members of the R2R3-MYB gene family from *Arabidopsis thaliana*. *Plant J* 16: 263–276
- Kumar S., Tamura K., Jakobsen I.B. and Nei M. (2001). MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* 17: 1244-1245
- Kusumi J., Tsumura Y., Yoshimaru H. and Tachida H. (2002). Molecular evolution of nuclear genes in cupressacea, a group of conifer trees. *Mol Biol Evol* 19: 736-747
- Lange B.M., Rujan, T., Martin, W. and Croteau R. (2000). Isoprenoid biosynthesis: The evolution of two ancient and distinct pathways across genomes. *Proc Natl Acad Sci* 97: 13172–13177
- Lee M. M. and Schiefelbein J. (1999). WEREWOLF, a MYB-related protein in *Arabidopsis*, is a position-dependent regulator of epidermal cell patterning. *Cell* 99: 473-483
- Lehti-Shiu M.D., Adamczyk B.J. and Fernandez D.E. (2005). Expression of MADS-box genes during the embryonic phase in Arabidopsis. *Plant Mol Biol* 58: 89-107.
- Li J., Yang X., Wang Y., Li X., Gao Z., Pei M., Chen Z., Qu L.J. and Gu H. (2006). Two groups of MYB transcription factors share a motif which enhances trans-activation activity. *Biochem Biophys Res Commun* 341: 1155-1163
- Lichtenthaler H.K. (1999). The 1-deoxy-D-xylulose-5-phosphate pathway of isoprenoid biosynthesis in plants. *Annu Rev Plant Physiol and Plant Mol Biol* 50: 47-65
- Martin D.M., Tholl D., Gershenzon J. and Bohlmann J. (2002). Methyl jasmonate induces traumatic resin ducts, terpenoid resin biosynthesis, and terpenoid accumulation in developing xylem of Norway spruce stems. *Plant Physiol* **129** : 1003-1018
- Mayrose M., Ekengren S., Melech-Bonfil S., Martin G. and Sessa G. (2006). A novel link between tomato GRAS genes, plant disease resistance and mechanical stress response. *Mol. Plant Pathol.* **7**: 593–604
- McGrath K.C., Dombrecht B., Manners J.M., Schenk P.M., Edgar C.I., Maclean D.J., Scheible W.R., Udvardi M.K. and Kazan K. (2005). Repressor- and activator-type ethylene response factors functioning in jasmonate signaling and disease resistance identified via a genome-wide screen of Arabidopsis transcription factor gene expression. *Plant Physiol* 139: 949-959

- Memelink J. and Gantet P. (2007). Transcription factors involved in terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. *Phytochem Rev* 6: 353-362
- Millar A.A. and Gubler F. (2005). The Arabidopsis GAMYB-like genes, MYB33 and MYB65, are microRNA-regulated genes that redundantly facilitate anther development. *Plant Cell* 17: 705-721
- Moore I., Samalova M. and Kurup S. (2006). Transactivated and chemically inducible gene expression in plants. *Plant J* 45: 651-683
- Müller D., Schmitz G. and Theres K. (2006). Blind homologous R2R3 Myb genes control the pattern of lateral meristem initiation in Arabidopsis. *Plant Cell* 18: 586-597
- Odom D.T., Dowell R.D., Jacobsen E.S., Gordon W., Danford T.W., MacIsaac K.D., Rolfe P.A., Conboy C.M., Gifford D.K. and Fraenkel E. (2007). Tissue-specific transcriptional regulation has diverged significantly between human and mouse. *Nat Genet* **39**: 730-732
- Patzlaff A., Newman L.J., Dubos C., Whetten R.W., Smith C., McInnis S., Bevan M.W., Sederoff R. R. and Campbell M.M. (2003a). Characterisation of PtMYB1, an R2R3-MYB from pine xylem. *Plant Mol Biol* 53: 597–608
- Patzlaff A., McInnis S., Courtenay A., Surman C., Newman L.J., Smith C., Bevan M.W., Mansfield S. D., Whetten R.W., Sederoff R.R. and Campbell M.M. (2003b). Characterisation of a pine MYB that regulates lignification. *Plant J* 36: 743–754
- Penfield S., Meissner R.C., Shoue D.A., Carpita N.C. and Bevan M.W. (2001). MYB61 is required for mucilage deposition and extrusion in the Arabidopsis seed coat. *Plant Cell* 13: 2777-2791
- Pervieux I.M., Bourassa F. Laurans R., Hamelin R. and Séguin A. (2004). A spruce defensin showing strong antifungal activity and increased transcript accumulation after wounding and jasmonate treatments. *Physiol Mol Plant Pathol* 64: 331-341
- Pickett F.B. and Meeks-Wagner D.R. (1995). Seeing double: appreciating genetic redundancy. *Plant Cell* 7: 1347-1356
- Rabinowicz P.D., Braun E.L., Wolfe A.D., Bowen B. et Grotewold E. (1999). Maize R2R3 Myb Genes: Sequence Analysis Reveals Amplification in the Higher Plants. *Genetics* 153: 427-444
- Riechmann J.L., Heard J., Martin G., Reuber L., Jiang C., Keddie J., Adam L., Pineda O., Ratcliffe O.J., Samaha R.R., Creelman R., PilgrimM., Broun P., Zhang J.Z.,Ghandehari D., Sherman B.K. and Yu G. (2000). *Arabidopsis* transcription factors: genome-wide comparative analysis among eukaryotes. *Science* 290: 2105-2110
- Rogers L.A. and Campbell M.M. (2004). The genetic control of lignin deposition during plant growth and development. *New Phytol* 164: 17–30
- Romero I., Fuertes A., Benito M.J., Malpica J.M., Leyva A. and Paz-Ares J. (1998). More than 80 R2R3 MYB regulatory genes in the genome of *Arabidopsis thaliana*. *Plant J* 14: 273-284
- Sacchettini J.C. and Poulter C.D. (1997). Creating isoprenoid diversity. *Science* 277: 1788–1789
- Scarpella E. and Meijer A. H. (2004). Pattern formation in the vascular system of monocot and dicot plant species. *New Phytol* 164: 209-242.

- Sharman B.C. (1943). Tannic acid and Iron alum with safranin and Orange G in studies of the shoot apex. *Stain Technol* 18: 105–111
- Schwechheimer C., Zourelidou M. and Bevan M.W. (1998). Plant transcription factor studies. *Annu Rev Plant Physiol Plant Mol Biol* 49: 127-150
- Smyth G.K. (2005). Limma: linear models for microarray data. In R Gentleman, V Carey, S Dudoit, R Irizarry, W Huber, eds, *Bioinformatics and Computational Biology Solutions using R and Bioconductor, Springer, New York*, p 397–420
- Song C.P., Agarwal M., Ohta M., Guo Y., Halfter U., Wang P. and Zhu J.K. (2005). Role of an Arabidopsis AP2/EREBP-type transcriptional repressor in abscisic acid and drought stress responses. *Plant Cell* 17: 2384-2396
- Stracke R., Ishihara H., Huep G., Barsch A., Mehrtens F., Niehaus K. and Weisshaar B. (2007). Differential regulation of closely related R2R3-MYB transcription factors controls flavonol accumulation in different parts of the Arabidopsis thaliana seedling. *Plant J* 50: 660-677
- Stracke R., Werber M. and Weisshaar B. (2001). The R2R3-MYB gene family in *Arabidopsis thaliana*. *Curr Opin Plant Biol* 4: 447–456
- **Theodoulou F.L. (2000).** Plant ABC transporters. *Biochim Biophys Acta* **1465**: 79-103
- Thompson J. D., Higgins D. G. and Gibson T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673-4680
- Topfer R., Maas C., Horicke-Grandpierre C., Schell J. and Steinbiss H.H. (1993). Expression vectors for high-level gene expression in dicotyledonous and monocotyledonous plants. *Methods Enzymol* 217: 67– 78
- Trapp S. and Croteau R. (2001). Defensive resin biosynthesis in conifers. *Annu Rev Plant Physiol Plant Mol Biol* 52 : 689-724
- van der Fits L. and Memelink J. (2000). ORCA3, a jasmonate-responsive transcriptional regulator of plant primary and secondary metabolism. *Science* 289: 295-297
- Verdonk J.C., Haring M.A., van Tunen A.J. and Schuurink R.C. (2005). ODORANT1 regulates fragrance biosynthesis in petunia flowers. *Plant Cell* 17: 1612-1624
- Vom Endt D., Kijne J.W. and Memelink J. (2002). Transcription factors controlling plant secondary metabolism: what regulates the regulators? *Phytochem* 61: 107-114
- Wang H-Y, Flatte M., Jakoby M., Bäumlein H., Weisshaar B. and Bauer P. (2007). Iron deficiency-mediiated stress regulation of four subgroup Ib BHLH genes in Arabidopsis thaliana. Plant 226: 897-908
- Xu Y.H., Wang J.W., Wang S., Wang J.Y. and Chen X.Y. (2004). Characterization of GaWRKY1, a cotton transcription factor that regulates the sesquiterpene synthase gene (+)-delta-cadinene synthase-A. *Plant Physiol* 135: 507-515
- Xue B., Charest P.J., Devantier Y. and Rutledge R.G. (2003). Characterization of a *MYBR2R3* gene from black spruce (*Picea mariana*) that shares functional conservation with maize C1. *Mol Genet Genomics* 270: 78-86
- Yanhui C., Xiaoyuan Y., Kun H., Meihua L., Jigang L., Zhaofeng G., Zhiqiang L., Yunfei Z., Xiaoxiao W., Xiaoming Q., Yunping S., Li Z., Xiaohui D.,

Jingchu L., Xing-Wang D., Zhangliang C., Hongya G. and Li-Jia Q. (2006). The MYB transcription factor superfamily of Arabidopsis: expression analysis and phylogenetic comparison with the rice MYB family. *Plant Mol Biol* 60: 107-123

- Yang Y.H., Dudoit S., Luu P., Lin D.M., Peng V., Ngai J. and Speed T.P. (2002). Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* 30: e15
- Yang Z., Tian L.-N., Latoszek-Green M., Brown P.D. and Wu K. (2005). Arabidopsis ERF4 is a transcriptional repressor capable of modulating ethylene and abscisic acid responses. *Plant Mol Biol* **58**: 585-596
- Yoo J.H., Park C.Y., Kim J.C., Do Heo W., Cheong M.S., Park H.C., Kim M.C., Moon B.C., Choi M.S., Kang Y.H., Lee J.H., Kim H.S., Lee S.M., Yoon H.W., Lim C.O., Yun D-J., Lee S.Y., Chung W.S. and Cho M.J. (2005). Direct Interaction of a Divergent CaM Isoform and the Transcription Factor, MYB2, Enhances Salt Tolerance in Arabidopsis. J Biol Chem 280: 3697-3706
- Zhang J.Z. (2003). Overexpression analysis of plant transcription factors. *Curr Opin Plant Biol* 6: 430-440
- Zhang P., Chopra S. and Peterson T. (2000). A segmental gene duplication generated differentially expressed myb-homologous genes in maize. *Plant Cell* 12: 2311-2322
- Zimmermann I.M., Heim M.A., Weisshaar B. and Uhrig J.F. (2004). Comprehensive identification of Arabidopsis thaliana MYB transcription factors interacting with R/B-like BHLH proteins. *Plant J* 40: 22-34

Chapitre VI

Etude de l'expression de *PtMYB14* dans les tissus vasculaires d'épinettes blanches sous le contrôle du promoteur du gène de l'*alcool cinnamylique deshydrogénase (CAD)*

VI-1-Avant-propos

Ce chapitre traite de l'analyse des lignées exprimant les constructions Pro_{CAD}GUS et Pro_{CAD}PtMYB14 dans de jeunes arbres en serre, et, par conséquent, il représente la suite du chapitre V. Ce chapitre n'a pas été soumis pour publication. J'ai réalisé l'exécution de l'ensemble des travaux et assuré la rédaction. Plus précisément, je suis intervenu au niveau de la mise en place des dispositifs expérimentaux (induction du bois de compression, traitements de blessures des arbres et d'application d'acide jasmonique), de la quantification des transcrits d'ARN par RT-PCR quantitative, de la microscopie (traitements des échantillons et colorations histologiques GUS) et de la détermination des activités GUS (avec l'aide de C. Levasseur). Les analyses transcriptomiques des transgéniques par microarrays ont été effectuée avec N. Dallaire. P.-O. Nadeau a fait toutes les analyses des fibres sur mes échantillons, et j'ai réalisé les analyses statistiques des données. Les analyses de chimie du bois ont été effectuées par les membres du laboratoire de Dr. C. Lapierre (INA-PG). La transgénèse et l'entretien des plants se sont effectués au Centre de Foresterie des Laurentides à Québec, sous la responsabilité de C. Levasseur au sein de l'équipe de Dr. A. Séguin.

VI-2-Résumé

Les résultats décrits dans le chapitre V suggèrent que PtMYB14 est impliqué dans des mécanismes de défense. Afin de limiter les effets pléiotropes et finalement létaux associés à l'utilisation d'un promoteur constitutif pour contrôler l'expression du transgène PtMYB14, nous avons choisi d'utiliser un promoteur qui dirige l'expression préférentiellement dans les tissus vasculaires. Notre choix basé sur la littérature s'est porté sur le promoteur de l'alcool cinnamylique deshydrogénase (Cinnamyl Alcohol Dehydrogenase, Pro_{CAD}). Ce promoteur n'ayant jamais été caractérisé chez un conifère, nous l'avons cloné chez l'épinette blanche (Picea glauca) et caractérisé. Nous avons ainsi montré qu'une séquence nucléotidique de 1,2 Kb, située en amont du gène CAD, est suffisante pour diriger l'expression du gène uid-A (GUS) dans les cellules du xylème en différenciation d'épinettes blanches en troisième cycle de croissance en serre. Les résultats obtenus par PCR quantitative confirment que le gène CAD est préférentiellement exprimé dans le xylème en différenciation. De plus, nous avons montré que le promoteur CAD est inductible par la blessure, et que le gène CAD est plus fortement induit par l'acide jasmonique que par la blessure. L'étude des arbres exprimant la construction Pro_{CAD}PtMYB14 a permis d'observer l'absence d'effets de ce MYB sur la formation du xylème secondaire. En dépit de l'absence de phénotype évident chez ces arbres, les analyses transcriptomiques réalisées sur des pousses terminales, sur du bois entier et du bois de compression, ont montré une dérégulation de l'expression de gènes impliqués dans les mécanismes de défense. Ce résultat renforce ceux obtenus chez les jeunes plantules. Enfin, certains des gènes dont le niveau d'expression est modifié par l'expression de PtMYB14 sont aussi dérégulés lors d'induction à la blessure ou à l'acide jasmonique dans de jeunes arbres non transformés.

VI-3-Introduction et mise en contexte

La surexpression ubiquitaire et élevée de *PtMYB14* avec le promoteur *UBI* (Pro_{UBI}) a un effet très fort, voire létal, sur le développement des plantules d'épinettes blanches, empêchant ainsi leur transfert en sol et leur étude pendant la formation du xylème secondaire (Chapitre V). Toutefois, l'expression du gène *PtMYB14* sous le contrôle du promoteur *CAD* n'a pas d'effet létal, et est suffisante pour provoquer la dérégulation de l'expression de certains gènes de la voie de biosynthèse du mévalonate et d'autres gènes liés aux mécanismes de défense. Le transfert en sol a été fait pour toutes les lignées « Pro_{CAD}PtMYB14 ». Par ailleurs, afin de caractériser l'activité du promoteur CAD, nous avons également mis en sol les lignées d'épinettes blanches exprimant la construction Pro_{CAD}GUS. Ainsi, nous avons pu étudier en parallèle et de façon complémentaire des épinettes blanches exprimant les constructions Pro_{CAD}PtMYB14 et Pro_{CAD}GUS, sur trois cycles de croissance. Les résultats complémentaires au chapitre V présentés ici traitent donc de l'isolement et de la caractérisation de la séquence du promoteur CAD et de son profil d'activité dans différents tissus et organes d'épinettes. Dans le but d'identifier des gènes qui pourraient être eux-mêmes soumis à la régulation par PtMYB14, nous avons étudié les gènes dont les niveaux d'expression sont dérégulés par la construction *Pro_{CAD}PtMYB14* dans la pousse terminale et dans le xylème secondaire en différenciation du bois normal et du bois de compression.

VI-4-Matériels et méthodes

VI-4-1-Isolement de 1,2 Kb de région génomique promotrice du gène *CAD* d'épinette blanche

Le promoteur du gène de l'alcool cinnamylique deshydrogénase (*CAD*) d'épinette blanche (*Picea glauca*) a été isolé par la technique de PCR basée sur une procédure utilisant la marche sur le chromosome ou « GenomeWalker Universal Kit » (BD Biosciences, USA). L'ADN génomique a été extrait des aiguilles d'épinette blanche avec la trousse « Genomic-Tip Kit » (Qiagen, Mississauga, Ontario). Deux amorces non chevauchantes et spécifiques ("Pro_{CAD}-GSP1": 5'-TGTAAGTGTAAGGGGACAAATGGCCAC-3' et "Pro_{CAD} -GSP2": 5'-AGTCCCGAGCTGCATATCCTGTAACAG-3') ont été dessinées sur la séquence ADNc du gène CAD d'épinette blanche issue du programme de séquençage d'ESTs et disponible dans la base de données Genbank. Deux cycles de réaction d'amplification par PCR ont été effectués en utilisant dans un premier temps l'amorce "Pro_{CAD}-GSP1" et une amorce spécifique de l'adaptateur. Puis dans un second temps une PCR nichée est réalisée avec l'amorce "Pro_{CAD}-GSP2" et un autre adaptateur spécifique. Ceci a permis d'isoler un fragment de 1,2 Kb issu d'une banque d'ADN génomique digérée par StuI et préalablement liguée à des séquences adaptatrices de la trousse « GenomeWalker ». Toutes les réactions PCR ont été effectuées avec le mélange « Advantage 2 Polymerase » (BD Biosciences, USA). Le produit PCR a ensuite été ligué au plasmide pCR2.1 avec la trousse de clonage « TA cloning » (Invitrogen, Carlsbad, CA) et complètement séquencé (Figure VI-1). A partir de cette séquence nous avons utilisé l'enzyme polymérase « high fidelity Taq DNA polymerase » (Invitrogen) pour amplifier le promoteur Pro_{CAD} sur de l'ADN génomique d'épinette avec les amorces "forward" et "reverse" contenant des sites de restriction ("Pro_{CAD}-Forward" (PstI): 5'-<u>CTGCAG</u>TACTATAGGGCACGCGTGGTCGACGGC-3', "Pro_{CAD}-Reverse" (EcoRI): 5'-GAATTCTTTTTTAAAACCACTTCACAGGATTCG-3'). La réaction de polymérase en chaîne a été faite dans 25 uL avec les étapes suivantes : pré-dénaturation à 94°C / 15 min. suivie par 34 cycles de 92°C / 1 min., 60°C /1 min., 72°C / 1 min. puis une étape d'élongation de 5 min à 72°C. Le produit PCR obtenu a finalement été cloné dans le plasmide pCR2.1 avec la trousse « TA cloning » (Invitrogen, Carlsbad, CA) et séquencé. En parallèle, le produit PCR correspondant à *Pro_{CAD}* - pCR2.1 et le plasmide pCAMBIA 1391Z ont été digérés avec les deux enzymes PstI and EcoRI (Roche) de façon séquentielle, extraits du gel (Qiagen), purifiés avec les colonnes QIAquick (QIAquick Nucleotide Removal Kit), et finalement ligués ensemble avec l'ADN T4 ligase (Invitrogen) toute une nuit à 14°C. Le produit de ligation Pro_{CAD} pCAMBIA 1391Z a ensuite été introduit par choc thermique dans des cellules compétentes TOP10F' (Invitrogen, Carlsbad, CA).

VI-4-2-Détermination du site d'initiation de la transcription dans le promoteur *CAD* par RACE-5'

Le site d'initiation de la transcription du promoteur *CAD* d'épinette a été déterminé par RACE-5' (5'-rapid amplification of cDNA ends). L'ARN total a été

extrait de pousses terminales (tige de la nouvelle croissance de l'année) d'épinette blanche en suivant la procédure de Chang et al. (1993) comme décrite dans Bedon et al. (2007). La synthèse de l'ADN complémentaire de type RACE-5' et les PCR ont été effectuées en utilisant la trousse d'amplification « SMART RACE cDNA Amplification Kit» (BD Biosciences Clontech, CA, USA) avec 1 µg d'ARN total. Un programme de type « Touchdown PCR » a été utilisé pour la réaction d'amplification RACE-5' sur la machine "DNA engine PTC-225 Thermal Cycler" (Biorad, Hercules, CA), avec les paramètres suivant: cinq cycles de deux étapes à 94°C pendant 30 sec. et 72°C pendant 3 min., cinq cycles de trois étapes à 94°C pendant 30 sec., 70°C pendant 30 sec. et 72°C pendant 3 min. et vingt-cinq cycles de trois étapes à 94°C pendant 30 sec., 68°C pendant 30 sec. et 72°C pendant 3 min. Les amorces PCR utilisées incluent les amorces universelles fournies par le fabricant et l'amorce "reverse" spécifique ("PgCAD-R5": 5'-CTTGAGTAGGGGTGCCGTCATGGTT-3'). Les produits PCR issus de la RACE-5' ont ensuite été clonés dans le plasmide pCR2.1 (Invitrogen, Carlsbad, CA, USA) et 8 clones positifs séquencés pour déterminer le site d'initiation de la transcription.

VI-4-3-Construction du plasmide vecteur d'expression et matériel végétal

Les caractéristiques du vecteur de transformation contenant le promoteur *CAD* de *Picea glauca* fusionné au gène *PtMYB14* (*Pro_{CAD}PtMYB14*) ainsi qu'au gène *uid-A* ou *GUS* (*Pro_{CAD}GUS*) sont précisées dans le chapitre V. La méthode de culture et de transformation génétique du tissu embryogène d'épinette blanche clone Pg653 par co-culture est détaillée dans les articles de Klimaszewska et *al*. (2001 et 2004) ainsi que dans les chapitre III et V de cette thèse.

Les plants d'épinette blanche utilisés dans cette étude sont tous issus du clone *Picea glauca* Pg653 du Centre de Foresterie des Laurentides de Québec (Québec, Canada). Les plantules transgéniques et témoins (non transformées) ont été transplantées en pot de 3 litres après dix à quatorze semaines passées en boîte de Pétri et un cycle de croissance dans des contenants d'un litre (voir chapitre II). Les cycles de croissance végétative étaient d'environ huit à dix semaines de

façon à obtenir des bourgeons bien clos et des arbres synchronisés dans leur croissance, puis suivis de cycles de dormance de six semaines minimum en chambre froide à 4°C. Les arbres de troisième cycle de croissance « $Pro_{CAD}PtMYB14$ » et « $Pro_{CAD}GUS$ » ont été positionnés de façon "randomisée" dans les serres et dans les chambres froides. Les arbres ont été fertilisés une fois par semaine avec 20g/L de N-P-K.

VI-4-4-Coloration histochimique et tests enzymatiques de la ßglucuronidase

La préparation des échantillons issus des arbres « Pro_{CAD}GUS » et des témoins de troisième cycle de croissance pour coloration histochimique est adaptée de l'article de Hawkins et al. (1997). Brièvement, des échantillons de racines, aiguilles, pousses terminales, sommets et bases de tiges principales (ou troncs) ont été sectionnés transversalement sur environ 5 à 50 mm de longueur selon l'organe puis prétraités 30 minutes dans de l'acétone froid 90% pour bloquer toute induction liée à la blessure (Hemerly et al., 1993) et faciliter la pénétration du substrat. Les échantillons ont ensuite été rincés deux fois avec du tampon phosphate de potassium à 100 mM (pH 7,0) et mis à incuber dans de l'acide 5-bromo-4-chloro-3-indolyl-ß-D-glucuronique à l'obscurité à 37°C et sous vide pendant 6 à 12 heures jusqu'à l'observation d'une coloration bleue suffisamment développée. Puis les échantillons ont été fixés dans du PBS 0,05 M final à pH 7,0 contenant 2% de paraformaldéhyde et 2,5% de glutaraldéhyde, pendant une nuit. Le tampon de fixation a ensuite été changé et les échantillons conservés à 4°C pour être par la suite sectionnés à main levée ou bien pour être fixés dans de la paraffine.

Les tests enzymatiques de l'activité de la protéine GUS ont été réalisés tel que décrits en détail dans l'article de Côté et Rutledge (2003). A la suite de la macération et du traitement de chaque tissu finement broyé (aiguilles et sections de troncs), les concentrations en protéines ont été déterminées en utilisant la trousse « Bio-Rad Bradford Protein Assay » (Bio-rad, Hercules, California) en fonction d'une gamme standard de la protéine BSA diluée. Les concentrations en ADN ont été déterminées en fonction d'une gamme standard d'ADN dilué (Lambda DNA, Roche, Indianapolis, Indiana). Les activités GUS ont été mesurées en double sur deux échantillons par arbre « Pro_{CAD}GUS » en troisième cycle de croissance, et ceci pour deux arbres d'une même lignée.

VI-4-5-Analyse de la qualité des fibres et analyses chimiques du bois

L'analyse des caractéristiques des fibres a été effectuée sur des sections d'environ un centimètre de long de section de tronc, sans écorce, issues d'épinettes blanches synchronisées en fin de troisième cycle de croissance en serre. Les sections ont été prélevées juste en dessous de la pousse terminale (formation du second cerne), et à la base de l'arbre, au dessus du collet (formation du troisième cerne). Les sections de bois ont été macérées dans des fioles de verre de 20mL avec la solution de Franklin dans un bain-marie à 65°C pendant 76 heures, puis les fibres ont été rincées deux fois avec de l'eau distillée et conservées dans du tampon Tris Phosphate 0,1 M à 4°C (Franklin, 1945). Douze arbres de chaque lignée transgénique (30-7 et 30-9) et témoin ont été utilisés. Un minimum de 8000 fibres individuelles par échantillons a été caractérisé. Les paramètres utilisés pour cette étude étaient les moyennes arithmétiques de longueur (Ln et Lw) et de diamètre des fibres (voir détails dans Pitre et *al.*, 2007). Les analyses ont été réalisées de façon automatique avec un appareil « Fiber Quality Analysis » utilisant la lumière polarisée circulaire (OpTest Equipment, Hawkesbury, Canada).

La détermination du contenu en lignine a été réalisée sur 150 mg de poudre de bois (ou résidu pariétal RP) selon la procédure Klason (Dence, 1992). Les dosages lignine Klason ont été effectués en double indépendant (pas le même jour) et sur trois lots de 12 arbres de troisième cycle de croissance pour la lignée transgénique exprimant la construction $Pro_{CAD}PtMYB14$ (30-9) et la lignée témoin non transformée. Les résultats sont exprimés en % pondéral du résidu pariétal (en % du RP).

VI-4-6-Isolement de l'ARN et analyses de l'expression des gènes par PCR quantitative

Les ARN totaux correspondant à chaque expérience et échantillon de cette étude ont été extraits selon le protocole de Chang et *al.* (1993). Ce protocole a été

adapté pour être réalisé en microtubes Eppendorf à partir de poudre finement broyée (Annexe I-1). La concentration et la qualité des ARN ont été déterminées avec un bioAnalyseur (model 2100, Agilent Technologies; RNA 6000 Nano AssayKit).

L'analyse de l'expression des gènes par PCR quantitative a été effectuée à partir de 1 microgramme d'ARN total de chaque échantillon converti en ADNc comme décrit dans le chapitre VI et en utilisant la trousse Qiagen et l'appareil Roche LC480. Les paires d'amorces ont été dessinées avec le logiciel Primer 3 (Rozen et Skaletsky, 2000).pour se fixer dans la région UTR-3' des gènes transcrits pour assurer la spécificité des amorces. Les amorces "forward" et "reverse" sont les suivantes (le numéro d'identification du clone est indiqué entre parenthèse ainsi que la longueur de l'amplicon), la signification des acronymes des gènes étant expliquée dans la légende des figures les utilisant:

PtMYB14 (les amorces s'hybridant dans la région du terminateur 35S de la construction *Pro_{UBI}PtMYB14* ont été utilisé, pas d'identifiant de clone, 101 pb): 5'-CAAAAATCACCAGTCTCTCT-3' et 5'-ACCCTATAAGAACCCTAATTCC-3'; PgMYB14a (pas d'identifiant de clone, 153 5'-AGGGATGAAGGGAACCATCG-3' 5'pb): et TGAATTATAGCGGTGGACTC-3'; PgMYB14b (pas d'identifiant de clone, 120 5'-AGGGATGAAAGGAACGCG-3' 5'pb): et ATAAGCAATCCATTAATCTTCC-3'; C4H (GQ0021b H10, 91 pb): 5'-TCTCACTTAAGTCGGTCCAG-3' et 5'-CTTCCTGGTAATGCAAAAAG-3'; DEF (GQ0131 B10, 169 pb): 5'-CTACTACTGTGATCTTCTCTGGTTT-3' et 5'-GTAAGTAAAAGATAACCTGAACCAC-3'; CAD (GQ00410 D06, 141 pb): 5'-CTGGACTACATCAATACTGC-3' et 5'-GATTTACTCATTCTGCACG-3'; CAMT (GQ0072_A09, 144 pb) : 5'-ATCATCAAGTTGCCTTTCC-3' et 5'-CGAGCAACTGAACTTATCAAC-3'; CDC2 (GQ0197 L17, 96 pb): 5'-GTGCAGAGAAAAAGTCGAAC-3' et 5'-CCACACCATATGTTCCTTCT-3'; 6PGD (GQ0068 D11, 183 pb): 5'-TCCTATAGAAGGGATACTTTGC-3' et 5'-ATATGGGCTTATTCTCTCTGG-3'; StrictS (GQ0193 F03, 131 pb): 5'-AGGTGTGAGAGAAGATGGAG-3' et 5'-GTGAGCGACCAATAGAGTTC-3'; SIP1 (GQ0014b H02, 141 pb): 5'-CTACACTCGGGAATCAGAAC-3' et 5'-GGAATCAAGATGCTGTGG-3'

BQR (GQ0131_B24, 97 pb): 5'-GTTTGTGATAGGGTACTTGTCC-3' et 5'-GGCAACTAACGAGAGTTTTG-3'; *DXS* (GQ0206.TB_O24, 108 pb): 5'-CTAATAGGCAAGCCTCAGAA-3' et 5'-AGGCTTTGCTTCTCTACCTT-3'; *RIBO* (GQ027119_G04, 190 pb): 5'-GCTTGTATTTCCTCCTTTTG-3' et 5'-AGTCAAATTGAATCGTCACC-3'.

VI-4-7-Expériences microarray sur des épinettes « Pro_{CAD}PtMYB14 » de troisième cycle de croissance et analyses statistiques

Les trois expériences microarray ont été réalisées sur des épinettes blanches en fin de troisième cycle de croissance et de façon randomisée. Pour l'expérience de bois de compression (expérience BC), quatre arbres Pro_{CAD}PtMYB14 (ligné 30-14) ont été maintenus inclinés à environ 45 degrés pendant 21 jours versus quatre arbres témoins (non transformés) verticaux. Le xylème en différenciation de la partie de l'arbre face au sol, représentant le bois de compression, a été raclé avec un « épluche patate », et, un des cotés des arbres verticaux a été choisi pour être prélevé et immédiatement congelé dans de l'azote liquide. Ensuite, les pousses terminales (expérience PT) de cinq arbres « Pro_{CAD}PtMYB14 » (ligné 30-9) et cinq arbres témoins (non transformés) ont été prélevées et les aiguilles enlevées après congélation dans l'azote liquide de façon à réduire la perception de stress. Enfin, quatre arbres « Pro_{CAD}PtMYB14 » (ligné 30-14) et quatre arbres témoins (non transformés) ont été entièrement écorcés (expérience BE) après incision longitudinale de l'écorce, la tige de bois sectionnée en morceaux de 2 cm de long puis rapidement congelés dans l'azote liquide. Les échantillons PT et BC ont été broyés dans de l'azote liquide à l'aide de mortiers et de pilons, les échantillons BE ont été broyés avec un « Mixer Mill MM300 » (Retsch), jusqu'à l'obtention d'une fine poudre. Puis les ARNs totaux ont été extrait pour chacun des échantillons et 1 microgramme d'ARN total a été utilisé pour une amplification indirecte de l'ARN avec la trousse « Superscript TM Indirect RNA Amplification System» (Invitrogen, Carlsbad CA) selon les instructions du fabricant. Ensuite, le couplage fluorescent a été effectué sur 5 microgrammes d'ARN amplifié (aRNA) avec les marqueurs fluorescent, soit le « Alexa Fluor® 555 » soit le « Alexa Fluor ® 647 » (Invitrogen, Carlsbad CA).

Le protocole d'hybridation des biopuces d'épinette blanche est décrit dans les chapitre III et V (Annexe III-2) de même que les méthodes d'analyses statistiques des données utilisées. Dans cette étude, nous nous sommes particulièrement intéressés aux gènes différentiellement exprimés qui avaient une *P*-value <0,01 pour les expériences BC et PT et une *P*-value <0,02 pour BE, à l'issue de l'analyse LIMMA, entre les individus témoins et transgéniques « $Pro_{CAD}PtMYB14$ ». Pour chaque expérience (BC, BE et PT), les résultats d'expression obtenus représentent les moyennes sur l'ensemble des répétitions biologiques.

VI-4-8-Application de jasmonate et blessure d'épinettes de deuxième cycle de croissance non transgéniques

Des épinettes blanches de génotype identique (Picea glauca clone 653) de deuxième cycle de croissance en serre ont été utilisées pour analyser les niveaux d'expression de plusieurs gènes en réponse à l'application d'acide jasmonique (AJ) et à la blessure (W). L'acide jasmonique a été dissous dans du méthanol et mis à la concentration de 800 µM avant d'être pulvérisé sur quatre arbres (un arbre par pot de 1 litre et environ 12 mL d'aspersion par arbre), les quatre arbres témoins étant vaporisés avec du méthanol uniquement et situés dans une serre différente avec des conditions de culture identiques. L'expérience de blessure a été effectuée en triturant quatre épinettes, de l'extrémité de la pousse terminale jusqu'à la base de l'arbre, avec une pince environ dix fois; les quatre témoins n'étant pas blessés et situés à une autre extrémité de la serre. Les deux expériences (AJ et W) ont duré exactement 24 heures, puis cinq pousses terminales avec aiguilles ont été récoltées suivi du prélèvement de la pousse terminale (xylème et écorce) pour chacun des arbres de l'une et l'autre des expériences. Les échantillons ont été immédiatement plongés dans l'azote liquide, les pousses terminales avec aiguilles broyées à l'aide de mortiers et de pilons, et les tiges principales broyées au «Mixer Mill MM300» (Retsch), jusqu'à l'obtention d'une fine poudre.

VI-5-Résultats

VI-5-1-Isolement et analyse des séquences en amont (5') du gène *CAD* d'épinette blanche (présumé promoteur du gène *CAD*, *Pro_{CAD}*)

Un fragment d'environ 1,2 Kb situé en amont (position 5') du gène CAD a été isolé par la technique de marche sur le chromosome ou «Genome Walking». L'analyse de la séquence en utilisant le logiciel PLACE (Higo et al., 1999) a révélé que ce présumé promoteur du gène CAD contient plusieurs séquences consensus qui sont des cibles de régulation chez les eucaryotes (Figure VI-1). La première boîte TATA se situe entre 30 et 22 nucléotides en amont du site d'initiation de la transcription qui se trouve lui même à 77 nucléotides de la méthionine du gène CAD. Le site d'initiation de la transcription (+1) du gène CAD a été identifié après le clonage et le séquençage de huit produits PCR différents issus de RACE-5'. Le premier nucléotide transcrit apparaît être une adénine entourée de deux pyrimidines, ce qui est en accord avec des études antérieures sur les sites d'initiation de la transcription (Joshi, 1987). Une boîte TATA alternative se trouve entre -504 et -498 pb. Des boîtes CAAT sont également identifiées à proximité de la boite TATA, les plus proches étant à 6, 30 et 63 pb en amont, mais une boite CAAT est également présente au niveau de l'extrémité UTR-5'. En plus des boites TATA et CAAT, plusieurs autres éléments régulateurs potentiels en lien avec l'expression du gène sont présents dans la région en amont du gène CAD. Cela inclut des boîtes pour la fixation de protéines de type WRKY impliquées dans les mécanismes de défense (4 boîtes), des sites MRE pour la fixation aux protéines MYB (4 boîtes dont une dans l'UTR-5') et un élément ERE de réponse à l'éthylène (1 boîte) (Figure VI-1).





La numérotation est relative au site probable d'initiation de la transcription désigné par +1 en gras italique, la région 5' transcrite non traduite est indiquée en italique jusqu'au codon méthionine (M). Les éléments *cis*-régulateurs potentiels ont été identifiés avec le logiciel PLACE (Higo et *al.*, 1999). Les boîtes probables TATAbox, CAAT-box (soulignés quand deux CAAT-box sont fusionnées en position sens et antisens) sont indiquées en gras, les CCAAT-box sont encadrées. Les sites probables de fixation des facteurs de transcription DOF (AAAG) sont soulignés. Les carrés en pointillés montrent des séquences probables de fixation de facteurs de transcription indiqués entre parenthèses: ERE, « Ethylene Responsive Element »; CCA1, « myb-related transcription factor CCA1 »; WRKY; MRE, « Myb Responsive Elements » (WAACCA est MYB1AT; CCWACC est MYBPZM).

VI-5-2-Phénotypes, niveaux d'expression et activités enzymatiques de la glucuronidase chez les arbres exprimant la construction moléculaire *Pro_{CAD}GUS*

La transformation génétique de l'épinette blanche avec la construction moléculaire contenant le promoteur CAD fusionné avec le gène rapporteur uid-A ou GUS (Pro_{CAD}GUS) a permis de générer plusieurs lignées d'arbres qui ont été cultivés pendant trois cycles de croissance. Ces arbres croissent et se développent de façon identique aux lignées témoin non transformées (Figure VI-2 A). L'activité transcriptionnelle in planta du promoteur CAD a été mesurée et validée par PCR quantitative avec des amorces ciblant le terminateur NOS du gène GUS à partir d'ARN messagers isolés de deux structures végétales très différentes : la base des troncs de jeunes arbres (écorce et bois) et des aiguilles, chez trois lignées transgéniques différentes et une lignée témoin. Le nombre de molécules de transcrits correspondant aux gènes GUS et CAD est du même ordre de grandeur dans les aiguilles et dans les sections de tronc, bien que sensiblement plus élevé pour le gène CAD dans les sections de troncs (Figure VI-2 B). L'activité enzymatique de la protéine GUS a également été mesurée dans les mêmes échantillons. Les activités enzymatiques observées dans les sections de troncs sont seulement informatives sur la fonctionnalité du promoteur CAD car nous avons observé une précipitation des protéines pendant les réactions enzymatiques qui ne se produit pas chez les aiguilles. De ce fait, les résultats montrent des activités enzymatiques GUS plus élevées dans les aiguilles que dans les sections de troncs, qu'ils soient verticaux ou inclinés (Figure VI-2 C).



Figure VI-2. Phénotypes, accumulation des transcrits et activités enzymatiques GUS chez les épinettes transgéniques exprimant la construction moléculaire $Pro_{CAD}GUS$.

A: Phénotype des trois lignées L-1, L-6 et L-7 exprimant $Pro_{CAD}GUS$, et la lignée témoin T non transformée, B: Nombre de molécules de transcrits des gènes GUS et CAD par nanogramme d'ARN total calculé selon le matériel et méthode du chapitre II, C: Activités enzymatiques de la protéine GUS dans les aiguilles et les sections de troncs (troncs verticaux et inclinés pendant 21 jours). Les sections de troncs ont été prélevées à la base de l'arbre au dessus du collet et incluent l'écorce et le bois ensemble sur environ 1,5 centimètres de long. Les moyennes et les barres d'erreurs ont été calculées sur deux échantillons par arbre « $Pro_{CAD}GUS$ » en troisième cycle de croissance, pour deux arbres d'une même lignée avec deux répétitions techniques par échantillon.

VI-5-3-Le promoteur du gène *CAD* dirige l'expression dans le xylème en différenciation

Dans le but de connaître dans quels tissus le gène CAD s'exprime préférentiellement, nous avons déterminé son profil d'expression par RT-PCR quantitative dans plusieurs tissus différents chez une épinette blanche âgée de trente ans (aiguilles; périderme, phloème en différenciation et xylème en différenciation issus du tronc; écorce et xylème en différenciation de racine). L'analyse de l'abondance des transcrits du gène CAD a révélé une préférence d'expression pour le xylème en différenciation de tronc et de racine (Figure VI-3). De plus, la localisation histochimique de l'activité GUS conduite par le promoteur CAD pendant la formation du bois a été observée dans le xylème en différenciation, la zone cambiale et le phloème en différenciation des pousses terminales en cours de lignification, c'est-à-dire au milieu du cycle de croissance (Figure VI-4-A et B). On remarque une absence de coloration dans la pousse terminale en début de cycle de croissance. L'étude histochimique de l'activité GUS au niveau de la formation du second et du troisième cerne de croissance montre également une localisation de l'expression du promoteur CAD dans les mêmes tissus (Figure VI-4 C, D et E) mais aussi dans les cellules de rayon (Figure VI-4 D). Des observations histologiques sur des sections transversales "déparaffinées" ont permis une localisation plus fine de l'expression GUS dans les différents types cellulaires du xylème et du phloème en différenciation, du cambium et des cellules de rayon (Figure VI-5). Le même profil d'expression vasculaire de l'activité GUS est observé dans les racines et les aiguilles de l'épinette blanche exprimant Pro_{CAD}GUS (Figure VI-6).



Figure VI-3. Niveaux d'expression du gène *CAD* de *Picea glauca* dans différents tissus et organes d'une épinette blanche de 30 ans.

Aig.: Aiguilles; Pér.: Périderme; Ph.: Phloème en développement (tissus raclé avec un scalpel); Xy.: Xylème en développement (récolté avec un « pelle-patate » pour avoir le tissu en croissance sans le bois).

Chapitre VI – Caractérisation de la construction Pro_{CAD}PtMYB14 dans de jeunes arbres





Figure VI-4. Localisation histochimique de l'activité GUS pendant la formation du bois chez des épinettes blanches transformées avec $Pro_{CAD}GUS$ (sections transversales à main levée).

Les analyses GUS ont été effectuées sur des arbres de troisième cycle de croissance cultivés en serre. Pa: parenchyme cortical; Ph: phloème; Xy: xylème; ZC: zone cambiale; CC: cycle de croissance végétative. Grossissement 20X. A, B : Pousse terminale ou tige primaire (formation du premier cerne); C, D : Stade de formation du second cerne (section en dessous de la pousse terminale); E : Stade de formation du troisième cerne (section à la base de l'arbre).





Les arbres sont dans leur troisième cycle de croissance. A (20X), B (40X), C (40X) et D (80X) sont des sections transversales de 8 micromètres issues de la base de l'arbre, au dessus du collet. D montre un élément de vascularisation dans l'écorce correspondant au point d'attachement d'une branche. E (20X) est une section de la pousse terminale avec un agrandissement en 60x en F. Xy: xylème. Ph:Phloème. BC: bois de compression avec des cellules à paroi épaissie après 30 jours d'inclinaison à 45 degrés. R: Cellules de rayon. Les traits en pointillés représentent la localisation de la zone cambiale.



Figure VI-6. Localisation histochimique de l'activité GUS dans les racines et les aiguilles chez des épinettes blanches transformées avec *Pro_{CAD}GUS*.

A. Racine entière montrant le profil vasculaire de l'expression de $Pro_{CAD}GUS$ (grossissement 4X); (B) et (C) indiquent les emplacements des sections transversales présentées respectivement en B (30X) et C (30X). B à E, sections transversales à main levée. D (50X) et E (50X) sont des sections transversales d'aiguilles du témoin et d'un transgénique, respectivement. Xy: xylème; Ph: phloème; PC: parenchyme cortical.

VI-5-4-Le gène *CAD* est inductible par la blessure et l'acide jasmonique

Nous avons cherché à savoir si le gène CAD de *Picea glauca* impliquée dans les processus de lignification et de développement normal de l'arbre, est aussi impliqué dans la formation de lignine à la suite de blessures ou d'attaque de pathogènes. La blessure pratiquée avec une pointe de « punaise à babillard » enfoncée à la base d'un arbre « $Pro_{CAD}GUS$ » jusqu'au xylème pendant trois jours consécutifs est suffisante pour induire une augmentation de l'activité GUS localement dans les cellules du xylème (Figure VI-7). De plus, l'analyse par PCR quantitative du gène *CAD* après blessure physique ou vaporisation d'acide jasmonique montre une augmentation significative du niveau d'expression de *CAD* autant dans les pousses terminales que dans le tronc entier, l'augmentation étant plus élevée avec l'application d'acide jasmonique (Figure VI-8).



Sans blessure

Trois jours après blessure

Figure VI-7. Localisation histochimique de l'activité GUS lors d'une blessure dans le tronc d'épinette blanche « Pro_{CAD}GUS ».

Le triangle en pointillé indique l'endroit où a été enfoncée une pointe de punaise à babillard pendant trois jours pour induire la blessure. Les arbres sont en fin de troisième cycle de croissance végétative. Les sections transversales (40X) sont faîtes à mains levées sur des blocs de tige prélevés à la base de l'arbre.





La quantification des ARN messagers a été déterminée par PCR quantitative sur des clones somatiques d'épinettes blanches de fin de deuxième cycle de croissance en serre (clone Pg653). Quatre arbres par traitement ont été utilisés et récoltés 24 heures après 1) vaporisation d'acide jasmonique (AJ) 800µM dissous dans du méthanol ou bien de méthanol seul pour les contrôles (C), 2) blessures en pinçant et tordant la tige principale, ou tronc, environ 10 fois (W) ou sans intervention (C). PT+Aig.: pousses terminales avec aiguilles, tronc: tige principale entière (écorce et xylème ensemble). *CAD: alcool cinnamylique deshydrogénase*, *CDC2: cycline 2*. Le niveau d'expression de *CDC2* est présenté dans la figure VI-13. Les étoiles indiquent le niveau de signification (*P*-value) du test de Student entre le contrôle (C) et l'essai (AJ ou W), une étoile :*P* < 0,05, deux: *P* < 0,01, trois: *P* < 0,001. D'autres gènes sont étudiés de la même façon (voir figure VI-13).

VI-5-5-Caractérisation des épinettes « Pro_{CAD}PtMYB14 » en troisième cycle de croissance

Aucune différence phénotypique n'a été observée sur les parties aériennes entre les arbres exprimant la construction *Pro_{CAD}PtMYB14*, et les témoins transformés avec le vecteur pCAMBIA vide (Figure VI-9). De même, les observations microscopiques n'ont pas montré de différences. Cependant, l'analyse de la qualité des fibres par « Fiber Quality Analyses » (voir méthodes) a montré une réduction très significative de la largeur des fibres d'environ 4 micromètres dans les arbres de la lignée 30-9 au niveau de la formation du second cerne (Tableau VI-1). Toutefois, cette différence n'a pas été observée au niveau du troisième cerne de croissance. De légères variations ont été observées au niveau de la longueur des fibres de la lignée 30-7 mais elles sont plus difficiles à préciser. La détermination du taux de lignine Klason montre une réduction significative mais très faible de 0,58 % du résidu pariétal dans le bois des arbres de la lignée 30-9 (Tableau VI-2). Mais l'analyse de la structure des lignines par thioacidolyse n'indique aucune différence de composition en lignine entre les bois des lignées 30-9 et témoin (données non présentées).





Figure VI-9. Croissance et morphologie des épinettes blanches transformées avec la construction *Pro_{CAD}PtMYB14*, en troisième cycle de croissance.

7-13 et 7-4 sont des témoins transformés avec le vecteur pCAMBIA vide; 30-10,
30-14 et 30-7, 30-9, 30-12 expriment la construction *Pro_{CAD}PtMYB14*.

		Formation du	second cerne	Formation du troisième cerne				
Caractères	Lignées	Moyennes	LSD	Moyennes	LSD			
Longueur (Ln)	Témoin	0,456 +/- 0,098	b*	0,571 +/- 0,01	ns			
(en millimètre)	30-7	0,494 +/- 0,014	a*	0,581 +/- 0,012	ns			
	30-9	0,45 +/- 0,011	b*	0,572 +/- 0,01	ns			
Longueur (Lw)	Témoin	0,602 +/- 0,009	a, b**	0,67 +/- 0,011	ns			
(en millimètre)	30-7	0,625 +/- 0,002	a**	0,681 +/- 0,011	ns			
	30-9	0,570 +/- 0,01	b**	0,674 +/- 0,011	ns			
Diamètre	Témoin	22,38 +/- 0,34	a***	20,39 +/- 0,39	ns			
(en micromètre)	30-7	22,35 +/- O,49	a***	21,08 +/- 0,47	ns			
	30-9	18,4 +/- 0,47	b***	19,72 +/- 0,43	ns			

Tableau VI-1. Analyse de la longueur et de la largeur des fibres chez les arbres transgéniques « Pro_{CAD}PtMYB14 ».

Des échantillons de fibres de douze arbres par lignée, en troisième cycle de croissance, ont été mesurés par FQA après traitement acide à la solution de Franklin. Les fibres proviennent de sections, sans écorce, de un centimètre de tige principale juste en dessous de la pousse terminale de l'année (formation du second cerne) et à la base de l'arbre (formation du troisième cerne). Les lignées transgéniques et témoin significativement différentes entre elles par analyse de variance sont indiquées, * : P < 0.05, **: P < 0.01, ***: P < 0.001.

Lignées	Essai 1	Essai 2	Moyennes
Témoin	29,27	29,32	29,35 +/- 0,28
(en % du RP)	29,45	29,75	
	28,89	29,4	
30-9	28,75	28,56	28,77 +/- 0,17
(en % du RP)	28,87	29,01	
	28,8	28,6	

Tableau VI-2. Contenu en lignine du bois des arbres « Pro_{CAD}PtMYB14 », déterminé par la méthode Klason.

Les résultats sont exprimés en % pondéral du résidu pariétal (en % du RP). L'écart-type prend en compte les six valeurs de mesures par lignée. Les moyennes présentées (Témoin non transformé et lignée 30-9) sont significativement différentes avec un test de Student pour une P-value inférieure à 0,001.

VI-5-6-Analyse du profil transcriptomique et identification des gènes dérégulés chez les arbres « Pro_{CAD}PtMYB14 »

L'analyse du transcriptome des épinettes blanches « Pro_{CAD}PtMYB14 » a été effectuée à l'aide de microarrays permettant d'étudier simultanément l'expression de 9000 gènes différents (voir chapitres III et V). Les organes et tissus comparés étaient : 1) les pousses terminales sans les aiguilles (PT) des lignées 30-9 et 7-4 (témoin), 2) les tiges principales des arbres sans l'écorce (BE), c'est-à-dire le bois des lignées 30-14 et 7-13 (témoin), et 3) le xylème en différenciation du bois de compression (BC) des arbres 30-14 et 7-13 (témoin) (voir chapitre II). Les résultats obtenus avec ces trois tissus sont plutôt différents que semblables. La distribution des seuils de signification de ces gènes en fonction des ratios d'expression différentiels indique que le BE ne donne presque aucun gène au seuil p<0,01 (Figure VI-10 A). Son profil de distribution est donc atypique comparativement à BC et PT. Les échantillons de BE semblent présenter une plus grande variation entre les répétitions biologiques. De ce fait, le seuil à été ajusté à p<0,02 dans le diagramme de Venn pour BE seulement (Figure VI-10 B). Un plus grand nombre de gènes dérégulés ont été détectés dans les pousses terminales (Figure VI-10 B), soit 124 gènes dérégulés. Parmi ceux-ci 24 gènes ont un ratio d'expression de +/-1,5. La liste complète des gènes dérégulés dans les expériences PT (p-value < 0.01), BC (p-value < 0.01) et BE (p-value < 0.02) est présentée dans l'annexe VI-1. Plusieurs de ces gènes sont communs entre les trois conditions expérimentales et les expériences réalisées sur les plantules « Pro_{UBI} » et « Pro_{CAD}PtMYB14 » du chapitre V (Tableau VI-3). On observe qu'un transcrit dont la séquence fortement similaire à la nuclease bifonctionnelle S1P1 est surexprimée dans les trois tissus des arbres « Pro_{CAD}PtMYB14 » alors qu'il était sous-exprimé dans les plantules avec le promoteur UBI (Chapitre V). Les autres gènes communs entre les expériences suivent en général le même profil d'expression, sauf pour quelques exceptions répertoriées dans le tableau VI-3.



Figure VI-10. Analyse par microarrays des épinettes « $Pro_{CAD}PtMYB14$ » en troisième cycle de croissance. Distribution des gènes dérégulés dans les trois expériences microarrays représentant le seuil de significativité (*P*-value) en fonction du ratio d'expression des transcrits (entre transgénique et témoin). B) Nombre de gènes dérégulés ([+] et [-]: gènes dont l'expression est respectivement augmentée et diminuée) qui ont une *P*-value < 0.01 pour les différents tissus analysés. Le nombre de gènes avec des ratios supérieurs à +1,5 ou inférieurs à -1,5 sont entre parenthèses. Les gènes issus de la PT et BC qui sont communs avec ceux de BE ont une *P*-value < 0.02. *: *nuclease bifonctionnelle S1/P1*. Les analyses microarrays ont comparé cinq arbres « $Pro_{CAD}PtMYB14$ » et témoin pour les PT, et quatre arbres pour les BE et BC.

				Pousses	s terminales	Bois	s écorcé	Bois de	compression	Plantu	ules R*
Spruce ID	BlastX Hit Versus Arabidopsis	AGI code	E-value	R	P-value	R	P-value	R	P-value	ProUBI	ProCAD
Processus liés au	ux acides nucléiques										
MN5233918	Helicase	AT4G00660	0,00E+00	1,77	1,51E-04						
MN5159570(a)	Bifunctional nuclease (S1P1)	AT1G68290	3,00E-77	1,59	1,45E-03	2,73	6,45E-03	1,79	2,35E-03	-2,32	
MN5171247(a)	Endoribonuclease (RIBO)	AT2G02990	5,00E-60			2,44	1,28E-02	1,90	4,74E-04		
MN5159358	Bifunctional nuclease	AT1G11190	2,00E-61					1,94	2,35E-03		
Métabolisme géi	néral										
MN5254928	Adenylosuccinate lyase	AT4G18440	6,00E-117	1,79	1,14E-06						1,63
MN5170468(a)	O-methyltransferase (CAMT)	AT4G35160	2,00E-18	1,54	4,35E-03			1,36	1,86E-03		
MN5207692	Acyl-CoA oxidase	AT5G65110	3,00E-61	1,52	4,39E-04						1,39
MN5255831	Oxidoreductase	AT4G21580	7,00E-105	1,45	1,54E-03						1,65
MN5234789	NAD-dependent malic enzyme	AT2G13560	9,00E-144	1,32	7,73E-03	1,60	1,26E-02				1,49
MN5159506	Polygalacturonase	AT3G62110	2,00E-145	1,25	4,34E-03						1,43
MN5196728	O-glycosyl hydrolase/transferase	AT1G55740	2,00E-138	1,22	8,66E-03					1,65	1,31
MN5239920	Asparaginase	AT3G16150	1,00E-125	-1,27	2,29E-03					1,78	
MN5196298	Phosphatidylinositol phosphatase	AT3G02870	5,00E-107	-1,31	8,24E-03						1,32
MN5239343	Beta-1,3-glucanase	AT5G56590	1,00E-73	-1,51	8,49E-03					3,83	
MN5241484	Xyloglucan endo-transglycosylase	AT2G36870	5,00E-82					1,78	8,23E-04		
Gènes reliés aux	stress oxidatifs et à la défense										
MN5237642	Disease resistance protein	AT3G04220	4,00E-08	1,49	5,85E-05						1,34
MN5232943	Dehydration-responsive protein (rd22)	AT5G25610	2,00E-49	1,41	8,99E-03						-1,37
MN5182640(a)	strictosidine synthase (StrictS)	AT3G51420	4,00E-64	1,39	1,70E-04			1,66	4,49E-03		1,85
MN5241957(a)	Phosphogluconate dehydrogenase (6PGD)	AT3G02360	6,00E-100	1,28	1,51E-04	1,52	1,28E-02				1,36
MN5235051	Pepsin A (aspartyl protease)	AT3G18490	2,00E-24	1,27	5,72E-03					2,28	
MN5250687(a)	Strictosidine synthase (StrictS)	AT3G51420	4,00E-64	1,25	1,54E-03			1,41	5,17E-03		1,63
MN5160251	Serine/threonine protein kinase	AT5G15080	6,00E-63	1,24	9,51E-03						1,26
MN5159512	NADPH-cytochrome P450 reductase	AT4G24520	4,00E-131	1,21	9,13E-03					-1,47	
MN5237498	Acid phosphatase	AT2G38600	8,00E-62			1,38	1,24E-02				
MN5241707	Disease resistance protein	AT1G69545	4,00E-07	-1,68	8,49E-03	-1,43	1,28E-02				
MN5249216	Superoxide dismutase	AT2G28190	2,00E-69	<u>-1,70</u>	8,24E-03						
MN5175638	Superoxide dismutase	AT1G08830	8,00E-64	<u>-1,77</u>	1,51E-04						
MN5249435	Superoxide dismutase	AT1G08830	4,00E-64	-1,90	3,57E-05						
MN5236092	18.2 kDa class I heat shock protein	AT5G59720	3,00E-52	-2,15	4,14E-06						
MN5177619	Thioredoxin	AT3G16110	5,00E-60					-2,13	2,35E-03		

Tableau VI-3. Liste des gènes les plus fortement dérégulés et communs entre les pousses terminales, le bois de compression, le bois entier écorcé d'épinettes blanches « Pro_{CAD}PtMYB14 » de troisième cycle de croissance et les plantules « Pro_{UBI} » et « Pro_{CAD}PtMYB14 » (suite et légende page suivante).

				Pousses terminales		Bois écorcé		Bois de compression		Plantules R*	
Spruce ID	BlastX Hit Versus Arabidopsis	AGI code	Evalue	R	P-value	R	P-value	R	P-value	ProUBI	ProCAD
Autres											
MN5182319	Zinc finger protein	AT5G25560	1,00E-65	2,28	1,25E-08						2,36
MN5162862	Ubiquitin-protein ligase/zinc ion binding	AT5G25560	6,00E-33	1,25	8,49E-03						1,83
MN5241101	Metal ion binding	AT2G18196	3,00E-34	-1,35	4,61E-03					-1,64	
MN5164432	Plantacyanin	AT2G02850	2,00E-27	-1,68	9,40E-03						
MN5232405	Zinc finger and C2 domain protein-like	AT5G47710	4,00E-18	-1,68	2,15E-03						
MN5193939	(NAD(P)H oxidoreductase isoflavone reductase	AT4G39230	3,00E-105			1,88	1,89E-02				
MN5260497(a)	Oxidoreductase (BQR)	AT4G27270	5,00E-89			1,50	1,24E-02	1,40	1,54E-03		
Protéines incom	nus										
MN5244157	unknown protein	AT1G79975	7,00E-09	1,63	1,62E-06			1,57	1,54E-03		1,67
MN5196215	unknown protein	AT4G37900	7,00E-20	1,31	1,45E-03						1,65
MN5249136	unknown protein	AT1G61930	2,00E-16			-1,58	1,40E-02				
MN5182745	unknown protein	AT1G03140	1,00E-05			1,82	1,28E-02	1,69	8,99E-04		
Protéines non ic	lentifiés avec Arabidopsis										
MN5163000	No significant match	n.a.	n.a.	2,09	1,43E-04						
MN5234405	No significant match	n.a.	n.a.	1,51	3,71E-03						1,62
MN5232404	No significant match	n.a.	n.a.	1,40	4,16E-05					-1,54	
MN5174245	No significant match	n.a.	n.a.	1,22	8,38E-03	-1,29	1,28E-02				
MN5244270	No significant match	n.a.	n.a.	-1,40	6,83E-03	-1,58	1,24E-02				<u>-1,88</u>
MN5256234	No significant match	n.a.	n.a.	<u>-1,63</u>	5,87E-04					<u>-3,02</u>	
MN5242914	No significant match	n.a.	n.a.	<u>-1,74</u>	3,20E-03					<u>-1,87</u>	
MN5158489	No significant match	n.a.	n.a.	<u>-1,77</u>	7,28E-05						
MN5248924	No significant match	n.a.	n.a.			2,27	1,37E-02	1,85	8,23E-04		
MN5259803	No significant match	n.a.	n.a.			-1,90	1,33E-02				
MN5175212	No significant match	n.a.	n.a.					-1,69	2,45E-03		

Tableau VI-3. (Suite) R: Ratio ou niveau d'expression du gène entre transgénique et témoin. Les deux dernières colonnes indiquent les ratio des gènes dérégulés dans les plantules in vitro « Pro_{UBI} » et « Pro_{CAD}PtMYB14 » du chapitre III (p-value < 0,01) dont les gènes sont également dérégulés dans les trois expériences PT, BE et BC des arbres « Pro_{CAD}PtMYB14 » de troisième cycle de croissance. Les lignes en gras montrent les gènes dont l'expression est dérégulée dans plus de deux expériences chez les arbres en serre. Les cellules des valeurs positives supérieures à 1,5 sont en sombre, et les valeurs négatives inférieures à -1,5 sont en clair et soulignées. (a): gènes dont l'expression a été analysée en PCR quantitative avec les acronymes entre parenthèses repris dans les figures

•

Nous avons aussi analysé par RT-PCR quantitative les transcrits du transgène *PtMYB14*, des endogènes et orthologues présumés *PgMYB14a* et 14b, des gènes dérégulés dans les expériences microarrays (6PGD, StrictS, S1P1, BQR, RIBO et CAMT), un gène codant pour une enzyme clef de la voie de biosynthèse des phénylpropanoïdes (C4H), le gène DXS impliqué dans la voie mévalonate plastidique, un gène de défensine répondant aux stress (DEF) et le gène CAD (Figure VI-11 et VI-12). On remarque que le profil d'expression de *PtMYB14* ne suit pas le même profil que celui du gène CAD à travers les différents tissus testés et plus particulièrement dans le bois écorcé (Figure VI-11). Cela pourrait être attribué au fait que le promoteur CAD que nous avons isolé ne contient pas tous les éléments cis régulateurs nécessaires. Alternativement, ce résultat pourrait être du à la position d'insertion de la construction Pro_{CAD}PtMYB14 dans le génome de l'épinette. De plus, PtMYB14 est trois fois plus exprimé dans les pousses terminales de la lignée 30-7 comparativement aux autres lignées 30-9, 30-10 et 30-14. Toutefois, cette différence affecte relativement peu l'expression des gènes testés (Figure VI-12). Parmi les gènes dérégulés dans les analyses microarray, nous avons validé environ quatre gènes sur les six testés, les meilleures validations étant pour les gènes de la strictosidine synthase (StrictS), nucléase bifonctionnelle (S1P1) et la 6-phosphogluconate deshydrogénase (6PGD) (Figure VI-11 et VI-12).



Figure VI-11. Niveaux d'expression de plusieurs gènes dans différents tissus et organes d'épinettes blanches « Pro_{CAD}PtMYB14 » de troisième cycle de croissance. (Légende page suivante)

PR

РТ

Aig.

BE

PPh

PR

PPh

РТ

Aig.

BE
Figure VI-11 (Légende). Les ARN messagers ont été quantifiés par RT-PCR quantitative sur trois lignées transgéniques: T (témoin non transformé), 30-10 et 30-14 expriment la construction Pro_{CAD}PtMYB14, avec 4 individus par lignée. Aig.: aiguilles, PT: pousse terminale sans aiguilles, BE: bois écorcé, PPh: périderme et phloème en développement, PR: pointes de racines, (a): indique les gènes dérégulés dans les analyses par microarray. 6PGD: 6-phosphogluconate déhydrogénase, StrictS.: Sstrictosidine synthase, C4H: cinnamate 4-hydroxylase, S1P1: bifunctionnal nuclease S1/P1, DEF: défensine, BQR: benzoquinone réductase, CAMT: RIBO: ribonucléase, CAD: orthométhyltransférase, alcool cinnamylique deshydrogénase, DXS: 1-déoxy-D-xylulose 5-phosphate synthase, CDC2: cycline 2. Les étoiles indiquent le niveau de signification (P-value) du test de Student entre le témoin et le transgénique, * : P < 0.05, **: P < 0.01, ***: P < 0.001.

Figure VI-12. Niveaux d'expression de plusieurs gènes dans les pousses terminales d'épinettes blanches « Pro_{CAD}PtMYB14 » de troisième cycle de croissance. (Figure page suivante)

La quantification des ARN messagers a été déterminée par RT-PCR quantitative sur trois lignées transgéniques: T (témoin non transformé), 30-9 et 30-7 expriment la construction *Pro_{CAD}PtMYB14*, avec 4 individus par lignée. Voir légende Figure VI-11.



Figure VI-12. Niveaux d'expression de plusieurs gènes dans les pousses terminales d'épinettes blanches « Pro_{CAD}PtMYB14 » de troisième cycle de croissance. (Légende page précédente)

VI-5-7-Etude de l'expression de gènes liés au stress oxydatif et aux mécanismes de défense chez des plants non transformés blessés ou traités par l'acide jasmonique

Des analyses antérieures ont montré que le gène PtMYB14 était induit par la blessure et l'acide jasmonique, et que PtMYB14 pouvait être impliqué dans les mécanismes de défense chez les conifères (Chapitre V). Les analyses par microarrays montrent aussi que la construction *Pro_{CAD}PtMYB14* entraîne la dérégulation de gènes liés aux stress oxydatifs et aux mécanismes de défense chez les jeunes arbres (Tableau VI-3), et chez les plantules in vitro (Chapitre V). Ainsi nous avons cherché à savoir si PgMYB14a et 14b ainsi que les gènes strictosidine synthase, nucléase bifonctionnelle SIP1 et 6-phosphogluconate deshydrogénase étaient eux aussi induit dans le cas de blessures ou d'application d'acide jasmonique. L'augmentation des transcrits du gène de la défensine sert de témoin positif montrant que l'induction des mécanismes de défense a fonctionné dans chacun des traitements. Nous avons aussi observé que l'expression des gènes PgMYB14a et 14b augmente en réponse aux deux traitements, mais leur réponse est environ trois fois plus forte à l'acide jasmonique, et ce à la fois dans les pousses terminales et dans la tige entière (bois et l'écorce) (Figure VI-13). Les gènes strictosidine synthase, nuclease bifonctionnelle SIP1 et 6phosphogluconate deshydrogénase ont été précédemment validés pour l'augmentation de leurs niveaux d'expression dans les arbres « Pro_{CAD}PtMYB14 »; ils sont aussi induits par les traitements de JA et blessures sur les arbres non transgéniques (Figure VI-13). Les transcrits de la strictosidine synthase et de la 6phosphogluconate deshydrogénase augmentent dans toutes les conditions, alors que la nucléase bifonctionnelle SIP1 augmente uniquement sous l'action de l'acide jasmonique dans la tige entière (Figure VI-13).





Figure VI-13. Niveaux d'expression de plusieurs gènes dans les pousses terminales avec aiguilles et dans la tige principale (tronc) après 24 heures d'induction à l'acide jasmonique et à la blessure.

La légende est identique à celle de la figure VI-8. Les acronymes des gènes sont définis dans la figure VI-11.

VI-7-Conclusions et perspectives

La caractérisation fonctionnelle d'un gène, ici PtMYB14, avec un promoteur préférentiel des tissus vasculaire (Pro_{CAD}) chez un conifère est, à notre connaissance, la première réalisation du genre chez une espèce ligneuse. En effet l'alcool cinnamylique deshydrogénase (CAD) qui catalyse la synthèse des alcools cinnamyliques, précurseurs immédiats de la lignine, est considéré comme un marqueur de la lignification. Ce promoteur de la CAD est donc particulièrement adapté pour étudier la différenciation du xylème (primaire et secondaire) et la lignification. Ainsi, les analyses histochimiques GUS sur les sections transversales de pousses terminales et de troncs permettent d'associer le gène CAD de conifère à la formation du bois. De même, la quantification des transcrits du gène CAD dans différents tissus du tronc d'épinettes blanches confirme que le xylème secondaire en différenciation est le lieu privilégié de l'expression du gène CAD. Les analyses d'activité \beta-glucuronidase ou de PCR quantitative sur des échantillons comprenant plusieurs types cellulaires (e.g. sections de tronc avec bois et écorce, aiguilles entières) ne sont pas assez précises pour mettre en évidence la spécificité d'expression au sein du tissu vasculaire qui est en effet présent dans ces deux organes. Elles permettent toutefois de conclure que la séquence de 1.2 Kb située en amont du gène CAD constitue une séquence promotrice fonctionnelle dans la construction Pro_{CAD}GUS. Par ailleurs, nous avons montré par histochimie que le promoteur CAD est induit par la blessure, et, par PCR quantitative que le gène CAD était également induit par la blessure et plus fortement à l'acide jasmonique. Ces observations semblent indiquer un rôle du gène CAD dans une lignification de type défense semblable à celui mis en évidence avec le promoteur Eucalyptus gunnii EgCAD2 (Lauvergeat et al., 2002).

Les épinettes blanches exprimant la construction $Pro_{CAD}PtMYB14$ ont un phénotype pratiquement identique aux arbres témoin, bien que quelques différences mineures aient été observées en terme de chimie du bois et diamètre des fibres. Ces différences mineures ne sont pas corrélées avec les niveaux d'expression du transgène, puisque la lignée 30-7 exprime PtMYB14 trois fois plus que la lignée 30-9 qui présente des réductions du diamètre des fibres. Cependant, au niveau transcriptomique, les résultats obtenus présentent des similarités évidentes avec les analyses antérieures

réalisées sur les plantules, notamment pour les gènes liés aux stress oxydatifs et aux mécanismes de défense. De plus, trois des gènes dont l'expression est augmentée dans les arbres « $Pro_{CAD}PtMYB14$ » (*6PGD*, *StrictS* et *S1P1*) sont également induits dans une expérience d'induction par la blessure et l'acide jasmonique. Par conséquent, PtMYB14 pourrait avoir un rôle dans les mécanismes de défense dans le xylème en différenciation bien que les effets phénotypiques de ce rôle aient été difficilement observés jusqu'à maintenant (e.g. cellules polyphénoliques, canaux résinifères traumatiques). Par conséquent, les analyses effectuées sur les arbres transgéniques « $Pro_{CAD}PtMYB14$ » contribuent à approfondir les connaissances sur le gène *MYB14* qui appartient à un sousgroupe (Sg-4C) fort intrigant des MYB de la classe R2R3. Par ailleurs, nos observations permettent de proposer des pistes d'analyses complémentaires. Par exemple :

- Analyser l'accumulation de la strictosidine, un précurseur des alcaloïdes (Dutta et al., 2005), car le niveau d'expression de la strictosidine synthase est supérieur dans les aiguilles des transgéniques comparativement aux témoins
- Mettre en évidence l'accumulation de substances en relation directe avec le stress et les mécanismes de défense tels que des molécules "signal" (e.g. acide salicylique) ou différents composés chimiques (e.g. terpènes, acide jasmonique)
- Finalement les niveaux d'expression des gènes dérégulés reste faibles avec peu de gènes ayant un ratio supérieur à 1,8 (P <0,01). L'utilisation d'un promoteur préférentiel à un tissu comme le xylème est prometteur. Toutefois augmenter la force de ce promoteur pourrait mener à des phénotypes plus nets.

VI-8-Références

- Bedon F., Grima-Pettenati J. et Mackay J. (2007). Conifer R2R3-MYB transcription factors: sequence analysis and gene expression in wood-forming tissues of white spruce (*Picea glauca*). *BMC Plant Biol* 7: 17
- Chang S., Puryear J. et Cairney J. (1993). A Simple and efficient method for isolating RNA from pine trees. *Plant Mol Biol Rep* 11: 113-116
- Côté C. et Rutledge R.G. (2003). An improved MUG fluorescent assay for the determination of GUS activity within transgenic tissue of woody plants. *Plant Cell Rep* 21: 619-214
- **Dence C. (1992).** Lignin determination. In C Dence, S Lin, eds, Methods in Lignin Chemistry 33-61

- Dutta A., Batra J., Pandey-Rai S., Singh D., Kumar S. et Sen J. (2005). Expression of terpenoid indole alkaloid biosynthetic pathway genes corresponds to accumulation of related alkaloids in Catharanthus roseus (L.) G. Don. *Planta* 220: 376-383
- Franklin G.L. (1945). Preparation of thin sections of synthetic resins and wood-resin composites and a new macerating method for wood. *Nature* 155: 51
- Hawkins S., Samaj J., Lauvergeat V., Boudet A. et Grima-Pettenati J. (1997). Cinnamyl alcohol dehydrogenase: Identification of new sites of promoter activity in transgenic poplar. *Plant Physiol* 113: 321-325
- Hemerly A.S., Ferreira P., de Almeida Engler J., Van Montagu M., Engler G., Inzé D. (1993). cdc2a expression in *Arabidopsis* is linked with competence for cell division. *Plant Cell* 5: 1711–1723.
- Higo K., Ugawa Y., Iwamoto M. et Korenaga T. (1999). Plant *cis*-acting regulatory DNA elements (PLACE) database. *Nucleic Acids Res* 27: 297-300.
- Joshi C. P. (1987). An inspection of the domain between putative TATA box and translation start site in 79 plants genes. *Nucleic Acids Res* 15: 6643-6653
- Klimaszewska K., Lachance D., Pelletier G., Lelu M.A. et Séguin A. (2001). Regeneration of transgenic *Picea glauca*, *P. mariana*, and *P. abies* after cocultivation of embryogenic tissue with *Agrobacterium tumefaciens*. *In Vitro Cell Dev Biol-Plant* 37: 748–755
- Klimaszewska K., Rutledge R.G. et Séguin A. (2004). Genetic transformation of conifers utilizing somatic embryogenesis. In L Peña, ed, Methods *in Molecular Biology*, Humana Press, Totowa, pp 151–164
- Pitre F. E., Cooke E.K. et MacKay J.J. (2007). Short-term effects of nitrogen availability on wood formation and fibre properties in hybrid poplar. *Trees* 21: 249-259
- Rozen S. et Skaletsky H. (2000). Primer3 on the WWW for general users and for biologist programmers. *In: S Krawetz and S Misener (Eds), Bioinformatics Methods and Protocols: Methods in Molecular Biology Humana Press, Totowa, NJ*:365-386

CHAPITRE VII

Discussion générale et perspectives

Les travaux présentés dans cette thèse s'inscrivent dans le cadre plus large de recherches (*i*) sur la compréhension des mécanismes moléculaires intervenant dans la formation du bois chez les conifères, et (*ii*) sur l'identification des déterminants génétiques majeurs impliqués dans ce processus. L'objectif général de ces recherches est d'identifier *via* une approche intégrée de génomique fonctionnelle et de génétique d'association des gènes marqueurs de la productivité et de la qualité du bois chez l'épinette blanche (*Picea glauca*), choisie comme système d'étude et d'application.

Mon projet de thèse s'insère dans l'approche de génomique fonctionnelle; il a d'abord consisté à identifier et à caractériser, au niveau moléculaire, plusieurs membres de la famille des facteurs de transcription *MYB-R2R3* chez les conifères *Picea glauca* et *Pinus taeda*. La plus grande partie de ma thèse concerne la caractérisation fonctionnelle de trois de ces facteurs de transciption, potentiellement impliqués dans la formation du bois. Ce projet de thèse a été effectué dans le cadre d'une co-tutelle entre les universités Paul-Sabatier, à Toulouse (France), et Laval, à Québec (Canada).

VII-1-Évolution et dénombrement des facteurs de transcription *MYB-R2R3* de conifères

L'identification et l'analyse des séquences de MYB-R2R3 chez l'épinette et le pin m'ont conduit à mieux caractériser une famille de facteurs de transcription très peu étudiée jusqu'à présent chez les conifères.

VII-1-1-Les séquences codantes

Diverses séquences nucléotidiques codantes des gènes MYB-R2R3 de conifères ont été isolées et caractérisées : 20 chez Picea glauca et 11 chez Pinus

taeda. La topologie des arbres phylogénétiques construits avec les séquences protéiques prédites de ces MYB de conifères indique la formation de groupes et de sous-groupes similaires à ceux observés chez des angiospermes dicotylédones comme *Arabidopsis* et monocotylédones comme le riz (Krantz et *al.*, 1998; Romero et *al.*, 1998; Jiang et *al.*, 2004a). Cependant des différences significatives apparaissent :

- Jusqu'à présent, aucune séquence appartenant au groupe B n'a été identifiée chez les conifères. Ce type de protéine MYB se fixe sur les éléments *cis* régulateurs de type I et II. Les groupes A, B et C sont les trois groupes majeurs identifiés chez les angiospermes (Romero et *al.*, 1998). L'absence de ce type de séquences chez les gymnospermes pourrait être le résultat d'une évolution génique distincte, suite à la séparation des gymnospermes et des angiospermes par exemple. Bien que l'on puisse supposer que ce type de séquences soit présent seulement chez les angiospermes, nos résultats basés sur le séquençage d'ESTs ne seront complètement validés que lorsqu'un génome entier sera séquencé. Entre temps, il serait néanmoins pertinent de tester cette hypothèse en poursuivant des recherches ciblées chez les conifères, soit dans les bases d'ESTs, soit par PCR directement sur de l'ADN génomique avec des amorces appropriées.
- Un sous-groupe de séquences MYB-R2R3 spécifique des conifères a été identifié chez *Picea* (10 séquences) et *Pinus* (7 séquences), et appelé Sg4-C. L'analyse des séquences et la construction d'arbres phylogénétiques indiquent des différences de motifs en acides aminés entre les angiospermes et les gymnospermes, ce qui suggère des différences fonctionnelles entre deux branches majeures du règne végétal. Le nombre élevé de séquences appartenant à ce sous-groupe est inattendu, il montre que la stratégie que j'ai mise en œuvre pour l'amplification spécifique de ses membres s'est avérée efficace. Le caractère particulier de ce sous-groupe sera discuté plus bas.

VII-1-2-Les séquences non codantes

Les informations apportées par la séquence des régions génomiques des *MYB*-*R2R3* sont importantes car selon Jiang et *al*. (2004a), la structure exon-intron est conservée entre les membres d'un même sous-groupe mais varie entre sous-groupes, chez Arabidopsis et le riz. De ce fait, sur le plan évolutif, il apparaît qu'il y a congruence entre les séquences codantes (e.g. motifs protéiques) et non codantes (e.g. structure exon-intron) pour la fonction des gènes MYB. Cette hypothèse mérite d'être approfondie, surtout en considérant que les introns sont impliqués dans l'épissage, l'échange de domaine protéique ou encore la régulation de l'expression de leur propre gène. Chez l'épinette blanche, les gènes des MYB-R2R3 présentent des structures exon-intron généralement semblables (nombre, position et phase des introns) à celles décrites chez des angiospermes. Par exemple, il y a soit un, soit deux introns dans la région correspondant au domaine de fixation à l'ADN chez l'épinette (chapitre II) et le pin (Annexe VII-1: PtMYB1, 4, 8 et 14), exactement comme chez les angiospermes. Cependant, j'ai mis en évidence un intron dans les régions carboxyterminales de PgMYB3 et PtMYB3 (Annexe VII-1 et VII-4), alors qu'à ma connaissance, la présence d'un intron dans cette région des MYB n'a jamais été rapportée, du moins chez Arabidopsis (Jiang et al., 2004a). Il serait donc pertinent de vérifier si l'ensemble des membres du sous-groupe 21 possède également un intron dans cette région, de façon à préciser une éventuelle spécialisation de cet événement chez les conifères (Fedorova et Fedorov, 2003). Les comparaisons de séquences introniques entre l'épinette et des angiospermes comme Arabidopsis ne sont pas pertinentes alors que c'est tout à fait possible entre les genres Picea et Pinus. Ces comparaisons ont révélé que généralement, entre les paires Pg/PtMYB1, 4 et 8, le second intron est plus conservé que le premier en terme de longueur et de pourcentage d'identité (Annexe VII-1). A l'inverse, les introns des membres du sousgroupe Sg4-C ont des longueurs variables, surtout pour les paires *PtMYB14/PgMYB14a* et 14b. Le plus intriguant dans l'étude des séquences introniques des gènes présumés homologues des MYB-R2R3 de Picea glauca et de Pinus taeda réside dans l'identification de séquences répétées présentes au sein d'un genre ou encore présentes dans les deux genres (Annexe VII-2 et VII-3). Il est intéressant de noter que Jiang et al. (2004a) ont également observé ce genre de répétitions dans les introns de certains MYB-R2R3 de maïs; mais aucune hypothèse quant au rôle de ces répétitions n'a encore été émise.

VII-1-3-Le cas du sous-groupe 4 de conifères (Sg4-C)

Les séquences MYB-R2R3 de conifère du sous-groupe 4 (Sg4-C) sont nombreuses et présentent des homologies de séquences élevées entre elles, suggérant que des duplications géniques récentes seraient à l'origine de l'expansion de ce sousgroupe chez les conifères, comme c'est le cas chez certains angiospermes (Braun et Grotewold, 1999; Krantz et al., 2000; Dias et al., 2003; Jiang et al., 2004b). De plus, il y a de fortes similarités de séquences dans les introns des gènes PgMYB5, 10, 13, 14a et 14b, avec plus particulièrement la conservation de blocks nucléotidiques qui suggèrent la possibilité d'une séquence ancestrale commune (Aannexe VII-2 et VII-3). Cependant, d'autres indices sont nécessaires pour mettre en évidence des duplications géniques récentes, comme par exemple, la présence d'un excès de substitutions non synonymes accompagnées de nouvelles fonctions géniques dans le cas de sélection positive liée à l'adaptation (Jia et al., 2004). En effet, les duplications géniques suivies de divergences fonctionnelles des gènes dupliqués sont les mécanismes majeurs de l'évolution pour l'acquisition de nouvelles fonctions (Nei, 1969). De plus, la détection d'un motif protéique spécifique aux membres du sousgroupe Sg4-C rejoint l'hypothèse de la sélection d'une nouvelle fonction suite à la duplication et à la néo-fonctionnalisation d'un gène ancestral commun aux membres du sous-groupe 4. Les études de caractérisation fonctionnelle et de transcriptomique présentées dans cette thèse constituent le point de départ d'analyses qui pourraient expliquer la présence et l'utilité des membres du sous-groupe Sg4-C absent des angiospermes, et appuyer encore davantage le rôle de PtMYB14, et possiblement les autres membres du Sg4-C, dans les mécanismes de défense biotiques et abiotiques chez les conifères.

VII-1-4-Représentation de l'ensemble des MYB chez l'épinette blanche *Picea glauca*

Les gènes *MYB-R2R3* que nous avons isolés proviennent de recherches effectuées au début de mon projet de thèse (2003) dans les bases de donnés de séquences contenant des étiquettes de séquences exprimées (ou ESTs pour Expressed Sequences Tags) d'épinette blanche ainsi que dans les bases de données de séquences

de Pinus taeda (http://web.ahc.umn.edu/biodata/nsfpine/). L'augmentation progressive du nombre d'ESTs d'épinette blanche disponibles dans les bases de données montre clairement que les séquences de gènes MYB identifiés lors de ma thèse ne représentent qu'une partie de l'ensemble des gènes MYB. En 2005, Pavy et al. rapportaient l'identification de 13 facteurs de transcription présentant un domaine de fixation à l'ADN de type MYB sur les 16 578 séquences consensus (ou contigs), représentant environ 10 000 gènes, obtenues à partir d'ESTs chez l'épinette blanche. Depuis, des analyses préliminaires informatiques au sein du projet ARBOREA II estiment que le nombre de gènes représentés a presque doublé (soit environ 18 000), et que le nombre de séquences différentes possédant un domaine MYB est d'environ 85 (Rigault et MacKay, communications personnelles). Cet inventaire suggère par ailleurs que le séquençage génomique apporte plus de séquences MYB que le séquençage d'ADN complémentaire (i.e. détection de gènes peu ou pas exprimés), étant donné que l'on retrouve 189 membres de la famille MYB chez la plante dicotylédone Arabidopsis thaliana et 182 membres chez la monocotylédone Oryza sativa (Schmid et al., 2005; Qu et Zhu, 2006). Chez le peuplier nouvellement séquencé (Tuskan et al., 2006), bien qu'aucune publication ne fasse l'inventaire des gènes MYB, il est possible d'identifier 221 séquences (E-value supérieure à 1.0e-05) avec un domaine MYB présumé (BLASTP avec la séquence du DBD de PgMYB5 sur le site http://www.jgi.doe.gov/) sur un total de 45 550 gènes supposés (Tuskan et al., 2006).

En l'absence de génome de conifère séquencé, les ESTs deviennent le seul moyen d'obtenir un nombre approximatif de gènes totaux ou bien un nombre de séquences de gènes *MYB*. Cette approche peut toutefois mener à des sous-estimations ou des surestimations importantes. Par exemple, la redondance de séquences entre les transcrits résultant d'épissage alternatif ou provenant du regroupement de plusieurs espèces, peut mener à une surestimation. Le nombre de séquences consensus présentant un domaine MYB dans « TIGR Gene Indices » (Quackenbush et *al.*, 2001) est de 135 pour le Pin (45 557 séquences uniques du genre *Pinus* au 19 juillet 2005), 107 pour l'épinette (59 303 séquences uniques du genre *Picea* au 21 juin 2006), 179

pour le peuplier (86 324 séquences uniques du genre *Populus* au 19 juin 2006) et 272 pour *Arabidopsis* (81 826 séquences uniques du genre *Arabidopsis* au 16 juin 2006). Ces données suggèrent clairement une surestimation pour *Arabidopsis*. Cependant, dans un grand nombre de cas, la faible diversité des banques d'ADN complémentaire séquencées alliée à une nombre de séquences disponibles limitées entraînent le plus souvent une sous-estimation de la taille des familles mutigèniques. La construction de banques d'ADNc avec des tissus et des organes de plusieurs stades de croissance (plantules, juvénile, mature), et conditions physiologiques différentes de l'arbre permettrait d'isoler, sans doute, davantage de séquences de facteurs de transcription chez une espèce dont le séquençage du génome n'est pas encore envisagé. De plus, les nouvelles méthodes de séquençage à très haut débit permettront vraisemblablement de séquencer chaque banque plus en profondeur.

VII-2-Démarches expérimentales pour la caractérisation des MYB de conifères

Les études des niveaux d'expression d'ARN messagers des facteurs de transcription *MYB-R2R3* ont porté sur différentes phases de la croissance et du développement ainsi que sur les réponses aux stress abiotiques. L'analyse des principaux tissus et organes de l'épinette blanche s'est faite à une échelle macroscopique et l'échantillonnage a mis l'emphase sur la formation du bois (i.e. la récolte d'un tissu unique : le xylème). L'analyse du xylème secondaire en différenciation a permis d'y associer une expression préférentielle de certains gènes *MYB-R2R3 (PgMYB2*, 4 et 8, et, dans une moindre mesure *PgMYB1* et 3).

J'ai étudié l'effet de stress abiotiques sur l'expression des gènes MYB. Par exemple, lors de la formation du bois de compression, j'ai mis en évidence l'induction des gènes PgMYB2, 4 et 8, et, dans une moindre mesure PgMYB9, 11 et 13. Des blessures mécaniques ou l'application d'acide jasmonique ont aussi induit plusieurs des Pg/PtMYB de Sg4-C y compris PtMYB14. Ces expériences ont mis en évidence l'inductibilité de certains facteurs de transcription MYB-R2R3 de conifère que ce soit à l'échelle d'organes, de plantules ou dans le xylème en différenciation

Deux types de modulation de l'expression des transgènes *PtMYB* (issus de *Pinus taeda*) surexprimés chez des épinettes transgéniques ont été utilisés pour explorer leurs rôles biologiques dans l'épinette blanche.

- La surexpression constitutive réalisée avec le promoteur du gène Ubiquitine de maïs (Christensen et al., 1992; Christensen et Quail, 1996). Ce promoteur a permis d'obtenir un niveau d'expression 10 fois supérieur au promoteur du virus de la mosaïque du chou-fleur CAMV-35S dans des protoplastes de maïs mais inférieur de 10 fois dans des protoplastes de tabac. Chez les plantules d'épinettes transgéniques, l'utilisation de ce promoteur a permis d'atteindre des niveaux d'expression élevés qui sont proches de ceux du facteur d'élongation *EF1-a* (un des transcrits les plus abondants chez l'épinette). Cependant, chez les arbres transgéniques de deuxième ou troisième cycle de croissance en serre, le niveau d'expression du transgène s'est révélé beaucoup plus faible (e.g. *PtMYB1*, chapitre III et chapitre IV). De plus amples études seront nécessaires pour comprendre ce phénomène mais il est possible que le niveau d'expression du promoteur *UBI* de maïs soit plus élevé dans des plantules en germination que dans des plantes plus âgées transférées en serre.
- Une démarche complémentaire de surexpression a ciblé le xylème en différenciation avec le promoteur du gène codant la cinnamyl alcool deshydrogénase (*Pro_{CAD}*) d'épinette blanche, impliqué dans la dernière étape enzymatique de biosynthèse des monolignols. L'utilisation d'un tel promoteur visait à contourner deux problèmes liés aux caractéristiques du promoteur *UBI* : un très fort niveau d'expression, du moins au stade plantule, et une expression dans tous les tissus végétaux. En effet, il s'avère que, selon le transgène *MYB* utilisé, une forte expression constitutive peut avoir des effets létaux pour les transgéniques, comme c'est le cas pour les plantules « Pro_{UBI}PtMYB14 ». Par contre, le promoteur du gène *CAD* conduit préférentiellement l'expression du transgène dans les tissus vasculaires (cellules du xylème primaire et secondaire, et du parenchyme de rayon) et avec un niveau d'expression physiologique vraisemblablement moins néfaste pour le développement de la plante. Ainsi, de par ses caractéristiques, l'utilisation du promoteur *CAD* semble adaptée dans le

cas de l'expression d'un transgène dont l'endogène n'est pas ou faiblement exprimé dans les mêmes tissus que le gène *CAD*, pour en étudier le rôle biologique dans les cellules du xylème par exemple.

VII-3-Rôles biologiques des facteurs de transcription MYB-R2R3 de conifère

La synthèse des analyses phénotypiques et transcriptomiques réalisées sur les transgéniques MYB aux stades plantule et « jeune arbre » nous ont permis de proposer des rôles biologiques majeurs pour certains des MYB étudiés, notamment dans la formation de la paroi secondaire lignifiée et dans les mécanismes de défense chez les conifères.

VII-3-1-Coordination transcriptionnelle de l'expression des gènes pour la formation de la paroi secondaire.

La formation de la paroi secondaire lignifiée intervient à la fin de l'expansion cellulaire et se caractérise par le dépôt de différents constituants pariétaux à l'intérieur de la paroi primaire : cellulose, hémicellulose et lignine. De ce fait, entre la fin de l'expansion cellulaire et le début de la formation de la paroi secondaire, une transition importante s'opère dans l'expression des gènes liés à la biosynthèse des composés pariétaux; transition qui doit être régulée très finement au niveau spatial et temporel notamment au niveau transcriptionnel.

La caractérisation fonctionnelle de facteurs de transcription couplée à des analyses transcriptomiques et phénotypiques s'est avérée un outil intéressant pour mettre en évidence la coordination de l'expression des gènes à l'origine de la formation de la paroi secondaire (Demura et Fukuda, 2007).

Dans ce travail, nous avons montré que certains *MYB* étaient préférentiellement (*PgMYB1* et *PgMYB3*) voire spécifiquement (*PgMYB2, PgMYB4* et *PgMYB8*) exprimés dans le xylème secondaire en différenciation de *Picea glauca* (Chapitre II). Ces résultats d'expression tissulaire suggèrent que *PgMYB1, 2, 3, 4 et 8* sont potentiellement impliqués dans la formation de la paroi secondaire des cellules du xylème. Il est intéressant de remarquer que *PgMYB1, 2 et 3* sont les orthologues potentiels respectivement de *AtMYB20, AtMYB103 et AtMYB52* (Chapitre II). Ces gènes d'*Arabidopsis* font partie des 52 spécifiquement exprimés dans le xylème secondaire selon Ko et *al.* (2006). De plus, *AtMYB20* et *AtMYB103* font partie des gènes cibles régulés par les *NAC, SND1 et NST1,* pour la formation de la paroi secondaire chez *Arabidopsis* (Zhong et *al.,* 2007). Chez *Arabidopsis,* la spécification des cellules du xylème semble être contrôlée par l'expression de gènes de type *NAC* (Kubo et *al.,* 2005; *NST1* et *NST2,* Mitsuda et *al.,* 2005; *SND1,* Zhong et *al.,* 2006; NST1 et NST3, Mitsuda et *al.,* 2007; revue par Demura et Fukuda, 2007). La surexpression de *SND1* et *NST1* chez *Arabidopsis* conduit à l'activation de la biosynthèse de la paroi secondaire (Zhong et *al.,* 2006). A l'inverse, la sous-expression conjointe par interférence d'ARN de ces deux gènes réduit très fortement la formation de la paroi secondaire de même que l'expression des gènes de biosynthèse des phénylpropanoïdes et de la cellulose (Zhong et *al.,* 2007).

Les analyses d'expression lors de la formation du bois de compression mettent en évidence deux groupes de MYB. Le premier comporte PgMYB1 et 3 dont les niveaux de transcrits ne sont pas affectés pendant la cinétique de traitement. PgMYB2, 4 et 8, qui sont spécifiques du xylème secondaire et induits dans le bois de compression appartiennent au deuxième groupe. Cette différence de comportement entre les deux groupes suggèrent des rôles différents dans la formation du xylème. Les études de plantes transgéniques ont permis de mieux appréhender ces rôles.

Ainsi, nous avons montré que la sur-expression constitutive de *PtMYB8* chez l'épinette conduit à une forte lignification ectopique. *PtMYB4* lui aussi membre du deuxième groupe provoque une lignification ectopique (Patzlaff *et al.*, 2003b). Le rôle de *PtMYB1*, membre du premier groupe, dans la formation de la paroi secondaire et la lignification reste à préciser. En effet, il est aussi capable d'activer la transcription du gène de la glutamine synthase (Gomez-Maldonaldo et *al.*, 2004) et sa surexpression augmente les niveaux d'expression de glutamine synthases dans des pousses terminales d'épinettes transgéniques (Chapitre IV). Les produits du gènes *PgMYB1* pourraient agir au niveau de la biosynthèse de la paroi secondaire par des mécanismes indirects, par exemple, *via* le contrôle des flux de carbone ou d'azote pour le bon déroulement de la xylogénèse.

Chez les plantules transgéniques, nous avons mis en évidence que la surexpression des gènes *PtMYB1* et *PtMYB8* pouvait mener à la dérégulation de l'expression de certains autres gènes *MYB*. Ainsi, il y a une légère sous-expression de *PgMYB1* et 2, et une sous-expression plus nette pour *PgMYB4* dans les lignées « Pro_{UBI}PtMYB1 ». Dans les lignées « Pro_{UBI}PtMYB8 », l'expression des gènes *PgMYB1, 2* et 3 augmente, et celle de *PgMYB4* et 8 diminue (Chapitre III). Ces résultats mènent donc à l'hypothèse d'un réseau d'interaction entre les facteurs de transcription MYB-R2R3 de conifère pour la gouvernance de la formation de la paroi secondaire des fibres du xylème qui rappelle celui décrit chez *Arabidopsis* avec les gènes *NAC*. Cette hypothèse reste à tester par l'expérimentation. Afin de vérifier s'il y a une hiérarchie potentielle entre différents types de facteurs de transcription pour la formation de la paroi secondaire comme chez *Arabidopsis*, il reste à analyser les profils d'expression (dans nos lignées transgéniques) et à réaliser la caractérisation fonctionnelle de gènes NAC et d'autres types de facteurs de transcription de conifère.

VII-3-2-Complexité et interdépendance des mécanismes de défense et de réponse aux stress chez les conifères : le cas de PtMYB14

Dans cette thèse, nous nous sommes aussi intéressé à *PtMYB14* qui appartient au sous-groupe 4 dont le membre le mieux caractérisé est *AtMYB4*, un répresseur de la voie des phénylpropanoïdes chez *Arabidopsis* (Jin et *al.*, 2000). Nous avons montré que ce sous-groupe possède une caractéristique spécifique des conifères avec un ensemble constitué de nombreux membres et un motif protéique propre, très proches phylogénétiquement mais présentant des profils d'expression très variés. Nous avons montré que la sur-expression de *PtMYB14*, tant constitutive que préférentielle des tissus vasculaires, a un impact relativement mineur sur la formation de la paroi autant chez les plantules que chez les jeunes arbres (Chapitre III et VI). Les analyses du transcriptome des plantules transgéniques ont montré que les gènes affectés étaient fort différents de ceux qui avaient été identifiés chez les transgéniques « Pro_{UBI}PtMYB1 » et « Pro_{UBI}PtMYB8 ». En effet, les niveaux de transcrits des gènes associés aux mécanismes de réponse aux stress et à la défense sont augmentés ainsi que les gènes de synthèse des flavonoïdes, alors que ceux de la voie de biosynthèse des phénylpropanoïdes sont diminués (Chapitre V). Les résultats de la surexpression de PtMYB14 semblent donc inversés par rapport à PtMYB1 et 8.

Une caractéristique commune à la surexpression de PtMYB14 avec les promoteurs des gènes UBI et CAD est l'augmentation de la synthèse des sesquiterpènes (α -farnesene et β -farnesene), en lien avec les niveaux d'expression du transgène. Nos résultats montrent aussi l'augmentation de la synthèse des monoterpènes (α -pinene et camphene), plus particulièrement avec la construction Pro_{UBI}PtMYB14. L'accumulation des deux types de terpènes reflète les niveaux d'expression différents des gènes correspondants observés en microarray et en PCR quantitative (Chapitre V). Cependant, aucune accumulation de diterpènes n'a été observée dans les plantules « Pro_{UBI}PtMYB14 », ce qui pourrait s'expliquer par la répression des transcrits du gène codant la GGPPS, dont le produit est directement impliqué dans la synthèse des diterpènes. Les mono- et sesquiterpènes sont impliqués dans les mécanismes de défense et sont induits par les blessures d'insectes et l'application de jasmonate (McKay et al., 2003; Keeling et Bolhmann, 2006; Phillips et al., 2007). Leur biosynthèse augmente suite au stress hydrique chez les conifères. Plus particulièrement, ce sont les monoterpènes comme l'α-pinene qui s'accumulent chez les conifères (Turtola et al., 2003; Ormeno et al., 2007). De plus, Singh et al., (2006) ont montré que l' α -pinene est capable de causer des dommages oxidatifs dans les tissus racinaires en provoquant l'accumulation d'espèces réactives de l'oxygène, entraînant la peroxidation des lipides, la perte d'intégrité des membranes et l'accumulation des enzymes antioxidantes.

La synthèse d'anthocyanes par les jeunes plantules exprimant PtMYB14 est aussi en accord avec l'accumulation de transcrits observée chez les deux types de transgéniques, dont principalement ceux de la chalcone synthase (voir Figure I-6, Chapitre V). Le niveau d'expression du transgène semble proportionnel à l'augmentation de la teneur en cyanidine 3-0-glucoside. La coloration rouge que nous avons observé dans les plantules d'épinettes « Pro_{UBI}PtMYB14 » semble se rapprocher de la couleur pourpre que prennent les aiguilles et les ramilles des jeunes plants de pin gris à l'automne ainsi que des hypocotyles rouges de graines récemment germées (Krugman, 1956; Nozzolillo et al., 1990; Nozzolillo et al., 2002). Ces tissus contiennent au moins cinq pigments anthocyanes (majoritairement de la cyanidine 3-0-glucoside) identiques à ceux que nous avons identifiés, et ce, dans des proportions très comparables. Parmi les explications apportées pour ce phénomène d'accumulation d'anthocyanes, la plus largement acceptée est celle de Krol et al. (1995) qui suggère que cet ensemble d'anthocyanes sert à protéger des effets dommageables de la lumière sur les jeunes tissus végétatifs. En effet, l'exposition de jeunes plantules « Pro_{UBI}PtMYB14 » à la lumière semble être requise pour provoquer le rougissement des hypocotyles (résultats non montrés).

La dérégulation de l'expression des gènes de la voie de biosynthèse des phénylpropanoïdes observée dans les plantules « $Pro_{UBI}PtMYB14$ » semble vraisemblablement être une conséquence indirecte de la surexpression liée à l'effet de dose très élevé du transgène *PtMYB14* ou encore au phénotype exacerbé de ces plantules transgéniques. En effet, les niveaux d'expression de gènes de type *O-méthyl transférase* (*COMT* et *CCoAOMT*) sont tous clairement diminués chez « $Pro_{UBI}PtMYB14$ », mais pas chez les transgéniques « $Pro_{CAD}PtMYB14$ ». Cette différence pourraît être due à un effet de dose puisque les transgéniques « $Pro_{CAD}PtMYB14$ » expriment le transgène environ cent fois moins que le promoteur *UBI*. L'analyse des pousses terminales et du bois des arbres « $Pro_{CAD}PtMYB14$ » n'indique pas non plus de répression des *OMT*. Cette dérégulation des *OMT* par *PtMYB14*. De la même façon, il n'y a pas de diminution de

l'expression des autres gènes de la voie de biosynthèse des phénylpropanoïdes dans les plantules « $Pro_{CAD}PtMYB14$ » qui expriment le transgène à un niveau plus proche du niveau physiologique normal. Cette conclusion est opposée à notre hypothèse initiale étant donné que plusieurs indices pouvaient associer PtMYB14 à la répression transcriptionnelle des gènes de le voie de biosynthèse des phénylpropanoïdes et des *OMT* en particulier. Par exemple, il y a un motif protéique potentiellement répresseur qui est commun au sous-groupe 4. Nous avons aussi observé une légère diminution de la quantité de lignine dans le bois des jeunes arbres « $Pro_{CAD}PtMYB14$ » en serre. Aussi, il a été attribué la fonction de répresseur des gènes des phénylpropanoïdes à plusieurs des membres du sous-groupe 4 chez d'autres espèces végétales (chez les angiospermes) pour lesquelles il a été démontré que la surexpression entraîne une réduction de la lignification (*Arabidopsis thaliana* MYB4, Jin et *al.*, 2000; AtMYB32, Preston et *al.*, 2004; *Anthirinum majus* MYB308/330, Tamagnone et *al.*, 1998; *Zea mais* MYB31 et 42, Fornalé et *al.*, 2006; *Eucalyptus gunni* MYB1, Legay et *al.*, 2007).

L'augmentation de la quantité d'amidon de presque 50% dans les plantules « $Pro_{UBI}PtMYB14$ », mais pas dans les plantules « $Pro_{CAD}PtMYB14$ », suggère aussi qu'il s'agit d'un effet indirect de la forte expression du transgène. Cette accumulation pourrait en effet être attribuée à la perturbation du métabolisme des isoprénoïdes, si la biosynthèse de la gibbereline (GA) s'en trouvait affectée tel que le suggère la répression des transcrits de la GGPPS. Puisque la GA est un régulateur de l' α -amylase (Jacobsen et Beach, 1985), une baisse du niveau de la GA pourrait potentiellement entraîner une accumulation de l'amidon. L'accumulation d'amidon a été également observée dans des embryons en germination d'épinettes blanches traités avec de la glutathion réduite (GSH) (Stasolla et *al.*, 2004). Les auteurs montrent aussi que l'addition de la GSH entraîne un phénotype hypertrophié des hypocotyles, similaire à ceux des plantules « $Pro_{UBI}PtMYB14$ ».

En conclusion, nos résultats indiquent que PtMYB14 agirait vraisemblablement dans la régulation du métabolisme des isoprénoïdes. Ainsi, il pourrait intervenir dans les mécanismes de défense et de réponse aux stress sousjacents à plusieurs processus physiologiques. Par exemple, la réponse terpénique est liée aux stress biotiques (e.g. attaques par des insectes) et aux stress abiotiques (e.g. déshydratation). Par ailleurs, son action sur la synthèse des flavonoïdes, qui agissent en tant qu'antioxidants, pourrait intervenir dans la réponse au froid. Notre approche expérimentale comparant l'effet de PtMYB14 via deux types de promoteurs indique que PtMYB14 agirait plus probablement en tant qu'activateur des mécanismes de défense plutôt que répresseur des gènes de la biosynthèse des phénylpropanoïdes comme décrit auparavant pour des membres de ce même sous-groupe 4 chez les angiospermes. Notre démarche comparative nous amène aussi à conclure que plusieurs des transcrits dérégulés et certains phénotypes des plantules « Pro_{UBI}PtMYB14 » résultent d'effets pléiotropes. Ce phénomène pourrait en partie s'expliquer par le fort niveau de surexpression combiné au caractère multigénique du sous-groupe 4 des conifères. Cet argument s'appuie sur trois considérations majeures. D'abord, les profils d'expression très variés des gènes du Sg4-C suite aux traitements d'acide jasmonique, de blessure et de froid (Chapitre V) indiquent une diversification et une spécialisation potentielle de leurs rôles biologiques. De plus, le site d'interaction avec les protéines bHLH présent dans tous les MYB-R2R3 du sousgroupe Sg4-C suggère leur participation à des interactions tripartites MYB-bHLH-WD40 (Ramsay et Glover, 2005), ce qui pourrait aussi contribuer à leur conférer des rôles diversifiés en accord avec leur profils d'expression. Finalement, la dose très élevée de PtMYB14 pourrait mener à l'interaction avec différents gènes cibles de façon plus ou moins spécifique, impliquant aussi des interactions plus ou moins spécifiques avec des protéines partenaires.

VII-4-Surexpression des facteurs de transcription pour étudier la formation du bois : Analyse critique de l'approche

L'observation de phénotypes évidents lors de la surexpression ubiquitaire Pro_{UBI}-PtMYB1, -PtMYB8 et -PtMYB14 indique que cette approche offre un certain potentiel pour analyser la fonction de ces gènes. Les limites de cette stratégie apparaissent toutefois lorsque l'expression du transgène compromet la viabilité de la plante où retarde son développement au point d'empêcher sa culture, et donc l'étude du xylème en croissance secondaire (e.g. « Pro_{UBI}PtMYB8 »). Par ailleurs, les plantes qui ont eu un développement plus normal après leur transfert en sol semblent exprimer le transgène beaucoup plus faiblement, c'est à dire à un niveau semblable au gène endogène. Ce phénomène pourrait expliquer l'absence de phénotype à ce stade, et même après plusieurs cycles de croissance. De même, je n'ai observé aucun phénotype lors de la surexpression du facteur de transcription *PtMYB3* (aucun résultat présenté). Le fait que l'expression des transgènes dans le xylème ne soit pas plus forte que l'endogène d'épinette pourrait s'expliquer par la forte expression naturelle de ceux-ci dans le xylème. Ainsi, l'approche expérimentale gain de fonction utilisant un promoteur constitutif fort au stade plantule et sans distinction des stades développementaux de l'arbre soulève différentes questions et demande une évaluation critique de la portée des résultats obtenus.

D'autres techniques de génomique fonctionnelle pour la caractérisation de facteurs de transcription *in planta* ont été utilisées chez *Arabidopsis*, mais très peu chez les arbres et encore moins chez les conifères (Schwechheimer et *al.*, 1998; Zhang et *al.*, 2003; Moore et *al.*, 2006). En effet, les phénotypes complexes analysés dans cette étude (faible lignification ectopique et réduction de la croissance racinaire chez « Pro_{UBI}PtMYB1 », plusieurs phénotypes chez « Pro_{UBI}PtMYB14 ») pourraient être éventuellement contournés ou simplifiés par des expériences de perte de fonction des gènes testés. Cependant, dans le cas de familles multigéniques comme les *MYB-R2R3*, chez un organisme non séquencé comme *Picea glauca*, cette approche risque souvent de ne donner aucun phénotype mutant à cause de la redondance de fonction totale, partielle ou inégale entre des gènes homologues, de séquences hautement

similaires (Bouché et Bouchez, 2001; Briggs et *al.*, 2006; Pickett et Meeks-Wagner, 1995).

La redondance génétique est un phénomène courant et complexe qui serait responsable de l'absence de phénotypes dans la majorité des mutants simples homozygotes de type perte de fonction (i.e. un seul gène muté) chez Arabidopsis (Briggs et al., 2006). De ce fait, il est fréquent et nécessaire de réaliser des doubles ou triples mutants, par croisement des plantes mutantes, afin de générer un phénotype. Par exemple, c'est l'analyse des doubles mutants de trois des gènes MYB-R2R3 du sous-groupe 14 chez Arabidopsis, un quatrième ne présentant pas de phénotype, homologues au gène Bl de tomate, qui ont permis de démontrer que ces gènes agissent selon un mode de redondance partielle pour la régulation de la formation du méristème axillaire (Müller et al., 2006; Feng et al., 2004; Schmitz et al., 2002). De même, les gènes AtMYB33 et AtMYB65 du sous-groupe 18 de Krantz et al. (1998) ne présentent aucun phénotype sous la forme de mutants simples mais le double mutant *myb33 myb65* présente des anomalies dans le développement des anthères (Millar et Gubler, 2005). L'approche expérimentale basée sur une perte de fonction simple s'est également montrée inefficace pour révéler des phénotypes chez les facteurs à domaine MADS (Lehti-Shiu et al., 2005), les bHLH (Wang et al., 2007) ou encore les protéines Aux/IAA (Overvoorde et al., 2005).

Dans le cas des conifères, la production de plantes mutantes demande une infrastructure plus lourde que celle utilisée pour *Arabidopsis*. L'obtention de doubles ou triples mutants par croisement chez les conifères est aussi irréaliste à cause de l'âge de la maturité sexuelle des conifères, porté vers 10-15 ans comparativement à environ trois semaines chez *Arabidopsis*. De ce fait, l'expérimentation par perte de fonction visant la diminution voire l'élimination de l'expression d'un gène endogène *via* des techniques de type interférence d'ARN (RNAi) aurait pu être utilisées et permettre d'identifier les gènes co-régulés par certains MYB (Cannon et *al.*, 1990; Pnueli et *al.*, 1994; van der Krol et *al.*, 1990). Toutefois,, la redondance de fonction qui contournerait la sous-expression pourrait être critique dans le cas des membres du sous-groupe Sg4-C incluant *PtMYB14*, de part leur nombre et leurs fortes similarités de séquences. Puisqu'elle est moins affectée par la redondance fonctionnelle, les techniques de surexpression représentent une alternative et une stratégie complémentaire à l'analyse de mutants et de sous-exprimants (Zhang 2003).

Plusieurs méthodes pour la caractérisation fonctionnelle de facteurs de transcription (FT) par surexpression seraient applicables chez les conifères afin de contourner l'utilisation de mutants et autres sous-exprimants. Par exemple, la surexpression peut se faire simultanément avec plusieurs gènes appartenant à la même famille (MADS : Pelaz et al., 2001; Honna et Goto, 2001) ou à des familles différentes (bHLH et MYB : Abe et al., 2003; Bovy et al., 2002) car les facteurs de transcription fonctionnent souvent de façon interdépendante (Singh, 1998). La surexpression peut aussi être réalisée avec des facteurs de transcription modifiés à cause de leur structure bipartite, composée d'un domaine de fixation à l'ADN et d'un autre domaine pour la modulation/régulation de la transcription. Ainsi, l'obtention de dominants négatifs ou positifs est possible via une délétion dans le domaine régulateur ou le DBD, ou bien en éliminant entièrement le domaine régulateur (Zhang, 2003). Des FT chimériques sont également concevables dans le but de produire des dominants négatifs par l'ajout d'un motif répresseur au FT (Hiratsu et al., 2003, méthode CRES-T ou « chimeric repressor silencing technology » utilisant le motif répresseur EAR; Matsui et al., 2004; Mistuda et al., 2006; Takase et al., 2007). Dans le cas de notre étude, par exemple, réaliser une chimère où la protéine PtMYB8 serait fusionnée au répresseur selon la méthode CRES-T conduirait à un dominant négatif et donc à une sous-expression potentielle des gènes de la lignification dans des plantules d'épinettes.

Une autre voie intéressante serait l'utilisation de promoteurs préférentiels, voire spécifiques de certains tissus, de certaines cellules et/ou de certains stades de développement, parce que la spécificité d'expression et le contexte définissent le rôle joué par le FT dans le développement de la plante et son interaction avec l'environnement (Doebley et Lukens, 1998). Dans certains cas, la surexpression avec

un promoteur constitutif fort conduit à un phénotype pléiotrope et difficile à interpréter. En contre partie, la surexpression d'un FT avec un promoteur spécifique du tissu où est normalement exprimé le FT permet de faciliter les analyses (e.g. Helariutta et *al.*, 2000). A l'inverse, pour connaître la contribution d'un certain tissu dans le contrôle de l'identité d'un organe, il est nécessaire d'exprimer le FT de façon ectopique dans ce tissu (e.g. Efremova et *al.*, 2001). Enfin, il faut considérer l'utilisation de promoteurs inductibles (par voie chimique, par exemple) qui seraient constitutifs, voire spécifiques de certains tissus afin d'être spécifiquement activés lors d'un stade développemental choisi (revue par Moore et *al.*, 2006).

Finalement, l'idéal pour étudier la formation du bois chez les conifères pourrait être de conduire l'expression du FT en ciblant très spécifiquement les cellules du xylème en croissance secondaire pour ne pas empêcher le développement normal de la plante jusqu'à ce stade. Dans cette perspective, la caractérisation du promoteur du gène *CAD* de *Picea glauca* présente certains avantages, mais la séquence présentée ici semble donner une expression plutôt faible, et n'est pas inductible par voie chimique, à notre connaissance.

VII-5-Les gènes de la famille *MYB-R2R3*: une nouvelle source de gènes candidats pour l'amélioration de la qualité du bois et un nouvel outil pour la conservation des forêts naturelles?

L'amélioration de la qualité du bois chez les conifères au Canada est un enjeu important aux niveaux économique et industriel mais aussi aux niveaux écologique et environnemental. En effet, les forêts canadiennes comptent jusqu'à 10% du couvert forestier mondial et les conifères en représentent 66%. Le Canada se classe au deuxième rang des producteurs de bois d'œuvre de résineux avec 19% de la production mondiale (Ressources Naturelles Canada 2005-2006). Parmi les 0,9 million d'hectares de forêts qui sont le théâtre d'une récolte annuelle, 53% se régénèrent naturellement, 43% sont replantés et 4% sont ensemencés (RNC 2005-2006). Les plantations forestières au Canada représentent donc environ 400 000 hectares par an, avec principalement l'épinette (*Picea glauca, Picea marianna*) et le pin (*Pinus banksiana*). Le Canada est le premier exportateur mondial de produits forestiers, avec 17,3% du commerce mondial. Les principaux produits exportés sont le bois d'œuvre et la pâte à papier. En 2005, l'industrie forestière a contribué à hauteur de 2,9% au produit intérieur brut du Canada avec 864 000 emplois, dont 339 900 emplois directs (2,1% de l'emploi total au Canada) et 524 100 emplois indirects et induits (RNC 2005-2006). Ainsi, les produits et les retombées de l'industrie forestière sont particulièrement importants pour les Canadiens et l'économie du pays.

Cependant, avec l'augmentation de la concurrence mondiale et la demande croissante en matériaux énergétiques et de construction, le bois devient une ressource naturelle dont la production doit être maîtrisée. Les facteurs essentiels à considérer sont, d'une part, le temps nécessaire à la croissance des arbres pour leur exploitation (40 à 80 ans pour les conifères) et, d'autre part, les propriétés physico-chimiques du bois qui ont un impact sur les différentes utilisations industrielles. Dans la perspective de plantations forestières productives, il devient donc indispensable de connaître les mécanismes génétiques liés à la formation du bois afin de sélectionner les arbres géniteurs des plantations de demain. Cet objectif pourra être atteint en mettant à profit la diversité génétique liée aux propriétés intrinsèques des arbres et du bois dans des populations de forêts naturelles.

VII-5-1-Des gènes candidats pour l'amélioration de la qualité du bois

Les gènes candidats (GC) pour l'amélioration de la qualité du bois sont des gènes potentiellement impliqués dans un caractère quantitatif important pour la valeur ajoutée de l'arbre en termes de propriétés physico-chimiques, comme la quantité et la qualité des lignines ou de la teneur en cellulose, par exemple. Les GC peuvent être identifiés par différentes approches telles que des études fonctionnelles (i.e. gain/perte de fonction), la co-localisation avec un caractère quantitatif ou QTL (e.g. quantité de lignine), ou le profilage d'expression des gènes dans différents tissus et organes d'un arbre ou dans différentes conditions physiologiques (e.g. spécifique au xylème secondaire, bois mature, bois de compression).

L'utilisation de gènes candidats est une stratégie prometteuse pour faire le lien entre les approches de génétique quantitative (i.e. détection de QTL) et de génétique moléculaire (i.e. fonctions des gènes) pour l'étude de caractères complexes (Pflieger et al., 2001). En effet, dans le cas de plantes qui possèdent de très grands génomes comme certaines céréales et surtout comme les arbres tels que les conifères, la génération de cartes moléculaires de liaisons basées sur des GC de fonctions connus peut être une façon d'identifier directement les déterminants génétiques plutôt que de chercher à affiner les cartes génétiques avec d'autres types de marqueurs moléculaires (Thornsberry et al., 2001; Palaisa et al., 2003). Ainsi, les facteurs de transcription comme les gènes MYB-R2R3 de cette étude pourraient représenter de meilleurs candidats conférant un caractère mesurable (i.e. niveaux de transcrits ou eQTL) que des gènes effecteurs. En effet, un simple facteur de transcription peut réguler de façon coordonnée plusieurs gènes influencant un caractère donné (e.g. quantité de lignine). Ainsi, récemment Goicoechea et al. (2005) ont mis en évidence que EgMYB2, un MYB-R2R3 d'Eucalyptus, co-localise avec un caractère quantitatif (contenu en lignine). Ces auteurs ont aussi montré que ce même gène était préférentiellement exprimé dans le xylème et qu'il agit comme activateur transcriptionnel de la formation de la paroi secondaire et de la biosynthèse de la lignine. Ainsi, l'approche gène candidat, incluant des gènes facteurs de transcription, offre de nouvelles perspectives pour la sélection génétique assistée par marqueurs (SAM) basée sur les gènes eux-mêmes et non pas sur des marqueurs anonymes ou proches des QTLs. La détection de caractères génétiques liés aux caractères phénotypiques d'intérêt permettrait la sélection précoce des arbres pour des caractères à expression tardive. Pour l'atteinte de cet objectif, le défi consiste à établir un lien entre les allèles et leurs effets, pour identifier les allèles dits "favorables" pour chacun des caractères quantitatifs d'intérêt. Ces analyses d'association doivent être faites dans diverses populations d'arbres, car la fréquence des allèles et leurs effets phénotypiques peuvent varier selon la population et les conditions édapho-climatiques (Neale et Savolaimen, 2004).

VII-5-2-Des gènes candidats pour la conservation des forêts naturelles

La régulation de la voie de biosynthèse de la lignine joue un rôle dans l'adaptation des arbres à leur environnement de part son importance dans le déterminisme des propriétés physico-chimiques du bois, son impact sur la croissance de l'arbre et la tolérance aux stress biotiques et abiotiques (Pot et al., 2002; Peter et Neale, 2004). Certains des facteurs de transcription que nous avons caractérisés dans cette étude pourraient être de bons candidats pour étudier la variation adaptative de l'épinette blanche. Le développement récent d'outils de génomique permet l'identification de marqueurs moléculaires de type polymorphisme nucléotidique (SNP) idéaux pour l'étude de la variation adaptative à cause de leur nombre relativement élevé dans les gènes et les génomes, leur présence dans des gènes identifiés et de fonctions connues, et surtout la possibilité d'identifier facilement des variations alléliques entre les individus (Gonzalez-Martinez et al., 2006). L'étude de la différenciation génétique parmi les populations naturelles (i.e. fréquence allélique) peut révéler des pressions de sélection locales résultant en des divergences adaptatives et identifier des endroits géographiques ayant favorisé l'apparition d'une mutation et son maintien via une pression de sélection naturelle (Storz et al., 2005). La préservation des différents polymorphismes génétiques adaptatifs des populations de forêts naturelles est un des objectifs majeurs de la génétique de conservation (Moritz, 2002).

L'identification et le choix de marqueurs moléculaires associés à des caractères phénotypiques adaptatifs aussi divers que les propriétés du bois, la tolérance à la sécheresse, la résistance au froid et aux maladies, ainsi que d'autres caractères phénologiques variés permettra de mieux cibler l'échantillonnage de populations pour l'établissement de réserves génétiques *in situ* ou pour mettre en place des plantations de conservation *ex situ*. Ces "collections" d'individus et de populations pourraient également constituer un réservoir de diversité génétique indispensable à la génération des futures plantations d'arbres répondant aux exigences économiques, écologiques et environnementales. En effet, la diversité génétique intraspécifique d'espèces d'arbres majeurs et de leurs espèces végétales

compagnes peut avoir de profondes conséquences sur les écosystèmes à travers l'étendue de leurs phénotypes (Whitham et *al.*, 2003) et, par le fait même, avoir des implications tout aussi importantes pour la conservation globale et l'équilibre de la diversité tant végétale que faunique.

VII-6-Références

- Abe H., Urao T., Ito T., Seki M., Shinozaki K. et Yamaguchi-Shinozaki, K. (2003). Arabidopsis AtMYC2 (bHLH) and AtMYB2 (MYB) function as transcriptional activators in abscisic acid signaling. *Plant Cell* **15**:63-78
- Bouché N. et Bouchez D. (2001). Arabidopsis gene knockout: phenotypes wanted. *Curr Opin Plant Biol* 4: 111-117
- Bovy A., de Vos R., Kemper M., Schijen E., Almenar Pertejo M., Muir S., Collins G., Robinson S., Verhoeyen M., Hughes S., Santos-Buelga C. et van Tunnen A. (2002). High-flavonol tomatoes resulting from the heterologous expression of the maize transcription factor genes LC and C1. *Plant Cell* 14: 2509-2526
- Braun E.L. et Grotewold E. (1999). Newly discovered plant c-myb-like genes rewrite the evolution of the plant *myb* gene Family. *Plant Physiol* 121: 21-24
- Briggs G.C., Osmont K.S., Shindo C., Sibout R. et Hardtke C.S. (2006). Unequal genetic redundancies in *Arabidopsis* a neglected phenomenon? *Trends Plant Sci* 11: 492-498
- Cannon M., Platz J., O'Leary M., Sookdeo C. et Cannon F. (1990). Organ-specific modulation of gene expression in transgenic plants using antisense RNA. *Plant Mol Biol* 15: 39-47
- Christensen A.H., Sharrock R.A. et Quail P.H. (1992). Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Mol Biol* 18: 675–689
- Christensen A.H. et Quail P.H. (1996). Ubiquitin promoter-based vectors for highlevel expression of selectable and/or screenable marker genes in monocotyledonous plants. *Transgenic Res* 5: 213-221
- **Demura T. et Fukuda H. (2007).** Transcriptional regulation in wood formation. *Trends Plant Sci* **12:** 64-70
- **Dias A.P., Braun E.L., McMullen M.D. et Grotewold E. (2003).** Recently duplicated maize *R2R3 Myb* genes provide evidence for distinct mechanisms of evolutionary divergence after duplication. *Plant Physiol* **131:** 610-620
- **Doebley J. et Lukens L. (1998).** Transcriptional regulators and the evolution of plant form. *Plant Cell* **10 :** 1075-1082
- Efremova N., Perbal M.C., Yephremov A., Hofmann W.A., Saedler H. et Schwarz-Sommer Z. (2001). Epidermal control of floral organ identity by class B homeotic genes in Antirrhinum and Arabidopsis. *Development* 128: 2661-2671
- Fedorova L. et Fedorov A. (2003). Introns in gene evolution. Genetica 118: 123-131

- Feng C., Andreasson E., Maslak A., Mock H.P., Mattson O. et Mundy J. (2004). Arabidopsis MYB68 in development and responses to environmental cues. Plant Sci. 167: 1099–1107
- Fornalé S., Sonbol F.M., Maes T., Capellades M., Puigdomènech P., Rigau J. et Caparrós-Ruiz D. (2006). Down-regulation of the maize and *Arabidopsis* thaliana caffeic acid O-methyl-transferase genes by two new maize R2R3-MYB transcription factors. *Plant Mol Biol* 62: 809-823
- Goicoechea M., Lacombe E., Legay S., Milhaevic S., Rech P., Jauneau A., Lapierre C., Pollet B., Verhaegen D., Chaubet-Gigot N. et Grima-Pettenati J. (2005). EgMYB2, a new transcriptional activator from *Eucalyptus* xylem, regulates secondary cell wall formation and lignin biosynthesis. *Plant J* 43: 553–567
- Gomez-Maldonaldo J., Avila C., de la Torre F., Canas R., Canovas F.M. et Campbell M.M. (2004). Functional interactions between a glutamine synthetase promoter and MYB proteins. *Plant J* 39: 513–526
- Gonzalez-Martinez S.C., Krutovsky K.V. et Neale D.B. (2006). Forest-tree population genomics and adaptive evolution. *New Phytol* 170: 227-238
- Helariutta Y., Fukaki H., Wysocka-Diller J., Nakajima K., Jung J., Sena G., Hauser M.T et Benfey P.N. (2000). The SHORT-ROOT gene controls radial patterning of the Arabidopsis root through radial signaling. *Cell* 101: 555-567
- Hiratsu K., Matsui K., Koyama T. et Ohme-Takagi M. (2003). Dominant repression of target genes by chimeric repressors that include the EAR motif, a repression domain, in *Arabidopsis*. *Plant J* 24: 733-739
- Honma T. et Goto K. (2001). Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. *Nature* **409**: 525-529
- Humphrey T.V., Bonetta D.T. et Goring D.R. (2007). Sentinels at the wall: cell wall receptors and sensors. *New Phytol* 176: 7-21
- Jacobsen J.V. et Beach L.R. (1985). Control of transcription of α-amylase and rRNA genes in barley aleurone protoplasts by gibberellin and abscisic acid. *Nature* 316: 275-277
- Jia L., Clegg M.T. et Jiang T. (2004). Evolutionary dynamics of the DNA-binding domains in putative R2R3-MYB genes identified from rice subspecies *indica* and *japonica* genomes. *Bioinformatics* 134: 575-585
- Jiang C., Gu X. et Peterson T. (2004a). Identification of conserved gene structures and carboxy-terminal motifs in the Myb gene family of *Arabidopsis* and *Oryza sativa* L. ssp. *indica*. *Genome Biol* **5**: R46
- Jiang C., Gu J., Chopra S., Gu X. et Peterson T. (2004b). Ordered origin of the typical two- and three-repeat Myb genes. *Gene* **326**: 13-22
- Jin H. L., Cominelli E., Bailey P., Parr A., Mehrtens F., Jones J., Tonelli C., Weisshaar B. et Martin C. (2000). Transcriptional repression by AtMYB4 controls production of UV- protecting sunscreens in Arabidopsis. *EMBO J* 19: 6150–6161
- Kaku S., Iwaza-Inoue M. et Toki K. (1992). Anthocyanin influence on water proton NMR relaxation times and water contents of leaves of evergreen woody plants during winter. *Plant Cell Physiol* 33: 131-137

- Keeling C. I. et Bohlmann J. (2006). Genes, enzymes and chemicals of terpenoid diversity in the constitutive and induced defence of conifers against insects and pathogens. *New Phytol* 170 : 657-675
- Ko J.H., Beers E.P. et Han K.H. (2006). Global comparative transcriptome analysis identifies gene network regulating secondary xylem development in *Arabidopsis thaliana*. *Mol Gen Genomics* 276: 517–531
- Kranz H.D., Denekamp M., Greco R., Jin H., Leyva A., Meissner R.C., Petroni K., Urzainqui A., Bevan M., Martin C., Smeekens S., Tonelli C., Paz-Ares J. et Weisshaar B. (1998). Towards functional characterization of the members of the R2R3-MYB gene family from *Arabidopsis thaliana*. *Plant J* 16: 263–276
- Kranz H., Scholz K. et Weisshaar B. (2000). c-MYB oncogene-like genes encoding three MYB repeats occur in all major plant lineages. *Plant J* 21: 231-236
- Krol M., Gray G.R., Hurry V.M., Öquist G., Malek L. et Huner N. P. A. (1995). Low-temperature stress and photoperiod affect an increased tolerance to photoinhibition in *Pinus banksianna* seedling. *Can J Bot* **73** : 1119-1127
- Krugman S. (1956). The anthocyanin and leuco-anthocyanins of sugar pine seedlings. *For Sci* **2**: 273-280
- Kubo M., Udagawa M., Nishikubo N., Horiguchi G., Yamaguchi M., Ito J., Mimura T., Fukuda H. et Demura T. (2005). Transcription switches for protoxylem and metaxylem vessel formation. *Genes Dev* 19: 1855-1860
- Legay S., Lacombe E., Goicoechea M., Brière C., Séguin A., MacKay J. et Grima-Pettenati J. (2007). Mplecular characterization of EgMYB1, a putative transcriptional repressor of the lignin biosynthetic pathway. *Plant Sci* (sous presse)
- Lehti-Shiu M.D., Adamczyk B.J. et Fernandez D.E. (2005). Expression of MADSbox genes during the embryonic phase in Arabidopsis. *Plant Mol Biol* 58: 89-107.
- Matsui K., Tanaka H. et Ohme-Takagi. M. (2004) Suppression of the biosynthesis of proanthocyanidin in Arabidopsis by a chimeric PAP1 repressor. *Plant Biotech J* 2: 487–493
- McKay S.A.B, Hunter W.L., Godard K-A, Wang S.X., Martin D.M., Bohlmann J. et Plant A.L. (2003). Insect attack and wounding induce traumatic resin duct development and gene expression of (-)-pinene synthase in sitka spruce. *Plant Physiol* 133: 368-378
- Millar A.A. et Gubler F. (2005). The Arabidopsis GAMYB-like genes, MYB33 and MYB65, are microRNA-regulated genes that redundantly facilitate anther development. *Plant Cell* 17: 705-721
- Mitsuda N., Hiratsu K., Todaka D., Nakashima K., Yamaguchi-Shinozaki K. et Ohme-Takagi M. (2006). Efficient production of male and female sterile plants by expression of a chimeric repressor in Arabidopsis and rice. *Plant Biotech J* 4: 325–332
- Mitsuda N., Seki M., Shinozaki K. et Ohme-Takagi M. (2005). The NAC transcription factors NST1 and NST2 of *Arabidopsis* regulates secondary wall thickening and are required for anther dehiscence. *Plant Cell* **17** : 2993-3006
- Mitsuda N., Iwase A., Yamamoto H., Yoshida M., Seki M., Shinozaki K. et Ohme-Takagi M. (2007). NAC transcription factors, NST1 and NST3, are key

regulators of the formation of secondary walls in woody tissues of *Arabidopsis*. *Plant Cell* (sous presse)

- Moore I., Samalova M. et Kurup S. (2006). Transactivated and chemically inducible gene expression in plants. *Plant J* 45: 651-683
- Moritz C. (2002). Strategies to protect biological diversity and the evolutionary process that sustain it. System Biol 51: 238-254
- Müller D., Schmitz G. et Theres K. (2006). Blind homologous R2R3 Myb genes control the pattern of lateral meristem initiation in Arabidopsis. *Plant Cell* 18: 586-597
- Neale D.B. et Savolainen O. (2004). Association genetics of complex traits in conifers. *Trends Plant Sci* 9: 325-330
- Nei M. (1969). Gene duplication and nucleotide substitution in evolution. *Nature* 221 : 5175-5177
- Newman L.J., Perazza D.E., Juda L. et Campbell M.M. (2004) Involvement of the R2R3-MYB, At MYB61, in the ectopic lignication and dark-photomorphogenic components of the det3 mutant phenotphype. *Plant J* 37: 239–250
- Nozzolillo C., Isabelle P., Andersen O.M. et Abou-Zaid M. (2002). Anthocyanins of jack pine (*Pinus banksiana*) seedlings. *Can J Bot* 80 : 796-801
- Nozzolillo C., Isabelle P. et Das G. (1990). Seasonal changes in the phenolic constituents of jack pine seedlings (*Pinus banksiana*) in relation to the purpling phenomenon. *Can J Bot* 68 : 2010-2017
- Ormeno E., Mévy J.P., Vila B., Bousquet-Mélou A., Greff S., Bonin G. et Fernandez C. (2007). Water deficit stress induces different monoterpene and sesquiterpene emission changes in Mediterranean species. Relationship between terpene emissions and plant water potential. *Chemosphere* 67: 276-284
- Overvoorde P.J., Okushima Y., Alonso J.M., Chan A., Chang C., Ecker J.R., Hughes B., Liu A., Onodera C., Quach H., Smith A., Yu G. et Theologis A. (2005). Functional genomic analysis of the AUXIN/INDOLE-3-ACETIC ACID gene family members in *Arabidopsis thaliana*. *Plant Cell* 17: 3282-3300
- Palaisa K.A., Morgante M., Williams M. et Rafalski A. (2003). Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phytoene synthase loci. *Plant Cell* 15:1795–1806
- Patzlaff A., Newman L.J., Dubos C., Whetten R.W., Smith C., McInnis S., Bevan M.W., Sederoff R.R. et Campbell M.M. (2003a). Characterisation of PtMYB1, an R2R3-MYB from pine xylem. *Plant Mol Biol* 53: 597–608
- Patzlaff A., McInnis S., Courtenay A., Surman C., Newman L.J., Smith C., Bevan M.W., Mansfield S.D., Whetten R.W., Sederoff R.R. et Campbell M.M. (2003b). Characterisation of a pine MYB that regulates lignification. *Plant J* 36: 743–754
- Pavy N., Paule C., Parsons L., Crow J.A., Morency M.J., Cooke J.K., Johnson J.E., Noumen E., Guillet-Claude C., Butterfield Y., Barber S., Yang G., Liu J., Stott J., Kirkpatrick R., Siddiqui A., Holt R., Marra M., Séguin A., Retzel E., Bousquet J. et MacKay J. (2005) Generation, annotation, analysis and database integration of 16,500 white spruce EST clusters. *BMC Genomics* 6: 144

- Pelaz S., Gustafson-Brown C., Kohalmi S.E., Crosby W.L. et Yanofsky M.F. (2001). APETALA1 and SEPALLATA3 interact to promote flower development. *Plant J* 26: 385-394
- Peter G. et Neale G. (2004). Molecular basis for the evolution of xylem lignification. *Curr Opin Plant Biol* 7: 737-742
- Pflierger S., Lefebvre V. et Causse M. (2001). The candidate gene approach in plantgenetics: a review. *Mol Breed* 7: 275-291
- Phillips M.A., Walter M.H., Ralph S.G., Dabrowska P., Luck K., Urós E.M., Boland W., Strack D., Rodríguez-Concepción M., Bohlmann J. et Gershenzon J. (2007). Functional identification and differential expression of 1-deoxy-D: -xylulose 5-phosphate synthase in induced terpenoid resin formation of Norway spruce (*Picea abies*). *Plant Mol Biol* 65: 243-257
- Pickett F.B. et Meeks-Wagner D.R. (1995). Seeing double: appreciating genetic redundancy. *Plant Cell* 7: 1347-1356
- Pnueli L., Hareven D., Broday L., Hurwitz C. et Lifschitz E. (1994). The TM5 MADS box gene mediates organ differentiation in the three inner whorls of tomato flowers. Plant Cell 6: 175-186
- Pot D., Chantre G., Rozenberg P., Rodriguez J.C., Jones G.L., Pereira H., Hannrup B., Cahalan C. et Plomion C. (2002). Genetic control of pulp and timber properties in maritime pine (*Pinus pinaster* Ait.). Ann For Sci 59: 563-575
- Preston J., Wheeler J., Heazlewood J., Li S. F. et Parish R. W. (2004). AtMYB32 is required for normal pollen development in *Arabidopsis thaliana*. *Plant J* 40: 979-995
- **Qu L.J. et Zhu Y.X. (2006).** Transcription factor families in Arabidopsis: major progress and outstanding issues for future research. *Curr Opin Plant Biol* **9**: 544–549
- Quackenbush J., Cho J., Lee D., Liang F., Holt I., Karamycheva S., Parvizi B., Pertea G., Sultana R. et White J. (2001). The TIGR Gene Indices: analysis of gene transcript sequences in highly sampled eukaryotic species. *Nucleic Acids Res* 29: 159-164
- Ramsay N.A. et Glover B.J. (2005). MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. *Trends Plant Sci* 10: 63-70
- Ressources Naturelles Canada. 2005-2006. Site internet. http://cfs.nrcan.gc.ca/sof/sof06/overview_f.html
- Romero I., Fuertes A., Benito M.J., Malpica J.M., Leyva A. et Paz-Ares J. (1998). More than 80R2R3 MYB regulatory genes in the genome of Arabidopsis thaliana. *Plant J* 14: 273-284
- Schmid M., Davison T.S., Henz S.R., Pape U.J., Demar M., Vingron M., Scholkopf B., Weigel D. et Lohmann J.U. (2005). A gene expression map of Arabidopsis thaliana development. Nat Genet 37: 501–506
- Schmitz G., Tillmann E., Carriero F., Fiore C., Cellini F. et Theres K. (2002). The tomato *Blind* gene encodes a *MYB* transcription factor that controls the formation of lateral meristems. *Proc Natl Acad Sci USA* **99**: 1064–1069
- Schwechheimer C., Zourelidou M. et Bevan M.W. (1998). Plant transcription factor studies. *Annu Rev Plant Physiol Plant Mol Biol* 49: 127-150

- Singh H.P., Batish D.R., Kaur S., Arora K. et Kohli R.K. (2006). alpha-Pinene inhibits growth and induces oxidative stress in roots. *Ann Bot (Lond)* 98: 1261-1270
- Singh K.B. (1998). Transcriptional regulation in plants: the importance of combinatorial control. *Plant Physiol* 118: 1111-1120
- Stasolla C., Belmonte M.F., van Zyl L., Craig D.L., Liu W., Yeung E.C., et Sederoff R.R. (2004). The effect of reduced glutathione on morphology and gene expression of white spruce (*Picea glauca*) somatic embryos. *J Exp Bot* 55: 695-709
- Storz J.F. (2005). Using genome scans of DNA polymorphism to infer adaptive population divergence. *Mol Ecol* 14: 671-688
- Takase T., Yasuhara M., Geekiyanage S., Ogura Y. et Kiyosue T. (2007) Overexpression of the chimeric gene of the floral regulator CONSTANS and the EAR motif repressor causes late flowering in Arabidopsis . *Plant Cell Rep* 26: 815-821
- Tamagnone L., Merida A., Parr A., Mackay S., Culianez-Macia F.A., Roberts K. et Martin C. (1998). The AmMYB308 and AmMYB330 transcription factors from *Antirrhinum* regulate phenylpropanoid and lignin biosynthesis in transgenic tobacco. *Plant Cell* 10: 135–154
- Thornsberry J.M., Goodman M.M., Doebley J., Kresovich S., Nielsen D. et Buckler E.S. 4th. (2001). Dwarf8 polymorphisms associate with variation in flowering time. *Nat Genet* 28: 286-289
- Turtola S., Manninan A-M, Rikala R. et Kainulainen P. (2003). Drought stress alters the concentration of wood terpenoids in scots pine and norway spruce seedling. *J Chem Ecol* 29: 1981-1995
- Tuskan G.A. et al. (109 co-auteurs). (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science* 313: 1596-1604
- Van der Krol A.R., Mur L.A., de Lange P., Gerats A.G.M., Mol J.N. et Stuitje A.R. (1990). Inhibition of flower pigmentation by antisense CHS genes: promoter and minimal sequence requirements for the antisense effect. *Plant Mol Biol* 14: 457-466
- Wang H-Y, Flatte M., Jakoby M., Bäumlein H., Weisshaar B. et Bauer P. (2007). Iron deficiency-mediiated stress regulation of four subgroup Ib BHLH genes in Arabidopsis thaliana. Plant 226: 897-908
- Whitham T.G., Young W.P., Martinsen G.D., Gehring C., Schweitzer J.A., Shuster S.M., Wimp G.M., Fisher D.G., Bailey J.K., Lindroth R.L., Woolbright S. et Kuske C.R. (2003). Community and ecosystem genetics: a consequence of the extended phenotype. *Ecol* 84: 559-573
- Zhang J.Z. (2003). Overexpression analysis of plant transcription factors. *Curr Opin Plant Biol* 6: 430-440
- Zhong R., Demura T. et Ye Z-H. (2006). SND1, a NAC domain transcription factor, is a key regulator of secondary wall synthesis in fibers of *Arabidopsis*. *Plant Cell* 18 : 3158-3170
- Zhong R, Richardson E.A. et Ye Z-H. (2007). Two NAC domain transcription factors, SND1 and NST1, function redundantly in regulation of secondary wall synthesis in fibers of Arabidopsis. *Planta* 255: 1603-1611

Annexe II-1. Protocole d'extractions des ARN messagers

Adapté de Chang, et al. (1993)

- 1. Ajouter 750 μ L de CTAB, contenant 2% de β -mercaptoéthanol, à 65°C dans chacun des tubes ependorfs contenant la poudre finement broyée (maximum de 400 μ L en volume du tube ependorf 1,5 mL).
- 2. Vortexer vigoureusement les tubes et les placer immédiatement dans le bainmarie à 65°C pendant 10 minutes. Vortexer périodiquement les tubes.
- 3. Ajouter 500μL de chloroforme : isoamyl (24 :1) par extrait, mélanger assez vigoureusement par retournement des tubes et les centrifuger à vitesse maximale (13 500 rpm) pendant 5 minutes à 4°C.
- 4. Prélever la phase aqueuse et la récupérer dans un nouveau tube.
- 5. Ajouter 500µL de chloroforme : isoamyl (24 :1) par extrait, mélanger assez vigoureusement par retournement des tubes et les centrifuger à vitesse maximale (13 500 rpm) pendant 10 minutes à 4°C.
- 6. Prélever la phase aqueuse et la récupérer dans un nouveau tube.
- 7. Répéter les étapes 5. et 6.
- 8. Ajouter l'équivalent du volume récupéré de LiCl-EDTA (7,5 M LiCl, 50mM EDTA), mélanger délicatement par inversion et placer les tubes au congélateur -20°C (ou -30°C) pendant 30 minutes.
- 9. Centrifuger pendant 15 minutes à vitesse maximale (13 500 rpm) à 4°C et éliminer le surnageant par pipetage.
- 10. Nettoyer le culot en ajoutant 500µL d'éthanol 80%. Retourner le tube de façon à faire décoller le culot de sur la paroi et centrifuger à 13 500 rpm pendant 15 minutes. La température de la centrifugeuse doit être ajustée à environ 20°C afin de mieux dissoudre les sels.
- 11. Retirer le surnageant par pipetage.
- 12. Laisser sécher le culot à l'air libre jusqu'à ce que le contour commence à devenir transparent (environ 30 minutes).
- 13. Dissoudre le culot avec de l'eau Nanopure autoclavée; entre 10 et $40\mu L$, selon la grosseur du culot.

Tampon d'extraction CTAB: 2% CTAB 2% polyvinylpyrrolidone (PVP)(MW 30000) 100 mM Tris-HCl pH 8.0 25 mM EDTA 2 M NaCl 0.5 g/L spermidine

Appendix II-2. Primers sequences of spruce *MYBs* used for genomic amplification and sequencing

MYB genes	Forward primer sequences	Tm Salt Adjusted	Reverse primer sequences	Tm Salt Adjusted	Amplified gDNA
names	Forward primer sequences	(0)	Reverse primer sequences	(0)	iengui (po)
PgMYBI	TTTAGAACGATCCCCGAATG	56	ATCTCATCCTCCTCCTCCAT	58	1336
PgMYB2	CAATGGGACGCCACTTATG	57	CTACCCTGTGAAAGACGAAGG	61	1898
PgMYB3	TTTTGAAATTGCGAACAGAATG	55	TATTTTTAGGCACCCGTTGATT	56	2430
PgMYB4	GCACGGCCATTCTATTCATT	56	CACCTGATCAAGGATCGACA	58	1503
PgMYB5	CCCTTGCAGCTAAGTTGCTT	58	TATGAATGCTTCGTGGTGGA	56	1243
PgMYB6	CTAATGGCGTCGATGAAAGG	58	GCAACCCCTGAACTCTGTGT	60	1007
PgMYB7	TGCAGAGGCTTGTTCAGAAATA	58	TTTTATTGGAAGGGCTGAGTGT	58	1220
PgMYB8	TCGATCAGAGTACGGGGATT	58	GGCTGCGAGTTTGTGTGAAT	58	1963
PgMYB9	TTCGATCCTGTTGTGAGCTG	58	CCAAGTGCCCCTTTATTTCA	56	1370
PgMYB10	CTTGGTCACGGTTTCCAATC	58	CCCATGGGATAACACACTCC	60	1022
PgMYB11	TTGCATTGTAGAGGGCCAAT	56	TCCATTCCATTCGATCTGTTC	57	2356
PgMYB12	TCAGAAGAACAGGGCACCA	57	ATAAGGGCAATGGCTGGACT	58	1507
PgMYB13	ACCTTGCACGGTCATTCTTC	58	ACACGACTTCGGGCAATTTA	56	938


Appendix II-3. Phylogenetic tree of MYBs from spruce, pine and nearest sequences from other species

Neighbour-joining tree method (1000 Bootstraps) based on an alignment derived using the Clustal W program of the complete coding sequences of the 18 MYB proteins identified in this study (spruce (Pg) and pine (Pt), with full and empty lozenge) and most similar sequences from other species. Figure and complete legend are available at <u>http://www.biomedcentral.com/1471-2229/7/17</u>.

Appendix II-4. Conifer MYB phylogeny based on partial sequences using a 31 amino acids region



The 31 amino acid region in the R3 MYB domain (shown in box) is present in all the MYB used here. The phylogenetic tree was based on an alignment derived using the Clustal W program and the neighbour-joining method (1000 Bootstraps). Spruce (Pg) and pine (Pt) MYBs are represented with full and empty lozenge, respectively. Figure and complete legend are available at <u>http://www.biomedcentral.com/1471-2229/7/17</u>.

Gene annotation	MNID ^a , NCBI ^b	Clone ID	Forward primer (5'-3')	Reverse Primer (5'-3')	Ref. ^c
Phenylalanine ammonia-lyase (PAL)	MN5159896	GQ0045_N21	TGGATTTGCATCCTACTG	TCCATCTTCAACTATAGGAC	-
Trans-cinnamate 4-hydrolase (C4H)	MN5158256	GQ0021b_H10	TCTCACTTAAGTCGGTCCAG	CTTCCTGGTAATGCAAAAAG	0
4-coumarate-CoA ligase (4CL)	MN5255003	GQ0044_C21	CATTCCTCAAAAGCATGAAGAG	ATCGCATCCACAAAGTACAC	-
p-coumarate-3-hydroxylase (C3H)	MN5158265	GQ0012a_A09	ATTGCCTTGCACAATGTAG	CCACTGCCAACTTTTCAC	0
Caffeic acid O-methyltransferase (COMT)	MN5170076	GQ0071_A01	TCATTGGGGCTTGTAATG	ACATGACTTCCAGAGATAAAGC	7
Caffeoyl-CoA O-methyltransferase (CCoAOMT)	MN5158324	GQ0012a_D03	ATTGAGATCAGCCAAATCC	GCGCTCTCCCTATAATCAG	-
Cinnamoyl-CoA reductase (CCR)	MN5259156	GQ00411_J17	CCCTCAGTATCCAATTCCTACC	CTGTTTCATACAGGCATTGTTTC	0
Cinnamyl alcohol dehydrogenase (CAD)	MN5259404	GQ00410_D06	CTGGACTACATCAATACTGC	GATTTACTCATTCTGCACG	-
Pinoresinol-lariciresinol reductase (PLR1)	MN5177685	GQ0044_N02	TCTATCCCGACGTTGAATAC	ATTGAAGCAGACAGATCGTT	0
Pinoresinol-lariciresinol reductase (PLR2)	MN5194637	GQ0047_N19	ACGTCCAATACACTACTGTCAA	GTTCCACGAAGTCAAATCTC	0
Glycosyl transferase 8 (GT8)	MN5178101	GQ0044_010	GGATGAACAGTGAGTAATGTCA	ACTAAGAAGGCAAACTACAAGG	0
3-DAHP synthase (DAHP)	MN5251829	GQ02010_L18	TCTTCATATGGAACATGCAC	AACTTTACATCCCATTGTCG	0
Xyloglucan endo-transglycosylase (XTH)	MN5255573	GQ0195_M16	AAATACACTACTGGGGGGGCTCAT	GAGTGGGAAATCTAACTGAATC	0
Cellulose synthase catalytic subunit (CesA))	MN5175268	GQ0043_K06	TAATCACGTCTGCCAATAAG	AGCTGCATATGACATTTACC	0
Arabinogalactan-like protein (AGP1)	MN5191579	GQ00612_D17	GCGTCCATTGTTTTAATGTAG	TGTATTTATCCCTCTGTCTGC	-
Dihydroflavonol 4-reductase (DFR)	MN5235944	$GQ0206_B08$	TGCAAGAGTGATTAATGTCG	TCAGTACAGCTGGAACAATC	0
Chalcone synthase (CHS)	MN5252787	GQ02011_A08	TTCCAGCAGTTCGGAATCTC	CCTCCACCTGATCCAGAATG	0
Sucrose synthase (SuSy)	MN5194616	GQ0045.B3_D20	ATGAGCAAAGATAGAATCAGG	GATTGGGCTCTTCTCTTTTAC	0
PgMYB1	MN5170681	GQ0072_J06	GATTGTACATTAACCCAGTAA	TAAACCATGTGGTATCTGTTA	-
PgMYB2	No EST	n.a.	TGGGTATTCTAGGTATTTCC	ATTAGGTAAGTATGCAGGG	
PgMYB3	No EST	n.a.	AGATCACGGACCCAGATCAAC	GAGCGAACGACCTCCTTCAG	-
PgMYB4	No EST	n.a.	GCAGTTTGAGTTTGAGTGTG	CTGGAGCATAGATTTGATGA	-
PgMYB8	No EST	n.a.	GGTGGACTCAGTTGTAATAA	GTATCTCACCTATTTACAGATCA	-
PtMYB1	AY356372		AGTGGAATGTCGATTTGG	TCATCGTCATCAAGAAAGG	ŝ
PtMYB8	DQ399057		GCAGACTGAAAGAAAACAACAC	CACACAGAAATCCAGAAGAATC	
Transgene (pMJM)	No EST	n.a.	CAAAATCACCAGTCTCTCT	ACCCTATAAGAACCCTAATTCC	0
EF1-a (100% homolog to AJ1325534)	MN5175100	GQ0041_B09	AACTGGAGAAGGAACCCAAG	AACGACCCAATGGAGGATAC	-

Appendix III-1. Primer sequences used for Real time RT-PCR (QPCR)

^aMNID, EST unique identifier at <u>http://ccgb.umn.edu:8088/ForestTreeDB/;</u> ^bNCBI, Genbank annotation; ^cReference: 1, Bedon *et al.* (2007), or Chapter II; 2, Chapter III; 3, Patzlaff *et al.* (2003a); n.a., not available.

Appendix III-2. Microarray manufacturing and quality control

The 9,503 amplicons spotted on the array represent a low redundancy set of cDNA clones selected among 14,225 consensus sequences (contigs and singletons) derived from large-scale EST sequencing of 13 different libraries. The libraries are among those described in Pavy *et al.* (2005) and are listed below (Table I). Their sequences and annotations may be found and queried in the database http://ccgb.umn.edu:8088/ForestTreeDB/ under the Dir7 EST assembly. Clone selection criteria were as follows: only a single clone was chosen to represent each contig or consensus sequence. Redundant contigs with identity up to 95% over 100 nt were eliminated following analysis by BLAST against other contigs and singletons. A size criterion of less than 2 kb and higher than 200 bp was chosen for clones having 5', and 3' sequence information, respectively. When 5' and 3' reads from one clone spanned different contigs, only one of these contig was included in the gene list. More detailed description of the array design and manufacture will be presented elsewhere.

Clones were robotically rearrayed from daughter glycerol stock 384-well plates into 96-well plates prefilled with 8% glycerol in 2XYT + ampicillin with a Microgrid I microarray printer (Biorobotics, Cambridge, UK), incubated overnight at 37°C, and checked for uniform optical density. Plasmid inserts cloned into pBluescrpt II SK+ were PCR-amplified in a MJ Tetrad PTC-205 thermocycler (Bio-Rad, Hercules, CA, USA) by using 1 µl overnight culture, 0.3 µM M13/pUC forward primer (5'-CCCAGTCACGACGTTGTAAAACG-3'), 0.3 µM M13/pUC reverse primer (5'-AGCGGATAACAATTTCACACAGG-3'), 2 mM MgCl₂, 10mM Tris-HCl, 50 mM KCl, 250 µM dNTPs, 1U AmpliTaq (Roche Diagnostics, NJ,USA), and nuclease-free H₂O (Qiagen, Valencia, CA, USA) to 100 µl. PCR conditions were as follows: 2 min at 95°C denaturation; 35 cycles of 30 sec at 95°C, 45 sec at 59°C, and 4 min at 72°C; and 7 min at 72°C. Amplicon specificity and yield was analyzed by capillary electrophoresis using the HT DNA SE 30 LabChip on Caliper AMS 90 system (Zymark -Caliper Life Sciences, MA, USA). Out of 9,357 clones, there were 6,369 strong single bands (68%), 722 absent (7%), and 2,265 multiple bands (24%). PCR products were robotically cleaned (Qiagen, Valencia, CA, USA) and consolidated into 384-well plates, lyophilized by speed-Vac, and resuspended in 15 µl 3x SSC. All cDNAs (average printing concentration of 180 ng l⁻¹) were printed as double, side-byside spots on EZ-Rays[™] aminosilane slides (Matrix Technologies, Hudson, NH, USA) with the Biorobotics Microgrid II microarray printer (Genomic Solutions, Ann Harbor, Michigan, USA). Microspot 10K quill pins (Biorobotics, Cambridge, UK) in a 48-pin tool were used to deposit 0.5 nl (0.1 ng cDNA) per spot onto the slide. The resulting microarrays have a 4 x 12 subgrid layout with 20 x 21 spots per subarray, each spot having approximate diameter and pitch of 90 μ m and 190 μ m, respectively. A 280-bp green fluorescent protein (GFP) cDNA was amplified from a GFP clone (BD Biosciences, Mountain View, CA, USA) by using the primers (5'-GAAACATTCTTGGACACAAATTGG-3') (5'and GCAGCTGTTACAAACTCAAGAAGG-3'), and printed in the subgrid corners (4 spots per subarray) to assist in placing on the grid.

The 20,160 spots (including controls) were distributed as follows: 9,053 unigenes printed in duplicate (18,106 spots) as well as 2,054 control spots including 1,546 negative and 508 positive controls. Buffer ("Blank well") and Glass ("Empty spot") controls were spotted 221 and 552 times, respectively. Positive controls were as follows: Ambion-1 to -8 (Ambion), NASC-sp1 to -sp10 (Nottigham Arabidopsis Stock Centre), and BRI-ece1 (BRI, Montreal) as spike-in controls; NASC-BT, -BAR, -GFP, -GLOBIN, -GUS, -HPH, -LUC, -NPTII, and pBluescript II KS as exogenous controls; and several conifer gene sequences whose expression was characterized by RT-QPCR including ptMYB1, ptMYB2, ptMYB3, ptMYB5, ptMYB7, ptMYB8. They were distributed in a randomized block design on 4 different locations on the array except for NASC-GFP which was spotted on 28 different locations. Transgenic controls were present as one doublet.

The slides were crosslinked in a UV Stratalinker 2400 (Stratagene, La Jolla, CA, USA) at 300 mJ. All slides were inspected manually to identify any major printing issues. One slide every 20 to 30 slides was hybridized with labeled random 9-mer oligonucleotide (SpotQC, Integrated DNA Technologies, Coraville, IA, USA) or SYBR Green 1 staining (Invitrogen, Carlsbad, CA, USA) and scanned using ScanArray Express scanner (Perkin Elmer, Wellesley, MA, USA). Presence/absence, shape, signal intensity vs. background, diameter and DNA binding site capability were measured for each spot on the slide using files generated by Imagene software (BioDiscovery Inc., El Secundo, CA, USA). Position and description of flagged spots (i.e., spots absent or thought to be unusable during post-hybridization analysis) as well as dust, sub-grid defects and other noticed irregularities were recorded.

Complete description of spruce cDNA libraries is available in ForestTree DB summaries at <u>http://ccgb.umn.edu:8088/ForestTreeDB/</u> (Pavy et *al.*, 2005).

Appendix V-1. Primers sequences generating DNA matrix for qRT-PCR amplification in pine.

	Amplicon	TIGR gene		
Gene name	lenght (nt)	index contig	Forward primers	Reverse primers
PtaHMGR	1070	TC58829	TTAGAGATGGGATGACCAGAGC	AGGAAACTTCCTTTTGGAGGAC
PtaMVAK	1219	TC67868	GATTTGAGAGACGTGAATGTGC	GCCTCCAGACAAAATCAGTACC
PtaFPPS	1094	TC66695	TTTATTTCTTTCCACGCTTTCC	GCAATGTTACTTCTGTCGCTTG
PtaDXPS	1616	TC67545	ACGAGGTATGATCAGTGGTTCC	ATCCAAGGAAAGCTGGTATGAC
PtaDXR	1519	TC74839	ACACACTGTTCCCTACAAGCAC	TCTTGTGGTTTTCCACAAACAG
PtaMECPS	827	TC57337	GGGAGCTCACAGCATATCAATC	CCCAAAGAAGTGAAGAACAAGG
PtaGGPPS	934	TC66873	TAATTCACGATGACTTGCCTTG	GCCAACAATATTTTCCAAGCTC
PtaAFS	805	TC68993	TATGACACTTTCGGAACAATGG	TGTGGGTGTGCCTACTAATGTC
PtaAG1	1049	TC73913	ATCCGGATTATGCATGTTTAGC	AAAGTGGGATGGGATCAATAAG

Primers sequences were designed with Primer 3 software (Rozen and Skaletsky, 2000) and are given in the 5' to 3' orientation (left to righ) (amplification procedure are given in methods). Meaning of abbreviations is given in figure V-10.

Appendix V-2. Primers sequences used for RT-qPCR in pine and spruce.

	Amplicon	AC number, or TIGR gene index contig, or			
Gene name	lenght (nt)	dir8:spruce contig	EST Id	Forward primers	Reverse primers
EF1-a	114	dir8: contig10829	GQ0203_014	AACTGGAGAAGGAACCCAAG	AACGACCCAATGGAGGATAC
CDC2	96	dir8: contig8413	GQ0197_L17	GTGCAGAGAAAAAGTCGAAC	CCACACCATATGTTCCTTCT
Transgene-					
Term-NOS	101	CAMV35S-term.		CAAAAATCACCAGTCTCTCTCT	ACCCTATAAGAACCCTAATTCC
Pinus taeda					
PtMYB5	182			CGGAGTTTGCAGTTAATAGAC	TAGTTGATGATACGGCACTG
PtMYB10a	92			AATTACAGGCATTCCAACTG	CAACGCTCAGATAACTACCC
PtMYB10b	99			TTTTTAAGGGTTTACATGTGC	CTGGTTTCTCTCTGGTTTTG
PtMYB13	112			TCTTTGATCCTCTGTAAATGG	ATGAGACCTCGAACTACGAC
PtMYB14	116	DO399056	NXSI 118 C02	GACAGTCCATAGAACCGTAGAC	ATGAGCCTGTAATGAGAGAAAC
PtMYB15	94			ACCAGAAATTAGTGCTTAGGAG	AATATTACAGCTTAGCCAGAGTG
PtMVB17	120			AATCTACAGTCCGAATAAATGG	CTTGAATTAGGGCATGAGAC
1011017	120				
PtaAACT	108	TC74130	NXCI-033-F04	TCTTTAGTAATTTGGTGTGGTG	CAACTCAAACAATCATCTCTTG
PtaHMGS	116	TC75195	NXCI-002-G02	TAAATCACTTCATTGCACTAGG	AGTTCTGATGAAATGAGAAGTAGAC
PtaMCT	123	TC66924	NXCI-054-A10	CAAATATTAAGGTGACTACACCAG	CTTCTAAGGGAACCTTATAGAGTAAC
PtaMVD1	138	TC59247	NXCI-124-C01	TGGCAATTGTTATGTTAGACC	TCACAAACCGATATCATAACC
PtaHMGR	125	TC58829	appendix V-1	AGGTCAATCAAGGATATCAAAG	CCCTAGCTTGATCAAATAAATC
PtaMVAK	108	TC67868	appendix V-1	TTTTGTCAGCTTTCTCCAAT	TACCATCGCCTTTACAAAAT
PtaFPPS	121	TC66695	appendix V-1	ATTATATGCCAATTATGGGAAG	TGAGCTCCTTATAACTAGTACGC
PtaDXPS	97	TC67545	appendix V-1	CATCCGATTATAAATGTCGAG	TCGGCAACTGTAACATAAATAC
PtoDVD	141	TC74920	appendix V-1	CARCATATTATCCTCCTTCC	CATGATAAAAGCATCCCTTC
PtaMECDS	141	TC 74037	appendix V-1	AGTTGGATTTCCATCTCAATAG	CATGATAAAAGCATCCCTTC
PtaCCPPS	122	TC66972	appendix V-1	GTACATTGCATTCAGACAAAAC	CATGACAACTCAAAGAGAAATG
Pta AFS	132	TC68003	appendix V-1	GACATTOTCAGACOTTTCTACC	TGACATACTTCTCATTCCAC
PtaAFS PtaAC1	144	TC72012	appendix V-1	GACATICIGAGAGETTICIACC	CACCTCTTCTCTCTACATTC
rtaAGI	132	10/3913	appendix v-i	GAGEIGAAITETETETTTEE	GAOCCICITCITCITCITACATIC
Dto DEE	122	TC57193	NVSI 042C12	A ACTTICA ACCGTCTTCTC	ACTTGAACTTCCGTCATCC
PtaDEF	133	TC//105	NXSL 0521104	ACTINGAAGCOIGIIGIG	
rtaAOC	119	1003070	NASI_055H04	AGIAAIGICIIIIIAAIICACIGCI	TAGATICACAAAGICAAGGIACAIA
D' 1					
Picea glauca	1.50	D C 2000 C 0			
PgMYB5	173	DQ399069	G00045 D2 G24	AATICIGGCAGCGAACIG	AATGCTTCGTGGTGGAATC
PgMYB5b	95	D COORDECT	GQ0045.B3_G24	GCGAGGIGAIAIAICGAAIG	CATTAATTIGATTITGCGATG
PgMYB10	133	DQ399064		GCIGIAITITAACAITICAIGG	
PgMYB13	86	DQ399061		AAATTACAGCTAGAGTGAGAGG	AACIIGAACCGIACACGAC
PgMYB14a	153			AGGGATGAAGGGAACCATCG	IGAATTATAGCGGTGGACTC
PgMYB14b	120			AGGGATGAAAGGAACGCG	ATAAGCAATCCATTAATCTTCC
PgMYB15	89			TAACCCATCACIGCTAATCAC	CGGATIACAGGCICITACIAIC
PgMYB16a	116			AAGGATGTCATGTTGTACCTG	AACACATCCAAAACCCTTC
PgMYB16b	94			CATAAGIGGCICAATAAGIGATAC	ACAGITCATAGAACAGITCAACC
PgMYB16c	111			GATATICAACATACGGCAGAG	TACACICCIGCAACAAAATTAC
PgHMGS	127	TC27328	GQ0024d_G04	TITAAGTAATGTGGAAGCATTG	GIGATITAATTICTICCCACIC
PgMVDI	113	TC13436	GQ0192_M16	AGATAAICCAIGAIAICGGIIC	
PgMECPS	125	1C508/	GQ0081_N19	AATIGCAAAATATCTCACTGC	GAATTAAACCICAATCAGCIIC
PgDXR	115	109956-1015767	GQ0163_P05	AICAACCAICAIIGGAAGAG	CITIGGIGIGCIAATICCIC
PgFPPS	117	TC21220	GQ0172_F22		CCATIGIIGIAAACACATICC
PgAACT	112	1013586	GQ0196_P02	TATAATAGTCGCTGTCATTIGG	AAACAATCATCTCTTGAACCTC
PgHMGR	138	TC41468	GQ0197_N01	IGATGACAGCCTCACATACTAC	GITTTAGATTICCICCAGAATG
PgAG1	104	1C10417-1C39627	GQ02010_L20	IGATATICAAGGATGGAGATG	AACIGATATCACATCACAGIGG
PgAFS	137	dir8:Contig11284	GQ0207_009	TICTAACCCTAACATCTCGTG	TGGGAAGTAGAGTTTATTGACC
PgMVAK	138	TC22986	GQ0207_P08	TATCGAGCATATCTTCCCTATC	ATTAACIGCATCAAAGACIGC
PgMCT	115	TC14243	GQ02790_E10	CTIGCAGAAAGAATTITAAACC	TAAACAAGCCTTGCACTTAAC
PgGGPPS	102	TC37669	GQ0134.B7_N09	ATAAGGACTGGATTTAGGTACAAG	CICIGGCAACAGAGIAACAIC
PgDXPS	108	TC12367	GQ0206.TB_024	CTAATAGGCAAGCCTCAGAA	AGGCTTTGCTTCTCTACCTT
PgDEF	169	TC32588	GQ0131B10	CTACTACTGTGATCTTCTCTGGTTT	GTAAGTAAAAGATAACCTGAACCAC
PgAOC	104	TC15161	GQ02777_G11	AAGCTGTACAGTGGCAATCT	GATTGCTTCATAACGGTCTC
PgAExp3	126	dir8:Contig8739	GQ0046.B3_C02	ATGGTGTACTGGGCTTATTG	TAGTATGTACAAAGCGACTTGC
PgPR4b	98	dir8:Contig3620	GQ0198.B3_D11	GAAGAAATATGGCTGGACTG	CTCCAGTGCCTCTGTTAGTC
PgCHIT	104	dir8:Contig5604	GQ0204.B3_D23	GATATGGAATGGTGACCAAC	GCAGAATGTCTGGTAGAAGC
PgGT	100	dir8:Contig4884	GQ0207_M01	AGATCCCACAATAGTGAAAAAC	GATGATAAAGAGTGCGGAAG
PgCP450	113	dir8:Contig3475	GQ02011_A16	GAGAAGTTCGGAGTTACTATGC	TCCTATTACTTGGACACAACATAC
PgCHS8	85	dir8:Contig10736	GQ02011_A08	TTCCAGCAGTTCGGAATCTC	CCTCCACCTGATCCAGAATG

Primers sequences were designed with Primer 3 software (Rozen and Skaletsky, 2000) and are given in 5' to 3' orientation (left to right) (amplification procedures are given in methods). Blanks indicate that contig or EST are not available or have not been identified. Contig from TIGR are available at the web site http://compbio.dfci.harvard.edu/tgi/plant.html and spruce dir8: contig at the ForestTreedb web site http://ccgb.umn.edu:8088/ForestTreeDB/ nimbus/project.do?project=ForestTreeDB. Meaning of abbreviations is given in figure V-10 and in appendix V-3.

Appendix V-3. Expression level (RT-qPCR) of *PtMYB14*, genes validating microarray and *R2R3-MYB* genes in hypocotyls of PtMYB14-OE plantlets, normalized with *CDC2* transcripts.



Pg653: spruce wild-type plantlets used for microarrays analyses, 7-3: spruce plantlets transformed with empty vector. Spruce plantlets with construct $Pro_{UBI}PtMYB14$ are lines 9-16 and 9-18 (grey bars), $Pro_{CAD}PtMYB14$ are lines 30-10 and 30-14 (bars with squares inside). CHIT, *chitinase*; PR4b, *pathogenesis related 4b*; AOC, *allene oxide cyclase*; 4CL, 4-coumarate CoA ligase; CAD, cinnamyl alcohol dehydrogenase; CHS8, chalcone synthase; GT, glucosyl-transferase; CP450, cytochrome P450; EXP, alpha expansin. Meaning of others abbreviations is given in figure V-10. In each graph, asterids indicate a significant difference between Pg653 wild-type and transgenics lines according to Student's test at 0.05 (*), 0.01 (**), or 0.001 (***).

Appendix V-4. Expression level (RT-qPCR) of *PtMYB14*, genes validating microarray and *R2R3-MYB* genes in roots of PtMYB14-OE plantlets, normalized with *CDC2* transcripts.



Legend is the same as in appendix V-2. One additional spruce transgenic line is added for *Pro_{UBI}PtMYB14* (line 9-3) and *Pro_{CAD}PtMYB14* (line 30-5).

						PT	В	Е	ā	ت ت
						0,01	0,	02	0,0	1
MN_id	BlastX_E	BlastX_Hit_Descrip	Pfam_p	Pfam_Hit_Descrip	Ratio	P-value	Ratio	P-value	Ratio	P-value
transport										
MN5237792	8,09E-63	Aquaporin	0	MIP; Major intrinsic protein					1.51	4,73E-03
MN5260562	4,68E-26	MATE efflux family protein	5,8E-10	MatE; MatE			1,34	1,24E-02		
MN5249302	3,63E-17	Putative protease inhibitor	3,3E-14	Tryp_alpha_amyl; Protease inhibitor/seed storage/LTP			-1,31	1,53E-02		
MN5158210	1,45E-41	Vacuolar proton-pumping PPase	1,2E-94	H PPase; Inorganic H+ pyrophosphatase	1,28	3,07E-03				
MN5255596	3,55E-118	Vacuolar ATP synthase catalytic subunit A	7,9E-43	ATP-synt_ab; ATP synthase alpha/beta family	1,27	9,68E-03				
MN5239669	4,40E-62	ADP-ribosylation factor 1	0	Arf; ADP-ribosylation factor family	0,79	9,28E-03				
MN5233224	8,49E-70	Alpha-soluble NSF attachment protein 2	0,00041	TPR_1; Tetratricopeptide repeat	-1,17	7,21E-03				
Proceene acco	ntiás à l'ADN	a st l'A BN								
MN5159358	3.75E-46	Bifunctional nuclease	2E-115	S1-P1 nuclease: S1/P1 Nuclease					1.94	2.35E-03
MN5171247	3,88E-16	RNase	2.9E-25	Ribonuclease T2; Ribonuclease T2 family			2,44	1.28E-02	1.90	4,74E-04
MN5159570	8,39E-58	Bifunctional nuclease	0	S1-P1 nuclease; S1/P1 Nuclease	1.59	1,45E-03	2,73	6,45E-03	1,79	2,35E-03
MN5244157	0		0,00202	mRNA_cap_enzyme; mRNA capping enzyme	<u>1,63</u>	1,62E-06			1.57	1,54E-03
MN5182493	8,23E-12	Putative AP2-related transcription factor	3,4E-39	AP2; AP2 domain			-1,48	1,53E-02		
MN5233918	0	Putative RNA helicase	7,9E-60	DEAD; DEAD/DEAH box helicase	1.77	1,51E-04				
MN5161487	6,86E-26	Arabidopsis thaliana genomic DNA clone:MCL19	4,7E-22	KH_1; KH domain	1,34	4,52E-03				
MN5196531	4,51E-15	Emb CAB41546.1	0,00141	ICL; Isocitrate lyase family	1,32	8,50E-03				
MN5164568	9,08E-58	Kelch repeat-containing F-box protein-like	8,9E-12	Kelch_1; Kelch motif	1,27	5,06E-03				
MN5257677	1,48E-09	THUMPD3 protein	9,1E-07	UPF0020; Putative RNA methylase family UPF0020	1,25	9,69E-04				
MN5241811	1,20E-39	Golden 2-like protein 1	2,9E-10	Myb_DNA-binding; Myb-like DNA-binding domain	1,23	8,49E-03				
MN5173059	6,93E-10	bZIP family transcription factor	2,4E-07	bZIP_1; bZIP transcription factor	1,22	5,06E-03				
MN5246266	0,00099346	Hypothetical protein	0,00637	ATX_III; Anemonia sulcata toxin III family	-1,31	3,07E-03				
Métabolisme g	général									
MN5237144	2,04E-34	GASA5-like protein	7,9E-83	GASA; Gibberellin regulated protein					1,49	9,03E-03
MN5236822	6,02E-100	Alpha galactosidase	3,5E-11	Melibiase; Melibiase					1,33	8,44E-03
MN5237079	2,18E-57	putative formamidase	5,5E-85	FmdA_AmdA; Acetamidase/Formamidase family					-1.53	1,36E-03
MN5234789	3,00E-101	NAD-dependent malic enzyme 62 kDa isoform	4E-111	Malic_M; Malic enzyme, NAD binding domain	1,32	7,73E-03	1,60	1,26E-02		
MN5195181	1,57E-88	Malate dehydrogenase	4,4E-75	Ldh_1_N; lactate/malate dehydrogenase			1,35	1,32E-02		
MN5177622	2,90E-52	Lysosomal Pro-X carboxypeptidase	3,5E-22	Peptidase_S28; Serine carboxypeptidase S28			1,54	1,89E-02		
MN5194333	1,78E-12	Putative myosin heavy chain	0,00128	UPAR_LY6; u-PAR/Ly-6 domain			1,39	1,24E-02		
MN5235739	1,62E-68	Proteasome subunit beta type 1	7,9E-38	Proteasome; Proteasome A-type and B-type			1,31	1,28E-02		
MN5254907	2,54E-64	Putative ribosomal protein L10a	1,7E-78	Ribosomal_L1; Ribosomal protein L1 p/L10e family			-1,38	1,53E-02		
MN5234024	0,00438429	NADH dehydrogenase subunit 4	0,00103	SlyX; SlyX			-1,67	1,24E-02		
MN5182319	6,21E-48	Zinc finger protein	2,5E-15	zf-CHY; CHY zinc finger	2.28	1,25E-08				
MN5254928	7,60E-74	Adenylosuccinate lyase	6,8E-42	Lyase_1; Lyase	1.79	1,14E-06				
MN5255831	6.45E-69	Putative NADPH auinone oxidoreductase	2.1E-29	ADH zinc N: Zinc-binding dehvdrogenase	1.45	1.54E-03			-	

Annexe VI-1. Liste des gènes dérégulés dans les expériences PT, BE et BC chez des épinettes blanches de troisième cycle de croissance exprimant la construction *Pro*_{CAD}*PtMYB14*.

PT, pousse terminale ; BE, bois écorcé ; BC, bois de compression. Les « P-value » associées aux gènes sont <0.01 pour PT et BC, et <0.02 pour BE. Les valeurs de ratio (entre transgéniques et témoins) soulignées sont <-1,5 ou > 1,5.

					ļ	ľ			ľ	
MN5177669	2,86E-88	putative receptor protein kinase tmk1 precursor	4,6E-31	Pkinase_Tyr; Protein tyrosine kinase	1,38	3,08E-03				
MN5239084	2,16E-27	Putative beta-mannosidase	0,01085	SlyX; SlyX	1,33	2,76E-03				
MN5235316	7,35E-113	Pectate lyase	3E-99	Pec_lyase_C; Pectate lyase	1,33	6,10E-03				
MN5237094	1,95E-102	Putative O-linked GlcNAc transferase	0,00257	MF_alpha; Yeast mating factor alpha hormone	1,31	9,53E-03				
MN5235051	2,27E-83	Nucleoid chloroplast DNA-binding protein-like	1,3E-07	Asp; Eukaryotic aspartyl protease	1,27	5,72E-03				
MN5175181	3,23E-44	Similar to A. thaliana protein BAB01483 similar	0,00014	TPR_1; Tetratricopeptide repeat	1,26	8,49E-03				
MN5162862	2,86E-22	Putative PGPD14 protein	0,00227	zf-C3HC4; Zinc finger, C3HC4 type (RING finger)	1,25	8,49E-03				
MN5159506	9,68E-101	Putative polygalacturonase	2,7E-25	Glyco_hydro_28; Glycosyl hydrolases family 28	1,25	4,34E-03				
MN5259929	2,08E-125	Cullin 1C	1E-121	Cullin; Cullin family	1,24	9,28E-03				
MN5160251	3,48E-44	Putative serine/threonine-specific protein kinase	8,1E-10	Pkinase; Protein kinase domain	1,24	9,51E-03				
MN5160228	3,44E-39	Putative UDP-glucuronate decarboxylase 3	1,3E-14	Epimerase; NAD dependent epimerase/dehydratase	1,23	8,39E-03	┢			
MN5246469	6,81E-92	Proteinase-like protein	2,5E-15	Peptidase_S8; Subtilase family	1,23	8,59E-03				
MN5191320	2,90E-64	Putative growth-on protein GRO10	1E-111	Peptidase_A22B; Signal peptide peptidase	1,22	9,28E-03				
MN5241096	0	Putative aspartate transaminase	1E-128	Aminotran_1_2; Aminotransferase class I and II	1,22	8,50E-03				
MN5171264	8,34E-65	Arm repeat containing protein homolog	2,2E-37	Arm; Armadillo/beta-catenin-like repeat	1,20	9,87E-03				
MN5235225	1,87E-29	Chloroplast nucleoid DNA binding protein-like	0,00152	WCCH; WCCH motif	1,20	8,38E-03				
MN5246036	2,25E-11	Ubiquitin ligase protein CIP8	5,1E-12	zf-C3HC4; Zinc finger, C3HC4 type (RING finger)	1,20	9,53E-03				
MN5191750	1,49E-116	Glyceraldehyde-phosphate dehydrogenase precursor	1E-119	Gp_dh_N; Glyceraldehyde 3-phosphate dehydrogenase	1,18	8,49E-03				
MN5173029	1,01E-44	T23G18.14	5E-06	Glyco_transf_29; Glycosyltransferase family 29	1,17	8,49E-03				
MN5239920	2,66E-72	OSJNBb0059K02.14 protein	5E-115	Asparaginase_2; Asparaginase	-1,27	2,29E-03				
MN5252889	4,02E-63	Putative ribosomal protein	3,1E-38	Ribosomal_S4; Ribosomal protein S4/S9	-1,30	8,50E-03			┝	
MN5164572	1,98E-55	Cytoplasmic ribosomal protein L18	4E-130	Ribosomal_L18e; Eukaryotic ribosomal protein L18	-1,31	6,10E-03				
MN5242126	9,61E-30	40S ribosomal protein S20	1E-56	Ribosomal_S10; Ribosomal protein S10p/S20e	-1,32	8,59E-03				
MN5181574	3,47E-31	Putative ubiquitin-conjugating enzyme	1,5E-57	UQ_con; Ubiquitin-conjugating enzyme	-1,34	3,07E-03				
MN5249880	4,25E-44	Putative ubiquitin conjugating enzyme	1,1E-76	UQ_con; Ubiquitin-conjugating enzyme	-1,35	8,38E-03				
MN5241101	1,96E-16	T2D23.3 protein	1,3E-14	HMA; Heavy-metal-associated domain	-1,35	4,61E-03				
MN5253965	3,40E-42	Putative phosphatase	1E-06	Hydrolase; haloacid dehalogenase-like hydrolase	-1,40	8,78E-04				
MN5255738	2,86E-24	ORFX	1,4E-61	PLAC8; PLAC8 family	-1,68	6,29E-03				
Métabolisme de	ss lipides									
MN5244783	4,71E-07	Catechol O-methyltransferase	0,04727	Antimicrobial_1; Frog antimicrobial peptide			1,65	1,40E-02		
MN5182776	7,90E-06	Alpha-carboxyltransferase aCT-1 precursor	0,04967	ACCA; Acetyl co-enzyme A carboxyltransferase	1,44	3,57E-05				
MN5239679	1,33E-54	Phospholipase D alpha 2 precursor	4,7E-11	PLDc; Phospholipase D Active site motif	1,39	4,34E-03				
MN5236213	4,26E-28	PSR9	0,00256	Antimicrobial_1; Frog antimicrobial peptide	1,35	1,34E-03				
MN5250504	4,25E-50	fatty acid beta- oxidation	1,6E-05	3HCDH; 3-hydroxyacyl-CoA dehydrogenase	1,21	9,87E-03				
MN5182379	2,03E-07	F1C9.18 protein	2,1E-14	GDPD; Glycerophosphoryl diester phosphodiesterase	-1,29	2,78E-03				
MN5254253	1,33E-58	Putative gamma-tocopherol methyltransferase	1,2E-07	CMAS; Cyclopropane-fatty-acyl-phospholipid synthase	-1,35	3,61E-03				
MN5158489	1,44E-30	Platelet-activating factor acetylhydrolase	0,01687	DUF1333; Protein of unknown function (DUF1333)	-1.77	7,28E-05				

Annexe VI-1. (suite)

Processus de c	létoxification	l (oxydoréduction)								
MN5159882	2,28E-40	F19G10.5 protein	1,2E-05	Cu-oxidase_2; Multicopper oxidase					-1,50	2,35E-03
MN5177619	9,27E-48	ESTs D22477	7,5E-08	Thioredoxin; Thioredoxin					-2.13	2,35E-03
MN5171335	1,42E-58	Putative quinone-oxidoreductase QR2	5,7E-14	Flavodoxin_1; Flavodoxin			1,62	1,24E-02		
MN5161625	9,24E-13	Putative mandelonitrile lyase	3,6E-18	GMC_oxred_C; GMC oxidoreductase			1.54	1,49E-02		
MN5241957	1,63E-66	6-phosphogluconate dehydrogenase,	0	6PGD; 6-phosphogluconate dehydrogenase	1,28	1,51E-04	1,52	1,28E-02		
MN5249107	6,33E-50	Peroxidase	3,5E-48	peroxidase; Peroxidase			1,33	1,28E-02		
MN5164520	6,02E-93	OSJNBa0019K04.9 protein	2,9E-27	PA; PA domain	1,24	8,49E-03				
MN5159512	1,67E-128	NADPH-cytochrome P450 reductase	8E-104	FAD_binding_1; FAD binding domain	1,21	9,13E-03				
MN5161834	6,58E-44	peroxisomal membrane protein-related	4E-06	Mpv17_PMP22; Mpv17 / PMP22 family	1,18	8,38E-03				
MN5254070	7,36E-29	Thioredoxin H-type	5,4E-36	Thioredoxin; Thioredoxin	-1,41	3,07E-03				
MN5164432	1,50E-18	Chemocyanin precursor	3,8E-34	Cu_bind_like; Plastocyanin-like domain	-1,68	9,40E-03				
MN5249216	1,89E-53	Cu-Zn-superoxide dismutase	6E-104	Sod_Cu; Copper/zinc superoxide dismutase (SODC)	-1,70	8,24E-03				
MN5249435	4,77E-44	Superoxide dismutase [Cu-Zn]	2E-111	Sod_Cu; Copper/zinc superoxide dismutase (SODC)	-1,90	3,57E-05				Í
Processus de 1	éponse aux s	tress et mécanismes de défense								
MN5248924	5,89E-11	Water deficit inducible protein LP3	8,5E-32	ABA_WDS; ABA/WDS induced protein			<u>2,27</u>	1,37E-02	1.85	8,23E-04
MN5176543	9,50E-12	Disease resistance response protein-like	6,4E-34	Dirigent; Dirigent-like protein					1,46	3,20E-03
MN5237498	7,15E-47	Putative acid phosphatase	1,1E-79	Acid_phosphat_B; HAD superfamily, subfamily IIIB			1,38	1,24E-02		
MN5235024	1,46E-25	TIR/P-loop/LRR	0,01428	UPF0060; Uncharacterized BCR			-1,37	1,82E-02		
MN5236305	0		0,00054	Defensin_2; Arthropod defensin			-1,37	1,26E-02		
MN5241707	7,76E-08	Disease resistance protein, putative	0,00368	A_amylase_inhib; Alpha amylase inhibitor	-1,68	8,49E-03	-1,43	1,28E-02		
MN5250687	1,39E-42	Mucin-like protein	1,4E-70	Str_synth; Strictosidine synthase	1,25	1,54E-03			1,41	5,17E-03
MN5237642	3,26E-11	NBS	0,05546	HC2; Histone H1-like nucleoprotein HC2	1,49	5,85E-05				
MN5251928	5,18E-38	Arabidopsis thaliana genomic DNA, clone: MVE11	0,00541	Defensin_2; Arthropod defensin	-1,16	8,49E-03				
MN5243402	2,66E-20	Putative ER6 protein	3,4E-18	Usp; Universal stress protein family	-1,19	7,38E-03				
MN5168599	2,96E-37	Abscisic acid-induced protein-like	8,5E-44	TB2_DP1_HVA22; TB2/DP1, HVA22 family	-1,27	3,97E-03				
MN5241301	3,88E-25	Putative ER6 protein	6,7E-22	Usp; Universal stress protein family	-1,31	4,01E-03				
MN5190704	8,15E-40	Heat shock protein 17.0	1,3E-44	HSP20; Hsp20/alpha crystallin family	-1,32	1,45E-03				
MN5249095	4,31E-31	Gb AAF26101.1	4,2E-10	Usp; Universal stress protein family	-1,35	8,52E-03				
MN5181585	0,00133613	17.7 kDa heat shock protein	1,2E-10	HSP20; Hsp20/alpha crystallin family	-1,41	9,69E-04				
MN5241580	4,58E-31	HVA22-like protein f	1, 1E-43	TB2_DP1_HVA22; TB2/DP1, HVA22 family	-1,41	3,07E-03				
MN5255771	1,02E-16	Putative gamma-thionin protein precursor	5E-33	Gamma-thionin; Gamma-thionin family	-1,42	2,31E-03				
MN5239343	3,31E-37	Beta-1,3-glucanase	5E-114	Glyco_hydro_17; Glycosyl hydrolases family 17	-1.51	8,49E-03				
MN5236092	6,74E-31	18.2 kDa class I heat shock protein	9,8E-50	HSP20; Hsp20/alpha crystallin family	-2,15	4,14E-06				
Métabolismes	des composé	s aromatiques et gènes reliés à la paroi végétale								
MN5241484	4,52E-59	Xyloglucan endo-transglycosylase	1,5E-86	Glyco_hydro_16; Glycosyl hydrolases family 16					1.78	8,23E-04
MN5170468	7,74E-14	Putative caffeic acid methyl transferase	6,9E-29	Methyltransf_2; O-methyltransferase	1.54	4,35E-03			1,36	1,86E-03
MN5193939	3,17E-96	Phenylcoumaran benzylic ether reductase	0	NmrA; NmrA-like family			1,88	1,89E-02		
MN5239070	4,74E-48	Putative prephenate dehydratase	6E-53	PDT; Prephenate dehydratase	1,29	1,61E-03				
MN5181530	3,25E-29	Caffeoyl-CoA O-methyltransferase	9,1E-66	Methyltransf_3; O-methyltransferase	-1,37	4,35E-03				
MN5177476	1,25E-60	Cycloartenol synthase	0,00742	AKAP_110; A-kinase anchor protein 110 kDa	1,25	9,13E-03				

Annexe VI-1. (suite)

MN5252038	2,41E-14	UDP-glucose glucosyltransferase	7,7E-16	UDPGT; UDP-glucoronosyl, UDP-glucosyl transferase	1,18	4,37E-03				
MN5193913	3,37E-77	Putative UDP-glucose glucosyltransferase	1,4E-23	UDPGT; UDP-glucoronosyl, UDP-glucosyl transferase	_	-	1,43	1,28E-02		
MN5260497	1,98E-58	1,4-benzoquinone reductase-like	1,5E-16	Flavodoxin_1; Flavodoxin			1.50	1,24E-02	1,40	1,54E-03
Messagers sec	ondaires									
MN5193322	4,95E-18	Putative annexin	1,6E-44	Annexin; Annexin					1,46	8,23E-04
MN5193086	4,76E-22	F18B13.6 protein	0,00133	7tm_5; 7TM chemoreceptor					1,44	1,34E-03
MN5253991	2,17E-29	Calreticulin 2 precursor	5E-54	Calreticulin; Calreticulin family			1,26	1,65E-02		
MN5191637	7,82E-23	Putative nodulin-like protein	0,01922	Ion_trans; Ion transport protein	1,37	5,57E-03				
MN5181763	1,00E-25	Nodulin-related protein-like	2E-06	MFS_1; Major Facilitator Superfamily	1,27	4,37E-03				Ì
MN5196728	1,33E-58	F20N2.14	3,5E-98	Raffinose_syn; Raffinose synthase or seed imbibition	1,22	8,66E-03				
MN5244669	4,10E-34	Putative vesicle-associated membrane protein	1,7E-39	Motile_Sperm; MSP (Major sperm protein) domain	-1,22	3,95E-03				
MN5196298	2,11E-71	L-galactose-1-phosphate phosphatase	2,8E-98	Inositol_P; Inositol monophosphatase family	-1,31	8,24E-03				
MN5253594	9,95E-19	Similarity to nodulin	7E-46	Cu_bind_like; Plastocyanin-like domain	-1,50	1,87E-03				
MN5232405	5,68E-09	Zinc finger and C2 domain protein-like	2,4E-26	C2; C2 domain	-1,68	2,15E-03				
Fonctions imp	récises									
MN5182745	0		0,00323	zf-C3HC4; Zinc finger, C3HC4 type (RING finger)			1.82	1,28E-02	<u>1,69</u>	8,99E-04
MN5182640	0		0,01981	BC10; Bladder cancer-related protein BC10	1,39	1,70E-04			<u>1,66</u>	4,49E-03
MN5193650	3,01E-80	Gb AAD25141.1	0	Phi_1; Phosphate-induced protein 1 conserved region					1,42	8,23E-04
MN5233833	1,74E-114	Pectate lyase	5,3E-96	Pec lyase C; Pectate lyase					1,34	2,87E-03
MN5232301	0		0,00355	Fokl N; Restriction endonuclease Fokl					-1,34	2,45E-03
MN5159343	0		0,06188	TIMP; Tissue inhibitor of metalloproteinase					-1,47	6,93E-03
MN5237624	0		0,00064	Ten_N; Teneurin Intracellular Region					-1,57	4,89E-03
MN5175212	0		0,00619	TNFR_c6; TNFR/NGFR cysteine-rich region					-1,69	2,45E-03
MN5170978	0		4,8E-05	Noggin; Noggin			1.55	1,28E-02		
MN5174245	0		0,00576	Antimicrobial_1; Frog antimicrobial peptide	1,22	8,38E-03	-1,29	1,28E-02		
MN5250669	0		0,0092	Hemagglutinin; Hemagglutinin			-1,43	1,28E-02		
MN5238948	0		0,01181	ATX_III; Anemonia sulcata toxin III family			-1,44	1,28E-02		
MN5244270	0		0,01138	Peroxin-13_N; Peroxin 13, N-terminal	-1,40	6,83E-03	-1.58	1,24E-02		
MN5177570	0		0,02804	ATX_III; Anemonia sulcata toxin III family	1,67	5,85E-05				
MN5207692	0		7,5E-05	Arteri_env; Arterivirus envelope protein	1.52	4,39E-04				
MN5234405	0,00444843	F4P13.22 protein	6,2E-05	DapB_N; Dihydrodipicolinate reductase, N-terminus	1.51	3,71E-03				
MN5232404	0		0,0643	LEF-8; Late expression factor 8 (LEF-8)	1,40	4,16E-05				
MN5191840	1,54E-07	Putative diaphanous 1	6E-06	FH2; Formin Homology 2 Domain	1,32	6,33E-03				
MN5175369	0		0,00365	SlyX; SlyX	1,32	1,54E-03				
MN5196215	0		3,6E-05	Noggin; Noggin	1,31	1,45E-03				
MN5239168	0		0,03799	Activin_recp; Activin types I and II receptor domain	1,26	4,63E-03				
MN5161827	0		2,7E-08	Vitellogenin_N; Lipoprotein amino terminal region	1,26	2,78E-03				
MN5194979	0		8,4E-05	7tm_5; 7TM chemoreceptor	1,22	3,07E-03				
MN5251444	0		0,00029	Tub; Tub family	1,20	4,74E-03				

Annexe VI-1. (suite)

																									1,40E-02	1,33E-02
																									-1.58	-1,90
4,37E-03	1,54E-03	4,37E-03	9,28E-03	8,99E-03	4,34E-03	6,10E-03	8,50E-03	4,40E-03	5,87E-04	3,20E-03	1,51E-04		1,43E-04	3,20E-03	3,07E-03	8,99E-03		6,10E-03	5,03E-03	8,99E-03	8,38E-03	4,70E-03	8,91E-03	4,65E-03		
-1,21	-1,23	-1,29	-1,30	-1,30	-1,32	-1,33	-1,45	-1.60	-1,63	-1.74	-1,77		2,09	1,61	1,43	1,41		1,33	1,31	1,23	1,18	1,18	-1,19	-1,37		
298 ANATO; Anaphylotoxin-like domain	946 Metallothio_11; Metallothionein family 11	10 CHCH; CHCH domain	1516 His_leader; Histidine operon leader peptide	078 NOA36; NOA36 protein	19 FARP; FMRFamide related peptide family	853 Ribosomal_L27e; Ribosomal L27e protein family		615 Herpes LAMP2; Herpesvirus Latent membrane protein 2	221 MF alpha; Yeast mating factor alpha hormone	371 Pro-NT_NN; Neurotensin/neuromedin N precursor	088 Tfb2; Transcription factor Tfb2		709 BURP; BURP domain	3-74 BURP; BURP domain	3-87 BURP; BURP domain	3-91 BURP; BURP domain		DUF1221; Protein of unknown function (DUF1221)	55 DUF810; Protein of unknown function (DUF810)	3-11 DUF604; Protein of unknown function, DUF604	3-57 DUF789; Protein of unknown function (DUF789)	3-14 DUF581; Protein of unknown function (DUF581)	1517 DUF1602; Protein of unknown function (DUF1602)	(633 DUF26; Domain of unknown function DUF26	3-33 DUF584; Protein of unknown function, DUF584	766 DUF1455; Protein of unknown function (DUF1455)
000	0,0	Emb CAB89373.1 6E-	0'0	000	Hypothetical protein OSJNBb0071021.10 [0,00	0,0		0,0	000	Glycine-rich protein homolog 0,03	0'0		0'0	Dehydration-responsive protein RD22 [1,3]	BURP domain-containing protein 2,41	BURP domain-containing protein 3,51		Plasma membrane H+-ATPase 0,00	expressed protein 1E-	Arabidopsis thaliana genomic DNA, clone:MYC6 [3,5]	AT4g03420/F9H3_4 [2,8]	3,61	unnamed protein product [0,00	0'0	F8K4.12 [1,8]	000
0	0	1,06E-13	0	0	9,52E-06	0	0	0	0	3,99E-10	0	pe BURP	0	4,00E-27	5,55E-44	1,25E-38	pe DUF	1,83E-53	6,22E-26	1,37E-08	2,01E-17	0	4,12E-12	0	8,81E-07	0
MN5245834	MN5171334	MN5244364	MN5169684	MN5162947	MN5231979	MN5182717	MN5164231	MN5246367	MN5256234	MN5242914	MN5175638	Protéines de tyl	MN5163000	MN5163790	MN5238980	MN5232943	Protéines de tyl	MN5249391	MN5193927	MN5158932	MN5236184	MN5259305	MN5164963	MN5237883	MN5249136	MN5259803

Annexe VI-1. (suite et fin)

Annexe VII-1. Longueurs et pourcentages d'identités de séquences des introns de quelques *MYB-R2R3* d'épinettes et de pin.

	Intron I longueur (nt)	Intron II longueur (nt)
PtMYB1	313	106
PgMYB1	83	101
PtMYB3	non séquencé	248
PgMYB3	1427	202
PtMYB4	257	259
PgMYB4	194	267
PtMYB8	115	96
PgMYB8	97	90
PtMYB14	94	293
PgMYB14a	133	444
PgMYB14b	133	448

Pourcentages d'identité

	Intron I	Intron II
PgMYB1 / PtMYB1	20,38%	69,56%
PgMYB3 / PtMYB3	non séquencé	67,33%
PgMYB4 / PtMYB4	59,53%	84,13%
PgMYB8 / PtMYB8	73,91%	79,38%
PgMYB14a / PtMYB14	59,39%	47,13%
PgMYB14b / PtMYB14	57,89%	47,79
PgMYB14a/PgMYB14b	97,74%	96,87%

Les pourcentages d'identités de séquences nucléotidiques ont été calculés avec la fonction « pairwise optimal global alignment » (Blosum62 matrix) du logiciel Bioedit 6.0.7. Tous les introns identifiés ici sont localisés dans le domaine MYB excepté pour les gènes Pg et PtMYB3 concernant les seconds introns présent en carboxy terminal.







Les carrés en pointillés montrent les nucléotidiques conservés entre les différents membres *MYB* et au sein du membre *PgMYB5*. L'alignement a été obtenu avec ClustlalW puis édité à la main. I : premier intron des gènes *MYB*. I1 et I2 correspondent respectivement à la première et à la deuxième partie du premier intron de *PgMYB5*. Les nucléotides conservés à plus de 50% entre les séquences sont en noir (BoxShade 3.21). Les chiffres indiquent la position du nucléotide dans l'intron. Pg : *Picea glauca*, Pt : *Pinus taeda*.

Annexe VII-3. Alignement de séquences des seconds introns de six membres du groupe Sg4-C



Les sections A, B et C montrent la conservation d'une séquences nucléotidiques de 62 nucléotides répétée cinq fois dans les seconds introns de PgMYB14a et 14b. Les nucléotides conservés à plus de 60% entre les séquences sont en noir (BoxShade 3.21). L'alignement a été obtenu avec ClustlalW puis édité à la main. II : second intron des gènes MYB. II1 et II2 correspondent respectivement à la première et à la deuxième partie du premier intron de PgMYB5. Voir légende annexe VII-2.

Annexe VII-4. Alignements de séquences des introns comparant *Pt/PgMYB1*, *3*, *4 et 8*.

PtMYB1 et PgMYB1

Comparaison de	es pi	remiers introns entre <i>Pt</i> et <i>PgMYB1</i>
INTRON-I-PtMYB1	1	GTAAATAAAA ATGTCATTTTCTCCACCTGGTTTCGATTCTCAAGTTATGACATGTCTGCT
INTRON-I-PgMYB1	1	GTAAATAAAA
consensus	1	*****
INTRON-I- <i>PtMYB1</i> INTRON-I- <i>PgMYB1</i> consensus	61 11 61	CCCTGAAAGCCTCTTGTAATGTCGACTAATAAAAAGATTAGTCCCTAGATTCCTTCC
INTRON-I-PtMYB1	121	ATAGCAGGAGCAGGACCGTCTGTGTGTGTGTGTGTTTTTTTAGATAAAAA TGCCATTGCAT
INTRON-I-PgMYB1	11	TGCCGTTGCAT
consensus	121	**** *****
INTRON-I- <i>PtMYB1</i> INTRON-I- <i>PgMYB1</i>	181 22	TGCGTGGAGATTCATGCAATTTGGAGTCTCTGTAAATTTCCCCATTGGTTTATAATTGCAG
INTRON-I- <i>PtMYB1</i> INTRON-I- <i>PgMYB1</i> consensus	241 26 241	TGGAGATTCATGCAATTTGGAGTCTCTGTAAATTTCCCATTGGTTTATAATTGCATTTGG TGGCCATACAGTCTCTGTAAATCCCATTTGTTTAATTGCATTTGG *** ************************************
INTRON-I-PtMYB1	301	ATAAATTTTGCAG
INTRON-I-PgMYB1	71	ATGAATTTTGCAG
consensus	301	** *******
Comparaison de	es se	econds introns entre <i>Pt</i> et <i>PgMYB1</i>
INTRON-II-PtMYB1		1 GTATATATAATGGAGCTGTAACTGTAAATTAATGGCGGTCATTGTGTATGATGTGTTGGT
INTRON-II-PgMYB1		1 GTATATATAATCGAGACCTCAT-AATGGTTCTGCGTGATATGTTGTT
consensus	-	L ***** *** * ** * ** ** ** ** ** ** **
INTRON-II-PtMYB1	63	1 TTTT-CTCTGAAATTTATGGGTTTTTTGTTGCTGTTTTCCACACCAG
INTRON-II-PgMYB1	4	7 TTTTTCTCTTTAAATTTATGGGGGTTTTTTTGTTGTTGTTGTTGTTGT
consensus	63	<u> </u> **** **** **************************

PtMYB3 et PgMYB3

Comparaison des	s see	conds introns entre <i>Pt</i> et <i>PgMYB3</i>
INTRON-II- <i>PtMYB3</i>	1	GTACTCCACTTTTCCTTGTCATCTACACTTTTACTAATCCTAGAGGACCATTTCTAAGGT
INTRON-II- <i>PgMYB3</i>	1	GTAATCCTCTTTTCCTTATCATCTAAACCTCTACTAACCCTATACAACCACTTCTAAG
consensus	1	*** *** ******************** ** * ******
INTRON-II- <i>PtMYB3</i>	61	ATTTGGTTTAAGACGCCGATGGCATTTTCAACGTGTCAAACGCC ATATTT GG TTTA A GAG
INTRON-II- <i>PgMYB3</i>	49	ATATTTAGGAG
consensus	61	***** ****
INTRON-II- <i>PtMYB3</i>	121	GCCCATGACATTTTCAACTTGTCAAATCCCATTAGGTAGTTGTGACAATTGTTATAGTTA
INTRON-II-PgMYB3	75	GCCTGTGGTATTTTCAACTTTTCAAATACCTTCAGGTGGTTGTGCCAATTGTTCTACTTA
consensus	121	*** ** *********** ****** ** * ***** ****
INTRON-II-PtMYB3	181	ATATAAAATTCCTAGATGGTTGATATTGGTTATAACTAATATTGAATTCTACATTGAATA
INTRON-II-PgMYB3	135	ATACGAAATTCTTATATTGTTGATATCGGTTATAACTAATATGGATTTCTGCATTGAATA
consensus	181	*** .***** ** ** ** ****** ************
INTRON-II- <i>PtMYB3</i>	241	ATTTGCAG
INTRON-II-PqMYB3	195	ATTTGCAG
consensus	241	*****

Voir légende page suivante.

Annexe VII-4. (suite) PtMYB4 *et* PgMYB4

Comparaison des premiers introns entre Pt et PgMYB4

INTRON-I- <i>PtMYB4</i>	1	GTAAGAACTCACCCACTTCCACTACTCTAATCGCATTCCCTGATCAGGTCTAGATTTTTG
INTRON-I- <i>PgMYB4</i>	1	GTAAGCACCTACCCACTTCCACTCCTCTAATCAGGTTT-
consensus	1	***** ** *********************
INTRON-I- <i>PtMYB4</i>	61	AAGGCCGCCACAAATCCAGACTGGTTCTTTCATTTCATGGAGAACTGGGGATTTTGCGCCCT
INTRON-I- <i>PgMYB4</i>	39	ACAATTCCAGACTGGTTCATTCATTGCGGGAACTGGGGATTGTAGGCCT
consensus	61	**** *******************************
INTRON-I- <i>PtMYB4</i>	121	TGAGGGTTGAAATATCTGTACGGATTTCCAATATTCTGCTATTCAATTCAGTACAATGCA
INTRON-I- <i>PgMYB4</i>	91	TGAGGGTTGAAATATCTCTA-GGTATTTC-GATAT-CAATACAATTCA
consensus	121	*******************************
INTRON-I- <i>PtMYB4</i>	181	ATTTAATTAAACGCATTTCGTGTGCGAAGGTTCAATGTGTCTTCTATTGACCCATACAATA
INTRON-I- <i>PgMYB4</i>	136	ATTTAATACAATGCAATTCAATTCAATGTAACTTATACAATA
consensus	181	******* ** *** *** ** *** *** *** ******
INTRON-I- <i>PtMYB4</i>	241	ATATGGCATTGTACAG
INTRON-I- <i>PgMYB4</i>	178	ATATGGCAATGTACAG
consensus	241	********

Comparaison des seconds introns entre Pt et PgMYB4

INTROÑ-II- <i>PtMYB4</i> INTRON-II- <i>PgMYB4</i> consensus	1 1 1	GTTTTTCTCCATTTCCAGAGCTCTTCTACCCTCTAATTTCTTTC
INTRON-II- <i>PtMYB4</i>	59	CCT-GTTGATTACCCGTTTCATTGTAGATAATTTAGATACTAGTACAAAGCC
INTRON-II- <i>PgMYB4</i>	61	CTTAGTTATCACATTCTGTACTCGTTCGTTGCAGACAACTTAGATACTAGAACAAATCC
consensus	61	* * *** *** *** *** *** *** *** *** *
INTRON-II- <i>PtMYB4</i>	110	CC-CTTGACCCGACCTTGAGAATGAAGGCATGTTCGTAGGTTTTTGCGCTGCTAGCGCAA
INTRON-II- <i>PgMYB4</i>	121	CCTCTTGACCCGACCTTCAGAAGGAAGGCATGTTTGTAGGTTTTTGAGCTGCAGGAGAAA
consensus	121	** ************* **** **********
INTRON-II <i>-PtMYB4</i>	169	AGGAAATGTGCAATGTATTAAGTTTACAGTGTTTAGGAGGACCTTAATAATAATAATACTTCA
INTRON-II <i>-PgMYB4</i>	181	AGGAAATGTACAATGTTTTAAGTTTACAGTGTATAGGAGGACCTTAATAATACTTCA
consensus	181	**********************************
INTRON-II- <i>PtMYB4</i>	229	ATTTCAATTAATCGGAGGCATGGCCACGCAG
INTRON-II- <i>PgMYB4</i>	238	ATTTCAATTAATCGG-GGCATGCCCGCGCAG
consensus	241	*****************************

PtMYB8 et PgMYB8

Comparaison des	s pi	remiers introns entre <i>Pt</i> et <i>PgMYB8</i>
INTRON-I-PtMYB8	1	GTAATGTTTGCTCAATGGTTTGATCGAAAAAAAAAAAAA
INTRON-I-PgMYB8	1	GTAAAGTTTGCTCAGTGGTTTGATGGAAGAAACAATCGGATTAAGTCGA
consensus	1	**** ******* ********* ***************
INTRON-I-PtMYB8	61	TTTCAGTTTGCTGAGCTGATAAATAATTTCCTTTTTGTACTTCTATTGTTCGCAG
INTRON-I-PgMYB8	50	TTTCAGTTTTGCTG-TAAATAATTTCCATTTT-TGG-TCTGTTATTCGCAG
consensus	61	******* ** *** ***********************

Comparaison des seconds introns entre Pt et PgMYB8

INTRON-II- <i>PtMYB8</i>	1	GTAATGTCAAAGACCCTGCAAGTAATGGCGGA-GTTCATAATAGTTGTGGATTGTTGGGT
INTRON-II- <i>PgMYB8</i>	1	GTAATGTCGCAGGCCCAAGTAATGACGAAAGTTCATAGTTGTTGAATGTTCCGT
consensus	1	******* ** *** ************************
INTRON-II- <i>PtMYB8</i>	60	GCAATTATATTGAAAAACAAAATTGCTTGCCACGCAG
INTRON-II- <i>PgMYB8</i>	55	GCAATTATATTGATAAACAAA-TTGTTTGCCACACAG
consensus	61	********** ****** *** *****************

I : premier intron, II : second intron. Les astérisques montrent les nucléotides identiques entre les paires de séquences de *Pinus taeda* (Pt) et *Picea glauca* (Pg).