



Université  
de Toulouse

# THÈSE

En vue de l'obtention du

## DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)

---

**Présentée et soutenue par :**

**Anna PLASENCIA CASADEVALL**

Le mardi 15 décembre 2015

**Titre :**

Transcriptional regulation of wood formation in Eucalyptus. Role of MYB transcription factors and protein-protein interactions.

---

ED SEVAB : Développement des plantes

**Unité de recherche :**

Laboratoire de Recherche en Sciences Végétales (LRSV) UMR 5546 UPS/CNRS

**Directeur(s) de Thèse :**

Jacqueline GRIMA-PETTENATI, Directrice de recherche CNRS, Toulouse  
Isabelle TRUCHET, Maître de conférences Université Toulouse III, Toulouse

**Rapporteurs :**

Marie BAUCHER, Senior research associate, Bruxelles  
David CAPARRÓS-RUIZ, Tenured researcher, Barcelone

**Autre(s) membre(s) du jury :**

Guillaume BÉCARD, Professeur Université Toulouse III, Toulouse  
Valérie LEGUÉ, Professeure Université Clermont Ferrand, Clermont Ferrand





Université  
de Toulouse

# THÈSE

En vue de l'obtention du

## DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)

---

**Présentée et soutenue par :**

**Anna PLASENCIA CASADEVALL**

Le mardi 15 décembre 2015

**Titre :**

Transcriptional regulation of wood formation in Eucalyptus. Role of MYB transcription factors and protein-protein interactions.

---

ED SEVAB : Développement des plantes

**Unité de recherche :**

Laboratoire de Recherche en Sciences Végétales (LRSV) UMR 5546 UPS/CNRS

**Directeur(s) de Thèse :**

Jacqueline GRIMA-PETTENATI, Directrice de recherche CNRS, Toulouse  
Isabelle TRUCHET, Maître de conférences Université Toulouse III, Toulouse

**Rapporteurs :**

Marie BAUCHER, Senior research associate, Bruxelles  
David CAPARRÓS-RUIZ, Tenured researcher, Barcelone

**Autre(s) membre(s) du jury :**

Guillaume BÉCARD, Professeur Université Toulouse III, Toulouse  
Valérie LEGUÉ, Professeure Université Clermont Ferrand, Clermont Ferrand



---

## ACKNOWLEDGEMENTS

Today, the day I begun working in this thesis seems like yesterday. I have always been told that this happens when you are having a good time and in my case, I think that it is in overall true. I feel really lucky to have sheared this time with such valuable people from who I have learned a lot from a scientific and a non-scientific point of view. So, thank you very much!

First of all, I am very grateful with my supervisors Jacqueline Grima-Pettenati and Isabelle Truchet for giving me the opportunity of doing this thesis. For their advice and help during this process and also for all the non-scientific moments we shared. It has really been a pleasure.

I would also like to thank the members of my thesis committee Susana Rivas, Christian Dubos and Laurent Deslandes. The meetings always became an interesting discussion, always trying to criticise from a constructive point of view and to give me ideas to improve my work.

I would also like to thank the reviewers of the manuscript and the members of the jury, for accepting the correction of this work.

And of course I would also like to thank all members of my hosting team. Thanks to Marçal Soler for helping me at any time, for teaching me so many things, for the unforgettable moments, for... It's been a pleasure to work with you. I think we are a really good team! Also thanks to Annabelle Dupas, always there when I needed a hand in the lab or a piece of chocolate! Thanks also to Nathalie, the master of RNA extractions! And thanks to the rest of the team Fabien Mounet, Raphael Ployet, Hong-Chien Nguyen, Athur Desplat, Hua Cassan-Wang, Christiane Marque and Chantal Teulières for the advices, the jokes and the words of encourage. I would also like to thank the internship students that I had the pleasure to work with Guilherme Silva-Martins, Antoine Pineau and Lone Marchesi; and to the former members of the international GFE office Hong Yu, Jorge Lepikson-Neto and Eduardo Camargo. I miss you and I wish you all a brilliant future. Last but not least, I am very grateful to Audrey Courtial, former colleague in GFE team and now a friend. Thank you for the after work times, you are one of the persons I am going to miss the most. And finally, I would like to thank Yves Barrière for the supply of interesting articles and delicious desserts!

I am very grateful to all the community of LRSV; especially to Christian Brière for helping me with tobacco infiltration, and Vincent Burlat for his help with paraffin inclusion and cutting of the roots. I would also like to thank to the "Myco team" in general but specially to Guillaume Bécard, Jean-Philippe Combier, Dominique Laouressgues, Jean-Malo Couzigou and Bruno Guillotin for their advices in *A. rhizogenes*-mediated transformation, cloning, RNA extraction from roots and root *in vitro* culture. And of course I would like to thank all the people from the common services of LRSV, the bioinformatics service, the Imagerie platform and the secretaries

team, whose work is sometimes not very visible but indispensable! I would like to highlight Julie Benedetti from the culture service for helping me during the interminable root samplings.

Thank you also to Charles Rosenberg for the advices about cloning and the supply of vectors and *A. rhizogenes* strains, Fernanda de Carvalho-Niebel for the advices on composite plants hardening as well as many other LIPM members who help me during this thesis.

I am also grateful to the people I meet during the KBBE project “Tree for joules”. Especially to José Carlos Rodrigues for his help in biochemical analyses of poplar plants. Also to Jorge Paiva and Victor Carocha who made me feel as home every time I went to Lisbon, and to Alicia Moreno and Isabel Allona who accomplished the same during the week I spend in GBGP while teaching me how to transform poplar. I hope we can keep in touch!

I am also grateful to Catherine Lapierre and Romain Larbat who performed most of the biochemical characterisation of *Arabidopsis* and *Eucalyptus* plants and Claudine Franche for her advice in *A. rhizogenes*-mediated transformation.

I would also like to thank the people of the Umeå Plant Science Centre. On one hand, to Edouard Pesquet for hosting me to his team and to Junko Takahashi and Andrés Gorzàs for helping me with Py-GC/MS and FTIR experiments even if some data are finally not included in the final manuscript. On the other hand, to the rest of the team and to the Spanish researchers from UPSC, Carolina, Ruben, Manolo, Gloria, etc... who really helped me during my internship at the North Pole!

Before finishing, I will like to thank the “non perms” of LRSV. Thank you for your patience in teaching me French (still some words are resisting: “peuplier”?, “bus”?, “dessous”?...), for the non-perms coffee break, for the minions caricatures, the barbecue times, football/rugby matches, beer times and support messages those last days. I wish you all the best!

Finally, I would like to thank my family and friends. Especialment als meus pares Manel i Rosa, al meu germà Albert, a en Marçal, la Dolors i en Josep, a la Montserrat Ferrer, la Mireia Lopez, la Marta Mallorquí, la Sandra Miquel, l'Àlex Sanchez, la Rosalia Trias, l'Alexandra Pérez, etc... Ha sigut un camí llarg però al final arriba la recompensa. Moltes gràcies per fer-me costat en aquest projecte i per compartir amb mi els bons i els mals moments! Per fer-me sentir, quan torno a Girona, com si no hagués marxat mai. Marçal, dels “encabezados”, als “peus de pàgina”!

Sens racinas pas de flors  
(Without roots there are no flowers)

Vous êtes mes racines / vosaltres sou les meves arrels / you are my roots







---

## ABSTRACT

*Eucalyptus*, the most planted hardwood worldwide for many industrial end-uses, is also the second forest tree whose genome has been fully sequenced. With the final goal of improving wood properties relevant to pulping or bioethanol production, we are focusing our efforts towards the functional characterization of transcription factors regulating the biosynthesis of lignified secondary cell walls (SCW) in *Eucalyptus*.

In this work, we have functionally characterised three *Eucalyptus* R2R3-MYB transcription factors (*EgMYB1*, *EgMYB2* and *EgMYB137*) using the dominant repression (DR) strategy in transgenic poplar plants. While confirming and deepening the antagonist roles of *EgMYB1* and *EgMYB2* in SCW formation, we have identified a new regulator of lignin biosynthesis (*EgMYB137*), not characterized yet in any species. Notably, *EgMYB1*-DR and mostly *EgMYB137*-DR lines showed an improved saccharification yield, and thus a better cellulose accessibility to degrading enzymes pointing them as promising candidates for the production of second generation bioethanol.

One of the main difficulties to use *Eucalyptus* as a model tree is the lack of a suitable transformation method. To overcome this limitation, we have optimised an *Agrobacterium rhizogenes*-mediated transformation of *E. grandis* in order to have a fast and efficient homologous transformation system. We have verified that hairy root morphology and vascular system anatomy were similar to wild type ones and proved that transgenic hairy roots could be used to functionally characterize candidate genes including those involved in SCW formation.

In order to get deeper insights into the regulation of the activity of *EgMYB1*, known to repress lignin biosynthesis, we have explored the role of its protein partners. We show here that *EgMYB1* interacts specifically with a linker histone (*EgH1.3*). This interaction reinforces the transcriptional repression of *EgMYB1* over lignin biosynthesis. The expression profiles of *EgMYB1* and *EgH1.3* overlap in xylem cells at the early stages of differentiation as well as in differentiated mature xylem cells with no or only thin lignified SCW. All together these results suggest that the *EgMYB1*-*EgH1.3* interaction might impair premature or inappropriate lignification of SCW by integrating developmental and position signals.

In conclusion, our results allowed the identification of a new transcription factor and a new protein-protein interaction participating in the regulation of SCW formation, increasing the complexity of the network regulating SCW but also opening new avenues to ultimately improve wood properties for industrial applications.



---

## RÉSUMÉ

*Eucalyptus*, le feuillu le plus planté au monde pour de nombreuses utilisations industrielles, est également le deuxième arbre forestier dont le génome a été séquencé. L'objectif de ce travail de thèse était de contribuer à mieux comprendre la régulation de la formation du bois chez *Eucalyptus*.

Nous avons tout d'abord entrepris la caractérisation fonctionnelle de trois facteurs de transcription de la famille MYB-R2R3 d'*Eucalyptus* (*EgMYB1*, *EgMYB2* and *EgMYB137*) en utilisant la stratégie de répression dominante (DR) chez des peupliers transgéniques. Nous avons ainsi pu approfondir les rôles antagonistes de deux masters régulateurs (*EgMYB1* et *EgMYB2*) de la paroi secondaire et identifier un nouveau régulateur de la biosynthèse de la lignine (*EgMYB137*), dont aucun orthologue n'avait encore été caractérisé. Les lignées d'*EgMYB1*-DR et surtout d'*EgMYB137*-DR montrent un rendement de saccharification plus élevé que les lignées contrôles, ce qui traduit un meilleur accès à la cellulose par les enzymes de dégradation. Ce sont potentiellement des bons candidats dans l'objectif de produire du bioéthanol de deuxième génération.

Une des difficultés majeures auxquelles on se heurte pour caractériser fonctionnellement des gènes chez *Eucalyptus* est l'absence d'une méthode de transformation performante. Pour contourner cette limitation, nous avons mis au point et optimisé une méthode de transformation efficace et rapide via *Agrobacterium rhizogenes*. Nous avons vérifié que la morphologie des racines transformées générées, nommées "hairy roots", et l'anatomie de leur système vasculaire sont similaires à celles des racines non-transformées. Nous avons aussi démontré que le système de «hairy roots» peut être utilisé pour la caractérisation fonctionnelle de gènes candidats, notamment ceux impliqués dans la formation ou la régulation de la paroi secondaire lignifiée.

Dans la dernière partie de ce travail de thèse, nous nous sommes intéressés à la régulation de l'activité d'*EgMYB1*, un répresseur de la voie de la biosynthèse de la lignine, via des interactions protéine-protéine. Nous avons montré qu'*EgMYB1* interagit spécifiquement avec une "histone linker" (*EgH1.3*). Cette interaction renforce la répression transcriptionnelle d'*EgMYB1* sur des gènes de biosynthèse de lignine. Les profils d'expression respectifs d'*EgMYB1* et d'*EgH1.3* coïncident dans les cellules du xylème à des stades très précoces de différenciation ainsi que dans cellules du parenchyme xylémien qui contiennent peu ou pas de parois secondaires lignifiées. Tous ces résultats suggèrent que l'interaction entre *EgMYB1* et *EgH1.3* peut prévenir la lignification prématurée des cellules du xylème en intégrant des signaux développementaux et positionnels.

En conclusion, nos résultats ont permis l'identification d'un nouveau régulateur de la paroi secondaire, et d'une nouvelle interaction impliquée dans la régulation fine de cette paroi. Ces découvertes augmentent le niveau de complexité des mécanismes de régulation qui contrôlent la formation de la paroi secondaire et offrent également des candidats prometteurs pour l'amélioration du bois utilisé en l'industrie.



---

# TABLE OF CONTENTS

<b>Acknowledgements</b> .....	<b>I</b>
<b>Abstract</b> .....	<b>V</b>
<b>Résumé</b> .....	<b>VII</b>
<b>Table of contents</b> .....	<b>IX</b>
<b>List of figures</b> .....	<b>XIII</b>
<b>List of tables</b> .....	<b>XV</b>
<b>Abbreviations</b> .....	<b>XVI</b>
<b>Objectives and organization of the manuscript</b> .....	<b>1</b>
<b>Bibliographic review</b> .....	<b>5</b>
<b>1.1 The vascular system</b> .....	<b>7</b>
1.2.1 The vascular system during primary growth .....	7
1.2.2 The vascular system during secondary growth .....	9
Secondary phloem .....	9
Secondary xylem or wood .....	11
<b>1.3 Wood formation: Xylogenesis</b> .....	<b>11</b>
1.3.1 Secondary cell walls .....	13
Biosynthesis of polysaccharide components .....	16
Biosynthesis and polymerisation of lignins.....	18
Biosynthesis of other phenolics compounds .....	20
1.3.2 Programmed cell death.....	21
<b>1.4 Wood plasticity</b> .....	<b>22</b>
<b>1.5 Regulation of wood formation</b> .....	<b>23</b>
1.5.1 Hormonal control of cambium activity and xylem specification and formation.....	24
1.5.2 Transcriptional regulation .....	25
First layer NAC master switches .....	26
Second layer master switch: AtMYB46/AtMYB83 and respective orthologs .....	28
Third layer of transcriptional regulation.....	29
1.5.3 Protein-protein interactions.....	32

**Functional characterization of three *Eucalyptus* MYB genes controlling wood formation and evaluation of their impacts on saccharification ..... 37**

**2.1 Preface ..... 40**

**2.2 Introduction..... 40**

**2.3 Results..... 44**

2.3.1 Phylogenetic relationship of EgMYB137 with close R2R3-MYB genes ..... 44

2.3.2 Expression profiling of the three studied MYB genes..... 46

2.3.3 Effect of the EgMYB1 and EgMYB2 dominant repression chimeric proteins on EgCAD2 promoter transcriptional activity ..... 47

2.3.4 Impact of EgMYB2-DR, EgMYB1-DR and EgMYB137-DR overexpression on poplar growth and development ..... 48

2.3.5 Histological and chemical analysis of the lignin content from transgenic poplar stems ..... 51

2.3.6 Secondary cell wall composition..... 53

Analytical Pyrolysis- Gas Chromatography / Mass Spectrometry (Py-GC/MS) ..... 53

Analysis of Lignin-Derived Monomers Released by thioacidolysis..... 55

2.3.7 Metabolite profiling ..... 56

2.3.8 Effects of EgMYB1-DR, EgMYB2-DR and EgMYB137-DR on saccharification yields..... 58

**2.4 Discussion..... 59**

**2.5 Material and Methods ..... 64**

2.5.1 Phylogeny reconstruction ..... 64

2.5.2 Microfluidic RT-qPCR and RNAseq ..... 64

2.5.3 Vectors and cloning ..... 64

2.5.4 Co-transfection experiments and fluorimetric GUS assays ..... 65

2.5.5 Plants transformation, production and sampling ..... 66

2.5.6 Biochemical analysis of SCW composition ..... 67

2.5.7 Metabolic profiling..... 68

Metabolite extraction..... 68

U-HPLC analysis ..... 68

2.5.8 Saccharification ..... 68

**2.6 Supplemental tables and figures..... 70**

***Eucalyptus* hairy roots, a fast, efficient and versatile tool to explore function and expression of genes involved in wood formation ..... 77**

**3.1 Preface ..... 80**

<b>3.2</b>	<b>Introduction</b> .....	<b>80</b>
<b>3.3</b>	<b>Results</b> .....	<b>82</b>
3.3.1	Optimization of <i>A. rhizogenes</i> -mediated transformation and generation of <i>E. grandis</i> composite plants .....	82
3.3.2	Construction of expression vectors for gene functional characterization and promoter activity analysis .....	86
3.3.3	<i>In vitro</i> hairy roots cultures and subcellular localization of proteins .....	87
3.3.4	Xylem structure from hairy roots.....	87
3.3.5	Tissue and cell expression pattern of the <i>EgCCR1</i> and <i>EgCAD2</i> promoters in hairy roots .....	88
3.3.6	<i>EgCCR1</i> silencing in hairy roots .....	90
<b>3.4</b>	<b>Discussion</b> .....	<b>93</b>
<b>3.5</b>	<b>Experimental procedures</b> .....	<b>96</b>
3.5.1	Plant material.....	96
3.5.2	<i>Agrobacterium rhizogenes</i> strains.....	97
3.5.3	Binary vector construction .....	97
3.5.4	Induction of hairy roots and generation of composite <i>E. grandis</i> plants.....	98
3.5.5	Hardening of <i>Eucalyptus</i> composite plants.....	98
3.5.6	Hairy root cultures .....	99
3.5.7	RNA extraction and reverse transcription - quantitative polymerase chain reaction (RT-qPCR).....	99
3.5.8	Biochemical lignin analyses.....	100
3.5.9	Histochemical gus assay .....	100
3.5.10	Microscopy and cell imaging .....	101
<b>3.6</b>	<b>Supplemental tables and figures</b> .....	<b>102</b>
<b>3.7</b>	<b>Supplemental information</b> .....	<b>105</b>
3.7.1	Information S1: Details on binary vector construction. ....	105
3.7.2	Information S2: RNA extraction optimisation and detailed RNA extraction protocol .....	105

**Wood formation in *Eucalyptus* is regulated by the interaction between the transcription**

<b>factor EgMYB1 and the histone linker protein EgH1.3</b> .....	<b>111</b>
<b>4.1 Preface</b> .....	<b>114</b>
<b>4.2 Introduction</b> .....	<b>114</b>
<b>4.3 Results</b> .....	<b>117</b>
4.3.1 EgMYB1 interacts specifically with a Histone linker protein in yeast.....	117
4.3.2 EgMYB1 interacts specifically with EgH1.3 in <i>N. benthamiana</i> cell nucleus .....	121

4.3.3	Transactivation assays in <i>N. benthamiana</i> leaves showed a strong repression of EgCCR and EgCAD gene promoters produced by EgMYB1 together with EgH1.3 .....	122
4.3.4	Transgenic <i>Arabidopsis</i> plants overexpressing together EgMYB1 and EgH1.3 showed strong reduction of secondary cell wall formation .....	124
4.3.5	<i>Arabidopsis</i> plants overexpressing both EgMYB1 and EgH1.3 showed an important reduction of the total lignin level and an alteration of their lignin monomeric composition .....	127
4.3.6	Transcriptomic analysis revealed a repression of lignin, cellulose and hemicellulose genes for 35S:EgMYB1/EgH1.3 <i>Arabidopsis</i> plants .....	130
4.3.7	Metabolite profiling analysis in stems revealed a strong induction of indole glucosinolates in 35S:EgMYB1/35S:EgH1.3 plants.....	131
4.3.8	Simultaneous overexpression of EgMYB1 and EgH1.3 in a <i>Eucalyptus</i> homolog transformation system validates the phenotype observed in <i>Arabidopsis</i> .....	133
4.3.9	Promoter expression analysis and qPCR revealed partial overlapping expression patterns between EgMYB1 and EgH1.3 .....	134
<b>4.4</b>	<b>Discussion.....</b>	<b>136</b>
<b>4.5</b>	<b>Material and methods.....</b>	<b>140</b>
4.5.1	Phylogeny reconstruction .....	140
4.5.2	Vector construction .....	140
4.5.3	Yeast-two-hybrid library construction and screening.....	141
4.5.4	Sub-cellular localization, co-localization and FRET-FLIM analysis.....	142
4.5.5	Transactivation assays .....	143
4.5.6	Plant material.....	144
4.5.7	Transcriptomic analysis.....	145
4.5.8	Histochemical analysis .....	146
4.5.9	Biochemical analysis .....	147
<b>4.6</b>	<b>Supplementary tables and figures .....</b>	<b>149</b>
	<b>General discussion and perspectives .....</b>	<b>161</b>
	<b>References .....</b>	<b>171</b>



---

## LIST OF FIGURES

Figure 1.1	Schematic representation of the anatomy of the vascular organization in the <i>Arabidopsis</i> inflorescence stem, hypocotyl and root.	8
Figure 1.2	Overview of procambial/cambial cell specification and xylem/phloem cell differentiation.	10
Figure 1.3	Xylogenesis or wood formation.	12
Figure 1.4	Schematic representation of the main polymers present in secondary cell wall (SCW) and their association in the SCW three-dimensional network.	14
Figure 1.5	Three-dimensional structure of the SCW layers of a tracheid (xylem cell).	15
Figure 1.6	General phenylpropanoid, and monolignol specific pathways for <i>E. grandis</i> .	19
Figure 1.7	Schematisation of the regulation of radial patterning.	25
Figure 1.8	Transcriptional network involved in the regulation of SCW formation.	26
Figure 2.1	Neighbor-Joining phylogenetic tree of R2R3-MYB proteins related to EgMYB137	45
Figure 2.2	Mean relative transcript abundance of EgMYB1, EgMYB2 and EgMYB137 genes.	46
Figure 2.3	Effects of EgMYB1 and EgMYB2 dominant repression (DR) chimeric proteins on the transcriptional activity of <i>EgCAD2</i> promoter.	47
Figure 2.4	Growth phenotypes of overexpressing EgMYB2-DR, EgMYB1-DR and EgMYB137-DR dominant repression poplar transgenic lines	50
Figure 2.5	Morphology of the leaves of transgenic poplar plants	51
Figure 2.6	Transversal sections of poplar stems overexpressing EgMYB2-DR, EgMYB1-DR and EgMYB137-DR	52
Figure 2.7	Klason lignin content in poplar stems.	52
Figure 2.8	Principal Component Analysis for Py-GC/MS data in poplar stems.	53
Figure 2.9	Box plots for the Py-GC/MS data in poplar stems.	54
Figure 2.10	Saccharification yields in poplar stems.	59
Figure S2.1	Mean relative transcript abundance of EgMYB1, EgMYB2 and EgMYB137 genes.	73
Figure S2.2	Principal component analysis for metabolic profiling data.	74
Figure 3.1	Developmental time course of <i>E. grandis</i> hairy roots expressing either the GFP or the DsRed fluorescent markers.	83
Figure 3.2	Sequential steps of the transformation of 14-days-old <i>E. grandis</i> seedlings by A4RS harbouring either a GFP or DsRed based binary vector.	85
Figure 3.3	Nuclear localization of histone linker fused to CFP in <i>E. grandis</i> root cells.	87
Figure 3.4	Comparison of xylem development and lignified secondary cell walls between transgenic and wild-type roots.	88
Figure 3.5	Histochemical localization of GUS activity in <i>E. grandis</i> hairy roots transformed with the <i>EgCCR1</i> and <i>EgCAD2</i> promoters.	89
Figure 3.6	Relative transcript levels of <i>EgCCR1</i> and SCW-related genes in <i>CCRAs</i> hairy root lines.	91
Figure 3.7	Transversal sections of <i>EgCCR1</i> down-regulated hairy roots.	92
Figure 3.8	Thioacidolysis lignin yields in <i>CCRAs</i> roots.	93
Figure S3.1	Co-transformation efficiency time course of <i>E. grandis</i> explants using <i>A. rhizogenes</i> strains A4RS transformed either with pHKN29 or with pGWAY vectors.	103
Figure S3.2	Maps of pGWAY-0 and pGWAY-1 destination vectors.	103
Figure S3.3	Linear growth of excised <i>E. grandis</i> hairy roots.	104
Figure S3.4	Relative transcript level of <i>EgCCR1</i> gene and <i>EgCCR1as</i> mRNA + <i>EgCCR1</i> mRNA	104
Figure S3.5	Partial thioacidolysis GC-MS chromatograms showing the <i>CCRAs</i> signature	104
Figure 4.1	Neighbor-Joining phylogenetic tree of histone linker proteins from <i>Eucalyptus grandis</i> , <i>E. gunnii</i> as well as from other 5 dicot- and 3 monocotyledonous plants.	119

Figure 4.2	Protein – protein interactions detected using Yeast-Two-Hybrid.	120
Figure 4.3	Effects of EgMYB1 and EgH1.3 on the transcriptional activities of the EgCAD2 and EgCCR promoters using transactivation assays in <i>N. benthamiana</i> leaves.	123
Figure 4.4	Macroscopic phenotype of <i>Arabidopsis</i> transgenic plants.	125
Figure 4.5	Xylem structure and secondary cell wall thickness from the inflorescence stem of two-month-old <i>Arabidopsis</i> transgenic plants growing in short day conditions.	126
Figure 4.6	Principal Components Analysis and hierarchical clustering of the relative amounts of each compound category detected by Pyrolysis – Gas Chromatography / Mass Spectrometry.	128
Figure 4.7	Relative transcript abundance of secondary cell wall biosynthetic genes measured by RT-qPCR.	131
Figure 4.8	Xylem structure of <i>Eucalyptus</i> hairy roots.	134
Figure 4.9	Expression pattern of EgMYB1 and EgH1.3 in <i>Eucalyptus</i> .	135
Figure S4.1	Subcellular localization of EgMYB1-CFP, AtMYB96-CFP and EgMYB137-CFP infiltrated together with EgH1.3-YFP in <i>N. benthamiana</i> leaves.	151
Figure S4.2	Comparison of the FRET-FLIM measures of the donor alone and in presence of the acceptor, expressed in frequency histograms according to the CFP mean lifetime ( $\tau$ ) intervals.	152
Figure S4.3	Effects of EgMYB1 $\Delta$ alone and with EgH1.3 on the transcriptional activities of the EgCAD2 promoter using transactivation assays in <i>N. benthamiana</i> leaves.	153
Figure S4.4	Effects of EgMYB1 and EgH1.3 on the transcriptional activities of the EgCAD2 and EgCCR promoters using transactivation assays in <i>N. benthamiana</i> leaves	154
Figure S4.5	Relative transcript abundance of transgenes in <i>Arabidopsis</i> transgenic plants.	155
Figure S4.6	Relative transcript abundance of transgenes in <i>Eucalyptus</i> hairy roots.	156
Figure S4.7	Map of pGWAY-MYB1 destination vector.	157

---

## LIST OF TABLES

Table 2.1	Thioacidolysis yields for extract-free wood from controls (wild type + empty vector lines) and transgenic poplar plants.	55
Table 2.2	Metabolic profiling of EgMYB2-DR, EgMYB1-DR and EgMYB137-DR poplar plants.	57
Table 2.3	Phenolic glycosides and oligolignols abundance for EgMYB2-DR, EgMYB1-DR and EgMYB137-DR poplar plants	58
Table 2.4	Accession number for the genes and promoters used in this study as well as primers used for cloning procedures	65
Table S2.1	Recovered peak names from pyrolysis analyses, their designed category and the identified compound.	70
table S2.2	Detailed information from the recovered peaks and the corresponding mass spectra from methanol soluble phenolic compounds.	72
Table 3.1	Mean efficiency of <i>in vitro</i> co-transformation of <i>E. grandis</i> explants with <i>A. rhizogenes</i>	84
Table S3.1	Thioacidolysis H, G and S monomers percentage and for S/G ratio from hairy roots.	102
Table S3.2	Primer sequences used for cloning.	102
Table S3.3	Primer sequences of genes used for RT-qPCR analyses.	102
Table 4.1	FRET-FLIM analysis showing the interaction between EgMYB1 with EgH1.3 in the nuclei of <i>N. benthamiana</i> transiently transformed epidermal cells.	122
Table 4.2	Thioacidolysis lignin structure in the hypocotyls of <i>Arabidopsis</i> transgenic lines.	129
Table 4.3	Lignin content and structure in inflorescence stems of <i>Arabidopsis</i> transgenic lines.	129
Table 4.4	Soluble metabolites in the inflorescence stem of the <i>Arabidopsis</i> transgenic lines.	131
Table S4.1	Relative amounts of the compounds categories detected in Py-GC/MS	146
Table S4.2	Primers designed in this study	147

---

## ABBREVIATIONS

<b>4CL:</b> 4-Hydroxycinnamoyl-CoA ligase	<b>HD-ZIP III:</b> Class III homeodomain-leucine zipper
<b>ABC:</b> ATP-binding cassette	<b>KNAT7:</b> KNOTTED1-like homeobox gene7
<b>AD:</b> Activation domain	<b>NAC:</b> (NAM, ATAF and CUC) transcription factors
<b>ARF:</b> Auxin response factors	<b>NST:</b> NAC Secondary wall thickening promoting factor
<b>Aux/IAA:</b> Auxin/Indole-3-Acetic Acid	<b>MP:</b> ARF5/MONOPTEROS
<b>BD:</b> Binding domain	<b>MS:</b> Murashige and Skoog media
<b>C4H:</b> Cinnamate 4-hydroxylase	<b>MYB:</b> plant orthologs of myeloblastosis oncogenes transcription factors
<b>C3H:</b> <i>p</i> -coumarate 3-hydroxylase	<b>OFP:</b> Ovate Family Protein
<b>CAD:</b> Cinnamyl alcohol dehydrogenase	<b>PAL:</b> Phenylalanine ammonia lyase
<b>CaMV:</b> Cauliflower Mosaic Virus	<b>PAT:</b> Auxin polar transport
<b>CCoAOMT:</b> Caffeoyl-CoA O-methyltransferase	<b>PCD:</b> Programmed cell death
<b>CCR:</b> Cinnamoyl-CoA reductase	<b>PCR:</b> Polymerase chain reaction
<b>cDNA:</b> Complementary deoxyribonucleic acid	<b>PIN:</b> PIN-FORMED
<b>CDS:</b> Coding sequence	<b>Py-GC/MS:</b> Pyrolysis-gas chromatography/mass spectrometry
<b>CesA:</b> Cellulose synthase A	<b>QDO/X/A:</b> auxotrophic media without tryptophan, leucine, adenine and histidine, supplemented with X- $\alpha$ -Gal and Aureobasidin.
<b>CFP:</b> Cyan Fluorescent Protein	<b>qRT-PCR:</b> Quantitative reverse transcription Polymerase chain reaction
<b>COMT:</b> Caffeic acid O-methyltransferase	<b>S:</b> Syringyl
<b>DDO:</b> auxotrophic media without tryptophan and leucine	<b>SCW:</b> Secondary cell wall
<b>DR:</b> Dominant repression	<b>TF:</b> Transcription factor
<b>EAR:</b> Ethylene-responsive element binding factor-associated Amphiphilic Repression	<b>VND:</b> vascular-related NAC-domain
<b>F5H:</b> Ferulate 5-hydroxylase	<b>X-<math>\alpha</math>-Gal:</b> 5-bromo-4-chloro-3-indolyl alpha-D-galactopyranoside
<b>FRET-FLIM:</b> Förster Resonance Energy Transfer – Fluorescence Lifetime Imaging Microscopy	<b>Y2H:</b> Yeast two-hybrid
<b>G:</b> Guaiacyl	
<b>GFP:</b> Green Fluorescent Protein	
<b>H:</b> <i>p</i> -hydroxyphenyl	





---

## OBJECTIVES AND ORGANIZATION OF THE MANUSCRIPT

The overall objective of my PhD was **to get deeper insights into the transcriptional regulation of wood formation in a woody angiosperm, *Eucalyptus***. Eucalypts, which are the target plants of my host team, are evergreen hardwood trees that belong to the *Myrtaceae* family. The *Eucalyptus* genus comprises more than 700 species (Myburg *et al.*, 2007). Most of them are native to Australia although, nowadays, they can be found in a wide variety of environments in more than 20 countries. Depending on the specie, they are cultivated in tropical (*E. grandis* and hybrids) or temperate zones (*E. globulus*, *E. gunnii*). Eucalypts wood is used to produce pulp, paper, timber or firewood, and its use as feedstock to bioethanol production is currently investigated. Due to their high adaptability, fast growth rate, and valuable wood and fibres properties, *Eucalyptus* species have become the most valuable and widely planted forest trees worldwide (Myburg *et al.*, 2007).

The genome of *E. grandis* has been recently sequenced and annotated (Myburg *et al.*, 2014). TF families involved in secondary cell wall (SCW) regulation such as MYB, NAC, ARF and AUX/IAA (Soler *et al.*, 2015; Hussey *et al.*, 2015; Yu *et al.*, 2014, 2015) have been characterised *in silico* and genes putatively implicated in SCW biosynthesis have been identified (Carocha *et al.*, 2015; Myburg *et al.*, 2014). These studies identified a considerable amount of candidate genes putatively regulating wood formation and gave valuable tools that make easier to use *Eucalyptus* as a model tree to study wood formation. The main handicap, however, is that transformation of *Eucalyptus* is long and tedious, which impairs the study of candidates genes in a homologous system.

In addition to improving our basic understanding of wood formation, an important developmental process in vascular plants, we aimed at **identifying key regulators that could impact wood traits such as lignin content and composition, and consequently secondary cell wall recalcitrance to degradation**. The recalcitrance of lignocellulosic biomass to degradation is a major obstacle for pulp and papermaking, which represents the major end-use of *Eucalyptus* industrial plantations with no less than 20 million ha worldwide. Recalcitrance to degradation is also impairing bioethanol production by limiting access to cellulose. In the context of rarefaction of fossil fuels, eucalypts, which are characterized by their rapid growth and wide adaptability, are promising feedstock in the emerging field of second-generation biofuels. Indeed, most of my PhD work has been performed within the frame of the Plant KBBE “Tree For Joules” project whose overall objective was to improve *Eucalyptus* and poplar wood properties for bioenergy.

This thesis manuscript comprises five main chapters. The **Chapter 1** consists in a bibliographic review on wood formation and its regulation at the transcriptional level.

The results are organized in three chapters each presented in the form of an article.

The main aim of the study described in **Chapter 2** was **to better understand the role of three MYB transcription factors (*EgMYB137*, *EgMYB1* and *EgMYB2*) in wood formation**. Thanks to the *E. grandis* genome sequence (Myburg *et al.*, 2014), a genome-wide survey of the R2R3-MYB genes family was performed (Soler *et al.*, 2015), pointing out *EgMYB137*, which has no clear ortholog in *Arabidopsis*, as a potential regulator of wood formation. *EgMYB1* (Legay *et al.*, 2010) and *EgMYB2* (Goicoechea *et al.*, 2005) were previously shown by my host team to be a repressor and an activator of SCW formation, respectively. We used the dominant repression strategy to evaluate the effects of modulating the expression of these three MYBs on SCW composition and on saccharification potential. Since *Eucalyptus* is very difficult to transform and regenerate, we used transgenic poplar to functionally characterize these three genes.

**Chapter 3** is an article accepted in Plant Biotechnology Journal. In this study, we aimed **to develop a fast, efficient and versatile homologous transformation system for *Eucalyptus***. We chose to optimise an *A. rhizogenes*-mediated transformation of *E. grandis* (the specie used to sequence the genome). Using this transformation system, chimeric plants with WT aerial parts and transgenic hairy roots are produced. Thus, we also tested if transgenic hairy roots can be used to explore function and expression of genes involved in wood formation using histochemical, transcriptomic and biochemical approaches. The method we described is timely since it will accelerate gene mining of the *Eucalyptus* genome both for basic research and industry purposes.

The objective of **Chapter 4** was **to better understand the regulation of the transcriptional activity of the master repressor *EgMYB1***. Because protein–protein interactions are crucial to fully understand how a transcription factor activity is regulated, my hosting team seek for protein partners of *EgMYB1* by screening a *Eucalyptus* xylem yeast-two-hybrid library. In the work presented in chapter 4, we aimed to demonstrate that *EgMYB1* interacts specifically with one of the obtained protein partners, the histone linker protein *EgH1.3*. Once verified, we aimed to study the effect of this interaction on lignin biosynthesis and SW formation. With this purpose, we used transgenic *Arabidopsis* overexpressing the two genes as well as *Eucalyptus* hairy roots transformed using the method developed in Chapter 3.

Finally the main results and perspectives are discussed in **Chapter 5**.







CHAPTER 1:

---

**BIBLIOGRAPHIC REVIEW**



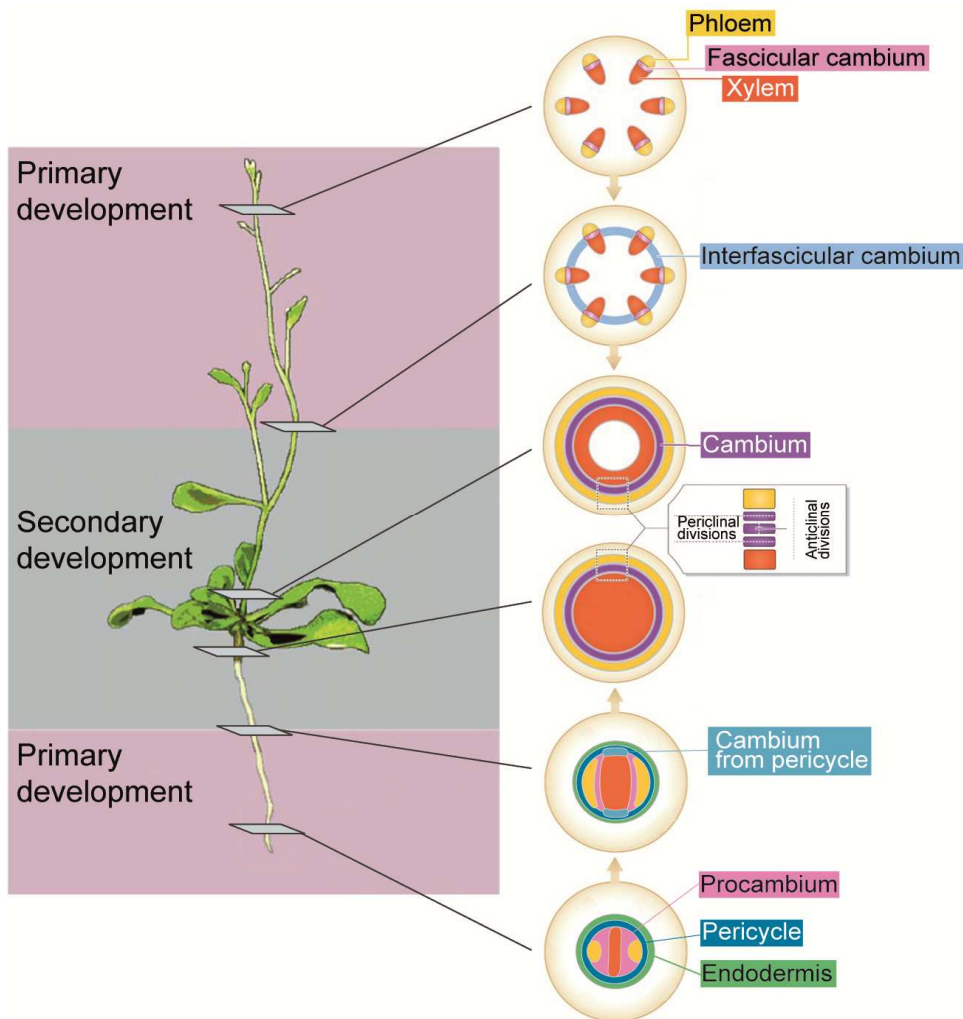
## 1.1 THE VASCULAR SYSTEM

From the numerous adaptations that land plants have developed during evolution, the acquisition of the vascular system some 400 million years ago have been a crucial event ensuring their successful earth colonization. The vascular system is composed of two main tissues, xylem and phloem. Xylem cells play crucial roles in water transport and mechanical support of the entire plant. Phloem is essential to the transport of photosynthate from photosynthetic tissues (source tissues) to developing tissues (sink tissues).

### 1.2.1 THE VASCULAR SYSTEM DURING PRIMARY GROWTH

Plant growth arises from mitotic cell divisions taking place in growth loci called meristems. During embryogenesis, two apical meristems are established and contribute to root (Root Apical Meristem, RAM) and shoot (Shoot Apical Meristem, SAM) development. Later on, but still on embryo stage, vascular plants develop a lateral meristem called procambium. During a first phase of growth, these meristems produce the primary plant body including the primary vasculature composed of primary xylem and phloem. Procambium cells undergo two types of division: the periclinal divisions parallel to the plant axis/surface, give rise to phloem and xylem precursor cells whereas the anticlinal divisions perpetuate the procambium tissue along the plant axis. In some species, primary vasculature, begin to differentiate from procambium in mature embryos, but in many others, vascular tissue differentiation only starts after the seed germinates (Esau, 1965).

In leaves, phloem and xylem generates towards abaxial and adaxial surfaces, respectively. In stems, primary vascular tissues are organised in discrete collateral vascular bundles separated by parenchyma cells, while in roots, the vascular tissue is organised in a bi- or multi-symmetric pattern (Figure 1.1).



**Figure 1.1: Schematic representation of the anatomy of the vascular organization in the *Arabidopsis* inflorescence stem, hypocotyl and root exhibiting either primary or secondary growth.** Xylem=red, phloem=yellow, procambium and fascicular cambium=pink, interfascicular cambium and cambium formed from pericycle=light blue, cambium=purple, pericycle=dark blue, endodermis=green. Adapted from Miyashima *et al.*, (2012); Nieminen *et al.*, (2015); Zhang *et al.*, (2011)

In both root and shoot vasculature, the most common organisation is, from the outside to the inside of the organ: phloem, procambium and primary xylem cells. Roots differ from stems by the absence of pith, and the presence of two additional single-cell ring-shape layers, the endodermis and the pericycle (Figure 1.1). Endodermis will further form the Casparian strip, essential for impairing the diffusion of solutes from the inside to the outside of roots. The pericycle is a single layer of meristematic cells, which are at the origin of lateral roots and have a role in secondary growth of roots (see below).

During this first phase of growth (primary growth), plants grow in the acropetal direction thanks to the activity of root and shoot apical meristems. With the notable exception of the

monocotyledons, many vascular plants undergo a second phase of growth (secondary growth), which implies growth in diameter or radial growth. This secondary growth can be limited to the hypocotyls as for instance in *Arabidopsis* or can be particularly important like in trees where it produces large amount of secondary xylem (wood).

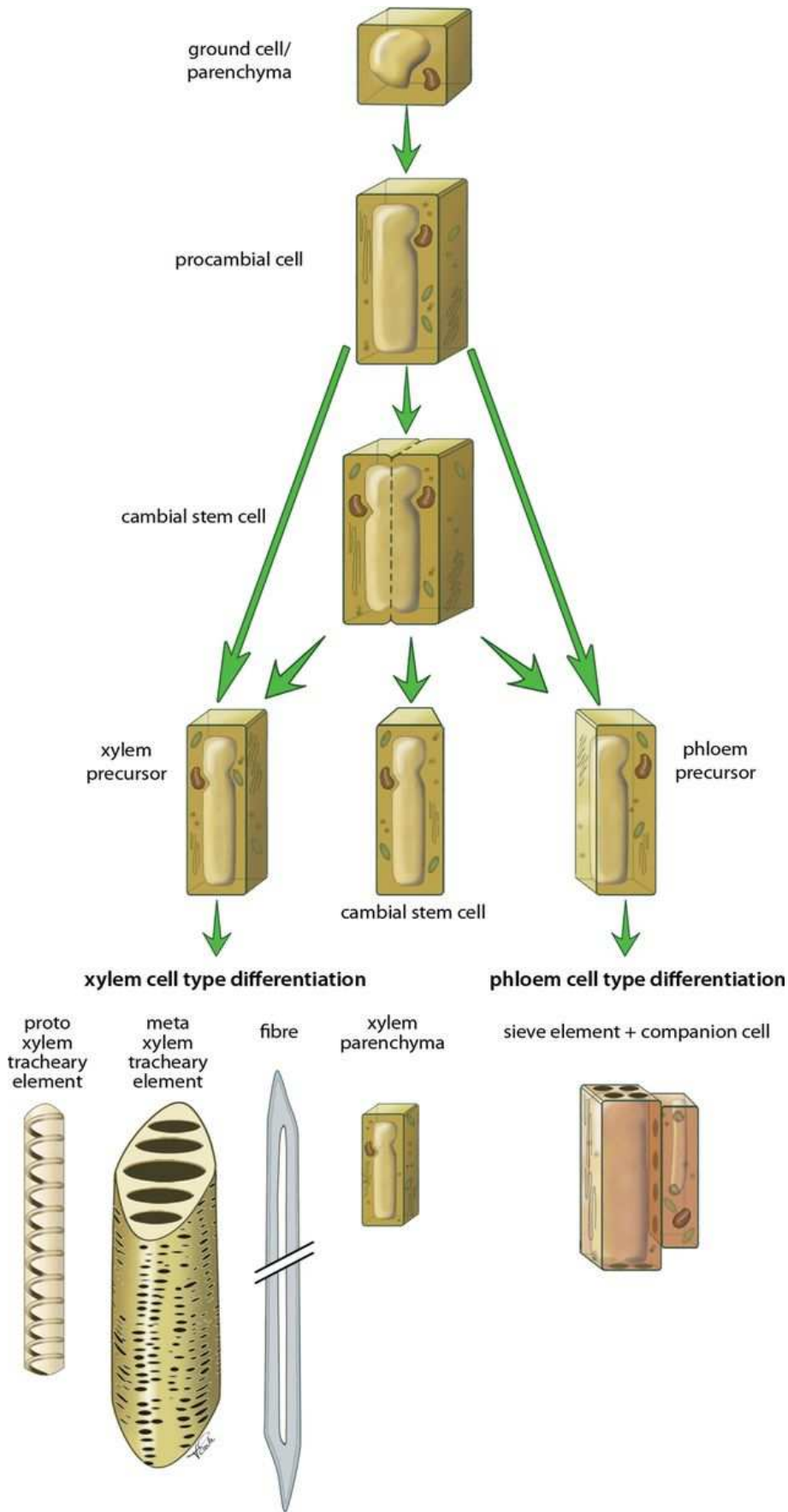
### 1.2.2 THE VASCULAR SYSTEM DURING SECONDARY GROWTH

During secondary growth in angiosperm stems, procambium cells will give rise to the fascicular cambium whereas the interfascicular cambium is thought to arise through the *de novo* recruitment of interfascicular parenchyma cells (Schuetz *et al.*, 2013). In roots, the interfascicular cambium differentiates from pericycle cells. The junction of these two populations of cambial cells makes the entire circular vascular cambium of mature tree trunks and roots (Figure 1.1).

The vascular cambium is made of two kinds of meristematic cells called the fusiform and the ray initials, which differ by their size and shape (Déjardin *et al.*, 2010). Depending on the type of initials and on the kind of division occurring, different cell types of phloem, xylem and also other vascular cambium undifferentiated cells or “stem cells” will be produced. Periclinal divisions of the fusiform initials will give rise to vessels, vessel-associated cells, fibres and axial parenchyma cells) and to phloem precursor cells (Figure 1.2). Radial isodiametric initials produce the radial cell system, consisting of parenchyma ray cells. Anticlinal divisions will generate more of the vascular cambium initials.

#### Secondary phloem

Phloem precursor cells differentiate into sieve elements and companion cells (Figure 1.2). The function of sieve tube elements is to transport metabolites from source tissues, such as leaves, to metabolic sinks, such as roots and seeds. Along their differentiation process, sieve tube elements lose their nuclei and most of their organelles, but they remain physiologically alive. Sieve tube elements are cytoplasmically coupled through plasmodesmata to companion cells which maintain the metabolic competency of the sieve tube elements, and also “load” molecular cargo into them for long-distance transport. In many species, phloem fibres with secondary cell walls (SCW, see below) are also present. They confer mechanical support (Esau 1965) and can be lignified in woody species for instance.



**Figure 1.2: Overview of procambial/ cambial cell specification and xylem/phloem cell differentiation.**

Extracted from Schuetz *et al.* (2013).



### Secondary xylem or wood

Wood is a complex three-dimensional and heterogeneous tissue composed of different cell types all harbouring lignified SCWs (Figure 1.2). The different cell types present have different shapes and functions. In higher plants xylem contains tracheids, vessels, fibres, parenchyma and ray cells

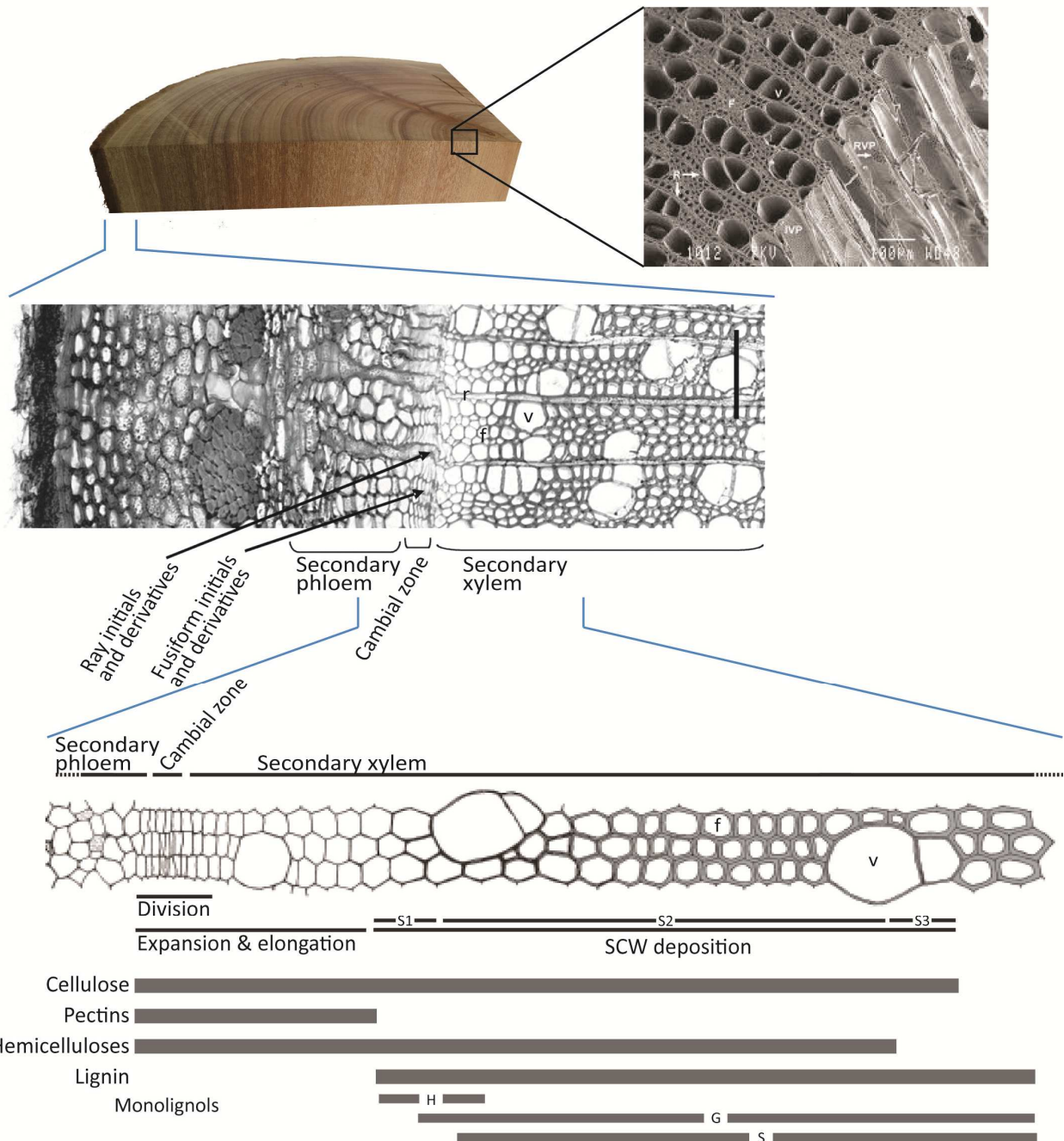
The tracheary elements include tracheids and vessels and are responsible for water and solutes conduction. Tracheids are elongated narrow tube like cells with hard thick and lignified walls with large cell cavity. Their ends are tapering or pointed. They are dead empty cells with their walls provided with one or more row of bordered pit. Tracheids occur alone in the wood of ferns and gymnosperms but in angiosperms they occur associated with vessels. Vessels are cylindrical tubular structure, formed of row of cells placed end to end from which the cross wall breaks down. The vessels members are connected by means of plates with pores and known as perforation plates through which water moves upwards. On the basis of the way SCWs are deposited, vessels are named as annular, spiral, scalariform, reticulate or pitted. SCW in tracheary elements also provide mechanical strength to plant organs. They differ notably at the organisation of their SCW (figure 1.2, see below). Vessel elements are the most evolutionarily advanced cell types, and are the more common tracheary elements in angiosperms (Růžička *et al.*, 2015).

The xylem fibres develop a thick and evenly deposited SCW. Their programmed cell death is delayed, resulting in a more extensive thickening and lignification of SCW, consistent with their primary role in providing mechanical support to the tree trunk (Růžička *et al.*, 2015).

The xylem ray cells perform radial transfer of assimilates between phloem and xylem, temporary storage as starch of lipids and their remobilisation at the new season (Déjardin *et al.*, 2010). During xylogenesis, they are involved in lignification of tracheary elements and fibres, and later on they are also involved in the formation of heartwood (Déjardin *et al.*, 2010).

## 1.3 WOOD FORMATION: XYLOGENESIS

Xylem differentiation is a complex developmental process. It involves five major developmental steps (Figure 1.3): cell division from the vascular cambium, cell expansion (elongation and radial enlargement), secondary cell wall deposition and programmed cell death (PCD).



**Figure 1.3: Xylogenesis or wood formation.** Three-dimensional view of xylem cells from poplar (extracted from Brunner *et al.*, 2004) and partial transversal cross-section of poplar stem (adapted from Baucher *et al.*, 2007), scale bar=100µm. A schematic representation of the cambial zone and the differentiating xylem is also included where the principal steps in xylogenesis are shown (adapted from Hertzberg *et al.*, 2001). The SCW sub-layers (S1, 2 & 3) as well as the deposition of the main polymers are indicated. PCD is not included because it takes place at different stages for vessels and fibres.

### 1.3.1 SECONDARY CELL WALLS

All plant cells contain primary cell walls, which determine the shape and the size of cells. At the end of their differentiation processes, some specialized cells also deposit SCW between the plasma membrane and the primary cell wall. Among them, there are phloem fibres and xylem cells, but also epidermal and pericarp cells, the endothecium of anthers and the endodermal Caspian strip in roots (Didi *et al.*, 2015). SCW deposition starts when cells have finished their expansion (Figure 1.3) and can continue even after those cells are dead after following a PCD. However, whether the cessation of growth is caused by the SCW deposition or by earlier events is not yet elucidated (Didi *et al.*, 2015).

Both, primary and secondary cell walls are composed of rigid cellulose microfibrils embedded in a gel-like matrix of non-cellulosic polysaccharides (hemicelluloses and pectins) and proteins. Cellulose is the most abundant biopolymer synthesized on land. It is composed only by one type of monomer, the glucose, a hexose type monosaccharide. Glucose molecules (approximately 10,000 in SCW, Timell T.E., 1967) are bound through  $\beta$ -1-4 linkages to form an unbranched and unsubstituted chain (Figure 1.4a). In turn, 18 of those chains are able to perform inter- and intramolecular hydrogen bonding and hydrophobic interactions to form 3 nm diameter microfibrils (Newman *et al.*, 2013). In contrast, the noncellulosic polysaccharides (hemicelluloses and pectins) are structurally more complex and show considerable heterogeneity. Hemicelluloses for example, contain a main chain constituted of xylose, a pentose monosaccharide, linked by  $\beta$ -1-4 (xyloglucans), or by  $\beta$ -1,4- mannan (called mannans, galactomannans and galactoglucomannans). These main chains are further modified and decorated with ramifications of arabinose or glucuronic acid among others.

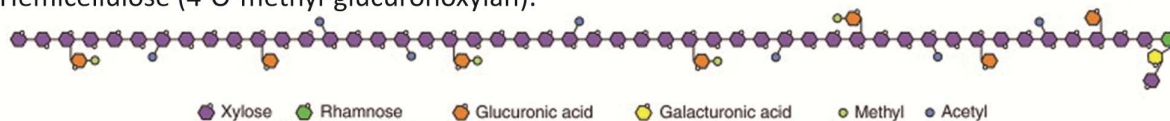
Huge differences between primary and secondary cell walls exist. In contrast to primary cell walls, most of SCW also contain high proportion of lignin, a complex aromatic polymer. The main units in the polymer, *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) units, are derived *p*-coumaryl, coniferyl and sinapyl alcohol (also called monolignols), respectively. The relative abundance of the three monomers will impact the properties of lignin.

**a Main secondary cell wall polymers:**

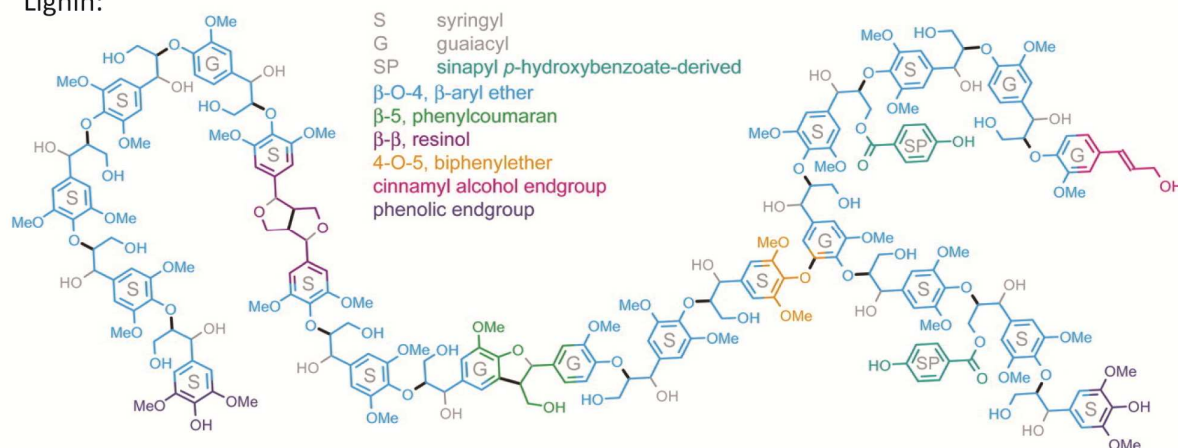


● Glucose

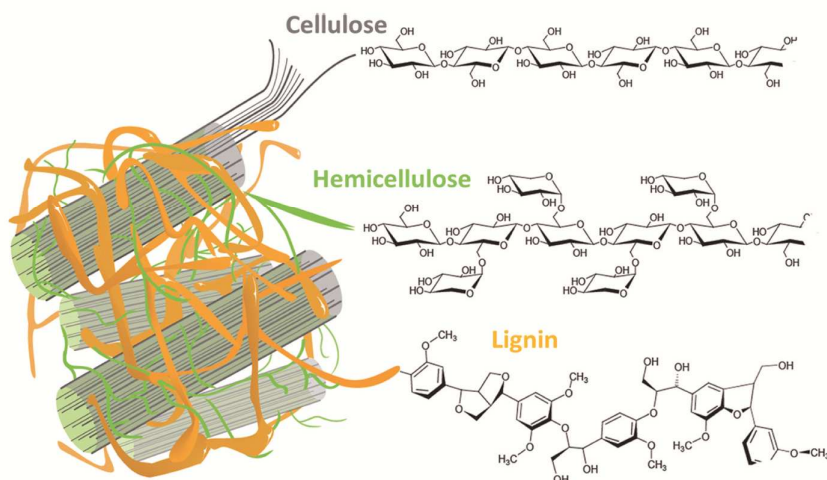
Hemicellulose (4-O-methyl-glucuronoxylan):



Lignin:



**b Secondary cell wall three dimensional network:**

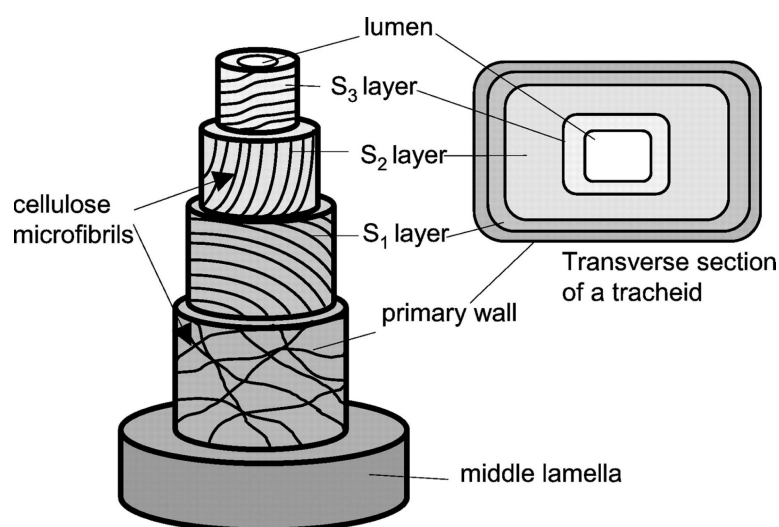


**Figure 1.4: Schematic representation of the main polymers present in SCW (a) and their association in the SCW three-dimensional network (b).** Extracted from Burton *et al.*, (2010); Rennie and Scheller, (2014); Vanholme *et al.*, (2010); and from the Ghent web site: <http://www.psb.ugent.be/bio-energy/313-lignin>.

The relative content of the different components (cellulose, hemicelluloses, and lignins) and the composition of the non-cellulosic polysaccharides are also different between primary and secondary cell walls. SCW, are composed of about 50% cellulose, 25% hemicelluloses and 25%

lignins with traces of pectins and proteins, impregnated with phenolics. These values can vary a lot between species and between individuals in the same tree species (Déjardin *et al.*, 2010). Regarding differences in composition, hemicelluloses from primary cell walls of dicotyledonous plants are mainly composed of xyloglucans and heteroxylans whereas SCW are mainly composed of 4-*O*-methyl-glucuronoxylans (Figure 1.4a) and glucomannans (Doblin *et al.*, 2010). Pectins are much less present in SCW than in primary cell walls. Thus, noncellulosic polysaccharides of SCW generally have lower degrees of branching or main chain substitution (Doblin *et al.*, 2010; Burton *et al.*, 2010). All these polysaccharides are intimately associated with one another, both non-covalently and covalently, and often with proteins and lignin, to form a three-dimensional network (Figure 1.4b).

Secondary cell walls of xylem cells contain several layers called S1, S2 and S3 (figure 1.5), each composed of a network of long bundles of cellulose microfibrils which are oriented at a fixed angle different in the S1, S2 and S3 layers. Together with the presence of low branching hemicelluloses, this feature facilitates intermolecular aggregation in aqueous media. The aggregates are stabilized through the formation of inter-molecular hydrogen bonding and hydrophobic interactions between sugar rings, which will reduce their solubility and enhance their ability to form gels. Besides, the presence of lignin, an insoluble polymer, results in a dehydration of the wall compartment, and consequently confers hydrophobicity to cells with SCW (Doblin *et al.*, 2010; Burton *et al.*, 2010).



**Figure 1.5: Three-dimensional structure of the secondary cell wall layers of a tracheid (xylem cell).** Extracted from Plomion *et al.* (2001).

S1, S2 and S3 layers differ in their thickness, chemical composition and cellulose microfibril angle (MFA, measured from the longitudinal axis) (figure 1.5). In particular, the alternation of high and low angles in the different layers of the secondary cell wall increases its rigidity and mechanical strength. This effect is again reinforced by the presence of lignin.

#### Biosynthesis of polysaccharide components

Biosynthesis of polysaccharide SCW components have been extensively studied and although some points need still to be elucidated, the main steps and the responsible enzymes have been described in several reviews (Cosgrove, 2005; Burton *et al.*, 2010; Doblin *et al.*, 2010; Zhong and Ye, 2015). Here I will focus on enzymes related to polysaccharides that are abundant in SCW: cellulose and 4-*O*-methyl-glucuronoxylans.

The key enzymes in wall biogenesis are glycosyltransferases that add glycosyl residues from activated sugar nucleotide donors to the elongating polymeric chain. Two major classes of these glycosyltransferases have been identified. One group, which is often referred to as the polysaccharide synthase group, is believed to catalyse the iterative addition of glycosyl residues to the end of the elongating polysaccharide chain, without releasing the chain between successive glycosylation events. These enzymes are encoded, among others, by members of the cellulose synthase (CesA) family genes.

The second group of glycosyltransferases is believed to add single glycosyl substituents to the main chain residues or to add short oligosaccharide chains that are appended to the main chain through single catalytic encounters with the nascent polysaccharide chain.

Cellulose biosynthesis takes place at the plasma membrane. The three CesA genes involved in cellulose biosynthesis during SCW formation in *Arabidopsis*, (*AtCesA4/ Irregular Xylem (AtIRX5)*, *AtCesA7/ AtIRX3* and *AtCesA8/AtIRX1*), have non-redundant functions and are different from those involved in the biosynthesis of cellulose deposited in primary cell wall. The respective orthologs of these genes are *PtrCesA4*, *PtrCesA7A*, *PtrCesA7B*, *PtrCesA8A* and *PtrCesA8B* in poplar (Song *et al.*, 2010) and *EgCesA2*, *EgCesA3* and *EgCesA1* in *Eucalyptus* (Ranik and Myburg, 2006). Their corresponding proteins are thought to interact to form a particle and in turn, six particles interact to form the rosette. As each CesA protein will synthesise one chain of cellulose, 18 chains will be simultaneously synthesised in the plasma membrane towards the cell wall and will interact to form the cellulose microfibrils (Newman *et al.*, 2013). In

*Arabidopsis*, other proteins such as COBRA-like 4 (*AtCOBL4*), chitinase-like protein 2 (*AtCTL2*) and *AtTED6* are thought to be implicated in cellulose biosynthesis and modulation of cellulose microfibrils crystallisation degree. In addition, a number of proteins with unknown functions have been shown to co-purify with cellulose synthase complexes in *Populus* xylem, suggesting their roles in cellulose biosynthesis (Song *et al.*, 2010). Both the SCW pattern and the cellulose microfibril angle (MFA) are thought to be directed by association of CesaA rosettes with microtubules. In primary cell wall biosynthesis of *Arabidopsis*, it takes place via a cellulose synthase-interacting protein 1 (*CSI1*). Although the functional roles of *CSI1* in secondary wall cellulose biosynthesis have not been investigated, it is likely that *CSI1* or *CSI* homologs may also guide the movement of secondary wall CesaA complexes (Zhong and Ye, 2015). Besides, it is thought that microtubules also direct the targeted transport of vesicles carrying the CesaA complexes, hemicelluloses and enzymes related to lignin polymerization, to specific plasma membrane domains, and thereby control the patterned deposition of secondary walls. For instance protoxylem vessels have an annular or helicoidal pattern that allows these cells to continue growing whereas metaxylem vessels have a pitted or reticulate pattern, which does not allow continue cell elongation.

The most abundant hemicellulose in SCW of dicotyledonous plants is xylan (Rennie *et al.*, 2014). It is synthesized in the Golgi apparatus by Type II membrane proteins anchored by a single N-terminal transmembrane domain and with their catalytic domains in the Golgi lumen. In *Arabidopsis* and poplar, it is composed of  $\beta$ -(1,4)-linked xylose residues substituted with acetyl, glucuronic acid (GlcA) or 4-O-methylglucuronic acid (Me-GlcA). In dicot, xylan molecules include the tetrasaccharide 4- $\beta$ -D-Xyl-(1-4)- $\beta$ -D-Xyl-(1-3)- $\alpha$ -L-Rha-(1-2)- $\alpha$ -D-GalA-(1-4)-D-Xyl at their reducing ends (Figure 1.4a). Several enzymes have been implicated in xylan synthesis (Rennie and Scheller, 2014; Zhong and Ye, 2015). Two members of the *Arabidopsis* Glycosyltransferase Family 43 (*AtGT43*), *AtIRX9* and *AtIRX14*, and one member of *AtGT47*, *AtIRX10*, encode putative xylosyltransferases required for synthesizing the xylan backbone. In addition, related genes *IRX9*-like (*AtIRX9-L*), *AtIRX10-L*, and *AtIRX14-L* appear to encode functionally redundant paralogs. The Glucuronic Acid (GlcA) substitutions on xylans in the *Arabidopsis* secondary cell wall are introduced by the action of *AtGUX1* and *AtGUX2*, both members of the GT8 family. 4-O-Methyl groups are transferred from S-adenosyl- methionine to GlcA residues by *GXMT1*, a protein containing a Domain of Unknown Function 579 (DUF579).

### Biosynthesis and polymerisation of lignins

Lignin is the generic term for a large group of aromatic polymers resulting from the oxidative coupling of phenolic metabolites, mainly *p*-hydroxycinnamyl alcohol monomers (the so-called monolignols): *p*-coumaryl, coniferyl and sinapyl alcohol.

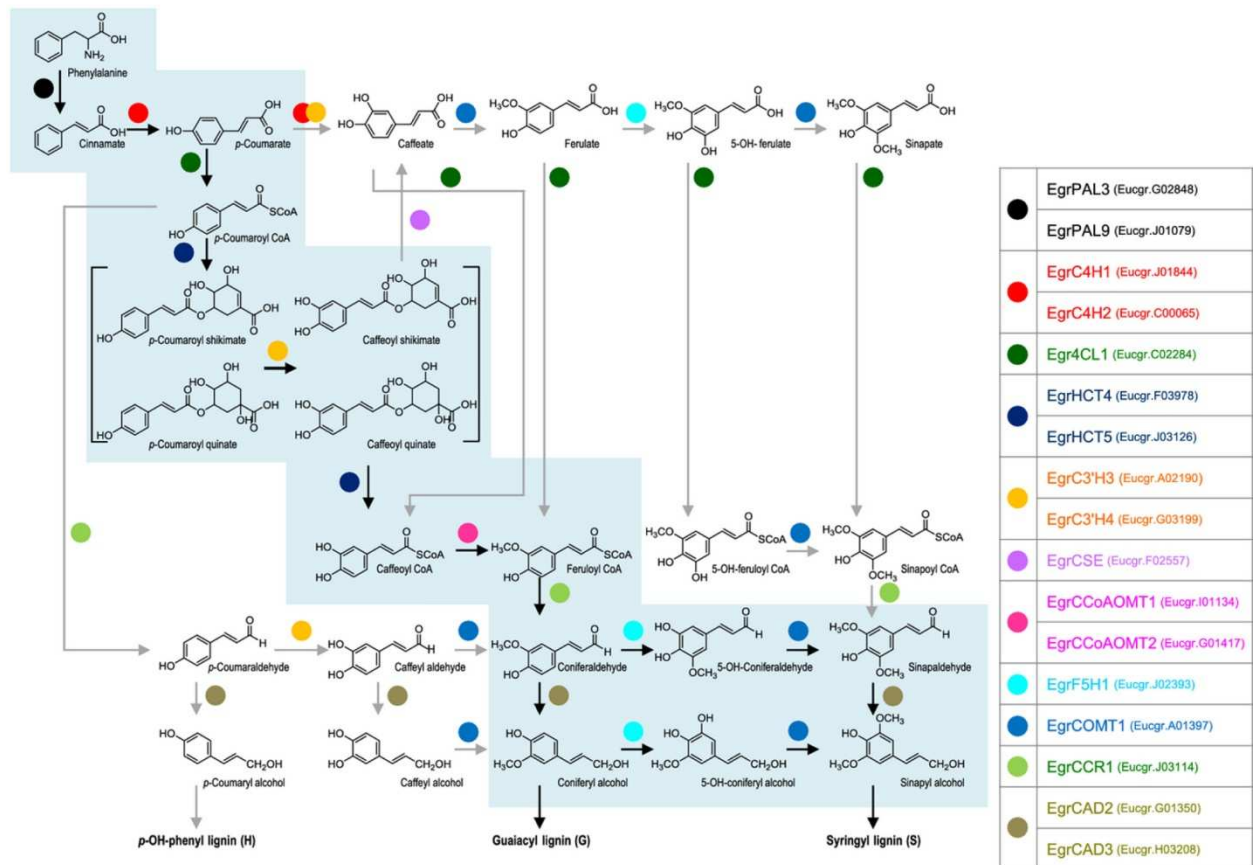
Lignin is deposited on SCW of developing xylem cells but its biosynthesis can also be induced upon wounding, pathogen infection, metabolic stress, and perturbations in cell wall structure. Most if not all genes involved in lignin biosynthetic pathway have been cloned and their impact on lignin amount and composition has been studied through mutants or reverse genetics in various species, mostly in *Arabidopsis* and poplar.

Biosynthesis of lignins starts with the shikimate pathway which produces chorismate. It will further be metabolised among other compounds to the aromatic amino acids: phenylalanine, tyrosine and tryptophan. Tryptophan is necessary for the production of auxin and secondary indolic metabolites, such as indolic glucosinolates in *Arabidopsis*. Phenylalanine is used as entry substrate for the general phenylpropanoid pathway and results, after seven steps, in feruloyl-CoA (Vanholme *et al.*, 2012, 2010). Involved enzymes are phenylalanine ammonia-lyase (PAL), cinnamate 4-hydroxylase (C4H), 4-coumarate: CoA ligase (4CL), hydroxycinnamoyltransferase (HCT), *p*-coumarate 3-hydroxylase (C3H) and caffeoyl-CoA *O*-methyltransferase (CCoAOMT) (Figure 1.6). Various pathways use phenylpropanoid pathway intermediates as starting compounds, including the pathways towards flavonoids, benzenoids, coumarins, and sinapate and ferulate esters (Vanholme *et al.*, 2012).

The monolignol-specific pathway includes four well-studied enzymatic steps that convert feruloyl-CoA into the monolignols coniferyl alcohol and sinapyl alcohol, and *p*-coumaroyl-CoA to *p*-coumaryl alcohol (Vanholme *et al.*, 2012). Reactions are catalysed by ferulate 5-hydroxylase (F5H), caffeic acid *O*-methyltransferase (COMT), cinnamoyl-CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD). F5H is also now often called coniferaldehyde 5-hydroxylase (CAld5H) to reflect its preferred substrate, coniferaldehyde. The preferred substrate of COMT is also now known to be 5-hydroxyconiferaldehyde and not 5-hydroxyferulic acid. Recently a new alternative route involving the caffeoyl shikimate esterase (CSE; (Vanholme *et al.*, 2013) has been discovered. This enzyme catalyzes the conversion of caffeoyl shikimate into caffeate, and together with 4CL, it can bypass the second HCT reaction thought



to convert the caffeate esters into caffeoyl CoA. Many of mentioned enzymes belong to multigene families. The members putatively involved in lignin biosynthesis on xylem SCW formation have been defined in *Arabidopsis* (Raes *et al.*, 2014), *P. trichocarpa* (Shi *et al.*, 2010) and *E. grandis* (Carocha *et al.*, 2015) (Figure 1.6). In some cases those genes are expressed also in non-lignified tissues because monolignols are also used to generate aromatic molecules other than lignin (oligolignols, lignans, monolignols glycosides, etc...).



**Figure 1.6: General phenylpropanoid, and monolignol specific pathways for *E. grandis*.** The name of the implicated enzymes and their substrates and products are shown. Reactions thought to be key steps in lignin biosynthesis are shaded in blue. Extracted from Carocha *et al.* 2015.

Monolignols are synthesised in the cytoplasm and translocated to the growing cell wall for polymerisation (Wang *et al.*, 2013). The mechanism by which monolignols are translocated to the growing cell wall is largely unknown. Alejandro and colleagues (2012) have shown that in *Arabidopsis*, the ABC-transporter (AtABCG29) pumps *p*-coumaryl alcohol across the plasma membrane. Thus, it has been postulated that the other lignin building blocks can be transported to the cell wall by a larger panel of ABC-transporters, or by ABC-transporters of

relaxed specificities (Sibout and Höfte, 2012). Alternatively, glycosylated forms of monolignols are transported to the vacuole for storage (Liu *et al.*, 2011). Those glycosylated forms may polymerise into lignin in the cell wall of vessels after cell death and membrane disruption (Barros *et al.*, 2015). Once the monolignol glycosides reach the cell wall, the present beta-glucosidases can hydrolyse them into monolignols, which could after be incorporated to the growing lignin polymer (Wang *et al.*, 2013).

Once on the cell wall, monolignols are oxidised by peroxidases and/or laccases to monolignol radicals that eventually polymerize into the lignin macromolecule via combinatorial radical-radical coupling reactions. Both, peroxidases and laccases belong to huge multifamilies and the involvement of one specific member to lignin polymerisation is very difficult to determine due to functional redundancy. Two laccases involved in lignification have been identified in *Arabidopsis* (Berthet *et al.*, 2011). Direct contact between the peroxidase/laccase and the substrate is not needed as radical-transfer reactions can also pass the radical from one molecule to another. All radical coupling reactions in lignification are termination events. Thus, continued lignification of cell walls requires new hydrogen-abstraction of monomers and of the growing lignin oligomer following each coupling reaction.

The moment when lignification starts is variable depending on the cell type and the tissue it belongs. In xylem elements, lignification starts as soon as SCW begins to be deposited by lignifying primary cell wall components (middle lamella and cell corners) (Wang *et al.*, 2013). Lignins associate gradually with hemicelluloses present in the cell wall, eliminate water and form a hydrophobic environment. The cells become both rigid and impermeable.

#### Biosynthesis of other phenolics compounds

The monolignols that result from the above-described monolignol-specific pathway are used for at least three different product classes apart of lignin (Figure 1.6): monolignol glucosides, (neo) lignans and oligolignols (Vanholme *et al.*, 2012). Oligolignols are small lignin polymers believed to polymerise similarly than lignin (Vanholme *et al.*, 2010; Morreel *et al.*, 2010), but that normally do not incorporate to the growing lignin polymer (Morreel *et al.*, 2004; Huis *et al.*, 2012). Monolignol glucosides, or 4-*O*-glucosylated monolignols, are thought to be storage forms that can eventually be finally polymerised into lignin after cell death. However this polymerisation can also take place after vacuole disruption occurs due to pathogen attack,

insect/herbivore feeding, cell freezing or desiccation, thus having a defence role (Wang *et al.*, 2013). Lignans and neolignans are monolignol dimers that differ in the way they are bound ( $\beta$ - $\beta$  for lignans and  $\beta$ -O-4- or  $\beta$ -5-for neolignans, Umezawa, 2003). They are natural antioxidants, possess anti-microbial activity and are believed to protect heartwood of many tree species against wood-rotting fungi (Wang *et al.*, 2013). Very recently the antioxidant role of a phenylpropanoid coupling product *in planta* has been demonstrated through deciphering the function of the Phenylcoumaran benzylic ether reductase (PCBER), one of the most abundant proteins in poplar xylem (Niculaes *et al.*, 2014). PCBER was shown to reduce phenylpropanoid dimers *in planta* to form antioxidants that protect the plant against oxidative damage.

The general phenylpropanoid- and monolignol-specific pathways also provide intermediates, such as hydroxycinnamic acids, that can be polymerised into the growing lignin polymer as alternative units. This has been shown in many cases when for instance an intermediate accumulates as a consequence of down-regulating a lignin biosynthetic enzyme (Baucher *et al.*, 2003; Boerjan *et al.*, 2003; Boudet *et al.*, 2003; Vanholme *et al.*, 2012). Intermediates of the phenylpropanoid pathway can also be further metabolised to sinapate esters (e.g. sinapoyl glucose and sinapoyl malate), flavonoids (flavonoid, flavonol, isoflavonoid, anthocyanin and proanthocyanin), benzenoids (e.g. gallic acid, salicylic acid, vanillin) or coumarins among others.

### 1.3.2 PROGRAMMED CELL DEATH

At the final stage of their differentiation, vessels and fibres follow a programmed cell death (PCD). Xylem cell death is an inseparable part of the xylem maturation program and current knowledge suggest that the necessary components for PCD are produced early during xylem differentiation and that cell death is prevented through the action of inhibitors (Bollhoner *et al.*, 2012). There are significant differences in the timing of PCD between the different xylem cell types on angiosperms. Vessel elements differentiate rapidly and die immediately after, whereas fibres stay alive much longer and the autolysis of their cell content is slower as shown in *Populus* stems (Courtois-Moreau *et al.* 2009).

The process has been extensively studied in *Zinnia elegans* culture system and described in detail (Ohdaira *et al.*, 2002). In early stages of PCD, a large array of hydrolytic enzymes are synthesised and transported to the vacuole where they are activated. Three hydrolase genes, XYLEM CYSTEINE PEPTIDASE1 (*AtXCP1*), *AtXCP2*, and BIFUNCTIONAL NUCLEASE1 (*AtBFN1*), have

been shown to be associated with programmed cell death during xylem differentiation in *Arabidopsis* (Funk *et al.*, 2002; Bollhoner *et al.*, 2012). The vacuole is thus converted to lytic vacuole similar to animal lysosomes. The autolysis of the cell starts with the rupture of the membrane of the vacuole. In addition, cytosolic pH decreases, activating cytosolic hydrolases. Nucleus, plastids, mitochondria disappear, and the cells are emptied of their cytoplasm.

Parenchyma cells, despite containing a lignified SCW, can stay alive and functional for decades showing that programmed cell death is not related to lignification. When they become dysfunctional, they die, liberate phenolics, and contribute to the formation of heartwood, the most internal part of the trunk, which often appears dark-coloured compared to the external part of the wood or sapwood. Sapwood is living part of the trunk where the water transport is still effective (Déjardin *et al.*, 2010).

#### 1.4 WOOD PLASTICITY

Trees are long-living organisms with a sessile lifestyle, which develop in a variable environment and are subjected to developmental control. As a consequence, wood is a complex and highly variable tissue, the formation of which is developmentally and environmentally regulated (Plomion *et al.*, 2001). Wood structure and composition is heterogeneous between species, developmental stages and between sections of a stem/trunk to respond to mechanical forces. The variability occurs at the tissue level (proportion of different cell types) as well as at the individual cell level (size, shape, wall structure, texture and chemical composition)(Plomion *et al.*, 2001). For instance, wood developed during spring (early wood) is different to that developed after (late wood) when cambium is less active. Later wood presents differences in density due to smaller fibres with thicker cell walls and smaller lumen (Déjardin *et al.*, 2010). There are also differences between juvenile wood (produced from trees from 5-20 years old, and from the youngest apical parts of older trees) and mature wood. Juvenile wood is less dense and has higher microfibril angle (MFA) of the cellulose microfibrils which allows him to bend without breaking. On the contrary, mature wood is denser and have lower MFA thus being mechanically more resistant (Barnett and Bonham, 2004).

Finally, plants must adapt to changes in their environment to maintain optimal growth. Ex: Wind, slope of the soil, light, etc... To respond to those needs, trees adjust their growth

following forced changes in orientation, to re-establish a vertical position. In angiosperms, this adjustment involves the differential regulation of vascular cambial activity between the lower (opposite wood) and upper (tension wood) sides of the leaning stem and results in asymmetric growth of the stem. On the tension wood side, cell divisions are activated, leading to high eccentricity on transverse section of the stem. Tension wood contains fewer vessels of smaller size than normal wood. It is characterised by the differentiation the G fibres which are fibres that are longer and contain a thick inner gelatinous layer (G layer). This G layer is almost only composed of pure cellulose in which microfibrils have a low MFA, and thus oriented almost or completely with the axis of the fibre (Barnett and Bonham, 2004).

## 1.5 REGULATION OF WOOD FORMATION

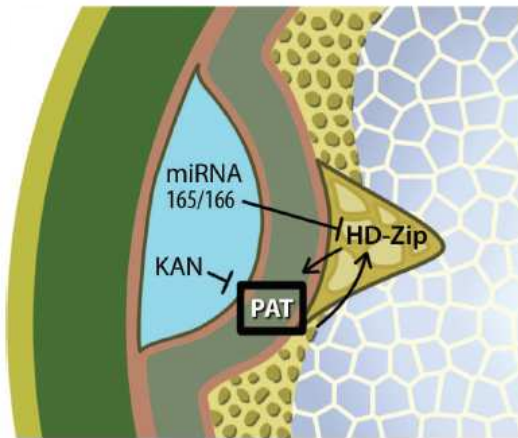
Xylogenesis or wood formation is a very complex and irreversible differentiation process. In one hand (pro)cambial development and activity should be regulated in order to form a continuous tissue which stays undifferentiated and to give rise to phloem stem cells towards the outer and xylem stem cells towards the inner part of stems and roots. On the other hand, xylem stem cells should follow the differentiation pathway corresponding to the proper cell type (vessels, fibres, axial and ray parenchyma). This means a tight and sophisticated regulation in the appropriate cell and at the appropriate time of the expression of genes and the activity of proteins (molecular effectors) involved in cell division, cell elongation, SCW biosynthesis and deposition and programmed cell death.

Xylogenesis is regulated by phytohormones, transcription factors and by post-transcriptional and post-translational mechanism such as those involving microRNA (miRNAs) and protein-protein interactions. Most of the studies have been performed using the model plant *Arabidopsis* enabling to get a general overview of xylogenesis regulation in this species. Some important insights however come from researches performed in trees species such as pine, poplar and eucalypts. Many reviews have addressed the regulation of xylogenesis including Růžička *et al.*, 2015; Schuetz *et al.*, 2013; Nakano *et al.*, 2015; Zhang *et al.*, 2014; Baucher *et al.*, 2007.

### 1.5.1 HORMONAL CONTROL OF CAMBIUM ACTIVITY AND XYLEM SPECIFICATION AND FORMATION

Plant hormones play important roles to regulate wood formation as reviewed in Fukuda (2004), Sorce *et al.* (2013), Ursache *et al.* (2013), Didi *et al.* (2015), and Mauriat *et al.* (2014). For example, ethylene and cytokinin acts as positive regulators of wood formation, promoting cambial cell proliferation, gibberellin promotes both cell division and xylem fibre elongation and auxin is involved in different steps: formation and maintenance of the pro(cambium), cell division and cell specification.

Auxin plays a central role in establishing vascular bundles, with auxin maxima driven by the PIN transporters predetermining the formation of leaf primordia in the shoot apical meristem. The auxin canalization hypothesis based on the assumption that locally elevated auxin concentration lead to increased auxin flow, describes a mechanism for the formation of vascular bundles. Polarization of auxin flow, indicated by the expression of the PIN transporters, precedes the formation of vascular bundles (Donner *et al.*, 2010). *ATHB8*, a class III HD-ZIP (homeodomain-leucine zipper) transcription factor, is required for preprocambial development and procambium differentiation, and its expression depends on the activity of the auxin response factor (ARF) *MONOPTEROS (MP)* (Donner *et al.*, 2010). *MP* is expressed in the procambial tissue of the hypocotyl and cotyledons and *mp* mutants show strong defects in vascular tissues. In addition to *ATHB8*, there are four other *HD-ZIP III* genes in *Arabidopsis*, *PHABULOSA (PHB/ATHB14)*, *PHAVOLUTA (PHV/ATHB9)*, *REVOLUTA/INTERFASCICULAR FIBRELESS (REV/IFL1)* and *CORONA (CNA/ATHB15)* that have a central role in xylem specification. They have been shown to regulate the number of procambium cells by promoting xylem differentiation during vascular development. Auxin positively regulates these genes which in turn, positively promotes polar auxin transport (PAT) in a positive feedback loop (Schuetz *et al.*, 2013). The activity of the class III HD-ZIP genes is opposed by the functionally related *KANADI* genes, MYB TFs that act as general repressors of canalised auxin flux (Fukuda, 2004). The microRNAs (miRNAs) 165 and 166 target the HD-ZIP mRNAs and enhance their degradation causing a phenotype close to the obtained by HD-ZIP loss of function mutants. All together, the model of Figure 1.7 has been proposed. In this model, *KANADI* is expressed in differentiating phloem tissue together with miRNA 165 and 166. This impairs auxin flow and limits HD-ZIP expression to cambial and xylem tissues.



**Figure 1.7: Cross-section of the fascicular cambium region of an *Arabidopsis* inflorescence stem where a schematisation of the regulation of radial patterning through polar auxin transport (PAT), HD-ZIP and KANADI (KAN) transcription factors and miRNA165 and 166 are shown. Epidermis and cortex are coloured in green, phloem tissue in blue, vascular cambium in khaki, xylem tissue in yellow, and pith cells in pale blue. Extracted from Schuetz *et al.*(2013).**

Plant hormones also play important roles to regulate secondary xylem (wood) formation. It has been shown for a long time that auxin concentration peaks at the cambial zone in tree stems, however auxin' possible mode of action during secondary growth is surprisingly poorly understood (Bhalerao and Fischer, 2014). Perturbing auxin signalling by overexpressing a mutated *Populus IAA3* gene resulted in reduced cell proliferation in cambium, and thus less wood formation (Nilsson *et al.*, 2008). Recently, my hosting team has shown that in *Arabidopsis* overexpressing a stabilized version of EgrIAA4 from *E. grandis*, the lignified secondary walls of the interfascicular fibres appeared very late, whereas those of the xylary fibres were virtually undetectable, suggesting that EgrIAA4 may play crucial roles in fibre development and secondary cell wall deposition (Yu *et al.*, 2015).

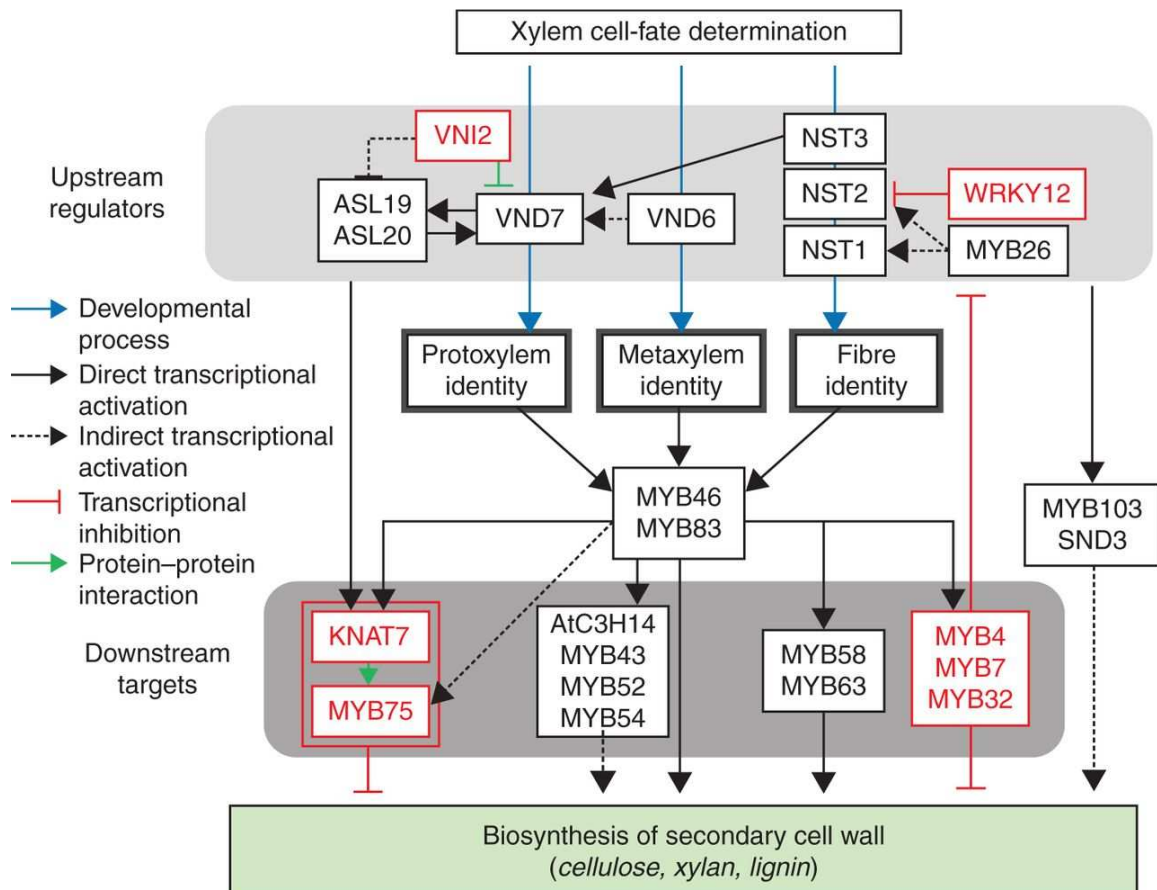
Cytokinins have also a critical role in cambium activity. The cytokinin receptors are expressed in the cambium, and their mutants show defects in cambium formation (Růžička *et al.*, 2015; and references therein). In addition, strong cytokinin synthesis mutants (quadruple isopentenyltransferase and septuple 5'-monophosphate phosphoribohydrolase lonely guy) show a striking phenotype in which the cambium is essentially absent (Růžička *et al.*, 2015; and references therein).

Giberellins are also known to stimulate cambial cell divisions and induce xylem fibre formation (Aloni, 2007). Recently, a role for strigolactones in tight interaction with auxin during cambium initiation was described in *Arabidopsis* (Agusti *et al.*, 2011).

## 1.5.2 TRANSCRIPTIONAL REGULATION

Transcriptional regulation of xylogenesis occurs through a hierarchical network with multileveled feed forward regulation where upstream TFs regulate the expression of

downstream TFs and both can regulate biosynthetic genes (Figure 1.8). For reviews see Nakano *et al.* (2015); Zhong and Ye (2015); Hussey *et al.* (2013); Grima-Pettenati *et al.* (2012); Didi *et al.* (2015); Růžička *et al.* (2015); Schuetz *et al.* (2013); Ko *et al.* (2014) and references therein. The involved TFs mainly belong to the NAC (N<sub>AM</sub>, A<sub>TAF</sub> and C<sub>UC</sub>) and the MYB (plant orthologs of *myeloblastosis* oncogenes) families, two of the largest plant transcription factor families. We will focus on key members of these two families.



**Figure 1.8: Transcriptional network involved in the regulation of SCW formation.** Adapted from (Ko *et al.*, 2014).

### First layer NAC master switches

TFs on the top of the cascade belong to the NAC family and are named master switches or master regulators. When overexpressed, these TFs can induce the xylogenesis developmental programme in any cell type. They control the downstream TFs as well as the molecular effectors involved directly in SCW biosynthesis. In *Arabidopsis* there are 10 master switches all belonging to NAC family, which have been also called Secondary Wall NAC master switches



(SWNs, Zhong and Ye, 2015). It has been shown that different kinds of NACs are specific of either vessel or fibre differentiation in *Arabidopsis*.

For instance, the development of vessels both in protoxylem and metaxylem is regulated by the VASCULAR-RELATED NAC-DOMAIN (VND) transcription factors which include 7 members (AtVND 1 to 7) all up-regulated during transdifferentiation of *Arabidopsis* suspension-cultured cells into tracheary elements (Kubo *et al.*, 2005; Zhou *et al.*, 2014). All of them are phylogenetically closely related and belong to subfamily “Ic” (Zhu *et al.*, 2012). When overexpressed, they induce ectopic deposition of vessel-patterned SCW and increase the transcript abundance of genes involved in SCW and xylem PCD. Actually, direct target genes of AtVND7 and AtVND6 include TFs involved in SCW regulation (AtMYB46, AtMYB83, AtMYB103, AtLBD15/ASL11 and AtLBD30/ASL19/JLO), genes related to cellulose and hemicelluloses biosynthesis, and PCD-related proteases (Yamaguchi *et al.*, 2011; Ohashi-Ito *et al.*, 2010). Interestingly, no lignin-related genes were detected with the exception of one laccase (AtLAC4) in AtVND7- overexpression experiments. VND1-7 seems to be expressed in xylary vessels but at different time points or organs (Kubo *et al.*, 2005; Zhou *et al.*, 2014; Yamaguchi *et al.*, 2008). Some hierarchical relationships seem to occur between AtVND members. All AtVNDs seem to activate AtVND7 expression by binding to SNBE/TERE-like motif in its promoter but AtVND7 is not able to significantly activate the expression of other AtVNDs in transactivation assays, with the exception of itself, suggesting that AtVND7 is a terminus of a transcriptional network within the VND family (Endo *et al.*, 2015).

The second group consists of NAC SECONDARY WALL THICKENING PROMOTING FACTOR 3/SECONDARY WALL-ASSOCIATED NAC DOMAIN PROTEIN 1 (NST3/SND1), NST1, and NST2, thought to act redundantly to activate SCW formation in both interfascicular and xylary fibres in stems and hypocotyls of *Arabidopsis* after the establishment of cell identity (Mitsuda *et al.*, 2007, 2005; Zhong *et al.*, 2006). Some evidences support this hypothesis: (i) they are not expressed in *Arabidopsis* cell cultures differentiating into tracheary elements (Mitsuda *et al.*, 2005), (ii) *Arabidopsis* double or triple mutants show very few or no SCW-containing fibres at all whereas vessel SCWs are virtually not affected, (iii) overexpression of each NST gene causes ectopic SCW deposition without changing cell shape but does not activate the expression of PCD-related or vessel marker genes. Besides, NST3 directly activates transcription factors regulating SCW formation (AtSND3, Class-II KNOTTED1-like homeobox gene AtKNAT7,

AtMYB20, AtMYB42, AtMYB43, AtMYB52, AtMYB54, AtMYB69 and AtMYB85) as well as transcription factors regulating lignin deposition (AtMYB63 and AtMYB58), SCW biosynthetic genes (AtCesA4, AtIRX9, AtPAL1, AtCCoAMT, At4CL3 as well as peroxidases and laccases) but not PCD-related genes (Zhong et al., 2010c; Ohashi-Ito *et al.*, 2010). However, the ectopic cells with SCW show tracheary element-like patterns of SCW deposition and NST1 is also expressed in early stages of vessel cells differentiation (Mitsuda *et al.*, 2005, 2007). Thus, a role in vessel SCW deposition of NST genes cannot be completely discarded.

In *Arabidopsis*, other NAC TFs, named SOMBRERO (SMB), BEARSKIN1 (BRN1) and BRN2, also can induce ectopic SCW deposition when overexpressed. However, their relationship with SCW regulation is not clear because in wild-type plants they are expressed in root caps regions where SCW is not deposited (Willemsen *et al.*, 2008; Bennett *et al.*, 2010).

All those NAC members are phylogenetically very close and thus they have been grouped in a NAC subfamily named VNS after “VND, NST/SND, SMB” -related proteins. Orthologs for VNS members are present in vascular plants as *E. grandis* (Hussey *et al.*, 2015), *P. trichocarpa*, *Vitis vinifera*, *Zea mays* and *Oryza sativa* among others, although only a few members have been characterised so far (Nakano *et al.*, 2015). It is the case of PtVNS07/PtrWND6A (ortholog of AtVND7) and PtVNS11/PtrWND1B (ortholog of AtNST3/SND1) in *P. trichocarpa* (Ohtani *et al.*, 2011). In all cases, overexpression caused ectopic deposition of SCW (Ohtani *et al.*, 2011). Differential expression of VNDs and NSTs in vessels and in fibres, respectively is observed in poplar primary xylem (Ohtani *et al.*, 2011) but, in contrast to *Arabidopsis* studies, it is not observed in secondary xylem in poplar nor in rice or maize (Zhong *et al.*, 2010b; Zhong *et al.*, 2011; Ohtani *et al.*, 2011). Poplar VNS genes target common downstream genes with their orthologs in *Arabidopsis*, but the differences observed in *Arabidopsis* on the targeting PCD genes (as AtXCP1 and 2) only by VND and not by NST genes were not observed in poplar (Ohtani *et al.*, 2011).

#### Second layer master switch: AtMYB46/AtMYB83 and respective orthologs

In *Arabidopsis*, all the tested VNS master switches have AtMYB46 (Zhong *et al.*, 2007) and its close homolog AtMYB83 (McCarthy *et al.*, 2009), as direct targets (Yamaguchi *et al.*, 2011; Ohashi-Ito *et al.*, 2010; Zhong et al., 2010c). Both genes are expressed in both proto- and metaxylem vessels and in xylary fibres as well as in interfascicular fibres and act redundantly in

activating SCW deposition. Their overexpressions activate the transcription of genes involved in cellulose, hemicellulose and lignin and induce ectopic deposition of SCW. Double mutants (*atmyb46-atmyb83*), but not simple mutants, show a severe decrease in SCW content in both xylem fibres and vessels and chimeric dominant repression proteins show the same phenotype. Among the direct target genes of AtMYB46 and AtMYB83 there are a number of downstream TFs that activate SCW biosynthesis such as AtMYB4, AtMYB7, AtMYB43, AtMYB52, AtMYB54, AtMYB58, AtMYB63 and AtKNAT7 but also SCW biosynthetic genes mainly for lignin but also for xylan, SCW modification, cytoskeletal organisation, vesicle transport and PCD genes (Zhong and Ye, 2012). Taking together, AtMYB46 and AtMYB83 are the main players of the second layer of master switches.

Actually, the orthologs of AtMYB46 and AtMYB83 were first described in *Pinus taeda* (PtMYB4, Patzlaff *et al.*, 2003) and *Eucalyptus* (EgMYB2, Goicoechea *et al.*, 2005). The corresponding *P. trichocarpa* orthologs are PtrMYB3 and PtrMYB20 (McCarthy *et al.*, 2010) and PtrMYB2 and PtrMYB21 (Zhong *et al.*, 2013). The functional orthologs of AtMYB46 have been also described in rice, OsMYB46, and maize, ZmMYB46 (Zhong *et al.*, 2011).

EgMYB2 overexpressing (EgMYB2-OE) tobacco plants which showed an increase in S/G ratio (Goicoechea *et al.*, 2005) due to an increment in S units and a higher increase of transcript abundance of genes involved in the monolignol-specific pathway and specially of those involved in ring modification, that is encoding C3H, F5H, CCoAOMT and COMT in tobacco. Besides EgMYB2 binds to EgCCR and EgCAD promoters in vitro and activates their transcription in transient expression assays (Goicoechea *et al.*, 2005) as well as activates both cellulose and xylan biosynthetic gene expression. EgMYB2 is also able to complement the *atmyb46-atmyb83* mutant (Zhong and Ye, 2009).

### Third layer of transcriptional regulation

The third layer of transcriptional regulation is composed of TFs that do not regulate other transcription factors (or at least not yet described) but molecular effectors that act over SCW formation. They mostly belong to MYB transcription factors but also there are some NAC (e.g. AtSND2, AtXND1, and AtSND3) as well as other kind of TF such as AtKNAT7. In some cases they are direct target genes of NAC master switches or AtMYB46 and AtMYB83. Here I will focus on MYB TFs.

Most of the direct target genes of NAC master switches have already been characterised. In some cases they seem to act over SCW biosynthetic genes like AtMYB52, AtMYB54 and AtMYB69. They all belong to subgroup 21 (Dubos *et al.*, 2010), which clusters, together with subgroups 22, 23 and 25, to a separate branch of the *Arabidopsis* MYB phylogenetic tree. Mutants of *Atmyb52* show a strong hyperlignification (Cassan-Wang *et al.*, 2013). This gene is co-expressed with cellulose and xylan biosynthetic genes as well as a transcription factor regulating only lignin biosynthesis (Cassan-Wang *et al.*, 2013). All together suggest that AtMYB52 is a repressor of the whole SCW program. Dominant repression of AtMYB54 and AtMYB69 showed a reduction of SCW thickness but overexpression did not show significant phenotypes (Zhong *et al.*, 2008). However, as they could activate the expression of *AtCesA8*, *AtIRX9* and *At4CL1* genes they were proposed to regulate the biosynthesis of all SCW components.

In most of the cases however, direct target genes of NAC master switches act over lignin biosynthesis and/or polymerisation. This is the case for instance of the *Arabidopsis* transcription factors AtMYB58, AtMYB63, AtMYB85 and AtMYB103. AtMYB58 and AtMYB63 are closely related and belong to subgroup 3 of R2R3MYB (Dubos *et al.*, 2010). They are both expressed in xylem fibres and vessels and their expression seem to be influenced or directly regulated by SND1/NST3 and NST1 (Ohashi-Ito *et al.*, 2010; Zhou *et al.*, 2009) and by AtMYB46/83 (Zhong and Ye, 2012). Overexpression causes ectopic lignification but not deposition of extra cellulose or hemicelluloses. Besides they activate lignin biosynthesis genes by binding their AC *cis*-elements, and not SCW *CesA* or xylan genes, (Zhou *et al.*, 2009). AtMYB85 is also expressed in both fibres and vessels of inflorescence stems. Dominant repression caused a reduction of SCW thickness and overexpression caused ectopic deposition of lignin but not cellulose nor xylan (Zhong *et al.*, 2008). AtMYB85 is phylogenetically close to AtMYB20, AtMYB42 and AtMYB43, also direct targets of NAC master switches, but their respective dominant repression in *Arabidopsis* showed no phenotype (Zhong *et al.*, 2008). The effect of AtMYB103 on SCW biosynthesis is even more restricted because it regulates specifically the ferulate-5-hydroxylase (F5H) gene and lignin composition is strongly affected in *atmyb103* mutants with a 70% decrease in S units (Ohman *et al.*, 2013). This is in agreement with the fact that SND master switches seem to regulate in *Arabidopsis* fibre SCW rich in S units as compared to vessels.

AtMYB4, AtMYB7 and AtMYB32 are directly regulated by AtMYB46 and AtMYB83 but not by NAC master switches. They belong to MYB subgroup 4 (Dubos *et al.*, 2010; Grima-Pettenati *et al.*, 2012) whose members are all characterized by an Ethylene-responsive element binding factor-associated amphiphilic repression-(EAR)-like repression motif in their C terminus (Kazan, 2006; Ohta *et al.*, 2001). EAR repressive domain, considered the most common mechanism of gene repression in plants, works through the recruitment of chromatin-remodelling factors to mediate gene silencing by changing the chromatin conformation (Kagale and Rozwadowski, 2011). AtMYB4 has been included in the hierarchical regulation network governing secondary cell wall biosynthesis although its role in regulating lignin biosynthesis has never been demonstrated. Actually, overexpressing plants and *atmyb4* mutants show no significant changes in lignin content but AtMYB4 overexpression plants showed a reduced content of sinapoyl malate, a product of phenylpropanoid pathway, as well as a repression of some genes from this pathway notably the C4H gene (Jin *et al.*, 2000). AtMYB32 repress the expression of genes involved in the phenylpropanoid and lignin pathways in flowers, affecting the composition of the pollen wall (Preston *et al.*, 2004). The mutation of AtMYB32 leads to aberrant pollen and partial male sterility. The *atmyb32* mutant shows increased expression of COMT and decreased expression of dihydro-flavonol 4-reductase (DFR) and anthocyanin synthase (AS). AtMYB4 is the closest *Arabidopsis* MYB gene of the *Eucalyptus* EgMYB1, although it is not a clear functional ortholog (Soler *et al.*, 2015).

EgMYB1, which is preferentially expressed in *Eucalyptus* xylem, is able to repress EgCCR1 and EgCAD2 expression by binding to their *cis*-regulatory elements and to decrease transcript abundance of genes related to cellulose, hemicelluloses and lignin biosynthesis. Overexpression of EgMYB1 in *Arabidopsis* and poplar showed thinner SCW containing less acid-insoluble lignin without changing the relative proportion of G and S units (Legay *et al.*, 2010). Functional characterisation of EgMYB1 orthologs in poplar, PtrMYB156 and PtrMYB221, has been performed (Lakhal, 2013). PtrMYB156 overexpression induces a reduction of lignin content in xylem cells, whereas only the dominant repression of PtrMYB221 showed a slight reduction of lignin content.

Members of subgroup 4 in monocotyledons form a separate cluster. Maize ZmMYB31 and ZmMYB42, repress non-redundantly the phenylpropanoid pathway (Fornalé *et al.*, 2010; Sonbol *et al.*, 2009; Fornalé *et al.*, 2006) and cause a strong reduction in lignin content (60% for

ZmMYB31 overexpression). ZmMYB42 but not ZmMYB31 reduces the lignin S/G ratio (Fornalé *et al.*, 2010). Like AtMYB4, both MYBs negatively regulate the accumulation of sinapoyl malate making plants more sensitive to UV radiation. ZmMYB42 also represses flavonol biosynthesis (Sonbol *et al.*, 2009) whereas ZmMYB31 increases the accumulation of anthocyanins suggesting that the reduction of lignin redirects the carbon flux towards the biosynthesis of anthocyanins (Fornalé *et al.*, 2010).

### 1.5.3 PROTEIN-PROTEIN INTERACTIONS

Protein-protein interactions seem to have a role in regulating VNS master switches activity. In *Arabidopsis*, it has been postulated that AtVND7 and AtVND6 may interact to form homo- and heterodimers with other AtVND proteins in a fine tune mechanism to regulate SCW formation (Kubo *et al.*, 2005; Yamaguchi *et al.*, 2008). Besides, the activity of AtVND7 and probably AtVND1-5 seem to be inhibited through interaction with another NAC, AtVNI2, in axial parenchyma closely associated with vessels (Yamaguchi *et al.*, 2010). This interaction would prevent the differentiation of these axial parenchyma cells into vessels.

Differences exist between *Arabidopsis* and other plants. For instance in poplar, PtVNS07/PtrWND6A (AtVND7 ortholog) activates the transcription of the ortholog of AtVNI2, a relationship that was not observed in *Arabidopsis*.

Examples of protein-protein interaction have also been found for other members of the SCW regulation network. For instance, AtMYB4 interacts with several proteins such as members of the BHLH family (BHLH012 and TT8) (Zimmermann *et al.*, 2004). An interaction with SAD2, an importin-like protein, is necessary to transport AtMYB4 to the nucleus (Zhao *et al.*, 2007). AtMYB4 is indeed able to repress its own promoter, but this ability lacks in the *sad2* mutant plants because the inability of AtMYB4 to migrate to the nucleus. Finally, the *Arabidopsis* Interactome Mapping Consortium (2011), found an interaction between AtMYB4 and the protein TOPLESS. Later AtMYB7 and AtMYB32, also containing an EAR motif, have been shown to interact with TOPLESS as well as with TOPLESS RELATED proteins (Causier *et al.*, 2012). TOPLESS is a co-repressor protein able to interact with the EAR domain to mediate repression via chromatin remodelling (Kagale and Rozwadowski, 2011).

The *Arabidopsis* KNOTTED1-LIKE HOMEODOMAIN PROTEIN 7 (AtKNAT7) interacts also with several TFs. AtKNAT7 is a transcriptional repressor directly regulated by AtMYB46 and AtMYB83 (Zhong and Ye, 2012). *Atknat7* mutants produce irx (irregular xylem) phenotype with collapsed vessels and thicker fibre cell walls (Li et al., 2011, 2012; Brown et al., 2005)(Romano et al. 2012). AtKNAT7 regulation of SCW formation is modulated by the interaction with the transcription factors AtMYB75 (PAP1), BEL1-LIKE HOMEODOMAIN6 (AtBLH6) and OVATE FAMILY PROTEIN1 (AtOFP1 and AtOFP4) (Liu et al., 2014; Bhargava et al., 2013; Li et al., 2011). The AtKNAT7 repressor activity is enhanced when interacts with AtMYB75/PAP1 (Bhargava et al., 2013). Alone, AtMYB75 (PAP1) acts as a positive regulator of anthocyanin biosynthesis and a repressor of SCW formation (Gonzalez et al., 2008; Bhargava et al., 2010). AtKNAT7 repressor activity is also enhanced when it interacts with AtBHL6 (Liu et al., 2014). Interestingly, both AtKNAT7 and AtBHL6 are able to bind to the promoter of AtREV/IFL1 and repress its expression. AtREV/IFL1 is one of the five *Arabidopsis* Class III HD-ZIP transcription factors, and has an important role in interfascicular fibres and xylem cell differentiation (Schuetz et al., 2013; Ohashi-Ito et al., 2005). Thus, AtKNAT7 and AtBLH6 can repress secondary cell wall formation through repressing this gene (Liu et al., 2014). AtOFP1 and AtOFP4 can interact with both AtKNAT7 and AtBHL6 (Liu and Douglas, 2015; Li et al., 2011), and enhance the repression activity of AtBHL6 supporting a role for these OFPs as components of a putative multi-protein transcription regulatory complex containing BLH6 and KNAT7 (Liu and Douglas, 2015). According to the authors, OFP1 and OFP4 may modulate the activity of the BLH6 - KNAT7 complex in regulating secondary wall formation in certain cell types or at specific developmental stages.









CHAPTER 2:

---

FUNCTIONAL CHARACTERIZATION OF THREE *EUCALYPTUS*  
MYB GENES CONTROLLING WOOD FORMATION AND  
EVALUATION OF THEIR IMPACTS ON SACCHARIFICATION



## Functional characterization of three *Eucalyptus* R2R3-MYB genes controlling wood formation and evaluation of their impacts on saccharification

Anna Plasencia<sup>1</sup>, Marçal Soler<sup>1</sup>, Jorge Lepikson-Neto<sup>1</sup>, Annabelle Dupas<sup>1</sup>, Romain Larbat<sup>4</sup>, Ana Alves<sup>3</sup>, Alicia Moreno-Cortés<sup>2</sup>, Natalie Ladouce<sup>1</sup>, Isabel Allona<sup>2</sup>, Isabelle Truchet<sup>1</sup>, José Carlos Rodrigues<sup>3</sup>, Fabien Mounet<sup>1</sup> and Jacqueline Grima-Pettenati<sup>1</sup>✉

<sup>1</sup> Plant Research Laboratory (LRSV), UMR5546, Toulouse III Paul Sabatier University - CNRS, Castanet Tolosan, France

<sup>2</sup> Centro de Biotecnología y Genómica de Plantas (GBGP), UPM/INIA, E-28223 Madrid, Spain

<sup>3</sup> Tropical Research Institute of Portugal (IICT), Centro de Estudos Florestais, 1349-017 Lisboa, Portugal

<sup>4</sup> Lorraine University (INPL) – INRA Agronomie et Environnement, UMR 1121, Vandœuvre-lès-Nancy, France

---

Not submitted

### Summary

Despite the huge economic importance of *Eucalyptus*, only a few regulatory genes controlling wood formation have been functionally characterized up to date. Among them, EgMYB1 and EgMYB2, were shown to be a repressor and an activator of secondary cell wall (SCW) formation, respectively. The main aim of this study was to better understand the role of these R2R3-MYB transcription factors and to functionally characterize EgMYB137, a new putative regulator of wood formation. We characterized the effects of overexpressing their dominant repression (DR) chimeric proteins over SCW composition in transgenic poplar plants. All transgenic lines obtained showed a decrease in lignin content although to different extents, being EgMYB137-DR plants the ones with the more dramatic decrease. They all also showed a decrease in oligolignols content not accompanied by an accumulation of glycosylated phenolics. Altogether, the results suggest that the whole monolignol biosynthesis is impaired with no accumulation of intermediates. Regarding lignin composition, EgMYB1-DR plants showed an increase in H units and EgMYB137-DR an increase in the S/G ratio. The observed changes in lignin content and composition in EgMYB1-DR and EgMYB137-DR lines caused significant increases in saccharification yields without pre-treatment. On the contrary, EgMYB2-DR plants showed a decrease in saccharification yield after alkali pre-treatment. In conclusion, we gained more insights into the roles of EgMYB1 and EgMYB2 and showed that EgMYB137 is a new R2R3-MYB transcription factor implicated in SCW regulation and a promising candidate for bioethanol production.

**Key words:** Secondary cell wall, saccharification, MYB genes, *Eucalyptus*

---

## 2.1 PREFACE

My contribution to this work was to perform the transactivation assays in tobacco, the poplar transformation and regeneration of the transgenic lines, as well as to produce the *in vitro* plant material. I have also taken care of the poplar lines during acclimation and hardening, and performed phenotypic observations and measurements to choose the lines for further characterization. I have coordinated the sampling and the different analyses (molecular, histological and biochemical analyses), performed statistical analyses and mined the obtained results. I also wrote the first draft of the manuscript

Marçal Soler selected EgMYB137, made all constructs, and participated to sampling and data analyses. Annabelle Dupas helped with plant transformation, production and sampling. Isabel Allona hosted me in her team in GBGP (Madrid) for one week where I was trained to poplar transformation by Alicia Moreno.

Jorge Lepikson-Neto prepared samples for biochemical analyses and performed saccharification assays. With Fabien Mounet, he analyzed lignin using Klason and thioacidolysis methods. He also performed pyrolysis analyses in IICT (Lisbon Portugal) with the collaboration of Anna Alves and José Carlos Rodrigues. Metabolite profiling was performed by Romain Larbat .

## 2.2 INTRODUCTION

Wood, also known as secondary xylem, is essentially formed by the thick secondary cell walls (SCW) from fibres and tracheary elements, which provide support to the aerial structures and allow an efficient water transport. SCW are mainly constituted of cellulose (50%), composed of linear chains of  $\beta$ -1,4 D-glucose units. Cellulose is embedded into a matrix of hemicelluloses, which are constituted of several kinds of branched polysaccharides that strengthen cell wall. However, further strengthening of this matrix happens when lignin, a complex phenolic polymer, is polymerized and crosslinked with the different polysaccharides within cell wall, thus conferring rigidity and imperviousness. SCW composition is highly variable between species, developmental stages (juvenile versus mature xylem), cell types and even between the different regions of a SCW from a single cell (Plomion *et al.*, 2001). The considerable SCW plasticity is also observed in responses to environmental cues such as for instance nitrogen

availability (Camargo *et al.*, 2014), or mechanical stress (Plomion *et al.*, 2001). For example, when woody angiosperms are submitted to bending, they can recover a vertical position by producing tension wood in the upperside of the leaning stem, while the wood produced in the underside is called opposite wood. Tension wood is strongly enriched in cellulose and generally contains less lignin (Paux *et al.*, 2004; and references therein).

Either used for pulping, sawn timber or as an energy source, wood is an important raw material for mankind reported to be the fifth most important world trade product (Plomion *et al.*, 2001). Depending on the industrial uses, different characteristics are required for wood feedstocks. For example, highly accessible cellulose is preferred for both pulp and bioethanol production, and this accessibility is greatly affected by lignin content and composition (Sykes *et al.*, 2015; Chen and Dixon, 2007). Lignin is mostly constituted of three main monomers, named guaiacyl (G), syringyl (S), and *p*-hydroxyphenyl (H) units. The precursors of these monomers are three monolignols (coniferyl, sinapyl alcohol and *p*-hydroxyphenyl alcohols, respectively) derived from the phenylpropanoid pathway even if other phenylpropanoid-derived units can also bind to the growing lignin polymer (for a review see Vanholme *et al.*, 2012; and references therein). The relative abundance of the different lignin units determines the types of linkages established between them, which in turn affect the lignin structure and SCW properties. For example, an increase in S units leads to a less condensed lignin with fewer highly stable carbon-carbon linkages such as 5–5 and  $\beta$ -5 linkages. This makes the lignin polymer more labile and facilitates access to cellulose (Studer *et al.*, 2011).

Xylem cells derive from vascular cambium, a lateral meristem that actively divides and produces daughter cells, which follow elongation, deposition of SCW, and programmed cell death to differentiate into the different xylem cell types. Cambial activity and xylem differentiation, including SCW deposition, are tightly regulated by a hierarchical network of transcription factors. Auxin transport and auxin-related genes control meristematic activity and cell fate, whereas NAC and MYB transcription factors, among others, transcriptionally regulate the activity of lignin, cellulose and hemicellulose biosynthetic genes (reviewed in Grima-Pettenati *et al.*, 2012; Nakano *et al.*, 2015; Schuetz *et al.*, 2013). The transcription factor cascade that regulates SCW formation is well studied in *Arabidopsis*. Basically, NAC genes are

top-level master switches expressed in xylem, in some cases specific from vessels or fibres. They can directly control the expression of SCW biosynthetic genes, but also of some MYB genes (Ohashi-Ito *et al.*, 2010; Yamaguchi *et al.*, 2011). In turn, MYB transcription factors can directly regulate the expression of SCW biosynthetic genes, but they can also regulate the expression of other MYB genes. Interestingly, some MYB genes like AtMYB46 and AtMYB83 have a general effect directly activating cellulose, hemicelluloses and lignin biosynthetic genes as well as other MYB regulators (Zhong and Ye, 2012). On the contrary, other MYB genes have a more restricted effect, like for example AtMYB58, AtMYB63 and AtMYB85, which only regulate lignin biosynthetic genes (Zhong *et al.*, 2008; Zhou *et al.*, 2009). In some cases the effect is even more limited, like AtMYB103. *Atmyb103* mutants contain a reduced amount of ferulate-5-hydroxylase (*AtF5H*) transcripts, a gene involved in the phenylpropanoid pathway, which causes a decrease in S lignin subunits (Ohman *et al.*, 2013).

Even if the hierarchical network of transcription factors regulating SCW formation is fairly well-known in *Arabidopsis*, also including predicted responses of some genes in response to abiotic stresses (Taylor-Teeples *et al.*, 2014), knowledge about SCW regulation in woody species is less studied. Many aspects of the regulation of xylem differentiation and secondary wall deposition are conserved between *Arabidopsis* and woody perennial plants (Chaffey *et al.*, 2002); however, a growing body of evidences highlight differences between *Arabidopsis* and trees. For example, in *Arabidopsis*, AtVNDs (VASCULAR-RELATED NAC-DOMAIN) are exclusively expressed in vessels whereas AtNSTs (NAC SECONDARY WALL THICKENING PROMOTING FACTORS) are expressed in fibres. However, in poplar, both kind of transcription factors are expressed in vessels and fibres, as also happens in the non-woody plants rice and maize (Ohtani *et al.*, 2011; Zhong *et al.*, 2010b; Zhong *et al.*, 2011). Moreover, some subgroups of R2R3-MYB genes expressed in the vascular cambium of trees are totally absent in herbaceous plants like *Arabidopsis* and rice (Soler *et al.*, 2015). Taking into account such differences and the fact that some wood features like the production of early and latewood, the spatial variation of juvenile wood and mature wood, and the development of complex vascular rays made of several specialized cell types are not occurring in an annual herbaceous like *Arabidopsis* (Chaffey *et al.*, 2002; Nieminen *et al.*, 2015), studying how secondary cell wall is regulated in woody species is necessary.



*Eucalyptus* species and hybrids are the most planted hardwoods worldwide because their fast growth, high adaptability and valuable wood properties, thus being one of the world leading sources of wood biomass (Myburg *et al.*, 2007). Moreover, since the genome of *E. grandis* is recently released (Myburg *et al.*, 2014), *Eucalyptus* trees constitutes an outstanding model for both understanding the regulation of wood formation and for improving industrial applications using tailored woods. In earlier works, we have characterized in *Eucalyptus* two antagonist regulators of the SCW, EgMYB1 and EgMYB2. EgMYB2 is a transcriptional activator of the lignin pathway (Goicoechea *et al.*, 2005), which was shown later to be a functional ortholog of the master regulator AtMYB46/83 (Zhong *et al.*, 2007; Zhong *et al.*, 2010a). EgMYB1 acts as repressor of the lignin pathway (Legay *et al.*, 2010, 2007). The closest ortholog of EgMYB1 in *Arabidopsis*, AtMYB4, was shown to be involved in the biosynthesis of phenylpropanoid-derived metabolites (Jin *et al.*, 2000) but not directly related to lignin biosynthesis. Overexpression of EgMYB1 was shown to reduce lignin content and to repress SCW biosynthetic genes in *Arabidopsis* and poplar plants without changing lignin composition probably by means of the EAR (Ethylene responsive element binding factor-associated Amphiphilic Repression, EAR) domain present at the C-terminal region of its sequence (Legay *et al.*, 2010, 2007). On the contrary, overexpression of EgMYB2 increases SCW thickness, changes lignin composition by increasing S/G ratio and activates the transcription of SCW-biosynthetic genes. However, the effects of the repression of these MYB genes in trees have not been yet studied. Furthermore, a genome-wide survey of the *Eucalyptus* R2R3-MYB family allowed the identification of 141 members in this specie (Soler *et al.*, 2015). Large scale expression studies allowed the authors to identify several *Eucalyptus* R2R3-MYB transcription factors highly and preferentially expressed in xylem as promising new candidates implied in the regulation of wood formation.

In this study, we deepened in the functional characterisation of EgMYB1 and EgMYB2. We have also used expression and phylogenetic data from Soler *et al.* (2015) to highlight EgMYB137 as a potential regulator of wood formation in *Eucalyptus* and we investigated for the first time its function. Because *Eucalyptus* transformation is a slow and inefficient process, we used another woody angiosperm, poplar, as a heterologous transformation system. For EgMYB1 and 2, we constructed dominant repression (DR) lines of these genes by overexpressing these protein fused with an EAR domain at their C-terminal region. This approach, first developed to facilitate

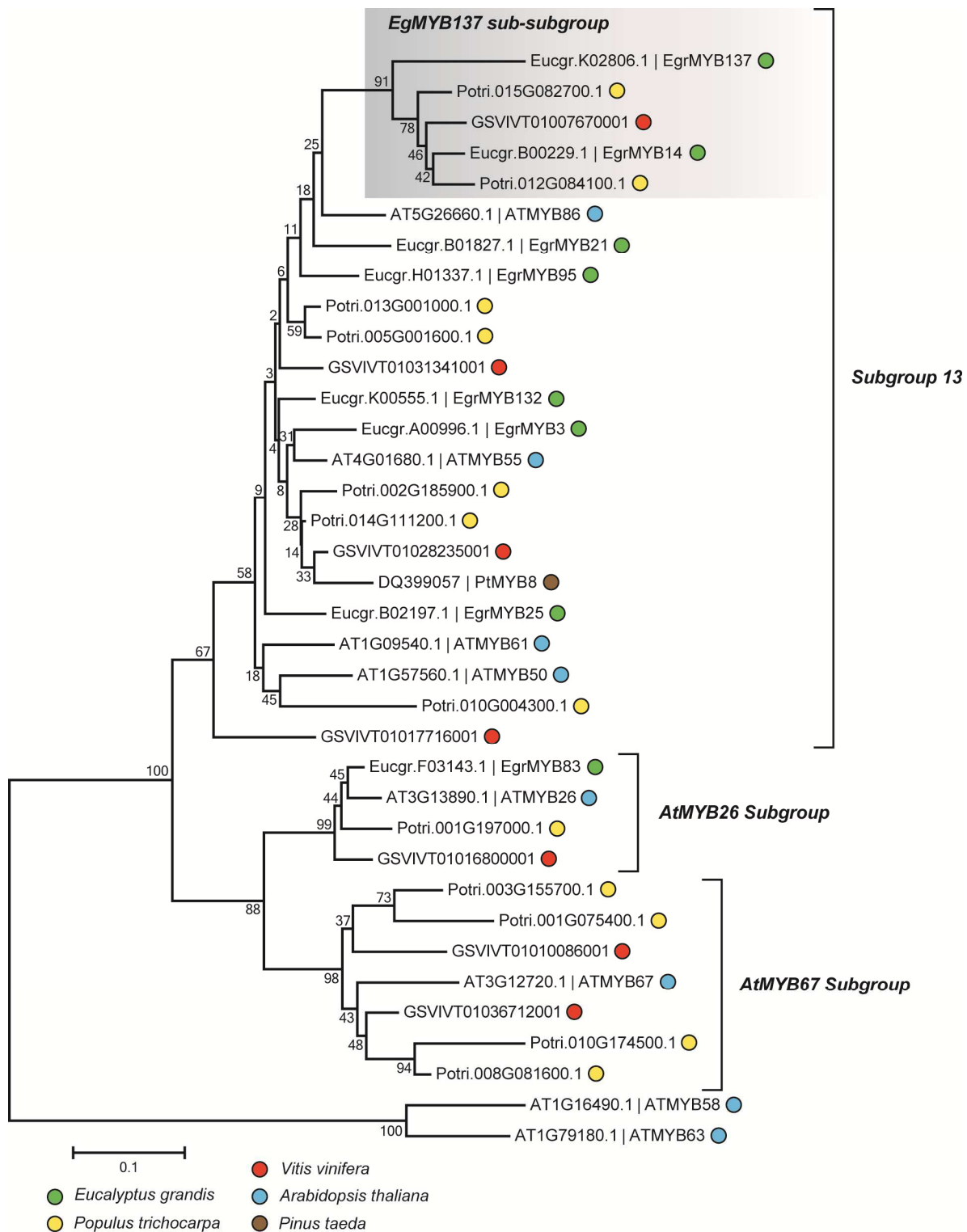
the analysis of functionally redundant transcription factors, for which gene-knock out or RNAi strategies often failed to exhibit an informative phenotype (Hiratsu *et al.*, 2003), was successfully applied to transcription factors involved in secondary wall biosynthesis (Kubo *et al.*, 2005; Mitsuda *et al.*, 2005, 2007; Zhong *et al.*, 2006, 2007). For EgMYB137, since it was never studied before, we generated both overexpressing (EgMYB137-OE) and dominant repression (EgMYB137-DR) lines, but unfortunately overexpressing plantlets were not able to produce roots and died during the plant regeneration stage. The effect of these genes in poplar transgenic lines were analysed mainly at the xylem secondary cell wall level using histochemical and biochemical methods as well as evaluating their saccharification potentials.

## 2.3 RESULTS

### 2.3.1 PHYLOGENETIC RELATIONSHIP OF EGMYB137 WITH CLOSE R2R3-MYB GENES

Soler *et al.* (2015) previously showed that EgMYB137 clusters inside the R2R3-MYB subgroup 13. We aimed to perform a more detailed phylogeny of R2R3-MYB genes of *E. grandis*, *Arabidopsis thaliana*, *Populus trichocarpa*, *Pinus taeda* and *Vitis vinifera* that are closely related to EgMYB137. We have performed a phylogenetic tree (Figure 2.1) with sequences belonging to subgroup 13+AtMYB26&AtMYB67 described in Soler *et al.* (2015) and a related R2R3-MYB sequence from *P. taeda*, PtMYB8, from Bedon *et al.* (2007).

The resulting phylogenetic tree showed that EgMYB137 had no clear ortholog in *Arabidopsis*. It was only distantly related to the already characterised R2R3-MYB genes *AtMYB61* and *AtMYB26* from *Arabidopsis* and to *PtMYB8* from *P. taeda*. The closely-related R2R3-MYB genes to EgMYB137 were *EgrMYB14* from *E. grandis*, Potri.015G082700.1 and Potri.012G084100.1 from *P. trichocarpa* and GSVIVT01007670001 from *V. vinifera*. To our knowledge, none of them have been characterised so far.

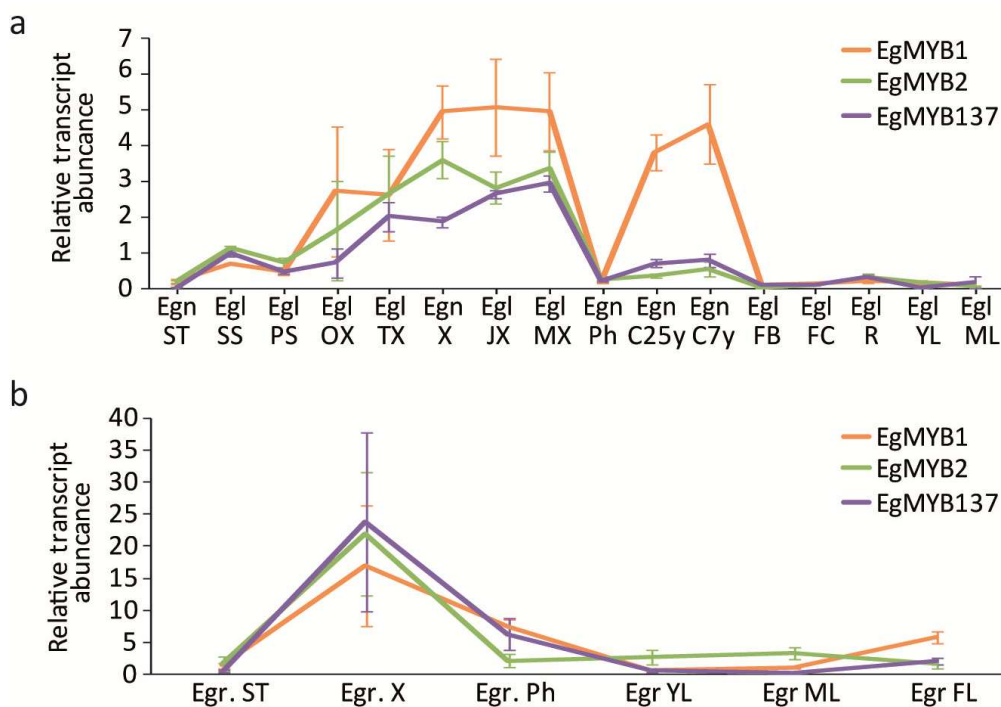


**Figure 2.1:** Neighbor-Joining phylogenetic tree of R2R3-MYB proteins from *Eucalyptus grandis*, *Arabidopsis thaliana*, *Populus trichocarpa*, *Vitis vinifera* and *Pinus taeda* closely related to EgMYB137. It also includes AtMYB58 and AtMYB63 R2R3-MYB proteins from *Arabidopsis* used to root the tree. Bootstrap values are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree (amino acid substitutions per site).

### 2.3.2 EXPRESSION PROFILING OF THE THREE STUDIED MYB GENES

Soler *et al.* (2015) performed a large expression profiling of the R2R3-MYB family in *Eucalyptus* using more than 16 distinct experimental sets covering 13 organs and tissues, as well as various developmental stages and environmental conditions. EgMYB1, EgMYB2 and EgMYB137 were classified inside a xylem supercluster of MYB genes because of their high and preferential expression in xylem tissues.

We extracted from this study, the relative transcript abundance of EgMYB1, EgMYB2 and EgMYB137 in a wide range of tissues and conditions using Microfluidic RT-qPCR or RNAseq (figure 2.2a and 2.2b). EgMYB2 and EgMYB137 were specific from xylem tissues whereas EgMYB1 was also highly expressed in cambium. Interestingly, EgMYB137 was significantly more expressed in tension than in opposite wood ( $p$ -value= 0.0175).

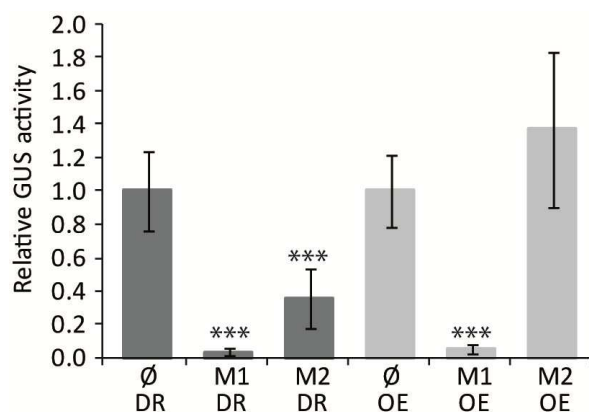


**Figure 2.2: Mean relative transcript abundance of EgMYB1, EgMYB2 and EgMYB137 genes.** Microfluidic (a) or RNAseq (b) transcript abundance for EgMYB1, EgMYB2 and EgMYB137 genes are shown. Mean values  $\pm$  standard deviation are expressed in relation to the mean of all samples (non  $\log_2$ -transformed), or as standardized fragments per kilobase of exon per million fragments mapped (FPKM) respectively. ST, shoot tips; SS, secondary stem; PS, primary stem; OX, opposite xylem; TX, tension xylem; X, xylem; JX, juvenile xylem; MX, mature xylem; Ph, phloem; C25y, cambium from 25-yr-old trees; C7y, cambium from 7-yr-old trees; FB, floral buds; FC, fruit capsule; R, lateral roots; YL, young leaves; ML, mature leaves; FL, flowers; Egl, *E. globulus*; Egn, *E. gunnii* x *E. dalrympleana*; Egr, *E. grandis*. Data extracted from Soler *et al.* (2015).

### 2.3.3 EFFECT OF THE EgMYB1 AND EgMYB2 DOMINANT REPRESSION CHIMERIC PROTEINS ON EgCAD2 PROMOTER TRANSCRIPTIONAL ACTIVITY

Both EgMYB1 and EgMYB2 are able to *in vitro* bind to cis-regulatory regions of the EgCAD2 promoter to transcriptionally repress or activate, respectively, its expression (Goicoechea *et al.*, 2005; Legay *et al.*, 2007). We further compared the transcriptional activities of the native forms (EgMYB2-OE and EgMYB1-OE), versus the dominant repression chimeric proteins (EgMYB2-DR and EgMYB1-DR) using the EgCAD2 promoter-GUS as reporter construct. This was not done for EgMYB137 because its target genes are unknown.

Forty-eight hours after co-transfection in *N. benthamiana* leaves, the activity of EgCAD2 promoter was quantified by measuring GUS activities. The results were compared to those obtained by co-transfecting the EgCAD2 promoter with empty vectors. Consistent with its repressor role (Legay *et al.*, 2007), EgMYB1-OE strongly repressed the EgCAD2 promoter resulting in a GUS activity of only 5.5% as compared to control (Figure 2.3). A slightly stronger repression was observed when using the EgMYB1-DR construct where only 3.6% of the GUS activity measured in the control was still detected. EgMYB2-OE behaved as a transcriptional activator as previously shown (Goicoechea *et al.*, 2005), causing an average increase of the GUS activity of 37% as compared to control. On the contrary, EgMYB2-DR caused a repression of 64.4% of the EgCAD2 promoter activity (Figure 2.3).



**Figure 2.3: Effects of EgMYB1 and EgMYB2 on the transcriptional activity of EgCAD2 promoter.** Results from co-transfection experiments in tobacco leaves. Overexpression of EgMYB1 or EgMYB2 either fused to the dominant repression (DR) peptide (M1DR and M2DR, respectively) or as the native form (M1OE and M2OE, respectively) were used as effectors. EgCAD2 promoter fused to GFP-GUS fusion protein was used as reporter. Data represent mean values  $\pm$  standard deviation (SD) of three independent experiments containing three technical replicates each. Relative GUS activity is normalised using the respective empty vector ( $\emptyset$ ). \*\*\**p*-value < 0.001.

These results demonstrated that the dominant repression constructs turned EgMYB2 into a repressor by the presence of the EAR motif fused to its C terminal region. For EgMYB1, the

strong repressing effect of the native form makes it difficult to see an enhancement of this effect when fused to an additional EAR motif.

#### 2.3.4 IMPACT OF EgMYB2-DR, EgMYB1-DR AND EgMYB137-DR OVEREXPRESSION ON POPLAR GROWTH AND DEVELOPMENT

We transformed poplar (*Populus tremula* x *P. alba*, clone INRA 717-1-B4) with six different constructions: three to overexpress the dominant repression forms of EgMYB1, EgMYB2, EgMYB137 (EgMYB1-DR, EgMYB2-DR, EgMYB137-DR, respectively), one to overexpress the native form of EgMYB137 (EgMYB137-OE) and two with corresponding empty vectors (control-OE and DR).

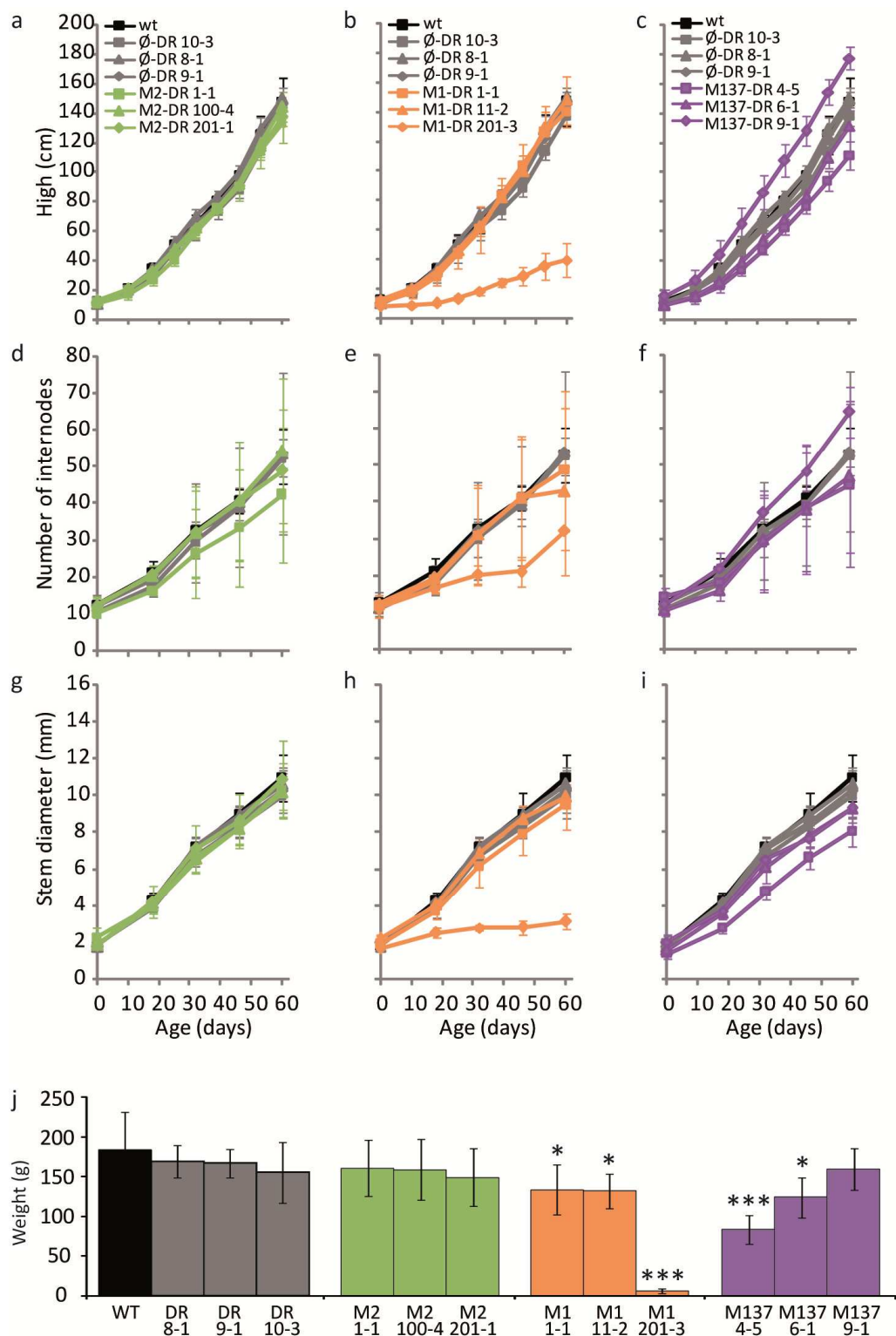
Although transformations of control-OE and EgMYB137-OE were performed at the same time, EgMYB137-OE transgenics did not regenerate whereas the control-OE plants regenerated without problems. This suggests deleterious effects of the EgMYB137-OE transgene expression during the regeneration process. Noteworthy, similar observations were previously made for the EgMYB2-OE construct in poplar plants (Legay, 2008) for which no transformants could be regenerated. The rest of the lines were regenerated without remarkable problems.

We did not observe any macroscopic phenotypes on the *in vitro* cultivated plantlets for any of the obtained lines. So, in order to select lines to be further analyzed, we estimated by RT-qPCR the steady state levels of the different transgenes in at least 14 plants per construction (Supplemental figure S2.1). Although all regenerated plants were resistant to the selective antibiotic (hygromycin 20 mg/L), only 50% or 67% of the tested plants expressed EgMYB137-DR or EgMYB2-DR respectively whereas all the tested plants expressed EgMYB1-DR transgene.

For each construction, we selected the 3 lines with the highest levels of transgene relative expression, i.e. lines 1-1, 11-2, and 201-3 for EgMYB1-DR; lines 1-1, 100-4, and 201-1 for EgMYB2-DR and lines 4-5, 6-1 and 9-1 for EgMYB137-DR. Ten plants per line were acclimated for three weeks and the five showing the most homogeneous phenotype were transferred to the greenhouse and grown in a fully randomized experimental design.

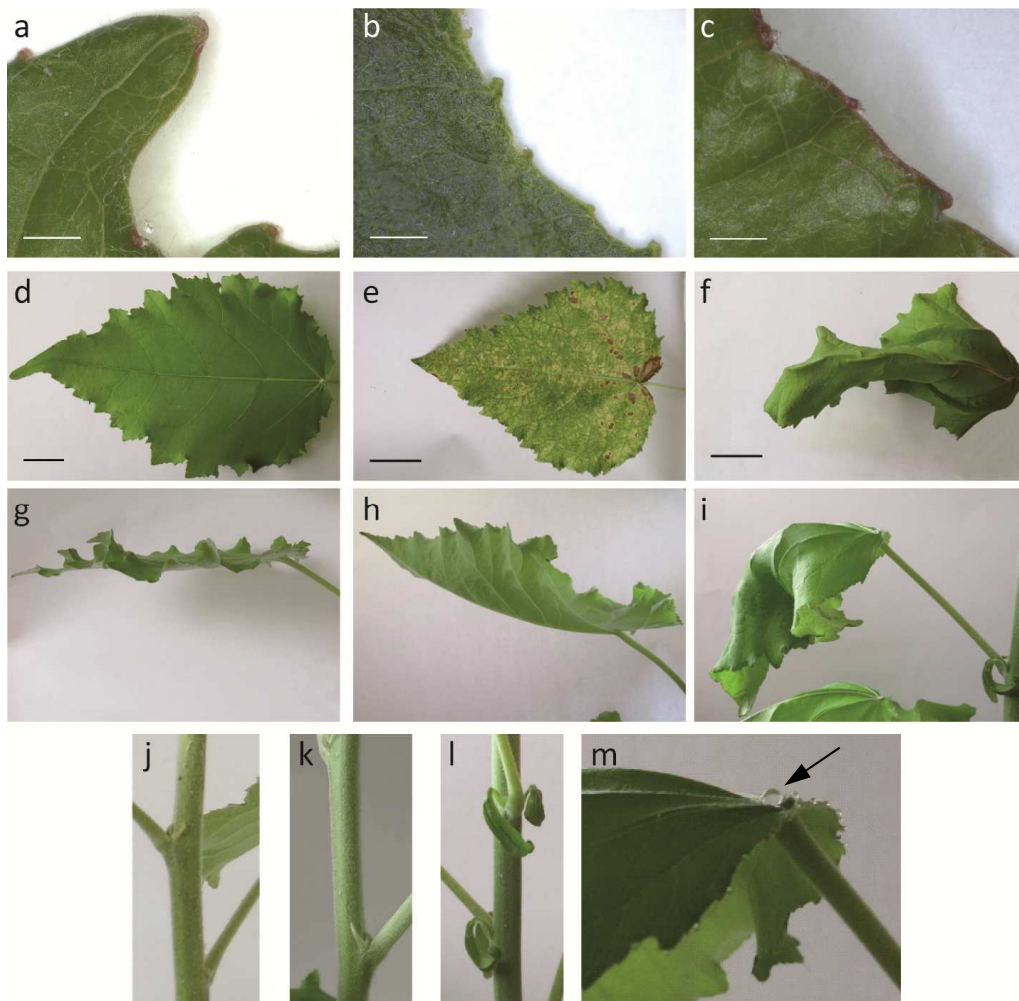
After two months of growth in greenhouse conditions, no obvious phenotype was observed for EgMYB2-DR lines as compared to controls. The three lines were indeed very homogeneous in terms of height, stem diameter, number of internodes and fresh biomass (Figure 2.4). For EgMYB1-DR, two lines were similar to controls in terms of height and stem diameter, whereas the plants from the third one (line 201-3) were much smaller with a 73% height reduction (p-value=  $6 \cdot 10^{-16}$ ) and 70% reduction of stem diameter (p-value=  $4 \cdot 10^{-13}$ ) at the end of the experiment (Figure 2.4). All three lines had lower fresh biomass although to a different extent (-96% for line 201-3, around -20% for lines 1-1 and 11-2). All three EgMYB1-DR lines also showed a consistent leaves phenotype (Figure 2.5). Young and mature leaves were smaller and curled upward in a cup shaped manner (Figure 2.5e, h). Old leaves underwent premature senescence (Figure 2.5e): they were usually more yellowish and felt prematurely as compared to control plants. In all cases, hydathodes were green coloured instead of red (Figure 2.5b). Because they produced tiny amounts of biomass, plants from line 201-3 EgMYB1-DR were not used for biochemical analyses.

EgMYB137-DR lines 4-5 and 6-1 were slightly but significantly smaller than controls after the two month of hardening (-24% and -10% respectively, p-values=  $10^{-6}$  and 0.0136 respectively), while line 9.1 was significantly taller (22% higher than controls, p-value= $2 \cdot 10^{-6}$ , Figure 2.3c). This was the tallest line measured in all experiments. All three lines also had smaller stem diameters (Figure 2.4i) with reductions of 23, 11 and 11% respectively (p-values=0.0001, 0.0399 and 0.0414, respectively). In addition, lines 4-5 and 6-1 showed a significant reduction of the final fresh biomass (-50 and -27%, respectively), while line 9.1 had no significant differences in comparison with controls (Figure 2.5j). EgMYB137-DR plants exhibited a consistent leaves phenotype visible in all lines. Leaves were curled downwards and their main vascular nervure was not straight (Figure 2.5f, i). In addition, guttation droplets were observed at the hydathodes (Figure 2.5m) and stipules were larger than control lines (Figure 2.5l).



**Figure 2.4: Growth phenotypes of overexpressing EgMYB2-DR, EgMYB1-DR and EgMYB137-DR poplar transgenic lines.** Mean values  $\pm$  standard deviation for height (a-c), number of internodes (d-f) and stem diameter (g-i) along the hardening process and final fresh weight (j) were calculated for each transgenic line and for wild type (WT) plants using five plants per line. Controls: wild type (black) and 3 transgenic lines transformed with empty-DR vectors (grey); M2-DR: EgMYB2-DR (green), M1-DR: EgMYB1-DR (orange); and M137-DR: EgMYB137-DR (purple) transgenic lines. \**p*-value < 0.05; \*\*\**p*-value < 0.001.

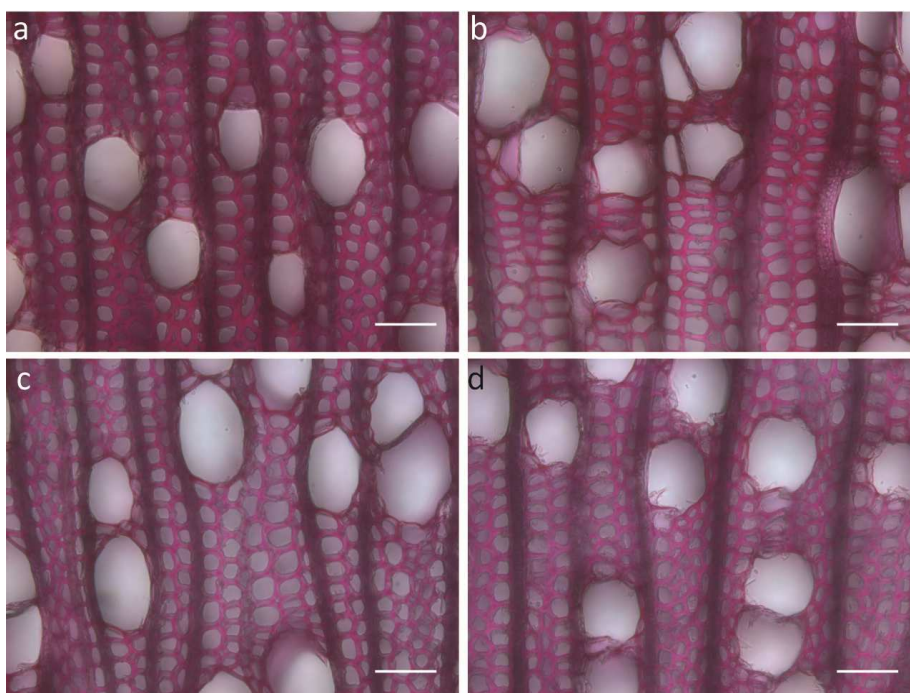




**Figure 2.5: Morphology of the leaves of transgenic poplar plants.** Representative pictures for controls (WT+empty vectors), EgMYB1-DR and EgMYB137-DR leaves. EgMYB2-DR pictures are not included because no significant differences with the controls were observed. A detail of hydathodes was taken with binocular loupe (a-c). Adaxial part of old leaves (d-f), a side view of mature leaves (g-i), a detail of stipules (j-l) and of a EgMYB137-DR leaf presenting guttation (arrow in m) are also shown.

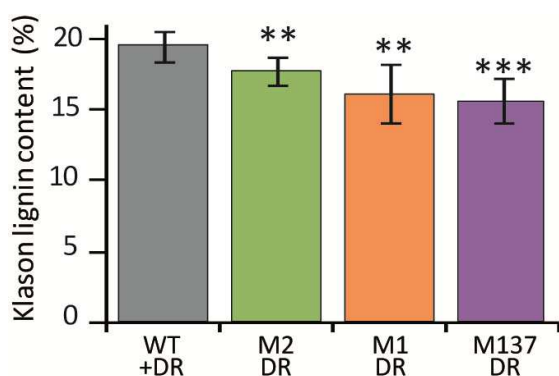
### 2.3.5 HISTOLOGICAL AND CHEMICAL ANALYSIS OF THE LIGNIN CONTENT FROM TRANSGENIC POPLAR STEMS

As a first step to study xylem structure and evaluate lignin content in the secondary cell walls (SCWs) of xylem cells, we stained stem sections with Phloroglucinol-HCl, a commonly used lignin stain. As shown in figure 2.6, xylem structure was similar between all studied transgenic lines and control plants, whereas all transgenic lines exhibited a weaker red staining suggesting lower lignin content. The most affected lines were those expressing EgMYB137-DR followed by the EgMYB1-DR lines. A slight difference in staining was observed between EgMYB2-DR and control sections.



**Figure 2.6: Transversal sections of poplar stems overexpressing EgMYB2-DR, EgMYB1-DR and EgMYB137-DR.** Sections were made at the bottom of the stem from WT (a) EgMYB2-DR (b), EgMYB1-DR (c) and EgMYB137-DR (d) poplar plants. Lignin is visualised in red by phloroglucinol staining; scale bars = 30  $\mu$ m.

This reduction of the lignin content was confirmed by quantifying the lignin content of the debarked stems using the Klason gravimetric method (Figure 2.7). In control lines, acid-insoluble lignin accounted for  $19.5 \pm 1.0\%$  of the extractive cell wall residue. The percentage of acid-insoluble lignin significantly decreased to  $17.8 \pm 1.0\%$  (-8.7%),  $16.2 \pm 2.1\%$  (-16.9%) and  $15.7 \pm 1.6\%$  (-19.5%) in EgMYB2-DR, EgMYB1-DR and EgMYB137-DR transgenic plants, respectively.



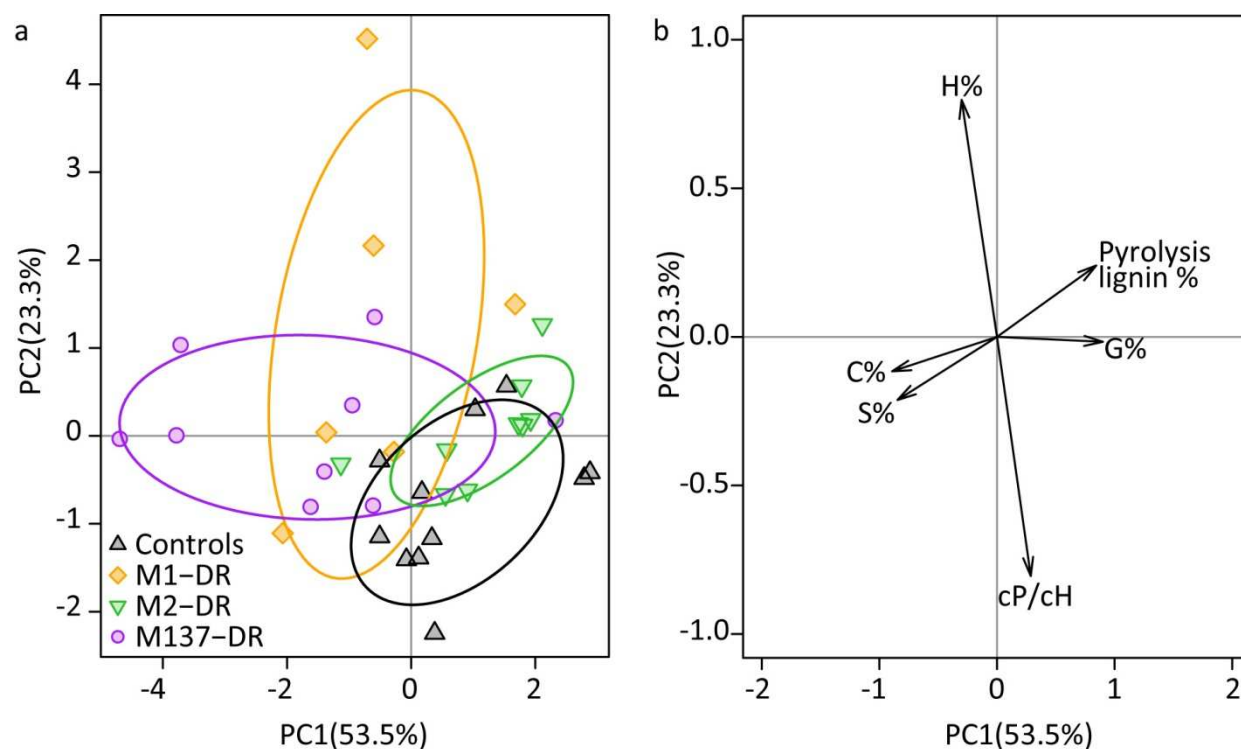
**Figure 2.7: Klason lignin content in poplar stems.** Quantification of acid-insoluble lignin, or Klason lignin, calculated as weight percentage of the extractive free wood. Data represent mean values  $\pm$  standard deviation (SD) of three lines containing three plants each. \*\**p*-value < 0.01; \*\*\**p*-value < 0.001.

### 2.3.6 SECONDARY CELL WALL COMPOSITION

#### Analytical Pyrolysis- Gas Chromatography / Mass Spectrometry (Py-GC/MS)

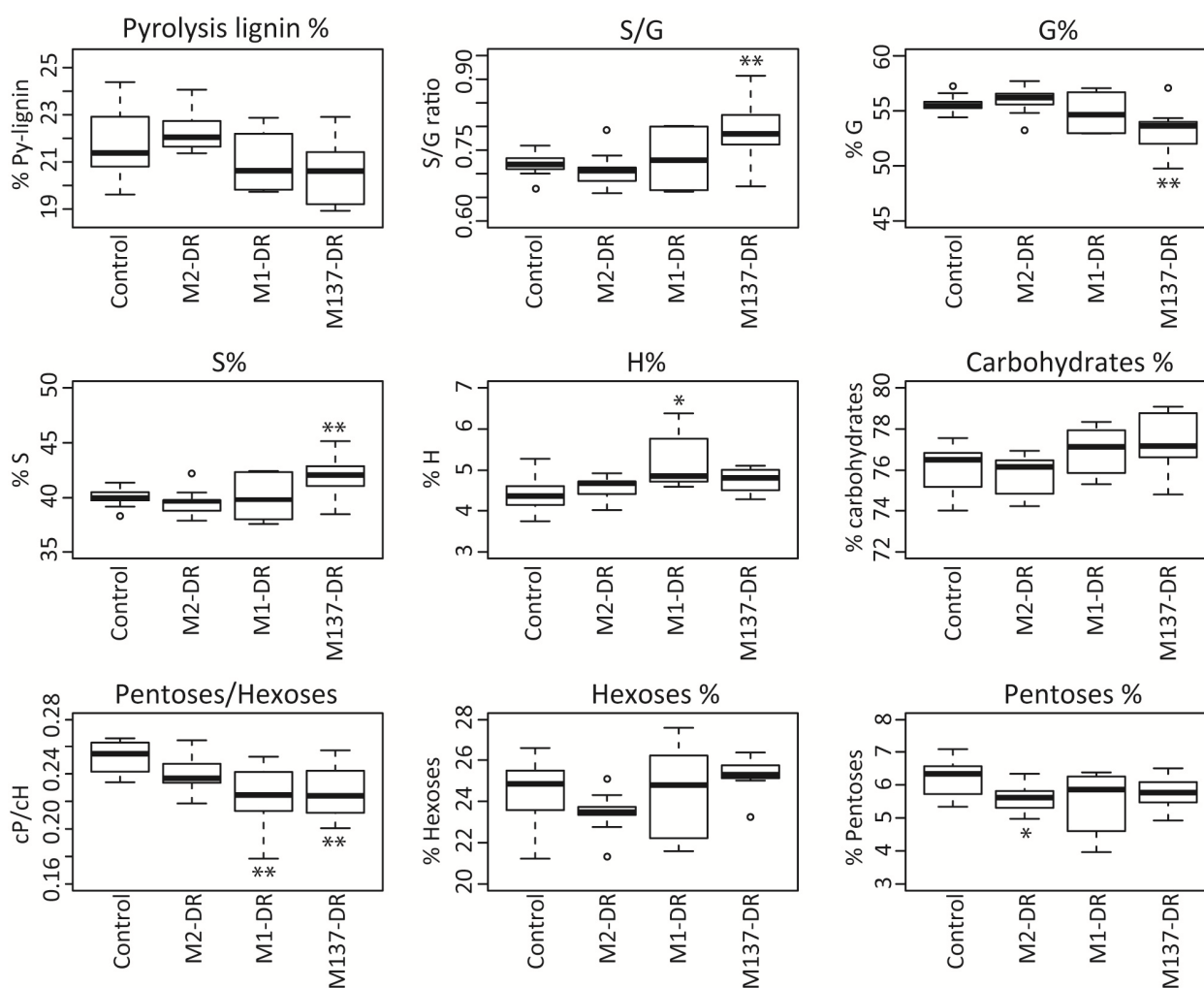
In order to get the most complete picture of the biochemical changes occurring in the SCWs, we first performed Pyrolysis - Gas Chromatography / Mass Spectrometry (Py-GC/MS) analyses of the debarked poplar stems. Pyrogram spectra were collected and the area of each peak was measured and assigned to the following categories: hexoses (cH), pentoses (cP), carbohydrates (cH + cP + unassigned carbohydrates, C), G lignin (G), S lignin (S), H lignin (H) and Pyrolysis lignin (G+S+H, Py-lignin). The identification of pyrolysis products used for quantification of the different categories is presented in supplemental table S2.1.

We used Py-lignin, G, S, H lignin and Carbohydrates (C) percentages as well as the Pentoses / Hexoses (cP/cH) ratio to perform a Principal Component Analysis (PCA) (Figure 2.8a).



**Figure 2.8: Principal component analysis (PCA) for Py-GC/MS data in poplar stems.** The percentage of Py-lignin, of the three lignin monomers (%H, %S, %G) and of carbohydrates (%C) as well as the ratio of pentoses vs hexoses -were used to perform PCA analyses. Both the score (a) and the loading (b) plots for the two main principal components are shown. Principal component 1 (PC1) explained 53.5% and PC2 explained 23.3%. Controls: wild type and 3 transgenic lines transformed with empty-DR vectors; M1-DR: Two EgMYB1-DR transgenic lines; M2-DR and M137-DR: Three EgMYB2-DR and three EgMYB137-DR transgenic lines, respectively.

The first principal component (PC1, 53.5%) separated EgMYB1-DR and EgMYB137-DR plants from controls (Figure 2.8a). In contrast, EgMYB2-DR plants were close to controls in the PCA score plot. To identify which categories of compounds were responsible for the differences between plants, we calculated the PCA loading plot (Figure 2.8b). We also calculated box plots for each of the categories used in the PCA and for the S/G ratio, the cP% and the cH% (Figure 2.9).



**Figure 2.9: Box plots for the Py-GC/MS data in poplar stems.** All data are shown as a percentage except the S/G and Pentoses/Hexoses (cP/cH) ratios. \**p*-value < 0.05; \*\**p*-value < 0.01; \*\*\**p*-value < 0.001. M2: EgMYB2; M1: EgMYB1; M137: EgMYB137.

Both EgMYB1-DR and EgMYB137-DR plants showed a decrease tendency for Py-lignin consistent with the observed decrease in Klason lignin. However, the effects of the two transgenes on lignin composition were clearly distinct. The SCWs of EgMYB137-DR plants

exhibited a slight but significant decrease in G percentage (-2.4%) and a concomitant increase in S percentage (2.1%) resulting in an increase of the S/G ratio (7.2%). On the contrary, in EgMYB1-DR plants, the G and S relative contents, and consequently the S/G ratio, remained unchanged indicating that the transgene expression similarly reduced the S and G content of the lignin polymer. Noteworthy, a significant increase in H content (8.1%) appeared to be a specific trait of EgMYB1-DR plants, separating them from the rest of the samples in the PCA analysis (Figure 2.8a and b). In plants overexpressing the chimeric protein EgMYB2-DR, no significant changes in lignin composition or content was noticeable.

Regarding the composition of carbohydrates, both EgMYB1-DR and EgMYB137-DR plants showed a decrease in their cP/cH ratio (-12% and -10%, respectively) mainly due to a reduction in pentoses abundance, likely reflecting a decrease in the hemicelluloses content. In EgMYB2-DR plants, only a significant decrease in pentoses percentage was observed (-9%).

#### Analysis of Lignin-Derived Monomers Released by thioacidolysis

The reaction of thioacidolysis primary cleaves the labile  $\beta$ -O-4 ether bonds, which are the major interunit linkages in native lignins (Lapierre, 1993). Thus, we have only measured the G and S units involved in the uncondensed lignin fraction (we could not measure the marginal H units involved in uncondensed fraction). The thioacidolysis yields (Table 2.1) were expressed relative to the extracted cell wall residue ( $\mu\text{mol/g CW}$ ) or to the Klason lignin ( $\mu\text{mol G+S/g KL}$ ).

**Table 2.1: Thioacidolysis yields for extract-free wood** from controls (wild type + empty vector lines) and transgenic poplar plants. Thioacidolysis lignin yields, are expressed in  $\mu\text{moles}$  per gram of extract-free sample ( $\mu\text{mol G+S/g CW}$ ) or in  $\mu\text{moles}$  per gram of acid insoluble lignin percentage ( $\mu\text{mol G+S/g KL}$ ,  $\mu\text{mol G/g KL}$  or  $\mu\text{mol S/g KL}$ ). Mean values  $\pm$  SD and t-test p-values are shown. KL: Klason, CW: extract-free sample (Cell Walls).

	$\mu\text{mol G+S/ g KL}$			$\mu\text{mol G+S/g CW}$			$\mu\text{mol G/g CW}$			$\mu\text{mol S/g CW}$			%G			%S			S/G		
	mean	SD	p-value	mean	SD	p-value	mean	SD	p-value	mean	SD	p-value	mean	SD	p-value	mean	SD	p-value	mean	SD	p-value
Controls	3166	$\pm$ 460		609	$\pm$ 84		234	$\pm$ 32		375	$\pm$ 56		38,4	$\pm$ 1,9		61,6	$\pm$ 1,9		1,61	$\pm$ 0,13	
EgMYB2-DR	3170	$\pm$ 392	0,9854	562	$\pm$ 62	0,1935	209	$\pm$ 27	0,0901	353	$\pm$ 39	0,3377	37,2	$\pm$ 2,0	0,1863	62,8	$\pm$ 2,0	0,1863	1,70	$\pm$ 0,14	0,1699
EgMYB1-DR	2943	$\pm$ 433	0,3641	<b>475</b>	<b><math>\pm</math> 95</b>	<b>0,0131</b>	<b>176</b>	<b><math>\pm</math> 32</b>	<b>0,0043</b>	<b>299</b>	<b><math>\pm</math> 65</b>	<b>0,0303</b>	37,5	$\pm$ 2,5	0,4333	62,5	$\pm$ 2,5	0,4333	1,69	$\pm$ 0,16	0,2782
EgMYB137-DR	3528	$\pm$ 944	0,3284	538	$\pm$ 176	0,2961	195	$\pm$ 74	0,1767	342	$\pm$ 102	0,4195	<b>35,7</b>	<b><math>\pm</math> 2,6</b>	<b>0,0320</b>	<b>64,3</b>	<b><math>\pm</math> 2,6</b>	<b>0,0320</b>	<b>1,81</b>	<b><math>\pm</math> 0,22</b>	<b>0,0328</b>

The total thioacidolysis yields ( $\mu\text{mol G+S/CW}$ ) were reduced in all transgenics as compared to controls but the reduction was found to be statistically significant only in EgMYB1-DR plants (-22%). However, none of the plants showed a significant reduction of the thioacidolysis yields when expressed relative to Klason lignin. Taken together these results suggest that the " $\beta$ -O-4

linked-G and S monomers" decrease is proportional to the reduction of Klason lignin. Finally, only EgMYB137-DR plants showed 12% increase in the S/G ratio.

### 2.3.7 METABOLITE PROFILING









We carried out metabolite profiling to identify differentially accumulating phenylpropanoid pathway intermediates and products other than lignin. Phenolic metabolites were extracted using methanol and analysed by Ultra-Performance Liquid Chromatography / Mass Spectrometry (U-HPLC/MS). Twenty five peaks were detected in the chromatograms, twelve of which could be identified using the corresponding retention time (RT) and mass spectra, and the majority of them corresponded to oligolignols or phenolic glycosides (Table 2.2, supplemental table S2.2). The sum of the recovered peak areas is indicative of the amount of the detected phenolic metabolites. EgMYB2-DR plants contained a similar amount of phenolic metabolites than the control whereas EgMYB1-DR and EgMYB137-DR plants had a reduction of 15% and 19% respectively (Table 2.2).

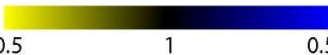
When a PCA using the recovered peak areas from samples was performed (Supplemental figure S2.2), the first principal component (PC1, 32.9%) clearly separates transgenics from control samples. Even though the same phenolic metabolites were detected in both types of samples, their relative abundance was lower in the transgenics except for sinapaldehyde and peaks 4, 5 and 9. Significant differences with controls were obtained for peaks highlighted in bold on table 2.2. Interestingly, all the transgenics have a significant reduction in oligolignols content (table 2.3) reaching a reduction of 17%, 20% and 34% for EgMYB2-DR, EgMYB1-DR and EgMYB137-DR, respectively. In EgMYB2-DR plants this reduction was mainly due to a decrease in two isomers of G(8-O-4)SP(8-5)G' (G: coniferyl alcohol, SP: sinapyl *p*-hydroxybenzoate, G': coniferaldehyde, -30% and -27%). EgMYB1-DR and EgMYB137-DR plants exhibited a significant reduction in G(8-O-4)S(8-5)G (S: sinapyl alcohol; -33% and -29% respectively). Additionally, EgMYB137-DR plants had a reduction of the two isomers of G(8-O-4)SP(8-5)G (-29% and -22% respectively) and of two unknown oligolignols corresponding to peaks 24 (-51%) and 25 (-31%). Concerning phenolic glycosides, transgenic plants did not have a reduction in their total amounts (table 2.3) although EgMYB1 showed a reduction of populoside C (-26%) and EgMYB137-DR plants did have a significant reduction of both populoside C and salicortin (-27% and -48%, respectively).

**Table 2.2: Metabolic profiling of EgMYB2-DR, EgMYB1-DR and EgMYB137-DR poplar plants.** Mean peaks area for 1µl of extract ± SD were calculated using one plant per line. Students test were performed to explore differences with control (WT and empty-DR vectors transgenic lines). p-values < 0.05 are highlighted in bold. A ratio in relation with the controls are shown in a colour code. Yellow squares indicate lower presence of the compound whereas blue squares represent an increment. The identified compounds are also shown. G(8-O-4)S(8-5)G = coniferyl alcohol(8-O-4)synapyl alcohol(8-5)coniferyl alcohol; G(8-O-4)S(8-5)G' = coniferyl alcohol(8-O-4)synapyl alcohol(8-5)coniferaldehyde; G(8-O-4)G(8-O-4)S(8-8)S = coniferyl alcohol(8-O-4)coniferyl alcohol(8-O-4)synapyl alcohol(8-5)synapyl alcohol; G(8-O-4)SP(8-5)G = coniferyl alcohol(8-O-4)synapyl *p*-hydroxybenzoate(8-5)coniferyl alcohol; G(8-O-4)SP(8-5)G' = coniferyl alcohol(8-O-4)synapyl *p*-hydroxybenzoate(8-5)coniferylaldehyde. Peaks 20-21 and 22-23 are two isomers of the same compound.

Identified compound	Peak ID	RT	EgMYB2-DR			EgMYB1-DR			EgMYB137-DR			Controls	
			Mean	SD	p-value	ratio	Mean	SD	p-value	ratio	Mean	SD	
	1	4,154	18216 ± 3246	0,1067	21452 ± 11918	0,3976	20658 ± 11689	0,3034	29883 ± 9683				
	2	4,562	51298 ± 4100	0,4059	49804 ± 2043	0,3949	56124 ± 9699	0,9248	56865 ± 9816				
	3	4,752	12295 ± 3689	0,7866	9776 ± 4415	0,6917	14180 ± 4942	0,9277	15143 ± 16582				
	4	5,194	55924 ± 9250	0,1891	44314 ± 6655	0,7720	49068 ± 8452	0,6669	46293 ± 7597				
	5	5,579	21744 ± 7474	0,2442	16641 ± 2130	0,9973	18213 ± 4352	0,5596	16634 ± 2371				
	6	5,765	51064 ± 4934	0,3325	51646 ± 14351	0,5226	45236 ± 7671	0,1526	59842 ± 13236				
Salicortin	7	6,160	35269 ± 5602	0,0623	35828 ± 4858	0,1325	<b>31249 ± 9733</b>	<b>0,0479</b>	59644 ± 16611				
sinapaldehyde	8	6,440	386062 ± 17049	0,0937	354897 ± 13880	0,6567	379628 ± 130812	0,6018	342872 ± 32470				
	9	6,595	14923 ± 4212	0,3863	22551 ± 23315	0,7369	39631 ± 24585	0,1490	18612 ± 5600				
Saillireposide	10	6,929	55784 ± 7990	0,2765	56668 ± 293	0,3933	59228 ± 13443	0,4305	70461 ± 19240				
Populoside C	11	7,246	103409 ± 9819	0,1819	<b>85972 ± 8465</b>	<b>0,0363</b>	<b>85938 ± 11752</b>	<b>0,0205</b>	116957 ± 12411				
	12	7,711	48381 ± 18396	0,6037	31284 ± 1604	0,1271	28940 ± 7309	0,0531	55810 ± 16995				
G(8-O-4)S(8-5)G	13	7,757	95111 ± 17128	0,2263	<b>74726 ± 3433</b>	<b>0,0246</b>	<b>79056 ± 8651</b>	<b>0,0169</b>	111055 ± 13646				
	14	7,863	<b>22982 ± 6572</b>	<b>0,0482</b>	25995 ± 245	0,1237	<b>24057 ± 3572</b>	<b>0,0431</b>	38874 ± 8827				
G(8-O-4)S(8-5)G'	15	8,107	62157 ± 11007	0,4439	51560 ± 8231	0,2291	58230 ± 19511	0,3757	72558 ± 19159				
	16	8,181	4328 ± 1129	0,6721	3340 ± 1757	0,4775	3538 ± 1922	0,4486	5132 ± 2879				
G(8-O-4)G(8-O-4)S(8-8)S	17	8,354	19449 ± 1658	0,3428	17609 ± 3818	0,2823	19135 ± 5152	0,4071	22574 ± 4856				
Tremulacin	18	8,835	29844 ± 7805	0,2953	47524 ± 1243	0,9829	31088 ± 3628	0,3163	47111 ± 24148				
	19	9,133	13212 ± 455	0,2895	12675 ± 412	0,2148	14072 ± 2326	0,9290	14201 ± 1361				
G(8-O-4)SP(8-5)G	20	9,708	75905 ± 19245	0,1737	91748 ± 27142	0,7835	<b>68173 ± 6120</b>	<b>0,0318</b>	96603 ± 15488				
G(8-O-4)SP(8-5)G'	21	9,852	38636 ± 7758	0,0927	50921 ± 14950	0,8825	<b>38535 ± 2015</b>	<b>0,0354</b>	49654 ± 6360				
G(8-O-4)SP(8-5)G'	22	9,989	<b>17140 ± 3116</b>	<b>0,0166</b>	22013 ± 1203	0,2555	16735 ± 6885	0,0837	24560 ± 2465				
G(8-O-4)SP(8-5)G'	23	10,110	<b>9215 ± 853</b>	<b>0,0118</b>	12534 ± 2341	0,9622	9976 ± 3106	0,1801	12605 ± 1307				
Putative oligolignol	24	10,994	162646 ± 19423	0,1581	142747 ± 28270	0,1073	<b>95781 ± 23551</b>	<b>0,0052</b>	196278 ± 30387				
Putative oligolignol	25	11,234	53734 ± 6063	0,9587	49906 ± 3215	0,3738	<b>37398 ± 6000</b>	<b>0,0107</b>	53962 ± 5079				
TOTAL AREA			1.46x10 <sup>6</sup> ± 1.03x10 <sup>5</sup>	0,0621	<b>1.38x10<sup>6</sup> ± 6.02x10<sup>5</sup></b>	<b>0,0267</b>	<b>1.32x10<sup>6</sup> ± 1.28x10<sup>5</sup></b>	<b>0,0129</b>	1.63x10 <sup>6</sup> ± 9.10x10 <sup>4</sup>				

**Table 2.3: Phenolic glycosides and oligolignols abundance for EgMYB2-DR, EgMYB1-DR and EgMYB137-DR poplar plants.** Mean peaks area  $\pm$  SD for phenolic glycosides (sum of peaks 7, 10, 11 and 18 corresponding to salicortin, salireposide, populoside C and tremulacin areas) and for oligolignols (sum of peaks 13, 15, 17, 20-25) were calculated. Students test were performed to explore differences with control (WT and empty-DR vectors transgenic lines). p-values < 0.05 are highlighted in bold. A ratio in relation with the controls are shown in a colour code. Yellow squares indicate lower presence of the compound.

	Phenolic Glycosides				Oligolignols			
	Mean	SD	p-value	ratio	Mean	SD	p-value	ratio
Controls	294173 $\pm$ 63990		1,0000		639848 $\pm$ 28216		1,0000	
EgMYB2-DR	224307 $\pm$ 12606		0,1281		<b>533993 <math>\pm</math> 63797</b>		<b>0,0294</b>	
EgMYB1-DR	225992 $\pm$ 12372		0,2308		<b>513762 <math>\pm</math> 68505</b>		<b>0,0258</b>	
EgMYB137-DR	207503 $\pm$ 22001		0,0787		<b>423018 <math>\pm</math> 63779</b>		<b>0,0016</b>	



0.5                      1                      0.5

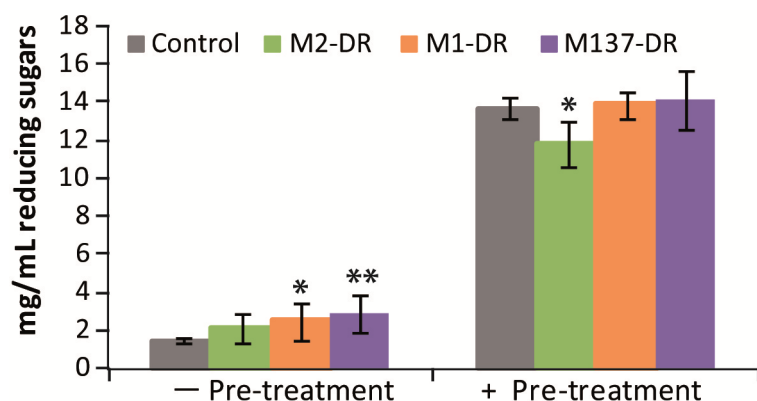
### 2.3.8 EFFECTS OF EgMYB1-DR, EgMYB2-DR AND EgMYB137-DR ON SACCHARIFICATION YIELDS

Biomass enzymatic digestibility (saccharification) measures the reducing sugars released by hydrolysis of lignocelluloses using a cellulase digestion. Before saccharification assessment, the samples could be exposed to an alkali (NaOH) pre-treatment that breaks the lignin structure to facilitate the access to sugars.

When performing saccharification without any pre-treatment, significant increases in saccharification yields of 65% and 96% were obtained for EgMYB1-DR and EgMYB137-DR samples, respectively (Figure 2.10). EgMYB2-DR plants did not show significant differences in saccharification yields as compared to controls.

We also evaluated the release of reducing sugars after the samples had been submitted to an alkali pre-treatment. As expected, the obtained values were much higher than the ones found without pre-treatment (Figure 2.10), but no differences with the controls were observed for EgMYB1-DR and EgMYB137-DR plants. Only EgMYB2-DR plants released significantly less reducing sugars (-14%) than the controls.





**Figure 2.10: Saccharification yields in poplar stems.** Saccharification yields after 48 h, without and with pre-treatment, of greenhouse-grown poplar plants. Mean values  $\pm$  standard deviation (SD) calculated for 3 lines per construction including at least 3 plants each are represented except for EgMYB1-DR where only 2 lines were used. \* $p$ -value < 0.05; \*\* $p$ -value < 0.01. M2: EgMYB2; M1: EgMYB1; M137: EgMYB137.

## 2.4 DISCUSSION

In this paper, we present further insights into the function of two *Eucalyptus* antagonist master regulators of wood formation (EgMYB1 and EgMYB2) as well as the first characterization of the function of a new SCW-regulator EgMYB137. Overexpression of EgMYB1 was already reported in poplar (Legay *et al.*, 2010), whereas overexpression of EgMYB2 was reported in tobacco (Goicoechea *et al.*, 2005).

In the present study, poplar plants overexpressing EgMYB137 could not be regenerated. Similar problems were encountered even after several attempts to obtain poplar plants overexpressing EgMYB2 (Legay, 2008). We also used the dominant repression strategy developed by Hiratsu *et al.* (2003) by fusing the EAR motif to the C termini of the three MYB transcription factors (EgMYB1, EgMYB2 and EgMYB137). The resulting chimeric MYBs are supposed to be transformed in dominant repressors and should inhibit the function of their homologs by competing with their cis-binding elements or interacting proteins (Hiratsu *et al.*, 2003; Kagale and Rozwadowski, 2011). Using the dominant repression strategy we were able to obtain transformed plants for all the tree genes. However, although all EgMYB1-DR regenerated plants were resistant to the selection antibiotic, only 50% and 67% of the tested plants expressed EgMYB137-DR and EgMYB2-DR, respectively. This suggests that silencing events occurred in a large part of the regenerated lines transformed with either EgMYB2 or EgMYB137, the two

same genes for which overexpressing poplar lines could not be obtained. The overexpression of their native forms, and to a lesser extent of their chimeric dominant negative ones, seems to be problematic for regenerating poplar plants.

EgMYB137-DR, EgMYB1-DR and EgMYB2-DR poplar lines showed lower acid insoluble lignin levels supported by lower phloroglucinol-HCl staining, even if the minor differences in lignin content present in EgMYB2-DR plants make difficult to be detected in phloroglucinol-HCl-stained sections. In all transgenic plants, lignin reduction was accompanied by a reduction in the levels of oligolignols, and plants with a higher reduction in acid insoluble lignin have also a higher reduction in total oligolignols quantity. Oligolignols are small lignin polymers, also derived from monolignols, believed to polymerise similarly to lignin (Morreel *et al.*, 2010; Vanholme *et al.*, 2010) but they are normally not incorporated into the growing lignin polymer (Huis *et al.*, 2012; Morreel *et al.*, 2004). It has been described that plants engineered to produce less lignin also have a concomitant reduction in oligolignols content, which is compensated by the production of glycosides of vanillic, caffeic, and sinapic acids (Meyermans *et al.*, 2000; Morreel *et al.*, 2004) or other detoxification products (Damiani *et al.*, 2005). In poplar, the role of glycosylated monolignols is not completely understood but it has been purposed as a storage or a detoxification mechanism for phenylpropanoids accumulation (Vanholme *et al.*, 2012; Meyermans *et al.*, 2000). In our case, none of the analysed plants (EgMYB2-DR, EgMYB1-DR and EgMYB137-DR) showed an accumulation of glycosylated monolignols. Instead, the levels of some specific glycosylated phenolics, like the populoside C in EgMYB1-DR and EgMYB137 or the salicortin in EgMYB137, are reduced. Altogether, this suggests that the whole monolignol biosynthesis is impaired and that there is not a sufficient accumulation of specific intermediates to induce a detoxification mechanism. Therefore, we can hypothesize that the dominant repression constructs of EgMYB1, EgMYB2 and EgMYB137 reduce the carbon flux directed to the phenylpropanoid metabolism, in line with a role as key regulators of the whole lignin biosynthesis process.

The three characterised dominant repression chimeric MYB genes have different effect over lignin composition. EgMYB137-DR plants showed a reduction in the percentage of G units detected by both Py-GC/MS and thioacidolysis, causing a rise in the S/G ratio. EgMYB2-DR

poplar plants did not show differences in lignin composition as compared to control, contrasting with EgMYB2 overexpressing plants, which showed an increase in S/G ratio (Goicoechea *et al.*, 2005) due to an increment in S units, even if these EgMYB2-OE results were obtained in a tobacco heterologous system.

EgMYB1-DR transgenic plants showed no differences in the S/G ratio in agreement with EgMYB1-OE poplar plants (Legay *et al.*, 2010). However, EgMYB1-DR plants had a significant increase in H units percentage in the lignin polymer evaluated by Py-GC/MS analyses. An increase in the percentage of H units was also obtained in the bark of EgMYB1-OE poplar plants (Legay *et al.*, 2008), but the percentage of H lignin in xylem cell walls was not significantly changed. This increase in H lignin could be due to a less pronounced decrease of H units as compared to G and S units or due to a real increase in H units. Thioacidolysis quantification of H units could help answering more precisely this question, unfortunately H quantification was not possible in our hands so far. Lignin enriched in H units are often produced in response to wounding and/or biotic stresses (Sattler and Funnell-Harris, 2013; and references therein), suggesting that the overexpression of EgMYB1-DR might induce a constitutive stress response in these transgenic lines.

The effects of EgMYB2-DR, EgMYB1-DR and EgMYB137-DR on saccharification yields were assessed. Without any pre-treatment, we observed significant increases in saccharification yields of 65% and 96% in EgMYB1-DR and EgMYB137-DR samples, respectively, whereas EgMYB2-DR plants had saccharification yields similar to controls. The saccharification yield therefore correlates negatively with the lignin content, being EgMYB137-DR lines those with lower lignin and higher saccharification yield, followed by EgMYB1-DR lines. Lignin content has been reported in many studies as a major determinant of lignocellulosic recalcitrance, impairing cellulose accessibility to degrading enzymes (Chen and Dixon, 2007; Sykes *et al.*, 2015; Van Acker *et al.*, 2013). Although lignin content is the most determinant trait affecting saccharification, lignin composition also play a role (Vanholme *et al.*, 2012). As G units are responsible for ramification of the lignin polymer, lignins with higher S/G ratio contain more linear parts and are less condensed, thus favouring accessibility to cellulose. A lower lignin

content and a higher S/G ratio are possibly the cause of the higher saccharification yield in EgMYB137-DR plants.

The alkali pre-treatment levelled off the differences between EgMYB137-DR, EgMYB1-DR and controls, but the yield of EgMYB2 was significantly lower. Alkali pre-treatment breaks the lignin polymer and increases the amount of cellulose accessible to degradation enzymes. A plausible explanation for the obtained reduced saccharification yield of EgMYB2-DR plants after alkali pre-treatment could be that the amount of cellulose was also decreased. Supporting this hypothesis, Py-GC/MS showed that EgMYB2-DR plants had the lowest percentage of hexoses (generally assumed to come from cellulose) from all the plants hereby measured. Moreover, several evidences point out that EgMYB2 and its orthologs in other species not only control lignin content but also the SCW polysaccharides, suggesting that EgMYB2 is not only a regulator of lignin biosynthesis, but also a regulator of the whole biosynthesis of SCW polymers (Goicoechea *et al.*, 2005; Zhong *et al.*, 2007).

Wood feedstocks with reduced or easily degradable lignin are interesting materials to be used in pulp and bioethanol industries. Plants with more accessible cellulose can reduce the severity of the pre-treatments needed and thus the whole industrial process is more ecological friendly and less expansive. EgMYB1 and mostly EgMYB137 are interesting targets since, when overexpressed as dominant repressors, they enhance sugar release without the need of pre-treatments and without severe penalty effects on growth and total biomass production.

The transient transactivation experiments using the EgCAD2 promoter confirmed that the dominant repression EgMYB2-DR construct worked efficiently. This transactivator able to activate the EgCAD2 promoter in its native form was indeed turned into a repressor by the presence of the EAR motif fused to its C-terminal region, thus repressing the EgCAD2 transcription. As mentioned above, poplar plants overexpressing EgMYB2 could not be obtained, and thus a direct comparison with poplar lines overexpressing EgMYB2-DR could not be performed. However comparison with the previously characterized tobacco lines overexpressing EgMYB2 (Goicoechea *et al.*, 2005) revealed contrasting phenotypes, with a higher lignification degree for EgMYB2-OE and lower in EgMYB2-DR plants.

For EgMYB1, the repressing effect of the native form on the activity of the EgCAD2 promoter measured in transient transactivation assays was so strong that it was difficult to see a potential enhancement of this effect with the dominant repressor form. This is in agreement with the fact that the poplar lines overexpressing EgMYB1-DR exhibited phenotypes similar to those overexpressing the native form of EgMYB1 (Legay *et al.*, 2010). Indeed EgMYB1 already contains an EAR repressor domain in its C-terminal region (Legay *et al.*, 2007). These results are in line with what was reported previously for another MYB repressor, AtMYBL2, for which plants overexpressing the dominant repression had similar phenotypes as the ones overexpressing the native form (Matsui *et al.*, 2008). However, this is not always the case. Dominant repression strategy has been also applied to AtMYB44, a transcription factor which contains, as EgMYB1, an EAR domain in its C-terminal region. In this case, plants overexpressing AtMYB44 and plants overexpressing the dominant repression form, AtMYB44-DR, show opposite phenotypes (Persak and Pitzschke, 2014).

We could not test in transactivation assays the activity of EgMYB137 since this transcription factor was only very recently reported and we did not have any clues about its potential target genes. In the absence of overexpressing and/or loss of function lines, the analyses of the sole EgMYB137-DR lines cannot tell us if EgMYB137 was an activator or a repressor, even if their lignin content and composition is severely altered. Indeed, previous conclusions based only on DR-lines (Zhong *et al.*, 2008) were proven later to be wrong as for example for KNAT7 (Li *et al.*, 2012) and MYB52 (Cassan-Wang *et al.*, 2013). Some indirect evidence suggests that EgMYB137 might be a repressor. For example, EgMYB137 is significantly more expressed in tension wood than in opposite wood, and tension wood is characterized by a very low lignin levels as compared to opposite wood. Besides, preliminary experiments in the lab showed that, in *Arabidopsis*, overexpression of EgMYB137 led to a minor decrease in lignin levels whereas dominant repression led to a strong lignin reduction (M. Soler, Personal communication). On the other hand, EgMYB137 is phylogenetically related, although distantly, to AtMYB61 and PtMYB8, all belonging to subgroup 13. Overexpression of both AtMYB61 and PtMYB8 were proposed to activate ectopic lignin deposition in *Arabidopsis* and *Picea glauca* plants, respectively, whereas loss-of-function mutants of AtMYB61 led to plants with less abundant vessels, with irregular shape and with thinner SCW (Newman *et al.*, 2004; Bomal *et al.*, 2008).

Thus, we cannot yet exclude that EgMYB137 could also be an activator of lignin biosynthesis turned into a repressor by the dominant repression strategy, and dedicated experiments are needed to determine if EgMYB137 is a repressor or an activator.

## 2.5 MATERIAL AND METHODS

### 2.5.1 PHYLOGENY RECONSTRUCTION

R2R3-MYB sequences from *E. grandis*, *A. thaliana*, *P. trichocarpa* and *V. vinifera* belonging to the “subgroup13 + AtMYB26 & AtMYB67” cluster defined by Soler *et al.* (2015) were retrieved from different plant genomes at Phytozome 10.3 database (<http://www.phytozome.net/>). A closely related sequence from *P. taeda* was obtained from Bedon *et al.* (2007).

Aminoacidic sequence alignment was performed using MAFFT with the FFT-NS-i method (Katoch *et al.*, 2002), and Neighbor-joining phylogenetic tree (Saitou and Nei, 1987) with 1000 bootstrap replicates was constructed with Mega5 (Tamura *et al.*, 2011) using the pairwise deletion method. The evolutionary distances were computed using the Jones–Taylor–Thornton substitution model and the rate variation among sites was modelled with a gamma distribution of 1.

### 2.5.2 MICROFLUIDIC RT-QPCR AND RNASEQ

Data was extracted from (Soler *et al.*, 2015). Differences in opposite vs tension and mature vs juvenile wood were statistically tested using student t-test.

### 2.5.3 VECTORS AND CLONING

The EgMYB1, EgMYB2 and EgMYB137 cDNA and *EgCAD2* promoter (Feuillet *et al.*, 1995) were amplified by PCR from *E. grandis* gDNA using PhusionTaq (Fisher Scientific) and gene-specific primer pairs (Table 2.4).

**Table 2.4:** Accession number for the genes and promoters used in this study as well as primers used for cloning procedures.

Gene	Accession	Primer forward	Primer reverse
EgMYB1	AJ576024.1	GGAGATAGAACCATGGGAAGGTCTCCTTGCTGCG	CAAGAAAGCTGGGTCGCTCCTATAATCCAAAACACTTGCC
EgMYB2	AJ576023.1	GGAGATAGAACCATGGGAATGGCAATGGGGATT	CAAGAAAGCTGGGTCGAAATCAAGAAAAGGGAAGGAGG
EgMYB137	Eucgr.K02806.1	CACCTTTCCTGCGCAGTTCTCAAT	GAAGATGTTGAAATGTGCCTGTA
PromEgCAD2	X65631.1	CACCTGAGCAAGTACCCACATCAA	TTTTGCTCAAAGATCCAAGC

The gene cDNA sequences were cloned into pDONR207 or into pENTR/D-TOPO vectors (Invitrogen) and verified by sequencing. Then, EgMYB1, EgMYB2 and EgMYB137 were recombined using LR Clonase II (Invitrogen) according to manufacturer's instructions to the following destination vectors. For transactivation assays, a reporter construct (promEgCAD2-GUS) was generated by inserting the EgCAD2 promoter into pKGWFS7 vector (Karimi *et al.*, 2002). For effector vectors, EgMYB1 and EgMYB2 genes were cloned into the binary vectors pFAST-G02 (Shimada *et al.*, 2010) vector for overexpression in the native form (EgMYB1-OE and EgMYB2-OE) and into pH35GEAR vector (kindly provided by Taku Demura) for overexpression as dominant repression chimeric proteins (EgMYB1-DR and EgMYB2-DR). EgMYB137 was also cloned into pH35GEAR vector, and all the dominant repression constructs (EgMYB1-DR, EgMYB2-DR and EgMYB137-DR) were also used to obtain transgenic poplar plants. EgMYB137 was also inserted into the pJCV53 vector (<https://gateway.psb.ugent.be/vector/show/pJCV53/>) to overexpress the gene in poplar transgenic plants, but the regeneration of transgenic plants was not possible for that construct.

These constructs were transferred into *Agrobacterium tumefaciens* strain GV3101/pMP90 (Koncz and Schell, 1986).

#### 2.5.4 CO-TRANSFECTION EXPERIMENTS AND FLUORIMETRIC GUS ASSAYS

The transformed *A. tumefaciens* GV3101/pMP90 bacteria were grown at 28 °C in liquid LB media supplemented with Rifampicin (50 mg/L), Gentamycin (20 mg/L) and Spectinomycin (100mg/L) until an optical density at 600 nm of 0.8. After centrifugation (10 min at 2400 xg), bacterial cells were resuspended using a freshly prepared infiltration media containing

MES/KOH 10 mM, MgCl<sub>2</sub> 10 mM and acetosyringone 100 μM, pH 5.6. The *A. tumefaciens* suspension containing the reporter (prom*EgCAD2-GUS*) binary vector was mixed with one containing an effector (EgMYB1-OE, EgMYB1-DR, EgMYB2-OE or EgMYB2-DR) binary vector to a final optical density at 600 nm of 0.5 each, and incubated for 3h at dark and at room conditions.

Leaves of 4- to 5-week-old *Nicotiana benthamiana* plants were infiltrated with the bacterial mixtures using 1 mL needless syringe. After 48h, 10 leaf discs were sampled and frozen with liquid nitrogen until further analysis. Protein extraction and quantitative GUS assays were performed as reported previously (Jefferson, 1987). Protein concentrations were determined by the Bradford method (Bio-Rad). GUS activities were carried out using 4-methylumbelliferyl-β-D-glucuronide as substrate and estimated as the mean of three independent assays, each containing three replicates.

### 2.5.5 PLANTS TRANSFORMATION, PRODUCTION AND SAMPLING

The hybrid poplar *Populus tremula* x *P. alba* INRA clone 717-1-B4 was used as wild-type (WT) and for plant transformation. Plants were maintained on MS-1B media (Duchefa) amended with 0.5mg/L of IAA and 0.5mg/L of IBA. Hybrid poplar was transformed via an *A. tumefaciens*-mediated protocol described previously (Gallardo *et al.*, 1999). Selection of transformants was conducted in hygromycin- (for dominant repression vectors) or in kanamycin- (for overexpressor vector) containing medium. Once whole plantlets had been regenerated, leaves from independent transformed individuals (lines) were screened by reverse transcription-quantitative polymerase chain reaction (RT-qPCR) for transgenes over-expression. Selected plants were acclimated for three weeks in a phytotron, further transferred to a greenhouse and grown for two month in a fully randomized experimental design.

Samples for RNA extraction and for histochemical and biochemical analyses were obtained from the stem of hardened poplar plants. Samples for histochemical analyses were obtained from the bottom of the tree stems and kept into ethanol (EtOH) 80% until further analyses. A stem portion placed 2 cm above the sample for histochemical analyses and until the 10<sup>th</sup> internode, was debarked and cut into pieces, frozen in liquid N<sub>2</sub> and further kept at -80°C until



lyophilisation. The freeze-dried stems were ground to powder in a ball-mill (MM400, Retsch, Germany) and used for biochemical analyses.

### 2.5.6 BIOCHEMICAL ANALYSIS OF SCW COMPOSITION

The milled samples were subjected to exhaustive extraction with water, then with ethanol, ethanol/toluene (50/50) and acetone in a Soxhlet apparatus. The recovered extractive-free samples were dried at 60°C overnight and used for the analyses of lignins by Klason and thioacidolysis. Quantification of acid-insoluble lignin was performed by the Klason procedure as previously described (Méchin *et al.*, 2014) using 50 mg of the ground extractive-free samples.

Thioacidolysis was performed as described previously (Méchin *et al.*, 2014) by using approximately 10 mg of sample and 1 ml of thioacidolysis reagent per mg of sample (for 100 mL of solution: 2,5 mL of Boron trifluoride, 10 mL of ethanethiol and 0.2 mL of tetracosane 1.25 mg/mL, up to 100 mL of dioxane). An aliquot of this solution (5 µL) was trimethylsilylated (TMS) with 100 µL of N,O-bis(trimethylsilyl)trifluoroacetamide and 10 µL of ACS-grade pyridine for 30 min at 60°C. The TMS samples were injected in a GC-MS (TSQ Quantu, Thermo Scientific) fitted with an autosampler, a splitless injector (280 °C), and an iontrap mass spectrometer operating in the electronic impact mode with a source at 220 °C, an interface at 280 °C, and a 50 to 650 m/z scanning range. The column was a ZB5-MSi column (Phénomenex) operated in the temperature program mode (from 45 to 180 °C at +30 °C/min, then 180 to 260 °C at +3 °C/min), with helium carrier gas at a 1.5 mL/min flow rate.

Prior to perform analytical pyrolysis, milled samples were sequentially extracted for 16 h with 95% ethanol followed by another 16 h in distilled water in a 125 mL Soxhlet apparatus. An aliquot of 30 mg of each extracted sample was further milled on a vibratory ball mill (Mixer Mill MM, Retsch) for 5 min and kept on desiccators. Analytical pyrolysis was performed using a CDS Pyroprobe 1000 with a coil filament probe connected to a GC (Agilent 6890) with flame ionization via a heated interface (270°C). The pyrolysis was carried out at 600°C for 5 s, using 75 µg of the extractive-free milled samples. Details of the conditions and quantification procedures have been published elsewhere (Alves *et al.*, 2006).

### 2.5.7 METABOLIC PROFILING

#### Metabolite extraction

Thirty mg of freeze dried stem powder were extracted with 1 mL of methanol (MeOH) 70%, blended for 1 min with an Ultra-Thurrax and centrifuged for 10 min at 10,000g. The supernatant was transferred to a new tube. One additional mL was added to the pellet, then vortexed and left 2 hours at room temperature. The mixture was centrifuged 10 min at 10,000g and the supernatant was recovered and pooled with the first one. The 2 mL solution was evaporated in a speed vacuum until dryness. The dried pellet was then dissolved in 500 $\mu$ L MeOH 70% and filtrated on 0.22  $\mu$ m filter before U-HPLC analysis.

#### U-HPLC analysis

Extracts (1  $\mu$ L) were analysed on an Ultra-Performance Liquid Chromatography (U-HPLC) system (Shimadzu) equipped with a PDA and a Mass Spectra (MS) detector. The sample was separated on a C18 Kinetex (100 mm  $\times$  2.1 mm) column (Phenomenex). The mobile phase consisted in 0.1% formic acid in ultra-pure water (solvent A) and 0.1% formic acid in methanol (solvent B).

The molecules were eluted through a gradient elution from 1 to 99% B for 13 min with a flow rate of 400 $\mu$ L/min and then 3 min with 99% B. The column was then re-equilibrated to 1% B prior to the next run. Mass analysis was carried out in ESI negative mode. Molecule identification was based on retention time,  $\lambda_{max}$ , exact mass and mass fragmentation data which were compared either to authentic standards, either to metabolomics data banks (Respect: <http://spectra.psc.riken.jp/>; Mass Bank: <http://www.massbank.jp/>; DNP: <http://dnp.chemnetbase.com>) and available data from the literature. Quantification was realized by measuring the area under peak at 280nm, 320nm or 350nm depending on the  $\lambda_{max}$  of each molecule.

### 2.5.8 SACCHARIFICATION

Saccharification yields of extractive-free samples were performed as (Bragatto *et al.*, 2012) and with most of the modifications performed by (Lepikson-Neto *et al.*, 2014). Briefly, approximately 10 mg of extracted and milled samples were washed with EtOH 70% overnight at 55  $^{\circ}$ C and further washed first with EtOH 70 % and after with acetone 1x. After drying at room

temperature, hydrolysis was performed on 400 µl of sodium acetate buffer pH 4,8 0.1M where cellulase (reference C2730, Sigma-Aldrich) has been added to a final concentration of 24 mg / mL. It was incubated for 48 h at 30°C and shaking at 750rpm in a comfort thermomixer (Eppendorf). Tubes were after centrifuged for 3 min at 16000 xg. The obtained supernatant was used to measure the reducing sugars as described in (Lepikson-Neto *et al.*, 2014).

Alternatively, an alkali pre-treatment was performed by incubating the EtOH and acetone washed samples with NaOH 0.25 %. Mixture was incubated for 3 hours at 70 °C and mixed at 750 rpm, and further washed three times with water. For these samples, enzymatic hydrolysis was performed at 50 °C instead of 30°C to maximize the amount of released sugars.

## 2.6 SUPPLEMENTAL TABLES AND FIGURES

**Table S2.1:** Recovered peak names from pyrolysis analyses, their designed category and the identified compound.

Compound	Category	Compound
C1.3	Carbohydrates	Acetaldehyde
C2.3	Carbohydrates	2-Propenal (acrolein )
C3	Carbohydrates	Propanal-2-one
C5.1	Carbohydrates	2,3-Butandione
Cu	Carbohydrates	Butanone-(2) or unknown
cH7	Carbohydrates and Hexoses	Hydroxyacetaldehyde
C11	Carbohydrates	Acetic acid
cH12	Carbohydrates and Hexoses	Hydroxypropanone
Cu	Carbohydrates	Unknown
C18	Carbohydrates	3-Hydroxypropanal
C19	Carbohydrates	3-Butenal-2-one
C20	Carbohydrates	(3H)-Furan-2-one
C25	Carbohydrates	2-Hydroxy-3-oxobutanal
C26	Carbohydrates	2-Furaldehyde, 2-furfural
C34	Carbohydrates	Dihydro-methyl-furanone
C35	Carbohydrates	Dihydro-methyl-furanone
cP36	Carbohydrates and pentoses	Isomer of 4-Hydroxy-5,6-dihydropyran-(2H)-one
C38	Carbohydrates	(5H)-Furan-2-one
C40	Carbohydrates	Gamma-Lactone and unknown
cP41	Carbohydrates and pentoses	4-Hydroxy-5,6-dihydropyran-(2H)-2-one
cH43	Carbohydrates and Hexoses	2-Hydroxy-1-methyl-cyclopenten-(1)-3-one
H45	H lignin	Phenol
G46	G lignin	Guaiacol
C50.1	Carbohydrates	Methyl-butylaldehyde derivative
G51	G lignin	3-Methyl guaiacol
H51.2 / H51.3	H lignin	<i>p</i> -Cresol / <i>m</i> -Cresol
C52	Carbohydrates	Gamma-lactone derivative
G54	G lignin	4-Methyl guaiacol
C54.1	Carbohydrates	Anhydrosugar
GC60/60.1+61	G lignin and Carbohydrates	Overlapping spectra; 4-ethyl-guaiacol
Cu	Carbohydrates	Unknown
Cu	Carbohydrates	Unknown
cH64	Carbohydrates and Hexoses	1,4:3,6-Dianhydro-glucopyranose
cP65	Carbohydrates and pentoses	1,5-Anhydro-arabinofuranose
G66	G lignin	4-Vinyl guaiacol
G68	G lignin	Eugenol
G68.1	G lignin	4-Propyl guaiacol
cH69	Carbohydrates and Hexoses	5-Hydroxymethyl-2-furaldehyde
C69.1	Carbohydrates	gamma-Lactone derivative
S70	S lignin	Syringol
G71	G lignin	Isoeugenol (cis)
cH73	Carbohydrates and Hexoses	Pyran-(4H)-4-one, 2-hydroxymethyl-5-hydroxy-2,3-dihydro
G75	G lignin	Isoeugenol (trans)
S76	S lignin	Syringol, 4-methyl-
cP76.1	Carbohydrates and pentoses	1,5-Anhydro-b-D-xylofuranose
G77	G lignin	Vanillin
G78.1	G lignin	Indene, 6-hydroxy-7-methoxy-, 1H-
G78.2	G lignin	Indene, 6-hydroxy-7-methoxy-, 2H-
G80	G lignin	Homovanillin

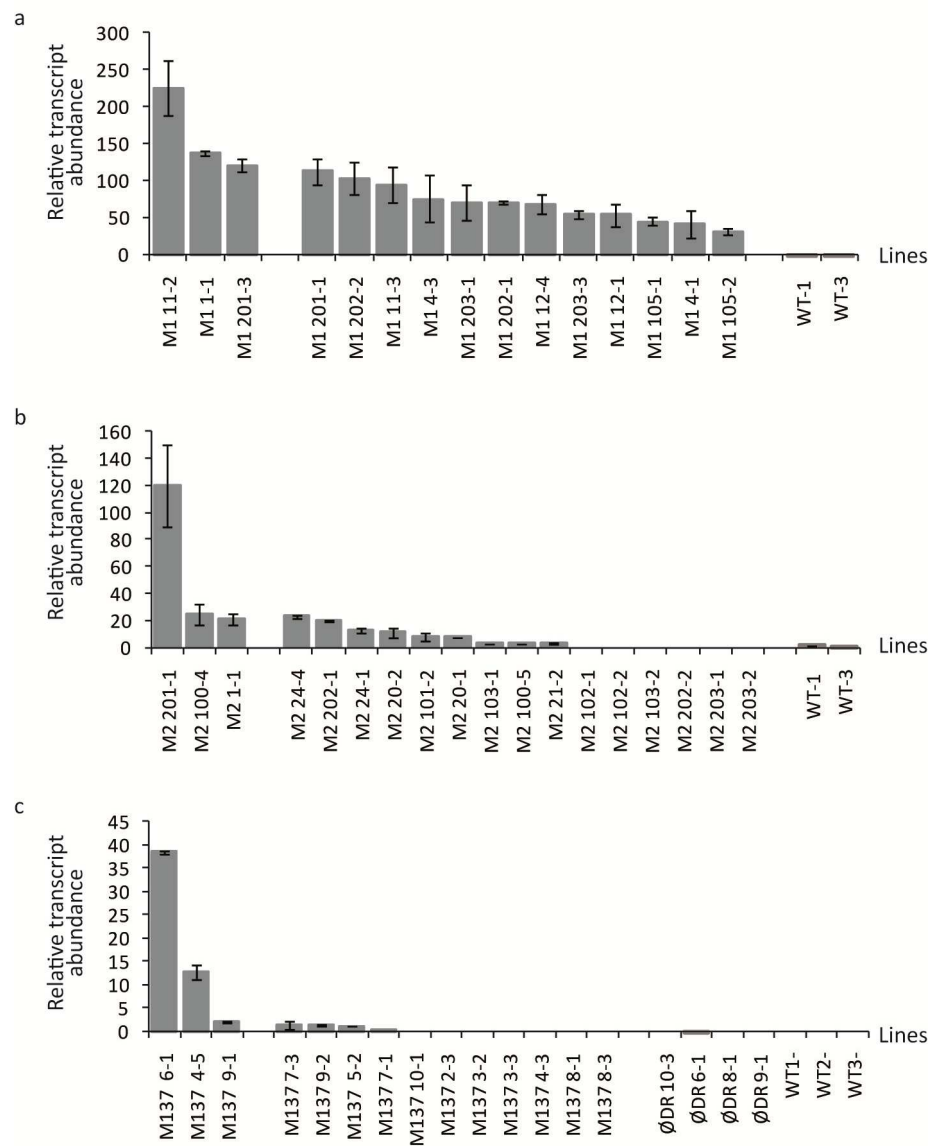
Chapter 2: Functional characterization of three *Eucalyptus* MYB genes controlling wood formation and evaluation of their impacts on saccharification

---

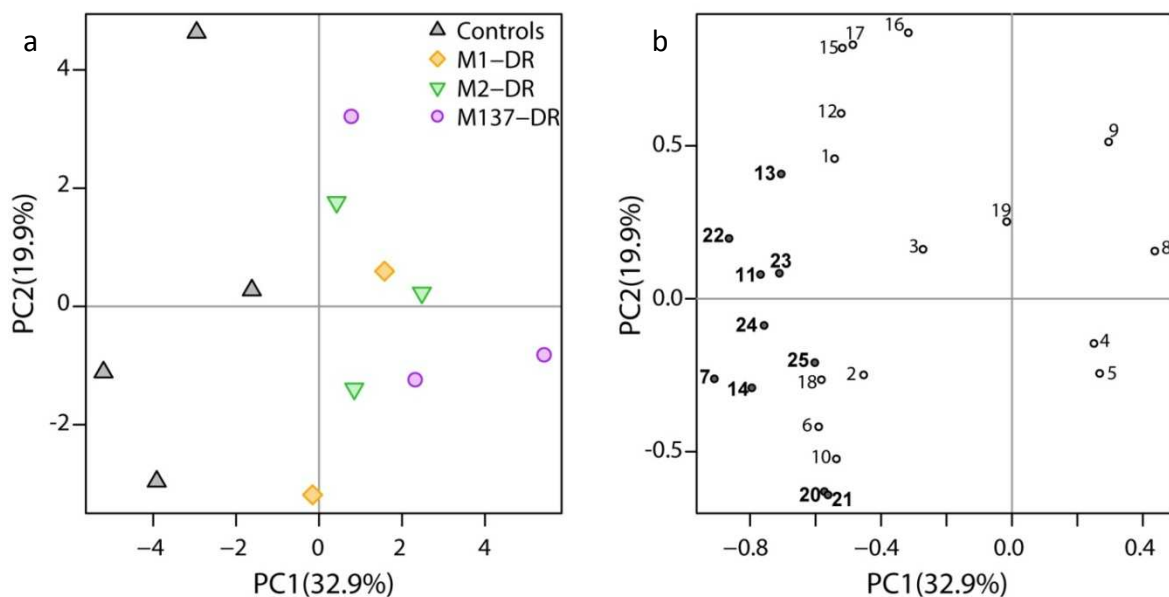
G83	G lignin	Acetoguaiacone
S86	S lignin	Syringol, 4-vinyl-
G87	G lignin	Guaiacyl acetone
G/S	G lignin and S lignin	Unknown
G89	G lignin	Propioguiacone
G92.1	G lignin	Isomer of coniferyl alcohol
G92/92.1	G lignin	G-CO-CH=CH <sub>2</sub>
G92.2	G lignin	G-CO-CO-CH <sub>3</sub>
cH96	Carbohydrates and Hexoses	1,6-Anhydro-b-D-glucofuranose (levoglucosan)
S97	S lignin	Syringol, 4-propenyl- (trans)
G98	G lignin	Dihydroconiferyl alcohol
S99	S lignin	Syringaldehyde
G99.1	G lignin	Coniferyl alcohol (cis)
S100	S lignin	Homosyringaldehyde
C101	Carbohydrates	Anhydrosugar: unknown
S102	S lignin	Acetosyringone
G103	G lignin	Coniferyl alcohol (trans)
G104	G lignin	Coniferylaldehyde
S107,1	S lignin	Isomer of sinapyl alcohol
S110	S lignin	Sinapyl alcohol (trans)
S110.1	S lignin	Sinapinaldehyde

**Table S2.2:** Detailed information from the recovered peaks and the corresponding mass spectra form methanol soluble phenolic compounds. Abbreviated compounds are as follows: G(8-O-4)S(8-5)G = coniferyl alcohol(8-O-4)synapyl alcohol(8-5)coniferyl alcohol; G(8-O-4)S(8-5)G' = coniferyl alcohol(8-O-4)synapyl alcohol(8-5)coniferaldehyde; G(8-O-4)S(8-8)S = coniferyl alcohol(8-O-4)coniferyl alcohol(8-O-4)synapyl alcohol(8-5)coniferyl alcohol; G(8-O-4)SP(8-5)G = coniferyl alcohol(8-O-4)synapyl p-hydroxybenzoate(8-5)coniferyl alcohol; G(8-O-4)SP(8-5)G' = coniferyl alcohol(8-O-4)synapyl p-hydroxybenzoate(8-5)coniferylaldehyde. Peaks 20-21 and 22-23 are two isomers of the same compound.

Peak number	Retention time	UV lambda max	M-	Fragment MS/MS	Formula	Molecule	Reference
1	4.13	280					
2	4.54	288, 321					
3	4.73	284					
4	5.19	280					
5	5.57	262, 297					
6	5.74	272					
7	6.14	270	423,1278	155 (100), 137 (90), 317 (40), 299 (40), 123 (40)	C <sub>20</sub> H <sub>24</sub> O <sub>10</sub>	Salicortin	Kamerer et al., 2005
8	6.44	343	207,0655	177 (100), 192 (20)	C <sub>11</sub> H <sub>12</sub> O <sub>4</sub>	sinapaldehyde	
9	6.57	281	477				
10	6.91	282/329	405,1172	283 (100), 121 (20), 163 (2), 181 (1)	C <sub>20</sub> H <sub>22</sub> O <sub>9</sub>	Salireposide	Kamerer et al., 2005; Ranocha et al., 2002
11	7.22	276	461,1071	389 (100), 283 (40)	C <sub>23</sub> H <sub>26</sub> O <sub>10</sub>	Populoside C	Boeckler et al., 2011
12	7.69	339					
13	7.74	276	583,2172	535 (100), 369 (40)	C <sub>31</sub> H <sub>35</sub> O <sub>11</sub>	G(8-O-4)S(8-5)G	Morreel et al., 2010
14	7.83	286, 310	435,1283				
15	8.07	280, 343	581,2022	533 (100), 367 (90), 355 (40), 551 (40)	C <sub>31</sub> H <sub>33</sub> O <sub>11</sub>	G(8-O-4)S(8-5)G'	Morreel et al., 2010
16	8.18	342					
17	8.34	339	809,3003	809 (100), 595 (40), 369 (1), 761 (1), 613	C <sub>42</sub> H <sub>49</sub> O <sub>16</sub>	G(8-O-4)G(8-O-4)S(8-8)S	Morreel et al., 2010
18	8.82	271	527,1553	405 (100), 155 (15), 299 (10)	C <sub>27</sub> H <sub>28</sub> O <sub>11</sub>	Tremulacin	Kamerer et al., 2005
19	9.10	273					
20	9.70	266	703,2379	703 (100), 655 (6), 561 (3), 503 (2), 369 (2), 667	C <sub>38</sub> H <sub>39</sub> O <sub>13</sub>	G(8-O-4)SP(8-5)G	Morreel et al., 2010
21	9.85	266	703,238	703 (100), 655 (6), 561 (3), 503 (2), 369 (2), 667	C <sub>38</sub> H <sub>39</sub> O <sub>13</sub>	G(8-O-4)SP(8-5)G	Morreel et al., 2010
22	9.98	339	701,2224	701 (100), 653 (10), 367 (10), 559 (3), 515 (2), 665	C <sub>38</sub> H <sub>37</sub> O <sub>13</sub>	G(8-O-4)SP(8-5)G'	
23	10.11	339	701,2224	701 (100), 653 (10), 367 (10), 559 (3), 515 (2), 665	C <sub>38</sub> H <sub>37</sub> O <sub>13</sub>	G(8-O-4)SP(8-5)G'	
24	10.97	326	727,238	727 (100), 551 (1), 501 (1)		Putative oligolignol	
25	11.21	272, 328	727,238	727 (100), 551 (1), 501 (1)		Putative oligolignol	



**Figure S2.1: Mean relative transcript abundance of EgMYB1 (a), EgMYB2 (b) and EgMYB137 (c) genes.** Mean values  $\pm$  SD RT-qPCR transcript abundance for EgMYB1, EgMYB2 and EgMYB137 genes. The three first lines were selected to be acclimated.



**Figure S2.2: Principal component analysis (PCA) for metabolic profiling data.** Peak areas from metabolic profiling were used in the PCA. Both the score (a) and the loading (b) plots for the two main principal components are shown. Each point from the score plot represents a plant. Each point on the loading plot represents a peak from U-HPLC/MS analysis. Numbers from loading plots represent peaks ID (for correspondence with identified compounds please refer to table 2.2). Peaks with area values significantly different with the controls are highlighted in bold. Principal component 1 (PC1) explained 32.9% of the variance whereas PC2 explained 19.9%.







CHAPTER 3:

---

*EUCALYPTUS* HAIRY ROOTS, A FAST, EFFICIENT AND  
VERSATILE TOOL TO EXPLORE FUNCTION AND EXPRESSION  
OF GENES INVOLVED IN WOOD FORMATION



## ***Eucalyptus* hairy roots, a fast, efficient and versatile tool to explore function and expression of genes involved in wood formation**

Anna Plasencia<sup>1\*</sup>, Marçal Soler<sup>1\*</sup>, Annabelle Dupas<sup>1</sup>, Nathalie Ladouce<sup>1</sup>, Guilherme Silva-Martins<sup>1</sup>, Yves Martinez<sup>2</sup>, Catherine Lapierre<sup>3</sup>, Claudine Franche<sup>4</sup>, Isabelle Truchet<sup>1</sup> and Jacqueline Grima-Pettenati<sup>1‡</sup>

<sup>1</sup> Plant Research Laboratory (LRSV), UMR5546, Toulouse III Paul Sabatier University - CNRS, Castanet Tolosan, France

<sup>2</sup> Cell Imaging Plateform, FRAIB, CNRS, Castanet Tolosan, France

<sup>3</sup> Jean-Pierre Bourgin Institute (IJPB), INRA / AgroParisTech, UMR1318, Saclay Plant Science, Versailles, France

<sup>4</sup> Rhizogenesis, UMR DIADE, IRD, Montpellier, France

\* These authors contributed equally to this work

---

Plant Biotechnology Journal  
*In press*

‡Correspondence

(Tel +33 (0) 5 34 32 38 13;

fax +33 (0) 5 34 32 38 02;

email [grima@lrsv.ups-tlse.fr](mailto:grima@lrsv.ups-tlse.fr))

**Key words:** *Agrobacterium rhizogenes*, *Eucalyptus*, hairy roots, xylem, secondary cell wall, lignin.

### **Summary**

*Eucalyptus* are of tremendous economic importance being the most planted hardwoods worldwide for pulp and paper, timber and bioenergy. The recent release of the *E. grandis* genome sequence pointed out many new candidate genes potentially involved in secondary growth, wood formation or in lineage-specific biosynthetic pathways. Their functional characterization is, however, hindered by the tedious, time-consuming and inefficient transformation systems available hitherto for eucalypts. To overcome this limitation, we developed a fast, reliable and efficient protocol to obtain and easily detect co-transformed *E. grandis* hairy roots using fluorescent markers, with an average efficiency of 62%. We set up conditions both to cultivate excised roots *in vitro* and to harden composite plants, and verified that hairy root morphology and vascular system anatomy were similar to wild type ones. We further demonstrated that co-transformed hairy roots are suitable for medium throughput functional studies enabling for instance, protein subcellular localization, gene expression patterns through RT-qPCR and promoter expression, as well as modulation of endogenous gene expression. Down-regulation of the *Eucalyptus Cinnamoyl-CoA Reductase1 (EgCCR1)* gene, encoding a key enzyme in lignin biosynthesis, led to transgenic roots with reduced lignin levels and thinner cell walls. This gene was used as a proof-of-concept to demonstrate that the function of genes involved in secondary cell wall biosynthesis and wood formation can be elucidated in transgenic hairy roots using histochemical, transcriptomic and biochemical approaches. The method described here is timely since it will accelerate gene mining of the genome both for basic research and industry purposes.

---

### 3.1 PREFACE

I have been involved in all the steps described in this chapter (except thioacidolysis analyses).

I have set up and optimized the *A. rhizogenes*-mediated transformation protocol of *E. grandis* and produced the necessary plant material with the great help of Annabelle Dupas and the internship students Guilherme Silva-Martins and Lone Marchesi. Claudine Franche guided us at this stage and provided us the strain A4RS.

Construction of the vectors and the sampling and the histochemical phenotyping of roots were done together with Marçal Soler. RNA extraction from roots was optimised by Marçal Soler and Nathalie Ladouce.

Electronic microscopy was performed with the help of Yves Martinez. Confocal microscopy was done together with Isabelle Truchet. Catherine Lapierre performed the thioacidolysis analyses.

Finally, I have done the data mining, statistical analysis and writing of the first draft together with Marçal Soler.

### 3.2 INTRODUCTION

Because of their rapid growth rate, broad adaptability to diverse edaphoclimatic conditions and their multipurpose wood properties, *Eucalyptus* species are among the leading sources of woody biomass worldwide (Myburg *et al.*, 2007; Paiva *et al.*, 2011). Their considerable economic importance has recently increased since traditional interests in pulp and paper production have been extended to the emergent areas of biofuels and biomaterials. Wood is formed of secondary cell walls (SCWs) mainly constituted of cellulose, hemicelluloses and lignins. The composition and structure of the SCWs, relying among interactions between these biopolymers, are major determinants of industrial processing efficiencies (Holladay *et al.*, 2007).

The huge economic importance of *Eucalyptus* wood has been a driving force to delineate the lignin pathway in this genus. More than 20 years ago, the gene encoding the Cinnamyl Alcohol Dehydrogenase (CAD), which catalyzes the last step in monolignol biosynthesis, was cloned in *Eucalyptus* right after that from tobacco among all plant species

(Grima-Pettenati *et al.*, 1993). The gene encoding Cinnamoyl-CoA Reductase (CCR), which catalyzes the penultimate step in monolignol biosynthesis, was first cloned in *Eucalyptus* (Lacombe *et al.*, 1997) and subsequently used to clone its orthologs in other plant species such as poplar (Leplé *et al.*, 1998), maize (Pichon *et al.*, 1998) and *Arabidopsis* (Lauvergeat *et al.*, 2001). Other secondary cell wall-related genes, including transcription factors, have been cloned in *Eucalyptus* but, due to the lack of an efficient stable transformation system, their functional characterization had been achieved mainly using heterologous systems such as poplar (Feuillet *et al.*, 1995; Hawkins *et al.*, 1997; Lauvergeat *et al.*, 2002; Legay *et al.*, 2010; Šamaj *et al.*, 1998), tobacco (Goicoechea *et al.*, 2005; Lacombe *et al.*, 2000) and *Arabidopsis* (Baghdady *et al.*, 2006; Creux *et al.*, 2008; Foucart *et al.*, 2009; Hussey *et al.*, 2011; Legay *et al.*, 2010).

The recent availability of the *Eucalyptus grandis* genome (Myburg *et al.*, 2014) has allowed genome-wide characterization of many gene families, notably those involved in the lignin biosynthetic pathway (Carocha *et al.*, 2015) as well as transcription factor families containing members known to regulate SCW formation such as the R2R3-MYB (Soler *et al.*, 2015), NAC (Hussey *et al.*, 2015), ARF (Yu *et al.*, 2014) and Aux/IAA (Yu *et al.*, 2015) among others. These studies have underscored many new candidates potentially regulating wood formation that need to be functionally characterized. Although stable transformation protocols have been established for several *Eucalyptus* species (Girijashankar, 2011; de la Torre *et al.*, 2014; Tournier *et al.*, 2003, and references therein), they are not suitable for medium/high throughput functional characterization of genes since they are tedious, time consuming and present low efficiencies. For these reasons, only very few functional studies have been performed in transgenic *Eucalyptus* (reviewed in Girijashankar, 2011).

To overcome these limitations, we developed an alternative stable transformation system using *Agrobacterium rhizogenes* that allow rapid *in vivo* analysis of transgenes. *A. rhizogenes*, a soil-born gram-negative bacterium discovered more than 80 years ago (Riker *et al.*, 1930), induces the production of numerous secondary functional roots called “hairy roots” upon wounding and infection of many plants. This is the consequence of the stable integration and expression into the host cell of a T-DNA encoding *root locus (rol)* genes carried on the root inducing (Ri) plasmid (Chandra, 2012; Chilton *et al.*, 1982). If in addition to the Ri plasmid, *A.*

*A. rhizogenes* harbours a binary vector, transgenic roots co-transformed with both the T-DNA from the Ri plasmid and from the binary vector can be obtained. *A. rhizogenes*-mediated root transformation has indeed been applied for a wide variety of purposes, ranging from studying nodulation, mycorrhization, interaction of plants with nematodes or pathogens, to the production of secondary metabolites, phytoremediation and rooting of recalcitrant species (reviewed in Christey, 2001; Georgiev *et al.*, 2012; Guillon *et al.*, 2006). Transformed hairy roots can be *in vitro* cultured in axenic conditions to produce secondary metabolites or recombinant proteins among others applications, but the development of composite plants (wild-type shoots with transgenic roots) has been a key milestone for functional characterization of genes (Hansen *et al.*, 1989). However, while hairy roots have been established in more than 100 species, the proportion of woody species transformed using this system is very low in comparison to annual plants (Christey, 2001). Besides, none of these studies fully explored the hairy root anatomy at the vascular tissue level, and whether the hairy roots are suitable to elucidate the function of genes involved in xylem or SCW formation is still an open question, particularly important for woody species.

In this paper, we describe a fast, efficient and reproducible *A. rhizogenes*-mediated transformation protocol for *E. grandis* that allows production of transgenic roots easily detectable by fluorescent markers. We show that *Eucalyptus* hairy roots are suitable for medium-throughput functional characterization of genes, enabling among others, protein subcellular localization, spatial and temporal patterns of gene expression and down-regulation of endogenous genes. Last but not least, we demonstrate taking CCR down-regulation as a proof-of-concept that the function of genes involved in SCW biosynthesis and wood formation can be elucidated using histochemical, transcriptomic and biochemical approaches.

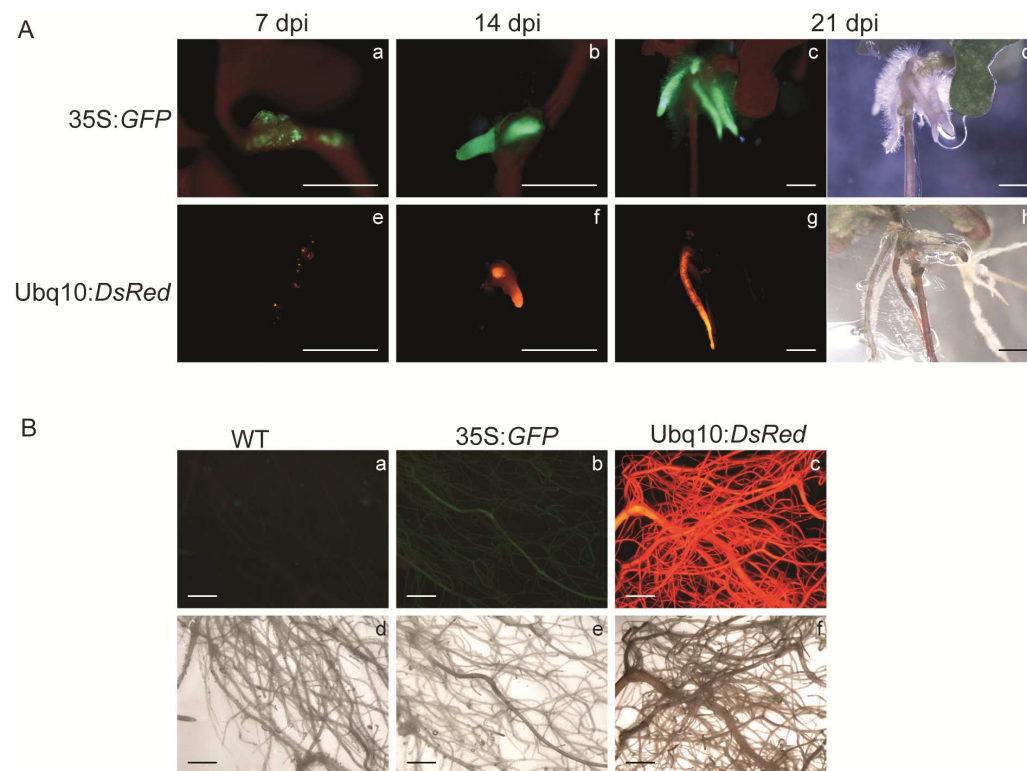
### 3.3 RESULTS

#### 3.3.1 OPTIMIZATION OF *A. RHIZOGENES*-MEDIATED TRANSFORMATION AND GENERATION OF *E. GRANDIS* COMPOSITE PLANTS

We first tested the hypervirulent *A. rhizogenes* strain A4RS previously shown to be the most effective in infecting other trees (Alpizar *et al.*, 2006; Bosselut *et al.*, 2011; Diouf *et al.*, 1995;



Gherbi *et al.*, 2008) as well as in another *Eucalyptus* species, *E. camaldulensis* (Balasubramanian *et al.*, 2011). We compared four protocols combining three types of *E. grandis* explants and three infection methods: (i) stabbing the hypocotyl or (ii) cutting and inoculating the base of the hypocotyl of 14-day-old seedlings, (iii) cutting and inoculating the radicle apex of 3-day-old germinating seeds, and (iv) stabbing the stem of an *in vitro* clonal line. The screening of the co-transformed hairy roots co-expressing both the T-DNA from the Ri plasmid and the T-DNA from a binary vector was performed using fluorescent markers (GFP or DsRed). The development time course of *E. grandis* transgenic hairy roots is described in Figure 3.1A. GFP or DsRed fluorescence were easily detected in calli neo-formed at wounding sites 7 days after infection and in emerging roots appearing in average one to two weeks later (Figure 3.1A, Supplemental figure S3.1). Twenty-one days after infection was found to be the optimal time to calculate co-transformation efficiencies (Supplemental figure S3.1).



**Figure 3.1: Developmental time course of *E. grandis* hairy roots expressing either the GFP or the DsRed fluorescent markers.** Panel A: stereomicrographs of roots emerging from infected sites on hypocotyls at 7 (a, e), 14 (b, f) or 21 (c, d, g, h) days after infection (dpi). GFP transformed calli or roots appeared in green whereas autofluorescence appeared in red (a, b, c). DsRed transformed calli or roots appeared in red without noticeable autofluorescence (e, f, g). Panel B: stereomicrographs of hairy roots from wild type (a, d) and composite plants overexpressing GFP (b, e) or DsRED (c, f) hardened for 1.5 month (75 dpi). Bright field was used to show root morphology and to localize untransformed roots (Panel A: d, h; Panel B: d-f). Scale bars = 2 mm.

Co-transformation efficiencies were variable depending on the plant material and the infection protocol (Table 3.1). The efficiency of the hypocotyl stabbing protocol of 14-day-old *E. grandis* seedlings (62% in average, reaching up to 75%) was found to be by far the highest.

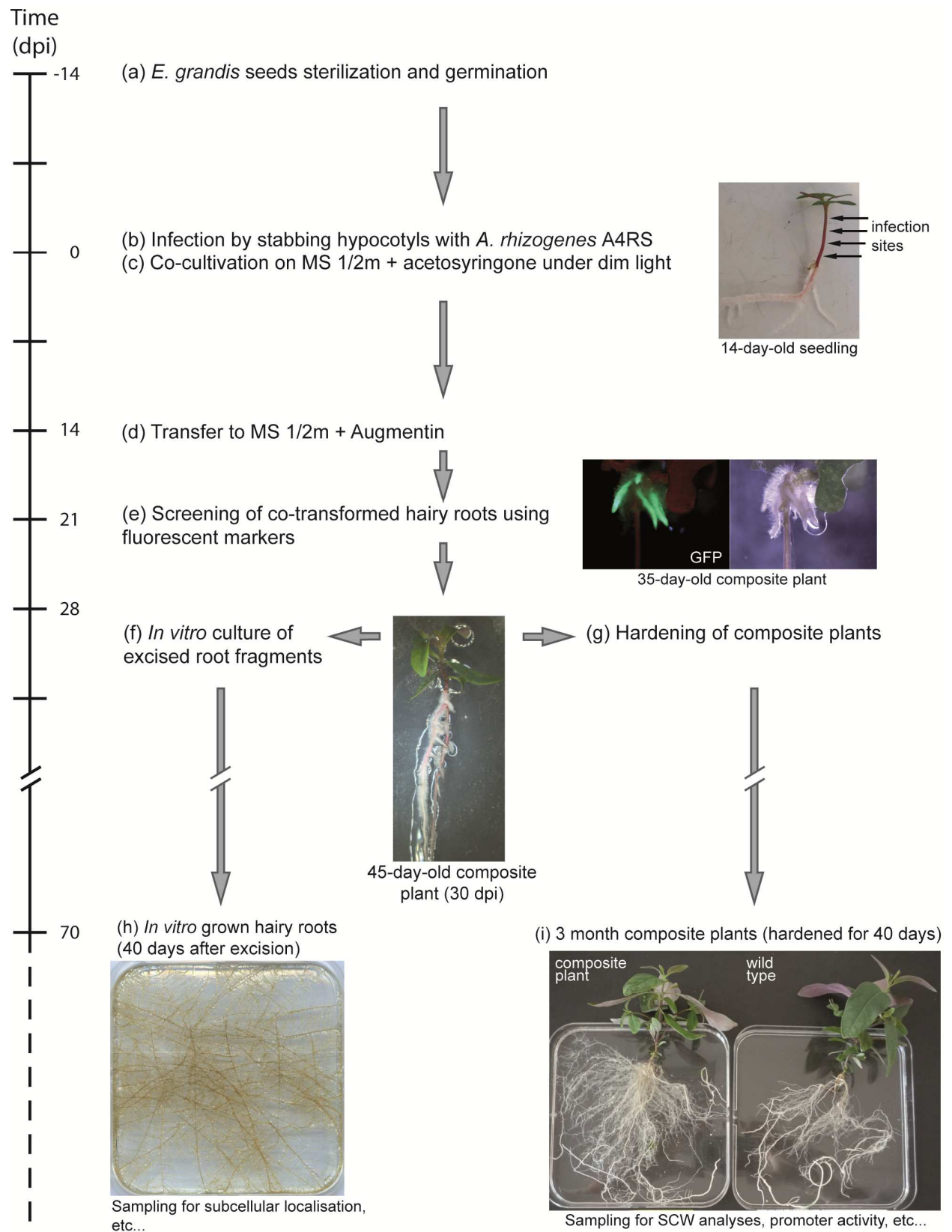
**Table 3.1: Mean efficiency of *in vitro* co-transformation of *E. grandis* explants with *A. rhizogenes*.**

<i>A. rhizogenes</i> strain	<i>E. grandis</i> Age (days)	<i>E. grandis</i> Material	Infection method	Number of analyzed plants	Co-transformed roots (%)*
A4RS	14	seedlings	Hypocotyl stabbing	96	62.2
A4RS	14	seedlings	Cut hypocotyl base	65	29.8
A4RS	3	seedlings	Cut radicle	30	12.5
A4RS	21	<i>in vitro</i> cultured clonal line	Stem stabbing	30	10.0
A4	14	seedlings	Hypocotyl stabbing	30	0.0
Arqua1	14	seedlings	Hypocotyl stabbing	30	4.0

\*Co- transformed roots efficiency calculated at 21 dpi as the percentage of plants presenting at least one fluorescent root over the total number of infected plants.

Having selected the best inoculation system and starting plant material, we tested two other *A. rhizogenes* strains bearing pA4-type Ri plasmids (A4 and ARqua1), known to elicit a limited number of Ri T-DNA-transformed roots with growth and morphology comparable to normal roots (Boisson-Dernier *et al.*, 2001; Chabaud *et al.*, 2006; Imanishi *et al.*, 2011; Quandt *et al.*, 1993). However, the infection of hypocotyls with ARqua1 led to transformation efficiency of only 4% whereas no transformed hairy roots at all were obtained with A4 (Table 3.1). As it usually happens due to the inherent recalcitrance of woody plants, only the most virulent strain (A4RS) gave satisfactory co-transformation efficiencies (Alpizar *et al.*, 2006). Therefore, all subsequent experiments were performed using the A4RS strain and the hypocotyl stabbing infection strategy following the protocol steps described in Figure 3.2.

**Figure 3.2: Sequential steps of the transformation of 14-days-old *E. grandis* seedlings by A4RS harbouring either a GFP or DsRed based binary vector.** (a) *E. grandis* seeds were sterilized and germinated in 1/4 strength MS medium. (b) 14 day-old seedlings were infected by stabbing the hypocotyl with a needle swabbed with *A. rhizogenes*. (c) Infected plants were co-cultivated with agrobacteria for 14 days on MS medium with 1/2 strength macroelements (MS 1/2m) supplemented with acetosyringone under dim light. (d) Plants were transferred to MS 1/2m medium supplemented with augmentin. (e) The hairy roots generated were examined at 21 days after infection (dpi) under a stereo fluorescence microscope. (f) Co-transformed roots were excised to be cultivated *in vitro* on MER media. (g) Before hardening, non-transformed roots were removed and the resulting composite plants were placed in pots and cultivated in a phytotron. (i) After 40 days of hardening, roots were sampled for SCW and promoter activity analyses, whereas *in vitro* grown excised roots (h) could be used for other purposes such as subcellular localisation of proteins.



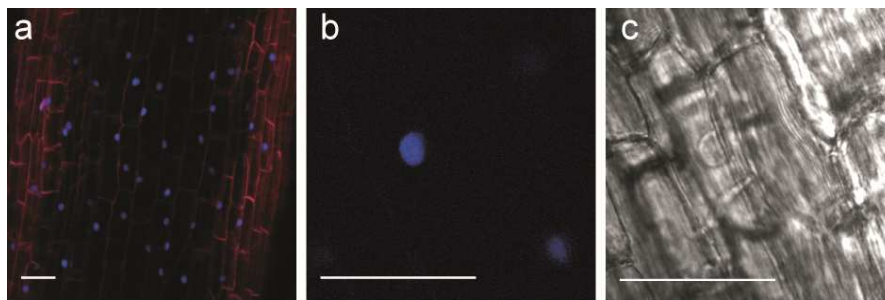
One month after infection, the transgenic hairy roots were ramified and reached a size between 4 and 8 cm allowing excision of a lateral root (2 cm long in average) for subsequent *in vitro* cultivation on solid media. The remaining fluorescent roots with the aerial part of the *in vitro*-grown composite plants were hardened for subsequent sampling and analyses (Figure 3.2f, g). Before hardening, we removed by cutting the non-transformed roots. All the composite plants survived the acclimation period and grew similarly to the non-transformed eucalypt plants. Whereas the length, thickness and branching of the composite *E. grandis* hairy roots were comparable to that of wild-type plant roots, one obvious difference was the presence of a high number of short and thin roots developing in the upper part of the hairy root system (Figure 3.2i). The fluorescence of either the GFP or the DsRed reporters was still easily detectable after 1.5 month (Figure 3.1B) and even after one year of growth. Notably, transgenic roots could restart growing after one to three sampling events when the older part (approximately one third) of the root system was left on the composite plant and still express stably the fluorescent reporter gene.

### 3.3.2 CONSTRUCTION OF EXPRESSION VECTORS FOR GENE FUNCTIONAL CHARACTERIZATION AND PROMOTER ACTIVITY ANALYSIS

Whereas the hairy roots expressing GFP were easily detectable under a fluorescent stereomicroscope, those expressing the DsRed were even easier to detect since there is no overlap between this fluorochrome and the natural autofluorescence of *Eucalyptus* roots (Figure 3.1B). For this reason, the Gateway-adapted binary vectors that we constructed for modulating gene expression (pGWAY-0, Supplemental figure S3.2) and/or enabling promoter studies (pGWAY-1, Supplemental figure S3.2) both contained the DsRed reporter gene. Using these new vectors, transformation efficiency dropped to 17.8% in average, which is lower than that obtained with the pHKN29 vector (62% in average). Estrada-Navarrette *et al.* (2007) also reported that co-transformation efficiencies vary with different binary vectors. Both vector backbone and size are parameters that can affect co-transformation efficiency. The pGWAY-0 and pGWAY-1 are bigger plasmids (17.5 Kb) derived from a pBIN19 backbone whereas the smaller pHKN29 (9 Kb) is derived from a pCAMBIA 1300 backbone.

### 3.3.3 *IN VITRO* HAIRY ROOTS CULTURES AND SUBCELLULAR LOCALIZATION OF PROTEINS

In order to determine the best culture conditions for excised transgenic hairy roots, we compared their growth curves in three solid media up to two months after excision (Supplemental figure S3.3). In the MER medium, the roots length increased slightly more than in the M medium and substantially more than in the MS medium, in which hairy roots virtually did not grow. Moreover, to demonstrate that *E. grandis* hairy roots can be used to study fluorescent-tagged proteins, we also used *in vitro* cultured roots transformed with a vector expressing an *E. grandis* histone linker (H1) fused to the cyan fluorescent protein (CFP). As shown in Figure 3.3, in all root cells the H1 protein was localized in the nucleus, consistent with its well-known subcellular localization.

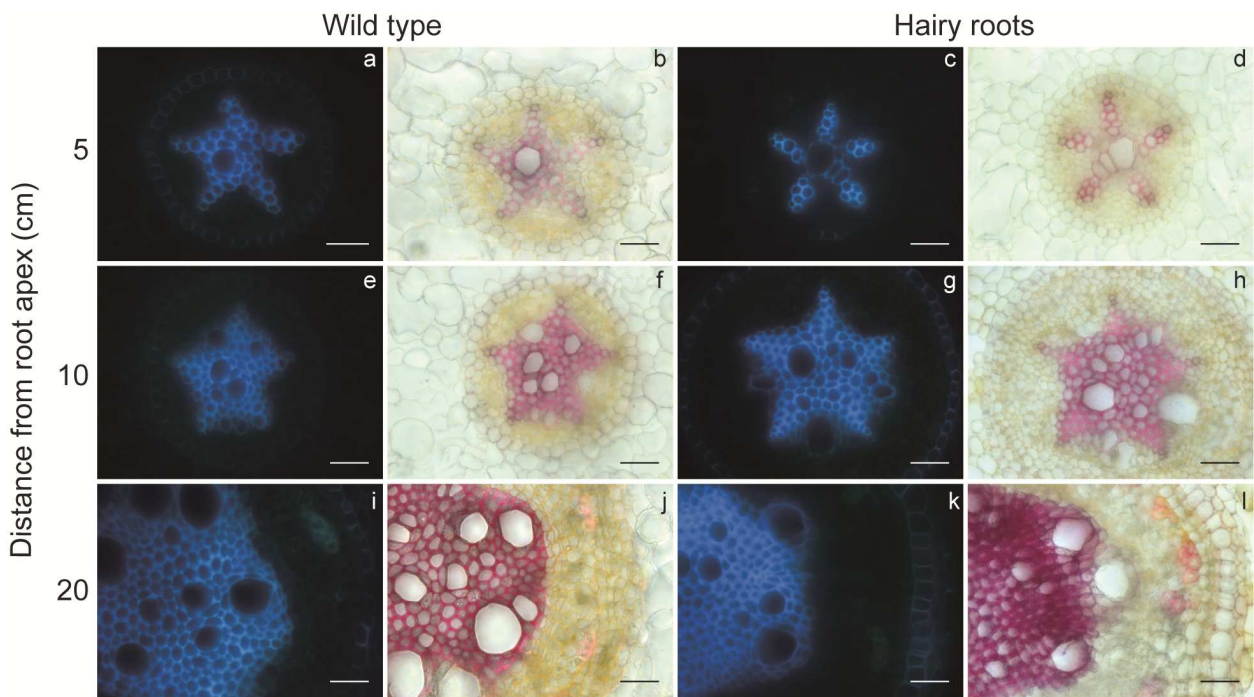


**Figure 3.3: Nuclear localization of histone linker (H1) fused to CFP in *E. grandis* root cells.** Confocal images of *E. grandis* hairy root cells expressing H1-CFP fusion protein. All nuclei fluoresced in blue and the cell walls in red (a). Detail of fluorescent nuclei under CFP emission (b) and bright field (c). Scale bars = 30 $\mu$ m.

### 3.3.4 XYLEM STRUCTURE FROM HAIRY ROOTS

To evaluate if transgenic hairy roots could be used as a proxy to explore the function of genes involved in xylem secondary cell wall formation, we compared the radial patterning and xylem anatomy of hairy roots relative to wild-type roots from the obtained composite plants. The observations under light microscopy and under UV light (exciting the natural lignin autofluorescence) showed that the development of the xylem was similar between hairy roots and non-transformed roots when sampling at the same distance from the apex (Figure 3.4). At 5 cm from the apex, primary xylem with four- to five-branches star pattern was observed in both types of roots (Figure 3.4 a-d), whereas at 10 cm xylem exhibited a nearly circular shape

implicating that secondary growth started to occur with vascular cambium derived from the procambium producing secondary xylem (Figure 3.4 e-h). At 20 cm, secondary xylem and secondary phloem were clearly distinguishable (Figure 3.4 i-l). In all stages, endodermis also fluoresced under UV light due to the presence of suberin. Both primary and secondary xylem developed in a similar way between transformed and non-transformed roots.

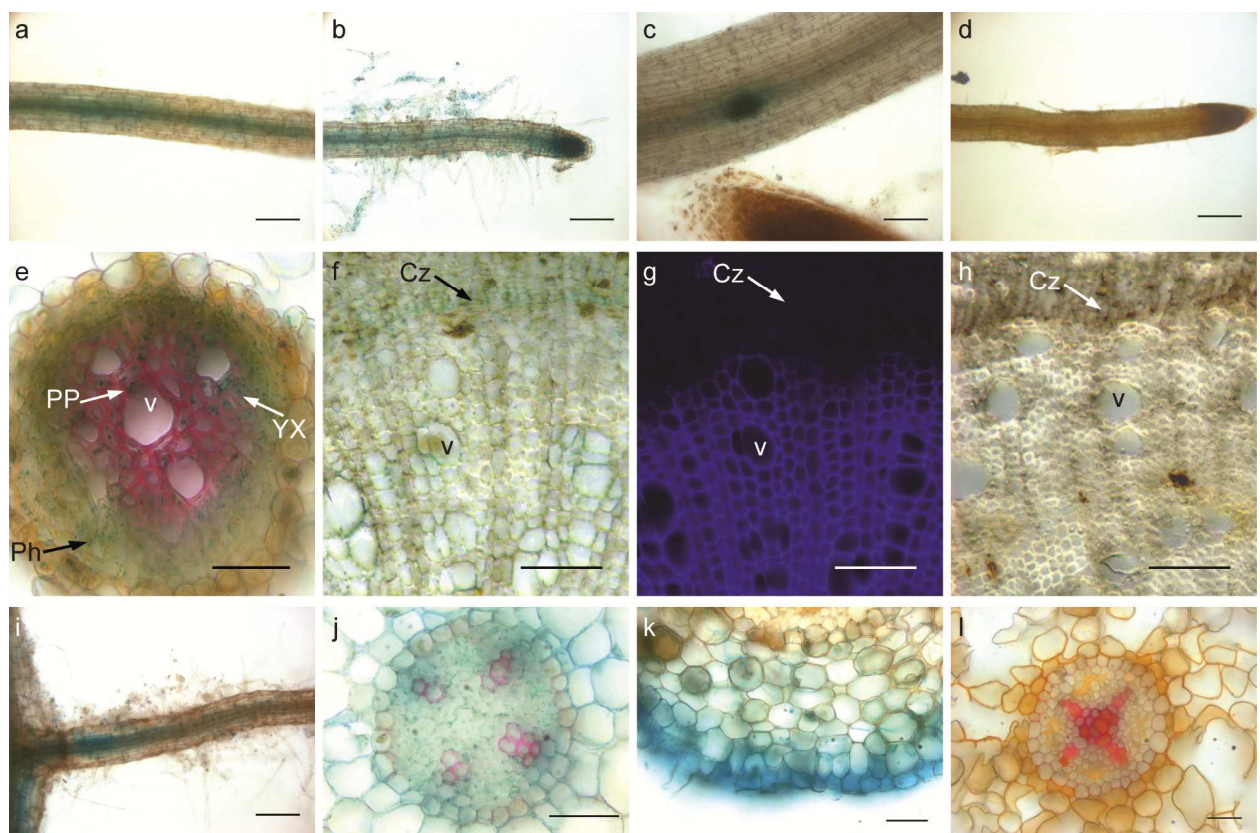


**Figure 3.4: Comparison of xylem development and lignified secondary cell walls between transgenic and wild-type roots.** Transversal root sections made at 5 (a-d), 10 (e-h) and 20 cm (i-l) from the root apex for wild-type (a, b, e, f, i, j) and hairy (c, d, g, h, k, l) roots. Lignified cell walls are visualized in blue by UV autofluorescence (a, c, e, g, i, k) and in red by using phloroglucinol-HCl (b, d, f, h, j, l). Scale bars = 30  $\mu$ m.

### 3.3.5 TISSUE AND CELL EXPRESSION PATTERN OF THE *EgCCR1* AND *EgCAD2* PROMOTERS IN HAIRY ROOTS

To validate that hairy roots were suitable to study the promoter activities of genes involved in xylem SCW formation, we produced chimeric *Eucalyptus* plants from which we analysed the transgenic hairy roots expressing the reporter gene GUS under the control of the *EgCCR1* or the *EgCAD2* promoter. Both promoters conferred a preferential expression in the vascular cylinder of hairy roots (Figure 3.5a-c, i-j). A strong GUS activity was also observed in root tips (Figure 3.5b) and at the emergence of lateral roots (Figure 3.5c). As observed in cross-sections performed at 15 cm from the root in apex of composite plants (Figure 3.5e), GUS staining was

observed in xylem lignifying cells still having a cytoplasm, which possibly are developing vessels prior to autolysis or fibres. The *EgCCR1* promoter was also expressed in paratracheal parenchyma cells located in vicinity of mature vessels elements that have already undergone autolysis and are not stained in blue. The paratracheal parenchyma cells are supposed to contribute to post-mortem lignification of vessels (Baghdady *et al.*, 2006; Pesquet *et al.*, 2013). In older roots exhibiting clear secondary growth (Figure 3.5f-g), GUS activity was also shown in cambial cells and epidermis (Figure 3.5k).



**Figure 3.5: Histochemical localization of GUS activity in *E. grandis* hairy roots transformed with the *EgCCR1* and *EgCAD2* promoters.** Roots from hardened composite plants (3 month-old) showing GUS activity for *EgCCR1* (a-c, e-g) and *EgCAD2* (i-k) promoters. Control roots are also shown (d, h, l). GUS activity is observed in the vascular cylinder (a, b, c, i), the root tip (b) and at the emergence of a lateral root (c). Transversal cross sections were performed at 2 (j, l), 15 (e, k) and 30 cm (f-h) from the root apex. Lignified cell walls are visualized in red using phlorogucinol-HCl performed after GUS staining (e, j, l) or, in the case of roots at 30 cm (f), are alternatively visualised under UV light (g). GUS activity is observed in secondary xylem cells (e, f), cells from the cambial zone (f) and cells from epidermis (k). No GUS activity is observed in control roots (d, h, l). YX: young xylem cells; PP: paratracheal parenchyma cell; Ph: Phloem; Cz: Cambial zone; v: vessel. Scale bars =100  $\mu$ m (a-d, f-i), and 30  $\mu$ m (e, j-l).

### 3.3.6 *EgCCR1* SILENCING IN HAIRY ROOTS

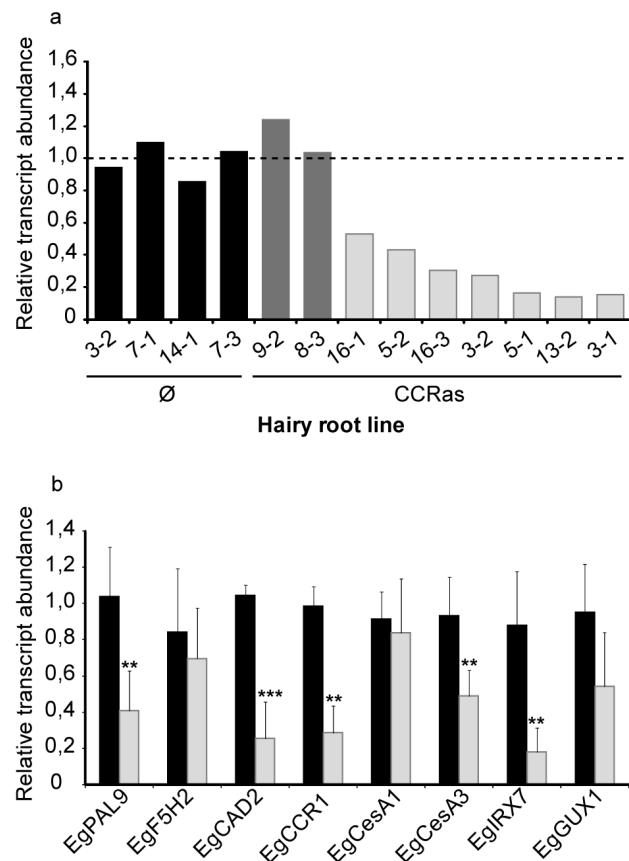
To further confirm that hairy roots were suitable to study the function of genes involved in xylem SCW formation, we performed, as a proof-of-concept, down-regulation of Cinnamoyl-CoA Reductase 1 (*EgCCR1*). We transformed the roots with an antisense *EgCCR1* construct cloned into the pGWAY-0 vector containing a DsRed marker. Four empty vectors and 10 antisense *EgCCR1* (*CCRAs*) composite plants were hardened and analysed. Some of the *CCRAs* hairy roots grew slower and showed a reduced size as compared to transgenic roots transformed with the empty vector, consistent with the reduced growth and size of the strongly *EgCCR1* down-regulated tobacco plants (Piquemal *et al.*, 1998).

Steady-state transcript levels of endogenous and antisense *EgCCR1* RNAs were analysed by RT-qPCR in the younger parts (0-5cm from the apex) of the co-transformed roots. All the transformants tested expressed the antisense *EgCCR1* transgene albeit to different levels (Supplemental figure S3.4). Three out of nine *CCRAs* lines (5-1, 13-2 and 3-1) showed a strong silencing of endogenous *EgCCR1* mRNA with residual transcript levels lower than 17% of those of control lines. Four *CCRAs* lines (16-1, 5-2, 3-2 and 16-3) expressed intermediate levels comprised between 53 and 28% of that of control lines (Figure 3.6a), whereas two lines (9-2 and 8-3) did not show endogenous *EgCCR1* silencing. This variability in the extent of down-regulation of a target gene is frequently observed with the antisense strategy (Piquemal *et al.*, 1998). Transcript level analyses were not performed for line *CCRAs* 5-3 because of the low quality of the RNA obtained, which could not be re-sampled and extracted a second time because the plant died after the first sampling. This is likely due to its very strong *CCRAs* phenotype, as shown below.

Transcript levels of five lignin (*EgPAL9*, *EgF5H1*, *EgF5H2*, *EgCAD2* and *EgCCR1*), two cellulose (*EgCesA1* and *EgCesA3*) and two hemicelluloses (*EgGUX1* and *EgIRX7*) biosynthetic genes were quantified in roots with either a strong or an intermediate level of *EgCCR1* silencing. All these genes are related to SCW biosynthesis (Carocha *et al.*, 2015; Myburg *et al.*, 2014) and, with the exception of *EgF5H1*, were expressed in young roots (Figure 3.6b). Among the analysed lignin-related genes, all but *EgF5H2* showed a significant transcript level reduction in *CCRAs* roots, meaning that lignin biosynthesis was also repressed at transcriptional level. Moreover, some of the cellulose and hemicelluloses biosynthetic genes analysed were also repressed.



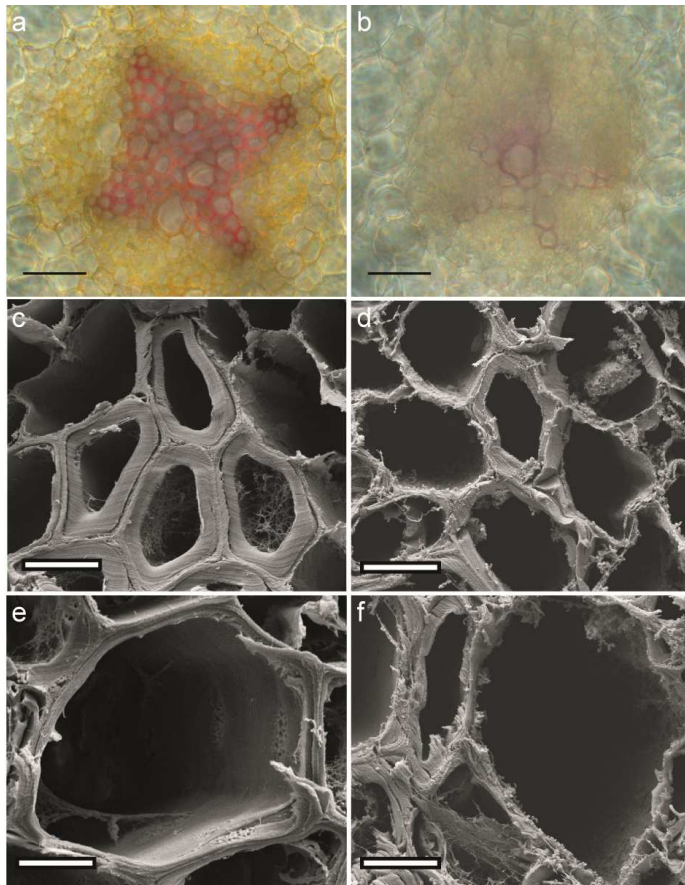
Transcriptomic analyses performed in *CCRAs* poplar plants also pointed out a reduced hemicelluloses biosynthesis (Leplé *et al.*, 2007).



**Figure 3.6: Relative transcript levels of *EgCCR1* and SCW-related genes in antisense *EgCCR1* (*CCRAs*) hairy root lines.** (a) Mean relative transcript levels of *EgCCR1* analysed by RT-qPCR. Black bars: control roots transformed with an empty pGWAY-1 vector ( $\emptyset$ ); dark grey bars: non-silenced *CCRAs* lines (9-2; 8-3), light grey bars: *CCRAs* lines exhibiting different degrees of silencing (16-1; 5-2; 16-3; 3-2; 5-1; 13-2; 3-1). (b) Transcript levels of several SCW-related biosynthetic genes were assessed in control (black bars) and silenced *CCRAs* roots (light grey bars). For each gene, mean relative transcript levels ( $\pm$  SD) for the four controls and for the seven silenced *CCRAs* lines are shown. \*\* $P < 0.01$ , \*\*\* $P < 0.001$ . *EgPAL9*: Phenylalanine Ammonia-Lyase 9, *EgF5H2*: Ferulate 5-Hydroxylase 2, *EgCCR1*: Cinnamoyl-CoA Reductase 1, *EgCAD2*: Cinnamyl Alcohol Dehydrogenase 2, *EgCesA*: Cellulose Synthase, *EgIRX7*: Glucuronoxylan glucuronosyltransferase, *EgGUX1*: glucuronyltransferase.

Histochemical observations of sections performed at 5 cm from root apex displaying a strong *EgCCR1* down-regulation showed that xylem cell walls reacted faintly to phloroglucinol staining as compared to roots transformed with an empty vector (Figure 3.7 a-b), which appeared strongly stained in red. These observations were consistently made on several silenced lines further supporting a strong reduction of the lignin content as a consequence of *EgCCR1*

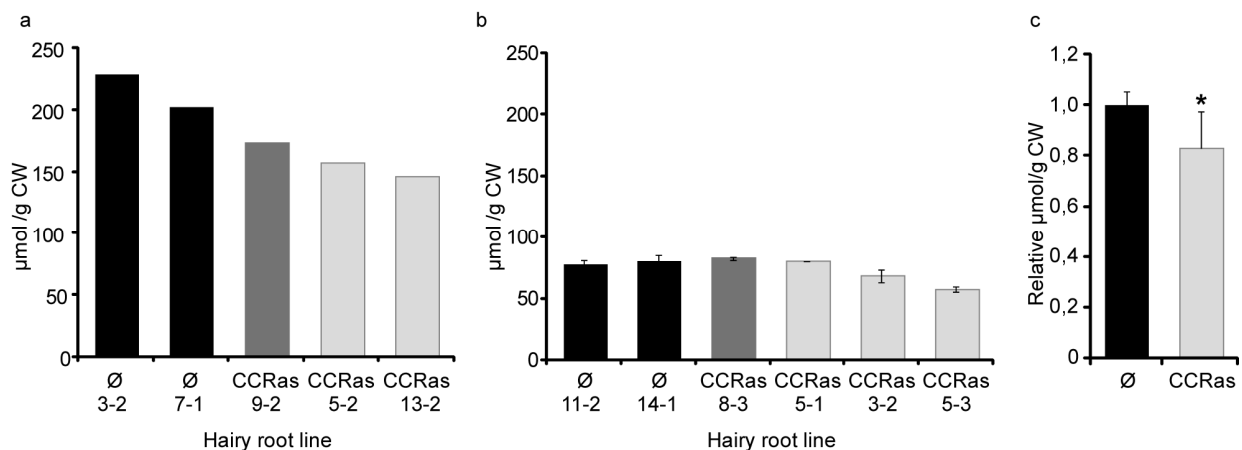
silencing. These roots observed under scanning electron microscopy (SEM) showed thinner cell walls and irregular shapes for both xylem fibres and vessels (Figure 3.7 c-f). The structure of the SCW was strongly altered resulting in collapsed vessels.



**Figure 3.7: Transversal sections of *EgCCR1* down-regulated hairy roots.** Sections were made at 5 cm from the root apex both for control roots transformed with pGWAY-0 empty vector (a, c, e) and for the *EgCCR1* down-regulated hairy roots (b, d, f). (a-b) Phloroglucinol staining; scale bars = 30  $\mu\text{m}$ . (c-f) SEM images of xylem fibers (c-d) and vessels (e-f); scale bars = 5  $\mu\text{m}$ .

Because of the low amount of available plant material, lignin evaluation was performed by thioacidolysis, a sensitive degradation method which specifically gives rise to *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) monomers from H, G and S lignin units only involved in labile ether bonds (Méchin *et al.*, 2014). The lignin-derived monomers were recovered from extractive-free samples prepared from the youngest parts of roots exhibiting only primary xylem as well as from more developed roots provided with secondary xylem. For each sample type, we analysed non-silenced, medium and strongly silenced lines. The total thioacidolysis yield, expressed in  $\mu\text{moles}$  (H+G+S) per gram of extract-free sample, gives an estimate of the amount of parent lignin structures. Consistent with their higher proportion of lignified cell walls, the older parts of the hairy roots gave rise to higher thioacidolysis yields that, in control lines, is 2.8-fold higher than those from younger ones (Figure 3.8a, b). Differences between

CCRas lines and controls were more obvious for older root fragments. When we normalized data together for the two kinds of roots used, thioacidolysis yields were significantly reduced in silenced lines ( $P < 0.05$ , Figure 3.8c). Interestingly, we also analysed by thioacidolysis samples from line CCRas 5-3 although we could not determine the extent of *EgCCR1* down-regulation as mentioned above. This line showed not only a strong phenotype (stunted roots), but also a lower thioacidolysis yield likely indicative of a lower lignin content. In addition, the percentage of G thioacidolysis monomers was reduced by 31% leading to a higher S/G ratio (Table S3.1) and, more importantly, the marker compound for CCR deficiency (G-CHR-CHR<sub>2</sub>, with R=Set) (Ralph *et al.*, 2008) was clearly observed on the corresponding GC-MS trace (Supplemental figure S3.5).



**Figure 3.8: Thioacidolysis lignin yields in CCRas roots.** Thioacidolysis lignin yields, expressed in  $\mu\text{moles}$  per gram of extract-free sample, for each individual line using older roots fragments (a) and young roots fragments (b). Normalised thioacidolysis lignin yields (mean values  $\pm$  SD) are also shown for control (4 lines) and down-regulated CCRas (5 lines) roots (c). Black bars: control roots transformed with pGWAY-0 empty vector ( $\emptyset$ ); dark grey bars, non-silenced CCRas roots; light grey bars, silenced CCRas roots. \* $P < 0.05$ .

### 3.4 DISCUSSION

Here, we describe an efficient, versatile and rapid method to obtain and screen *Eucalyptus* transgenic hairy roots enabling medium throughput functional gene characterization and gene function hypothesis testing. This homologous transformation system enables gene silencing, as shown for the *EgCCR1* gene as a case study, and compensates the lack of mutant collections. Besides, it is worth noting that for the first time the *EgCCR1* gene, cloned initially in *Eucalyptus* in 1997 (Lacombe *et al.*, 1997) and studied since in many plant species (Leplé *et al.*, 2007; Mir

Derikvand *et al.*, 2008; Piquemal *et al.*, 1998), is functionally characterized in a *Eucalyptus* homolog system.

The down-regulation of *EgCCR1* using an antisense construct led to a decrease in lignin content, an alteration of SCW structure and the presence of collapsed vessels as found in other species (Goujon *et al.*, 2003; Jones *et al.*, 2001; Leplé *et al.*, 2007; Piquemal *et al.*, 1998). Nearly all the lignin-related genes were down-regulated similarly to what was found in *CCRAs* tobacco plants (Dauwe *et al.*, 2007), with the exception of the *EgF5H2* gene. Since ferulic acid is known to accumulate in *CCRAs* plants (Mir Derikvand *et al.*, 2008), one can hypothesize that *EgF5H2* could be involved in detoxification processes via metabolising ferulic acid into sinapic acid as suggested by Meyermans *et al.* (2000), Leplé *et al.* (2007).

We further showed that the co-transformed hairy roots are a versatile system suitable for many applications, such as protein subcellular localization and analysis of promoter activity pattern. With that tool, we demonstrated that both the *EgCCR1* and the *EgCAD2* promoters conferred a preferential expression in the vascular cylinder of hairy roots. This result is in agreement with previous studies performed with the same promoters in heterologous transformation systems such as poplar (Feuillet *et al.*, 1995; Hawkins *et al.*, 1997; Lauvergeat *et al.*; 2002; Šamaj *et al.*, 1998), grapevine (Lauvergeat *et al.*, 2002), tobacco (Lacombe *et al.*, 2000; Lauvergeat *et al.*, 2002) and *Arabidopsis* (Baghdady *et al.*, 2006), all showing vascular preferential expression in both stems and roots. Consistent also with these previous studies, a strong GUS activity was observed in root tips and at the emergence of lateral roots. GUS activity was also detected in cambial cells and epidermis of older roots as previously described in poplar (Hawkins *et al.*, 1997; Lauvergeat *et al.*, 2002; Šamaj *et al.*, 1998). These results confirm that the fine spatial and temporal dissection of these *Eucalyptus* promoters was accurately studied in heterologous plant systems, in line with the highly conserved network regulating SCW biosynthesis and deposition (Zhong *et al.*, 2010a). However, the expression patterns driven by the *Eucalyptus cellulose synthase (CesA)* promoters showed differences between homologous (*Eucalyptus*) and heterologous (poplar) transformations (Creux *et al.*, 2013), further supporting the importance of homologous transformation systems such as the one presented here.

Traditional *A. tumefaciens*-mediated stable transformation and regeneration systems of *Eucalyptus* are long, tedious and low-efficient procedures (Girijashankar, 2011; de la Torre *et*

*al.*, 2014, and references therein). One major advantage to use *A. rhizogenes*-mediated transformation is the time needed to obtain material suitable for functional analyses, which is considerably reduced. One month after infection, composite plants are ready to be hardened and root fragments can also be excised to be cultivated *in vitro*. Transgenic roots from hardened composite plants developed secondary xylem similar to that of untransformed roots. In addition, after harvesting, hairy roots can restart growing from the remaining upper part (at least 15 cm), providing biological replicates. Roots can also be excised from *in vitro* composite plants and grown on solid medium to perform analysis in a more controlled environment. In addition, it is in theory possible to regenerate the whole plant from a transformed hairy root, as previously done for instance for *Allocasuarina verticillata* (Phelep *et al.*, 1991).

The *Eucalyptus* hairy root system developed in this study has also several advantages as compared to another alternative method called “Induced somatic sector analysis” or “ISSA”, previously developed by Spokevicius *et al.* (2005) and Van Beveren *et al.* (2006) to overcome the low efficiency and time-consuming *A. tumefaciens*-mediated transformation and regeneration of *Eucalyptus* plants. This method is based on *A. tumefaciens* transformation but restricted to wood sectors of stem from actively growing trees, where the transgene is transferred into actively dividing cambial, xylem, phloem and ray initial cells. Being successfully applied in the analysis of transgenes in woody stem tissues of *Pinus*, *Populus* and *Eucalyptus*, ISSA can only be used to functionally characterize genes involved in the formation of vascular tissues and is particularly useful for promoter analyses (Creux *et al.*, 2013; Hussey *et al.*, 2011). In contrast, the hairy root system can be used for a vast panel of genes, thus being a more versatile system, with the only limitation that for getting the most accurate results, genes of interest should be expressed in roots. Another important advantage of the transgenic hairy root system in comparison to ISSA is the quantity of transgenic material obtained. ISSA allows obtaining very limited regions of transformed cells surrounded by untransformed cells, thus making difficult the harvesting and analysis of transgenic wood. In contrast, transgenic hairy roots are homogeneously transformed and easily detected with fluorescence, and provide sufficient quantities of material for molecular biology, histochemical and biochemical approaches, as illustrated for the characterization of *EgCCR1* down-regulated hairy roots.

To the best of our knowledge, less than a handful of studies were dedicated to *Eucalyptus* hairy roots. With the aim of improving the rooting capacities of some *Eucalyptus* clones, the early work of MacRae and van Staden (1993) had evaluated their susceptibility to several *A. rhizogenes* strains. Later on, a protocol for generation of composite plants of *E. camaldulensis* using a GFP-based screening was briefly described (Balasubramanian *et al.*, 2011) but no functional gene characterization was achieved. In that study, 68% of plants showed root induction, but only 36% of those were co-transformed, resulting in a co-transformation efficiency of 25%. Using the same GFP-based vector (PHKN29), we obtained much higher co-transformation efficiencies (62%), while using the larger pGWAY vectors, the efficiencies were of the same order of magnitude of that obtained by Balasubramanian and colleagues (2011) with the small pHKN29 vector. A recent study reported gene silencing in *E. urograndis* (*E. grandis* × *E. urophylla*) hairy roots of STOP1, a transcription factor regulating aluminium and proton tolerance in plants (Sawaki *et al.*, 2014), but the *A. rhizogenes* transformation method used was not detailed and surprisingly only referred to that of tobacco (Sawaki *et al.*, 2013).

In conclusion, we demonstrated here the powerfulness of the hairy roots as a homologous, versatile, rapid and efficient transformation system enabling medium throughput functional characterization of genes, especially those involved in xylem and secondary cell wall formation, which are of particular importance in woody plants. The development of such a tool that can be used easily and in routine is timely, since it will accelerate gene mining of the recently sequenced *Eucalyptus* genome both for basic research and industry purposes. Moreover, the composite plants as well as the excised roots open a whole range of new applications such as, for instance, the study of interactions with root pathogens and mycorrhizal fungi.

## 3.5 EXPERIMENTAL PROCEDURES

### 3.5.1 PLANT MATERIAL

Commercial *E. grandis* seeds (W. Hill ex Maiden, cultivar LCF A001) purchased at IPEF (Instituto de Pesquisas e Estudos Florestais, Piracicaba, Brazil) were surface-sterilized by 15 min treatment in a 1% sodium hypochlorite solution containing Twin-20. Germination was carried out on ¼-strength Murashige and Skoog (MS medium, Sigma-Aldrich, St. Louis, USA) solidified

with Phytigel 5.5 g/L (Sigma-Aldrich) at 25°C in the dark for three days. To obtain 0.5 to 1 cm long radicle, germination was performed at 25°C for three days in dark. We inverted the plates to allow growth of the radicle towards the exterior of the media. To obtain *in vitro* plantlets around 8 mm long with the hypocotyls fully expanded and the first two leaves just appearing, we germinated seeds inside plates at normal position at 25°C for three days in dark followed by 11 more days in light (12  $\mu\text{mol m}^{-2} \text{s}^{-1}$ , 8-16h photoperiod, 50% humidity). We also used an in-house *E. grandis* line obtained by *in vitro* clonal propagation of one seedling, grown in M media (Tournier *et al.*, 2003) for three weeks in conditions described above.

### 3.5.2 *AGROBACTERIUM RHIZOGENES* STRAINS

Three strains of *A. rhizogenes* were tested: (i) the A4 agropine mannopine-type strain originally isolated from roses exhibiting hairy root symptoms (Moore *et al.*, 1979), (ii) the A4RS strain, an A4-derived strain selected as resistant to the antibiotics rifampicin and spectinomycin (Jouanin *et al.*, 1986) and (iii) the low virulence ARqua1 strain, a Smr-derivative of *A. rhizogenes* strain A4T (Boisson-Dernier *et al.*, 2001; Quandt *et al.*, 1993).

The binary vectors were introduced into *A. rhizogenes* strains by electroporation (Nagel *et al.*, 1990) and transformants selected on 50 mg/L kanamycin. Prior to infection, ARqua1, A4 and A4RS strains were grown for 72 h at 28°C in LB, 2xYT and AG (Franche *et al.*, 1997) solid media, respectively, supplemented with acetosyringone 100  $\mu\text{M}$  and kanamycin 50 mg/L. In the case of A4RS strains, all media used were also supplemented with spectinomycin 100 mg/L.

### 3.5.3 BINARY VECTOR CONSTRUCTION

To optimize the transformation protocol, we used pCAMBIA-based binary vectors expressing either the GFP (pHKN29, Kumagai and Kouchi, 2003) or the DsRed fluorescent proteins (pCAMBIA 2200 derived vector, Fliegmann *et al.*, 2013), allowing easy detection of co-transformed hairy roots.

We constructed two Gateway-based binary vectors pGWAY-0 and pGWAY-1 (Supplemental figure S3.2) to overexpress genes of interest under the 35S CaMV promoter, and to investigate promoter activities, respectively. The construction of the vectors is described in Supplemental

figure S3.2 legend. The *EgCCR1* cDNA (Lacombe *et al.*, 1997), the *EgCAD2* promoter (Feuillet *et al.*, 1995), the *EgCCR1* promoter (Gago *et al.*, 2011) and the Histone linker protein (Eucgr.I02364) were amplified using primers listed in Table S3.2 and cloned into binary vectors as described in Supplementary Information S1.

#### 3.5.4 INDUCTION OF HAIRY ROOTS AND GENERATION OF COMPOSITE *E. GRANDIS* PLANTS

We tested four different transformation protocols using at least 30 plants for each. We removed the tip of the emerging radicle of 3-day-old seedlings and infected the wound surface as previously described (Boisson-Dernier *et al.*, 2001). We also inoculated 14-day-old seedlings by cutting and infecting the base of hypocotyls based on the work of (MacRae and Van Staden, 1993). We finally used the hypocotyl/stem-stabbing infection protocol as reported (Balasubramanian *et al.*, 2011) using an infected 0.45 mm-thick needle to inoculate 14-day-old seedlings and 3 week-old clonal lines of *in vitro* grown *E. grandis* plantlets.

In all cases, co-culture was performed after infection for 14 days in MS with ½ strength of macroelements supplemented with 30 g/L of sucrose and 100 µM acetosyringone, at 22-20°C, 40% humidity and under dim light ( $7 \mu\text{mol m}^{-2} \cdot \text{s}^{-1}$ , 8-16h photoperiod conditions). Plants were further transferred on the same medium without acetosyringone and supplemented with 300 mg/L augmentin (amoxicilin/ clavulanic acid) to prevent agrobacteria growth (Ieamkhang and Chatchawankanphanich, 2005). Plants were grown in controlled conditions ( $12 \mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ , 8-16h photoperiod, 40% humidity) and the screening for co-transformed hairy roots was performed 21 days after infection (21 dpi). GFP or DsRed fluorescence indicating co-transformed roots was detected using a stereomicroscope. Plants with at least one co-transformed root (containing both Ri and a binary vectors) were considered as positive. To calculate transformation efficiencies, from two to five independent transformation experiments were performed.

#### 3.5.5 HARDENING OF *EUCALYPTUS* COMPOSITE PLANTS

In order to obtain roots containing enough xylem, 45 day-old composite plants were then cultivated in 200 mL pots on Oil-Dri US-Special Substrate (Type III / R; Damolin, Fur, Denmark) in a phytotron ( $130 \mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$ , 8-16h photoperiod, 25-22°C, 80% humidity). Plants were



currently watered with tap water and once a week using MS with ½ strength of macroelements solution. After 4-8 weeks, GFP or DsRed fluorescence was verified again. Fluorescent roots were then sampled for RNA extraction (segments of five first cm from root apex) and for biochemical analyses (many segments from 0-10 cm from root apex for young material, or a single thick segment 10 cm long from 20-30 cm from the main root for old material). Samples were immediately frozen with liquid nitrogen and further kept at -80°C until analysis. Root segments (at least 20 cm from the root apex) were fixed with ethanol 80% for histology analyses or directly placed into GUS buffer (see below) for promoter activity analyses. After sampling, aerial part was partially pruned to prevent imbalance between aerial and root parts, and plants were maintained in the phytotron.

### 3.5.6 HAIRY ROOT CULTURES

Fluorescent root tips at least 2 cm long were excised from *in vitro* composite plants at 30 dpi and grown with augmentin on solid M media (Bécard and Fortin, 1988) containing 0.3% phytigel and subsequently sub-cultured in order to obtain clonal lines. After at least 2 transfers in M media, three media all supplemented with sucrose 30 g/L were tested to optimize root growth: (i) M medium (Bécard and Fortin, 1988), (ii) MS medium with ½ strength of macroelements and (iii) MER medium (modified M media by increasing KNO<sub>3</sub> to a final concentration of 2.5 mM, and adding NaH<sub>2</sub>PO<sub>4</sub> to 0.9 mM). Elongation rate of roots was studied by placing ten 2 cm-tips of lateral roots belonging to four different transformation events on each medium for 60 days in Petri dishes. The linear elongation of each individual root was measured every 10 days.

### 3.5.7 RNA EXTRACTION AND REVERSE TRANSCRIPTION - QUANTITATIVE POLYMERASE CHAIN REACTION (RT-QPCR)

We tested several segments of the roots and several protocols of RNA extraction detailed in S2. The best RNA yield and quality were obtained using 0-5 cm roots grinded by Fastprep (MP Biomedical, Santa Ana, USA) extracted with modifications of the protocol of (Muoki *et al.*, 2012) detailed in Supplemental Information S2.

Remaining traces of DNA were removed with Turbo DNA-free DNase I (Ambion, Austin, USA). RNA quality and quantity was checked with a Nanodrop ND-1000 Spectrophotometer (Thermo Fisher Scientific, Waltham, USA). First-strand cDNA synthesis was performed with 1 µg of total RNA using High Capacity cDNA RT Kit (Applied Biosystems, Foster City, USA). RT-qPCR was performed in technical triplicates using ABI 7900HT fast real-time PCR system (Applied Biosystems) with the Power SYBR Green PCR Master Mix (Applied Biosystems). Relative Transcript Abundance was calculated using the  $2^{-\Delta\Delta C_t}$  method (Livak and Schmittgen, 2001) with the housekeeping genes *PP2A-1* and *PP2A-3* to normalize data (Cassan-Wang *et al.*, 2012) and the mean of the control samples to standardize results.

*EgCCR1* transcript level was assessed using primers annealing at the 3'UTR region, which do not amplify the *EgCCR1 antisense* messenger, and primers hybridizing against the CDS region of the *EgCCR1* gene, thus quantifying both the *EgCCR1* gene and the *EgCCR1 antisense* mRNA levels (Table S3.3). We also measured relative transcript abundance for some genes involved in lignin, cellulose and hemicellulose biosynthesis, selected according to their expression pattern in *Eucalyptus* (Myburg *et al.*, 2014; Carocha *et al.*, 2015). Genes selected were *EgPAL9*, *EgF5H1*, *EgF5H2* and *EgCAD2* involved into the lignin biosynthetic pathway, *EgCesA1* and *EgCesA3* involved in cellulose biosynthesis, and *EgFRA1/IRX7* and *EgGUX1* related to hemicellulose biosynthesis. All primers are listed in Table S3.3.

### 3.5.8 BIOCHEMICAL LIGNIN ANALYSES

Young root samples from (0-10 cm from root apex) and old root samples (single thick segment 10 cm long from 20-30 cm from the main root) were ground to a fine powder using a ball-mill (MM400, Retsch, Haan, Germany). These samples were subjected to exhaustive water, then ethanol extraction in a Soxhlet apparatus. The recovered extract-free samples were dried at 40°C overnight before analysis by thioacidolysis as previously described (Méchin *et al.*, 2014).

### 3.5.9 HISTOCHEMICAL GUS ASSAY

Histochemical GUS assay was based on methods described by (Hawkins *et al.*, 1997). Briefly, transformed roots sections were incubated at 37 °C for 12 hours in 100 mM sodium phosphate (pH =7.0), 10 mM EDTA, 0.5 mM  $K_3[Fe(CN)_6]$ , 0.5 mM  $K_4[Fe(CN)_6]$  and 1 mM 5-bromo-4-chloro-

3-indolyl glucuronide (X-Gluc, Euromedex, Souffel Weyersheim, France). Samples were further fixed using 3.7% formaldehyde, 5% acetic acid and 50% ethanol buffer.

### 3.5.10 MICROSCOPY AND CELL IMAGING

GFP or DsRed fluorescence indicating co-transformed roots was detected using a stereomicroscope Axiozoom V16 (Zeiss, Marly le Roi, France) equipped with a colour CCD camera (ICC5, Zeiss) and with filter sets for GFP (525/50nm) and for DsRed (607/80nm). CFP fluorescence was analyzed with a Confocal Laser Scanning Microscope (TCS SP2-SE, Leica, Wetzlar, Germany).

Transverse sections (50  $\mu$ m thick) of roots embedded in 5% low gelling point agarose (Sigma-Aldrich) were obtained using a vibratome (VT 100S, Leica) and observed using an inverted microscope (DM IRBE, Leica) equipped with a CDD colour camera (DFC300 FX, Leica), or using the stereomicroscope. Lignified SCW were visualised either in red by Phloroglucinol-HCl staining or in blue due to autofluorescence under UV light.

Sections were also dehydrated in an ethanol series, submitted to a critical point dry with CO<sub>2</sub> as a transitional fluid using a CPD300 unit (Leica) and, once dried, coated with nickel (5 nm) in a EM MED020 coating system (Leica) and analysed using a scanning electron microscope (ESEM Quanta 250 FEG, FEI, Mérégnac, France) at 5KV.

### **Acknowledgements**

The authors are grateful to C. Rosenberg for kindly providing a pCAMBIA-derived vector as well as for useful advice while starting this work, F. Carvalho-Niebel and the Endomycorrhizal symbiosis and Cell signaling team (LRSV) for their precious advice, L. Marchesi for her help with transformation, R. Ployet for providing primers for *GUX1* gene. We also thank the TREEFORJOULES project (ANR-2010-KBBE-007-01), the CNRS, the Ministry of Education, Research and Technology, France (fellowship to A.P.) and the Toulouse III University (UPS) for their financial support.

The authors declare that they have no conflict of interest

### 3.6 SUPPLEMENTAL TABLES AND FIGURES

**Table S3.1:** Mean values  $\pm$  SD for H, G and S monomers percentage and for S/G ratio from young roots of control ( $\emptyset$ , empty pGWAY-0 vector) or CCRas lines. Three technical replicates were performed for each sample.

Sample	%H	%G	%S	S/G
$\emptyset$ 11-2	0.2 $\pm$ 0.0	47.4 $\pm$ 0.8	52.4 $\pm$ 0.9	1.11 $\pm$ 0.04
$\emptyset$ 14-1	0.3 $\pm$ 0.0	48.4 $\pm$ 0.5	51.3 $\pm$ 0.5	1.06 $\pm$ 0.02
CCRas 8-3 (not silenced)	0.6 $\pm$ 0.0	45.2 $\pm$ 0.7	54.2 $\pm$ 0.7	1.20 $\pm$ 0.03
CCRas 5-1	0.2 $\pm$ 0.0	42.0 $\pm$ 0.3	57.8 $\pm$ 0.3	1.37 $\pm$ 0.02
CCRas 3-2	0.5 $\pm$ 0.0	47.8 $\pm$ 1.2	51.8 $\pm$ 1.2	1.08 $\pm$ 0.05
CCRas 5-3 (stunted roots)	0.1 $\pm$ 0.0	33.1 $\pm$ 0.2	66.9 $\pm$ 0.2	2.02 $\pm$ 0.02

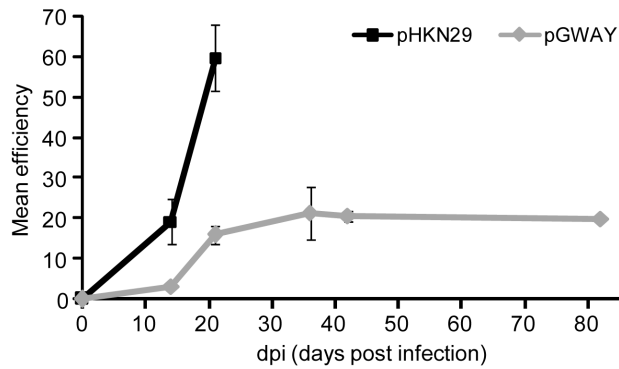
**Table S3.2:** Primer sequences used for cloning

Gene	Accession	Primer forward	Primer reverse
<i>EgCCR1as</i>	Eucgr.J03114.1	CACCCGCACCGTGATGGATCTAA	CACCTCCTGAACCCCTCTC
Prom <i>EgCAD2</i>	X65631.1	CACCTGAGCAAGTACCCACATCAA	TTTTGCTCAAAGATCCAAGC
<i>EgH1</i>	Eucgr.I02364.2	GGGGACAAGTTTGTACAAAAAAGCAGG CTTCGCCATGTCGACCACTGTAGA	GGGGACCACTTTGTACAAGAAAGCTGGGT TAGCAGCAGCTACCTTCTTGG

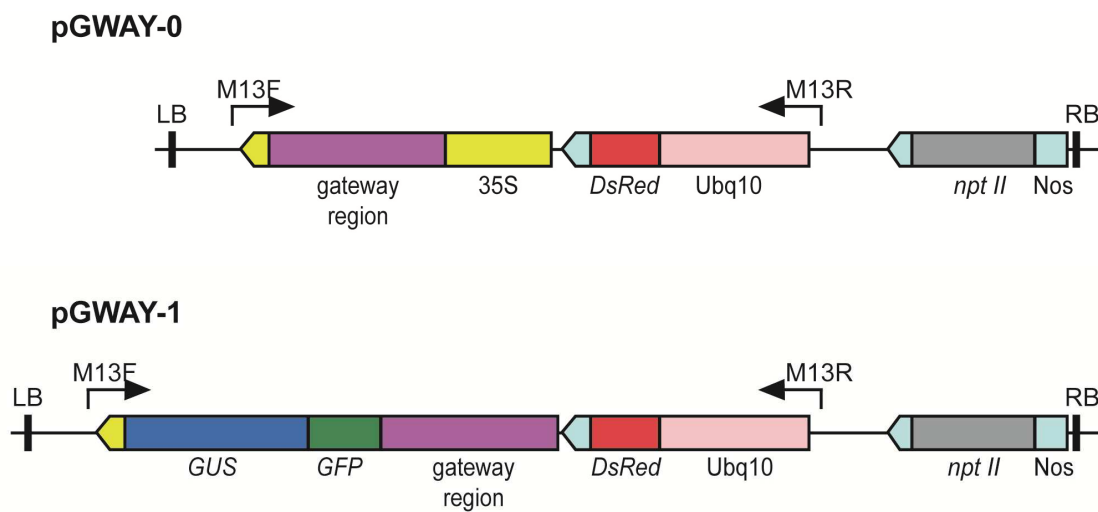
**Table S3.3:** Primer sequences of genes used for RT-qPCR analyses

Gene	Accession	Primer forward	Primer reverse
<i>EgCCR1-3'UTR</i>	Eucgr.J03114.1	TCCGGAGAAATGAGAGAAACATG	AATGCAGAACTCAGGGTTGG
<i>EgCCR1-CDS</i>	Eucgr.J03114.1	GCGATGTGGTGGAAATCCTTGC	GGGTTACCTCATCAGAGCACTTG
<i>EgPAL9</i>	Eucgr.J01079.1	TGGAAGGAGCCAGAATCATGCC	TTGTGGACCGACCATAATGTGC
<i>EgF5H1</i>	Eucgr.J02393.1	AAGCAAATGGAGGGTCGGGTTG	TCCAAATCTTGCAGCCCTCCTG
<i>EgF5H2</i>	Eucgr.I02371.1	ATTGAGACGAGGCATCGAACCG	CGCATGCAAAGCCAGCCATTTT
<i>EgCAD2</i>	X65631.1	TTACCTGGCCTTGTGAAGC	TGACTTCTCCCAAGCATAACC
<i>EgCesA1</i>	Eucgr.D00476.1	AAGTGGACGACGCTGCTGATAC	AACGACACCCACCATGTTACAG
<i>EgCesA3</i>	Eucgr.C00246.1	GTCCTGAGGATGACGATAACGC	ATTGGGAACCTGCCACTAACAG
<i>EgIRX7</i>	Eucgr.J00384	ACTGTACAGGCTGTGAGCATCG	TCCACGCATAAGCTCTCAATCCC
<i>EgGUX1</i>	Eucgr.H04942	GGCCAAATCTGAGTGTCTGAG	TCCAGTGTA AACCCGTTCTCTG

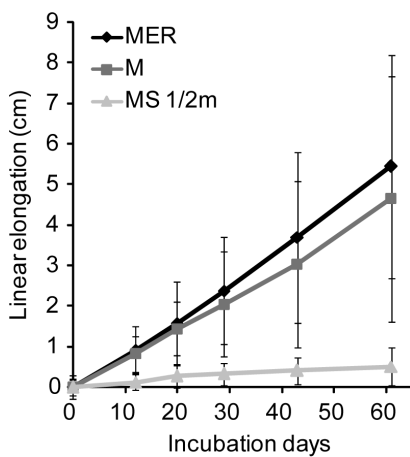
Primers for *EgCCR1-CDS*, *EgPAL9* and *EgF5H2* were designed by Carocha *et al.* (2015). Other primers were designed using Quantprime (Arvidsson *et al.*, 2008) or Primer3 (Untergasser *et al.*, 2012).



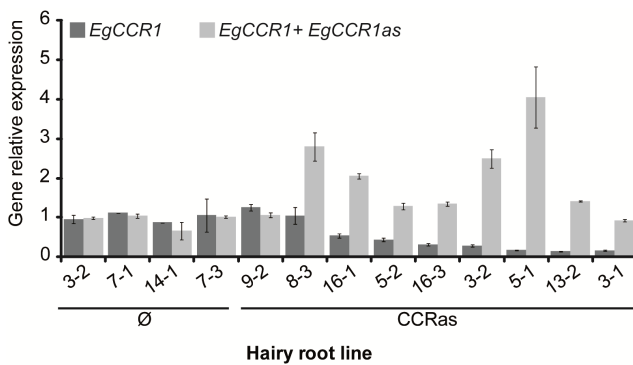
**Figure S3.1:** Co-transformation efficiency time course of *E. grandis* explants using *A. rhizogenes* strains A4RS transformed either with pHKN29 or with pGWAY vectors. Mean values and standard errors are represented. At least two independent experiments including at least 30 plants for each time point were used except for time 80 dpi, where a single experiment including 60 plants was used.



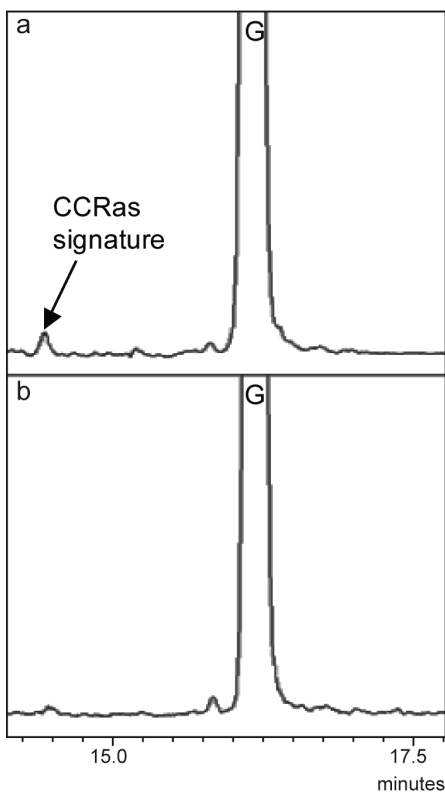
**Figure S3.2:** Maps of pGWAY-0 and pGWAY-1 destination vectors. LB: T-DNA left border; M13F and M13R: Hybridization sites for primers M13 forward and M13 reverse respectively; 35S: *CaMV* 35S promoter; *GUS*:  $\beta$ -glucuronidase; *GFP*: enhanced green-fluorescent protein (GFP) linked to the endoplasmic reticulum-targeting signal; gateway region: GATEWAY™ cassette (including attR1, CmR, ccdB, attR2 features); *DsRed*: *DsRed* fluorescent protein; *Ubq10*: *Arabidopsis thaliana* Ubiquitin-10 promoter; *npt II*: neomycin phosphotransferase II; *Nos*: *Nopaline synthase* promoter; RB: Right border. 35S and *Nos* terminators are shown as yellow and light blue arrows respectively. To construct the pGWAY-0, a 3 Kb fragment consisting of the 35S *CaMV* promoter, the gateway cassette and the 35S *CaMV* terminator was excised from the pK7WG2D vector (Karimi *et al.*, 2002) by digestion with Kpn I/HindIII and cloned into the binary vector pBIN19 that confers both plant and bacteria kanamycin resistance (Bevan *et al.*, 2010). A fragment containing the *A. thaliana* ubiquitin-10 (*Ubq10*) promoter, the *DsRed* cDNA and a *Nos* terminator from pRedRoot (Limpens *et al.*, 2004) was excised by digestion with HindIII, cloned in pGEM®-T Easy vector (Promega Corp, Madison, USA) and inserted into the modified pBIN19 vector. To construct the pGWAY-1 vector, a 4.5 kb fragment consisting of the gateway cassette, the cDNA encoding the fusion protein eGFP-*GUS* (enhanced green-fluorescent protein -  $\beta$ -glucuronidase genes) and the 35S *CaMV* terminator was amplified by PCR from pKGWFS7 (Karimi *et al.*, 2002), using the primer pair 5'-ACTGAGAagcttGCGGCCGACTAGTGATA-3' and 5'-CTGAGggtaccCCTGCAGGTCAGTGGATTTT-3' containing a HindIII and a KpnI restriction sites, respectively (in lower case). The PCR fragment was cloned into the pGEM®-T Easy vector (Promega), excised by digestion with KpnI/HindIII and cloned into the pBIN19 vector. The *DsRed* cassette was added to the vector as mentioned above for the pGWAY-0 vector.



**Figure S3.3:** Linear growth of excised *E. grandis* hairy roots. Roots were cultivated on three different media: MER (black), M (dark grey) or MS ½ strength macroelements (light grey). Mean linear growth  $\pm$  SD of 4 distinct hairy roots lines with at least 2 roots per line transformed with an empty pGWAY vector are shown.



**Figure S3.4:** Relative transcript level of *EgCCR1* gene (dark grey) and *EgCCR1* antisense mRNA + *EgCCR1* mRNA (light grey) for each line of controls (4 lines) and *CCRas* (7 silenced lines and 2 non-silenced lines) hairy roots. Transcript accumulation was quantified by RT-qPCR. Mean relative expression levels ( $\pm$  SD) were calculated using *EgrPP2A1* and *EgrPP2A3* as references. Roots transformed with pGWAY-0 empty vector construction were used as controls. Three technical replicates were performed.



**Figure S3.5:** Partial GC-MS chromatograms reconstructed at m/z 269 showing the guaiacyl (G) thioacidolysis monomers released from 10 cm segments of young roots of silenced *CCRas* 5-3 (a) and control lines (b). The vertical axis has been zoomed so as to clearly observe the *CCRas* signature from the silenced *CCRas* line while present as a trace component from the control line.

## 3.7 SUPPLEMENTAL INFORMATION

### 3.7.1 INFORMATION S1: DETAILS ON BINARY VECTOR CONSTRUCTION.

The *EgCCR1* cDNA (Lacombe *et al.*, 1997) was amplified to be cloned in antisense (as) orientation into the pENTR/D-TOPO vector (Invitrogen, Paisley, UK) and then transferred into the pGWAY-0 vector using the LR clonase II (Invitrogen). Similarly, the sequence of the *EgCAD2* promoter (Feuillet *et al.*, 1995) and the sequence of the *Histone linker* (Eucgr.I02364.2) were also amplified and cloned into the pENTRY/D-TOPO and pDONR207 (Invitrogen), respectively, to be subsequently transferred respectively to the pGWAY-1 and pBin19-35S-GW-CFP vector (Froidure *et al.*, 2010) using LR clonase II. Primers used for cloning are described in Table S3.2. For the *EgCCR1* promoter, we modified the pBIN19-derived binary vector from (Gago *et al.*, 2011) by inserting the DsRed cassette using the procedure described for constructing the pGWAY-0 vector (Fig S3 legend).

### 3.7.2 INFORMATION S2: RNA EXTRACTION OPTIMISATION AND DETAILED RNA EXTRACTION PROTOCOL

For RNA extraction, we tested several segments of the roots and several protocols. The tested root segments comprised 0-5, 0-10 and 0-15 cm from the apex, and we used around 100 mg of fresh material for extraction. The tested protocols for RNA extraction included a CTAB-based protocol with LiCl precipitation (Southerton *et al.*, 1998), a fast CTAB-based protocol without LiCl precipitation (Muoki *et al.*, 2012), RNAeasy plant mini kit (Qiagen, Hilden, Germany) and Power plant RNA isolation kit (MoBio, Carlsbad, USA). Because of the abundance of phenolic compounds in *E. grandis* roots, extraction using kits failed in all cases. In contrast, both CTAB-based methods worked pretty well in root fragments from 0-5 cm, but they failed in more than

half of the samples when segments from 0-10 and 0-15 were used. In addition, when a root fragment had a strong brown or reddish colour, extraction could fail even using 0-5 cm root tips segments. We finally chose the method of Muoki *et al.* (2012), which was more rapid and gave slightly better yields than the method of Southerton *et al.* (1998), but with small modifications further detailed. We also tested grinding roots before extraction into a fine powder under liquid nitrogen in a ball-mill (MM400, Retsch, Haan, Germany) and a less strong procedure grinding roots mixed with extraction buffer using Fastprep (MP Biomedicals, Ill Kirch, France). Both procedures gave satisfactory results, and even if in general ball-mill grinding gave somewhat better yield, using Fastprep is faster and allows RNA extraction even from only 4-5 root tips of 5 cm each. The final method we currently used is detailed as follows:

1. Take 4-5 young and white-coloured root tips of 5 cm each inside a 2 mL tube and immediately froze in N<sub>2</sub>.
2. Add 2% polyvinylpolypyrrolidone (PVPP) and 2% β-mercaptoethanol to the extraction buffer I and warm it at 65° C. Extraction buffer I composition is 2% cetyltrimethylammonium bromide (CTAB), 100 mM Tris-HCl pH 8.0, 25 mM EDTA pH 8.0 and 2M NaCl.
3. Add 900 µL extraction buffer and one ¼ inch ceramic bead (MP Biomedical, Santa Ana, USA) inside the 2 mL tubes containing root samples. Immediately grind samples using the Fastprep (MP Biomedical) at speed 4 m/s during 20 s. Repeat grinding 3-4 times for better results.
4. Incubate tubes at 65° C for 15 min. Mix by shaking every 5 min.
5. Add 900 µL of chloroform:isoamylalcohol (CIA) 24:1. Vortex thoroughly for 20 s and centrifuge at 16.000 g for 10 min at room temperature.



6. Transfer the aqueous phase to a fresh 2 mL tube and repeat step 5. Ceramic beads can be recovered from the organic phase, washed and reused in following extractions.
7. Transfer the aqueous phase to a fresh 2 mL tube. Add 1 mL of extraction buffer II and vortex thoroughly for 20 s. Extraction buffer II is composed of phenol saturated with 0.1 M citrate buffer pH 4.3 containing sodium dodecyl sulfate (SDS) 0.1%, sodium acetate 0.32 M and EDTA 0.01 M.
8. Add 200  $\mu$ L of chloroform, vortex thoroughly another 20 s and keep it for 10 minutes at room temperature. Centrifuge at 16000 g for 10 minutes at 4° C.
9. Transfer aqueous phase to a fresh 1.5 mL tube, add 0.6 volumes isopropanol, vortex briefly and keep it for 10 min at room temperature. Then, centrifuge at 16000 g for 30 min at 4° C.
10. Discard supernatant and wash pellet with ethanol 70%.
11. Dissolve pellet in 15  $\mu$ l of DEPC water and store at -80°C.







CHAPTER 4

---

**WOOD FORMATION IN *EUCALYPTUS* IS REGULATED BY THE  
INTERACTION BETWEEN THE TRANSCRIPTION FACTOR  
EGMYB1 AND THE HISTONE LINKER PROTEIN EGH1.3**



## Wood formation in *Eucalyptus* is regulated by the interaction between the transcription factor EgMYB1 and the histone linker protein EgH1.3

Marçal Soler<sup>1\*</sup>, Anna Plasencia<sup>1\*</sup>, Romain Larbat<sup>2</sup>, Cécile Pouzet<sup>3</sup>, Alain Jauneau<sup>3</sup>, Susana Rivas<sup>4</sup>, Edouard Pesquet<sup>5</sup>, Catherine Lapierre<sup>6</sup>, Isabelle Truchet<sup>1</sup>, Jacqueline Grima-Pettenati<sup>1</sup>

<sup>1</sup> Plant Research Laboratory (LRSV), UMR5546, Toulouse III Paul Sabatier University - CNRS, Castanet Tolosan, France

<sup>2</sup> INRA UMR 1121 « Agronomie & Environnement » Nancy-Colmar, TSA 40602, 54518 Vandoeuvre Cedex, France - Université de Lorraine UMR 1121 « Agronomie et Environnement », TSA 40602, 54518 Vandoeuvre Cedex, France.

<sup>3</sup> Cell Imaging Plateform, FRAIB, CNRS, Castanet Tolosan, France

<sup>4</sup> Laboratoire des Interactions Plantes-Microorganismes (LIPM), UMR 0441/2594 CNRS/INRA, F-31326 Castanet-Tolosan, France.

<sup>5</sup> Umeå Plant Science Centre (UPSC). Department of Plant Physiology, Umeå University, Umeå, Sweden.

<sup>6</sup> Jean-Pierre Bourgin Institute (JJPB), INRA / AgroParisTech, UMR1318, Saclay Plant Science, Versailles, France.

\* These authors contributed equally to this work

---

Not submitted

Submission scheduled on  
December 2015  
In Plan cell

**Key words:** Xylogenesis, cambium, protein interactions, secondary cell walls, wood, lignin biosynthesis, phenylpropanoids metabolism, *Eucalyptus*

### Summary

Xylogenesis is a differentiation process tightly regulated spatially and temporally by a complex network of transcription factors that leads meristematic cells to divide into xylem cells with thick lignified secondary cell walls. We have previously shown that EgMYB1 was a key transcriptional repressor of lignified secondary cell wall biosynthesis in *Eucalyptus* xylem. Here, we demonstrated that EgMYB1 interacts specifically with the histone linker protein EgH1.3. Histone linker proteins are known to interact with other proteins to modulate chromatin organisation, but a phenotypic effect of an interaction between a transcription factor and a histone linker has not been shown before in plants. Our data showed that EgMYB1-EgH1.3 interaction reinforces the transcriptional repression of EgMYB1 towards lignin biosynthesis and secondary cell wall formation, as demonstrated in tobacco transactivation assays, but also in both transgenic *Arabidopsis* plants and *Eucalyptus* hairy roots overexpressing the two genes. In addition, an important increase in soluble secondary metabolites was observed in *Arabidopsis* plants, suggesting a cross-talk between lignin biosynthesis and indole-derived secondary metabolites mediated by the EgMYB1-EgH1.3 interaction. Expression profiles of EgMYB1 overlaps that of EgH1.3 mostly in the cambial zone, in xylem cells at early stages of their differentiation as well as in mature parenchymatous xylem cells, which have no or only thin lignified secondary cell walls. Taken together, these data suggest that the EgMYB1-EgH1.3 interaction could integrate developmental and position signals to prevent premature deposition of secondary cell walls in cells at the first differentiation stages of xylogenesis.

---

## 4.1 PREFACE

In this paper, I am 1<sup>st</sup> co-author with Marçal Soler. The interaction between EgMYB1 and EgH1.3 was found prior to the beginning of my PhD. Marçal Soler performed the construction of the eucalypts yeast two hybrid (Y2H) library and the screening using EgMYB1 as a bait. I have contributed to the verification of protein-protein interactions using targeted Y2H with Marçal Soler and to the validation of the interactions *in planta* using FRET-FLIM with Marçal Soler, Isabelle Truchet, Cécile Pouzet and Alain Jauneau. I have also performed the transactivation assays and generated the pGWAY and pGWAY-MYB1 vectors. I have transformed *E. grandis* and hardened composite plants with the help of Annabelle Dupas. I did the sampling and the histochemical phenotyping of roots together with Marçal Soler. RNA extractions from roots and qRT-PCR were performed by Nathalie Ladouce. Marçal Soler did the cloning, produced the *Arabidopsis* transgenic plants and performed most of their phenotyping.

During a one-month internship in the group of Edouard Pesquet (UPSC), I have performed pyrolysis-gas chromatography/mass spectrometry. Catherine Lapiere performed thioacidolysis and acetyl bromide analyses. Romain Larbat realised the metabolic profiling. Susana Rivas kindly provided the vectors for Y2H and FRET-FLIM as well as precious advice.

Finally, I have done the data mining, statistical analysis and writing of the first draft together with Marçal Soler.

## 4.2 INTRODUCTION

The appearance of the vascular system around 400 million years ago allowed the colonisation of earth by plants. By connecting the shoot with the roots, it enables long-distance transport of water and sap while providing mechanical support. In most vascular plants a lateral meristematic tissue, the vascular cambium, produces secondary phloem towards the outside of the stem and secondary xylem (commonly known as wood) towards the inside. The activity of this secondary meristem plays a main role in the secondary (diametral) growth of plant axes, allowing the development of tree-like forms and, therefore, it produces the major part of the biomass on Earth.



Wood cells contain secondary walls composed of three main polymers, cellulose, hemicelluloses and lignin, which are intimately cross-linked within a supramolecular network. Wood is highly recalcitrant to degradation, mainly because of its phenolic component, lignin. Lignin is mainly built up by *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) units, which derive from the biosynthesis of the respective monolignols, produced within the phenylpropanoid metabolism. These H, G and S units are connected by a multiplicity of intermonomeric linkages, thus creating a highly resistant barrier.

Wood is widely used by humankind for lumber and paper manufacture, and is increasingly exploited as an environmentally cost-effective, renewable source of energy. The economic importance of wood has focused considerable research attention on the activity of the vascular cambium and the xylem differentiation (xylogenesis) process, which includes cell division, elongation, deposition of lignified secondary walls (at different extents depending on the developmental status and cell type), and programmed cell death (for a review see Schuetz *et al.*, 2013 and references therein). An extremely subtle and sophisticated network of spatial and temporal regulation events is necessary to coordinate the expression of the several hundreds of genes involved in xylem differentiation. During the last decade, research conducted mainly in *Arabidopsis* have revealed that the deposition of secondary walls is controlled by a complex hierarchical regulatory network of transcription factors (Kubo *et al.*, 2005; Zhong *et al.*, 2006, 2007; Zhong and Ye, 2007). In this transcriptional network, NAC and MYB transcription factors act as first and second-level master switches, respectively. Together, they regulate a battery of downstream transcription factors (most of them also belonging to the MYB family) and secondary cell wall biosynthesis genes (for recent reviews see Grima-Pettenati *et al.*, 2012; Hussey *et al.*, 2013; Zhong and Ye, 2015; Nakano *et al.*, 2015). Recently, a high-throughput yeast-one-hybrid screening followed by computational prediction analysis led to a more complete and complex picture of the network, highlighting for instance multiple feed-forward loops to control both transcription factors and biosynthetic genes in normal and abiotic stress conditions (Taylor-Teeples *et al.*, 2014)

Whereas it is well known that transcription factors work as transcriptional complexes, the role of protein–protein interactions in the secondary cell wall transcriptional control network is just beginning to be investigated. Studies performed hitherto have mainly focused on KNOTTED

*ARABIDOPSIS THALIANA7* (KNAT7), a repressor of secondary cell wall formation (Li *et al.*, 2012) whose repression activity can be modulated by interacting with other transcription factors. One of these genes is AtMYB75 (also named as PRODUCTION OF ANTHOCYANIN PIGMENT 1 or PAP1), initially identified as an activator of anthocyanin biosynthesis (Borevitz *et al.*, 2000) that, through an interaction with KNAT7, modulates carbon allocation and represses secondary cell wall biosynthesis (Bhargava *et al.*, 2013, 2010). However, KNAT7 is also able to interact physically with OVATE FAMILY PROTEINS (OFP1 and 4) (Li *et al.*, 2011) on one hand, and with BEL-LIKE HOMEODOMAIN 6 (BLH6) (Liu *et al.*, 2014) on the other hand. These interactions with KNAT7 enhance the repression of secondary cell wall formation, and further physical interactions between OFP1 and/or OFP4 with BLH6 precisely modulates the repression activity of the protein complex (Liu and Douglas, 2015). Another important example of the involvement of protein-protein interactions in regulating xylem differentiation is that of the NAC transcription factor VASCULAR-RELATED NAC-DOMAIN 7 (VND7), a major inductor of xylem vessels differentiation whose activity is blocked when it interacts with the transcriptional repressor VND-INTERACTING 2 (VNI2; Yamaguchi *et al.*, 2008, 2010).

Although many aspects of the regulation of xylem differentiation and secondary wall deposition are conserved between *Arabidopsis* and woody perennial plants (Lev-Yadun, 1994; Chaffey *et al.*, 2002; Nieminen *et al.*, 2004; Melzer *et al.*, 2008; Lens *et al.*, 2012), a growing body of evidences highlighted the differences among woody and herbaceous plants, thus supporting the need for deeper investigations in trees (Chaffey *et al.*, 2002; Soler *et al.*, 2015; Nieminen *et al.*, 2015). *Eucalyptus* is one of the most planted trees worldwide because of their rapid growth and their high-quality wood. However, and despite the recent release of the *E. grandis* genome (Myburg *et al.*, 2014), only a few transcription factors related to wood formation have been characterized to date (Goicoechea *et al.*, 2005; Legay *et al.*, 2007, 2010; Hussey *et al.*, 2011; Yu *et al.*, 2015). One of these genes, EgMYB1, is a key MYB transcription factor that represses the lignin biosynthetic pathway (Legay *et al.*, 2007, 2010).

Several lines of evidence suggest that EgMYB1 repressing activity could be precisely regulated by protein-protein interactions. First, the cis-regulatory regions of the promoters of two lignin biosynthetic genes (EgCCR and EgCAD2), to which EgMYB1 binds to repress their transcriptional activities, are recognized by nuclear protein complexes (Lacombe *et al.*, 2000; Lauvergeat *et al.*,

2002; Rahantamalala *et al.*, 2010). Second, the R3 domain of EgMYB1 contains a conserved bHLH interaction motif (Zimmermann *et al.*, 2004) suggesting that it might interact with a bHLH (basic helix-loop-helix) protein possibly within a ternary complex involving WD40 (Legay *et al.*, 2007). Third, AtMYB4, the closest ortholog to EgMYB1 in *Arabidopsis*, contains the same motif and can bind in yeast to the bHLH proteins TT8 and bHLH012 (Zimmerman *et al.*, 2004). And fourth, AtMYB4, is also able to interact with other proteins, like the importin SAD2 to migrate into the nucleus (Zhao *et al.*, 2007) and the co-repressor TOPLESS to repress target gene expression (*Arabidopsis* Interactome Mapping Consortium, 2011).

In order to fully understand the function of EgMYB1, how its activity is regulated and how it represses target genes in appropriate cellular types, we have undertaken the search for its protein partners in *Eucalyptus* developing xylem, the tissue where EgMYB1 is preferentially expressed. Here, we demonstrate that EgMYB1 physically and specifically interacts with a histone linker protein (EgH1.3) and that this interaction is able to enhance EgMYB1's transcriptional repression activity, leading *in planta* to a stronger reduction of both xylem secondary cell wall thickness and total lignin content while increasing the soluble secondary metabolites. Since the stable transformation of *Eucalyptus* is low efficient and time-consuming, the effects of the interaction *in planta* were evaluated using *Arabidopsis* and further validated using a powerful homologous system that we have recently developed consisting of *E. grandis* transgenic hairy roots (Plasencia *et al.*, 2015). Taking into account the overlap between the expression profiles of EgMYB1 and EgH1.3, we propose a model in which the interaction between EgMYB1 and EgH1.3 is a fine tuned mechanism that increases the repression activity of EgMYB1, thus preventing secondary cell walls to be deposited in meristematic cells and ensuring a metabolic homeostasis.

## 4.3 RESULTS

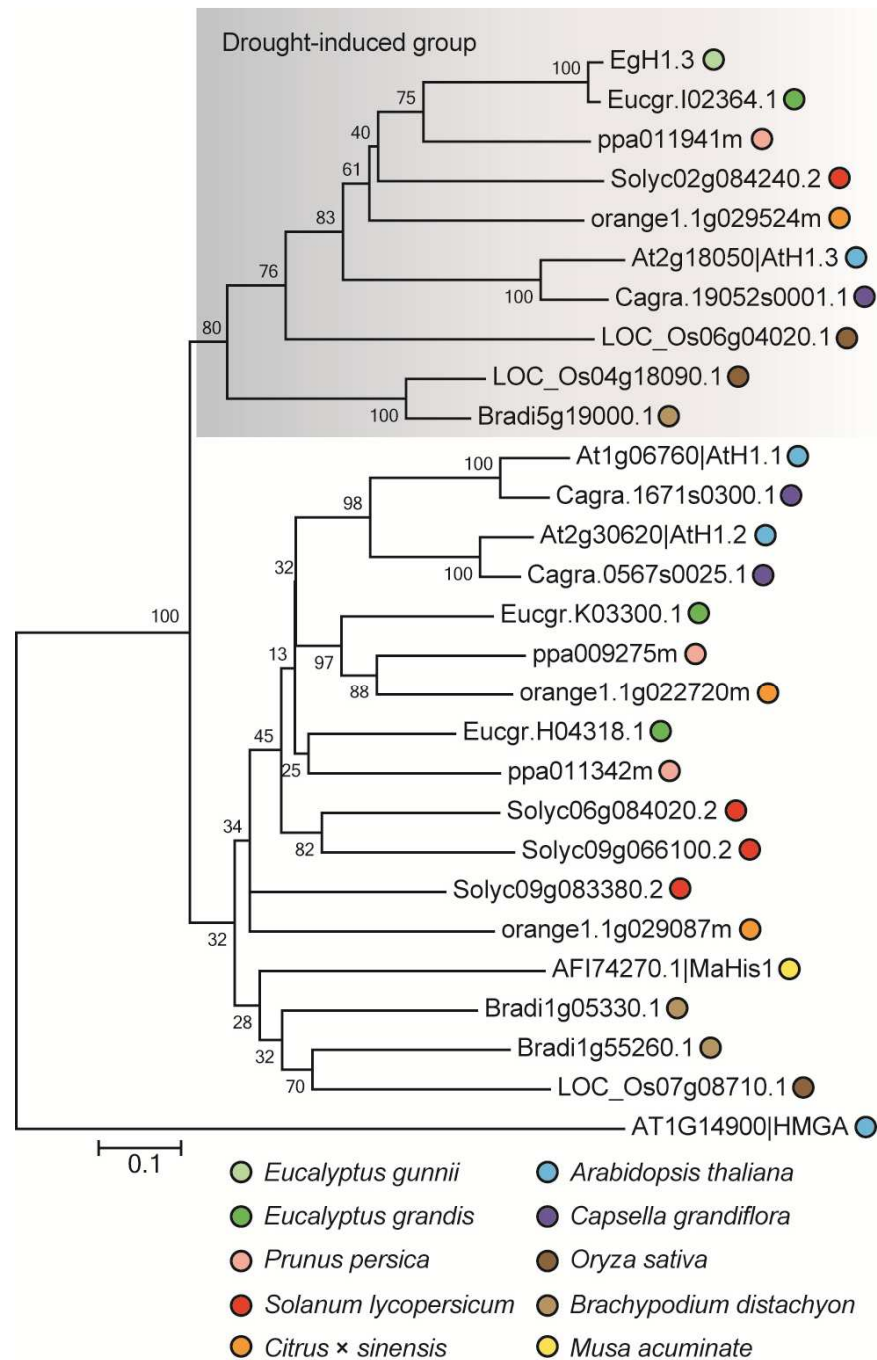
### 4.3.1 EGMYB1 INTERACTS SPECIFICALLY WITH A HISTONE LINKER PROTEIN IN YEAST

In order to identify proteins interacting with EgMYB1 in *Eucalyptus* secondary xylem, we first constructed a yeast-two-hybrid library using a mixture of cDNAs from cambium and xylem samples at different development stages (juvenile and mature) and submitted or not to a

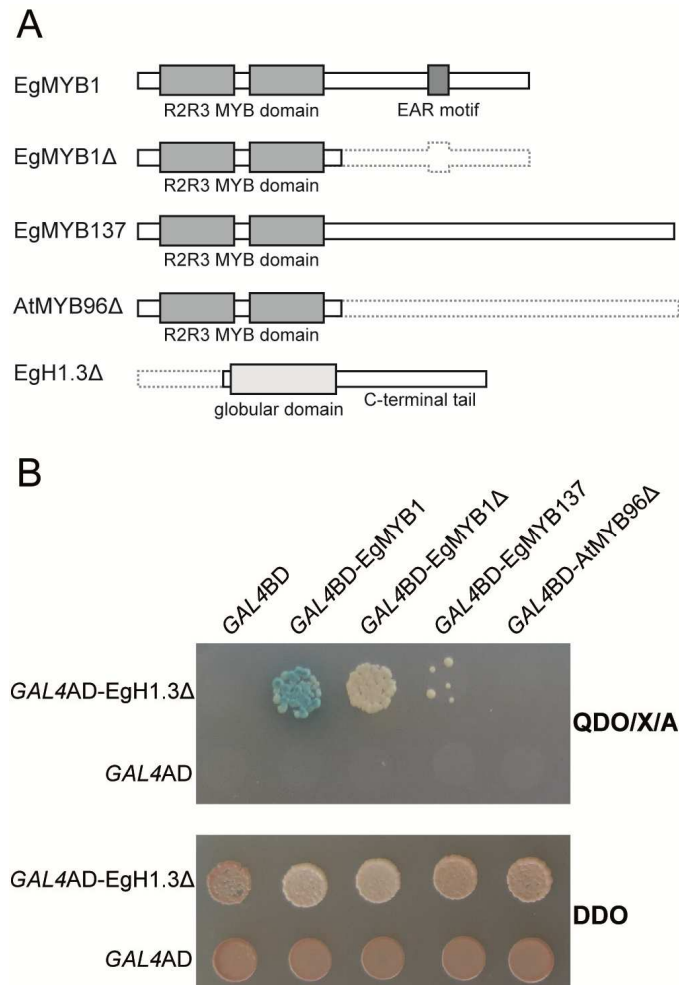
mechanical stress (tension, opposite and straight xylems). Because EgMYB1 contains a repressing domain at its C-terminal region that could inhibit the detection of protein-protein interactions through yeast-two-hybrid approach, we screened the library using as bait a truncated version containing only the R2R3 domain (EgMYB1 $\Delta$ ). We obtained several prey candidates, but the most abundant one (7% of the sequenced clones) corresponded to partial sequences matching a *Eucalyptus* histone linker protein (Eucgr.I02364).

With the three existing histone linker proteins from *A. thaliana* (Ascenzi and Gantt, 1997), we retrieved their potential orthologs in the genome of *E. grandis* and other plant genomes using BlastX. With the obtained sequences we constructed a phylogenetic tree, showing that *E. grandis* has also three histone linker proteins, and that our candidate is likely the ortholog of *Arabidopsis* AtH1.3 (Figure 4.1). Therefore, we named it as EgH1.3. Putative orthologs of EgH1.3 from other tree species (*Prunus persica* and *Citrus x sinensis*), annual plants (*Arabidopsis thaliana*, *Capsella grandiflora* and *Solanum lycopersicum*) and monocotyledonous plants (*Oryza sativa*, *Brachypodium distachyon* and *Musa acuminata*) cluster in a robust phylogenetic group with EgH1.3 and AtH1.3, which clearly separates from the rest of histone linker variants (Figure 4.1).

We isolated the plasmid with the largest sequence found of EgH1.3 (named EgH1.3 $\Delta$ , containing the 510 last bp of the open reading frame and 110 bp of the 3'UTR, which includes the globular and the C-terminal domain of the protein, figure 4.2A) and we validated the interaction with EgMYB1 and EgMYB1 $\Delta$  in a yeast-two-hybrid targeted approach (Figure 4.2). Unexpectedly, the strength of the interaction was higher with EgMYB1 than with EgMYB1 $\Delta$ , as shown by the blue colour indicating  $\alpha$ -galactosidase activity present only in yeast cells containing both EgH1.3 $\Delta$  and EgMYB1. Therefore, even if the R2R3 domain of EgMYB1 is sufficient for interacting with EgH1.3, the interaction is strengthened in the yeast in presence of the C-terminal domain of EgMYB1. We further confirmed the specificity of the interaction by showing that unrelated MYB proteins such as EgrMYB137 (Eucgr.K02806), with a tissue expression pattern very similar to EgMYB1 (Soler *et al.*, 2015), or the R2R3 domain of AtMYB96 (AT5G62470), did not interact with EgH1.3 $\Delta$  (Figure 4.2).



**Figure 4.1: Neighbor-Joining phylogenetic tree of histone linker proteins from *Eucalyptus grandis*, *E. gunnii* as well as from other 5 dicotyledonous and 2 monocotyledonous plants.** It also includes the histone linker protein from banana tree MaHis known to interact to a WRKY transcription factor, and a High-Mobility Group A (HMGA) from *Arabidopsis* protein used to root the tree. Bootstrap values are shown at the internodes. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree (amino acid substitutions per site). The cluster of putative orthologs for the histone liker variant EgH1.3 is named “drought-induced” and is highlighted in grey.



**Figure 4.2: Protein – protein interactions detected using Yeast-Two-Hybrid (Y2H).** **A.** Schematic representation of the proteins tested in the targeted Y2H screening. Dotted lines represent the missing parts of the truncated proteins ( $\Delta$ ). The rectangles in grey correspond to protein domains or motifs. **B.** Growth of yeast cells expressing the proteins from panel A fused either to GAL4AD (*Gal4* Activation Domain) or to GAL4BD (*Gal4* Binding Domain). Transformed yeast cells where protein-protein interactions occur, grow on the highly selective medium QDO/X/A (quadruple drop-out auxotrophic media lacking tryptophan, leucine, adenine and histidine), supplemented with X- $\alpha$ -Gal [5-bromo-4-chloro-3-indolyl alpha-D-galactopyranoside] and Aureobasidin A). Blue colour coming from the hydrolysis of X- $\alpha$ -Gal is an additional indication of the occurrence of protein-protein interaction. Yeast cells transformed with both GAL4AD and GAL4BD plasmids grow on the selective medium DDO (double drop-out auxotrophic medium without tryptophan and leucine).

#### 4.3.2 EGMYB1 INTERACTS SPECIFICALLY WITH EGH1.3 IN *N. BENTHAMIANA* CELL NUCLEUS

We had previously shown by transient expression assays in tobacco leaves that EgMYB1 was localized in the nucleus of epidermal cells (Legay *et al.*, 2007). Here, we showed the nuclear localisation of EgH1.3 using a 35S:EgH1.3-YFP construct and further demonstrated that, when both proteins were transiently expressed in the same cell (35S:EgMYB1-CFP + 35S:EgH1.3-YFP), they perfectly co-localized in the nucleus (Supplemental figure S4.1).

Interaction between EgMYB1 and EgH1.3 was validated *in planta* using Förster Resonance Energy Transfer – Fluorescence Lifetime Imaging Microscopy (FRET-FLIM). This non-invasive quantitative technique measures CFP lifetime, which allows the monitoring of the fluorescence energy transfer from the CFP (donor) to the YFP (acceptor) molecules, fused to EgMYB1 and EgH1.3, respectively. When both proteins interact, the distance between the donor and the acceptor is so narrow that, after a pulse CFP excitation, part of the energy is transferred to the extremely close YFP molecule, causing a decrease in the CFP lifetime. As compared to EgMYB1-CFP expressed alone, we found a significant decrease of the CFP lifetime when the full-length sequence of EgMYB1-CFP was co-transformed with the full-length sequence of EgH1.3-YFP, thus validating *in planta* the interaction found in yeast (Table 4.1, Supplemental figure S4.2). Noteworthy, with FRET-FLIM the interaction was confirmed using the full-length protein of EgH1.3 instead of the partial sequence EgH1.3Δ. We also found a significant diminution of the CFP lifetime when co-transforming EgMYB1Δ-CFP with EgH1.3-YFP, thus confirming that the R2R3 domain of EgMYB1 interacts with EgH1.3. The specificity of the interaction was also validated *in planta* using either the full-length version of EgrMYB137 or AtMYB96 in presence of EgH1.3. These proteins co-localized in the nuclei with EgH1.3 (Supplemental figure S4.1), but no CFP lifetime decrease was observed. Similarly, no CFP lifetime decrease was found when EgMYB1-CFP was co-transformed with 35S:YFP, (Table 4.1, Supplemental figure S4.2).

**Table 4.1.** FRET-FLIM analysis showing the interaction between EgMYB1 with EgH1.3 in the nuclei of *N. benthamiana* transiently transformed epidermal cells.

Donor	Acceptor	Lifetime <sup>a</sup>	SD	N <sup>b</sup>	E <sup>c</sup>	P value <sup>d</sup>
EgMYB1-CFP	-	2.580	0.276	107	-	-
<b>EgMYB1-CFP</b>	<b>EgH1.3-YFP</b>	<b>2.161</b>	<b>0.395</b>	<b>58</b>	<b>16.24</b>	<b>2.8 x 10<sup>-13</sup></b>
EgMYB1-CFP	YFP	2.567	0.221	56	-	0.757
EgMYB1Δ-CFP	-	2.879	0.203	82	-	-
<b>EgMYB1Δ-CFP</b>	<b>EgH1.3-YFP</b>	<b>2.491</b>	<b>0.318</b>	<b>59</b>	<b>13.48</b>	<b>3.8 x 10<sup>-15</sup></b>
AtMYB96-CFP	-	2.648	0.396	54	-	-
AtMYB96-CFP	EgH1.3-YFP	2.603	0.282	29	-	0.595
EgMYB137-CFP	-	3.016	0.206	60	-	-
EgMYB137-CFP	EgH1.3-YFP	3.007	0.154	60	-	0.798

<sup>a</sup> Mean CFP lifetime ( $\tau$ ) in nanoseconds.

<sup>b</sup> Number of measured nuclei.

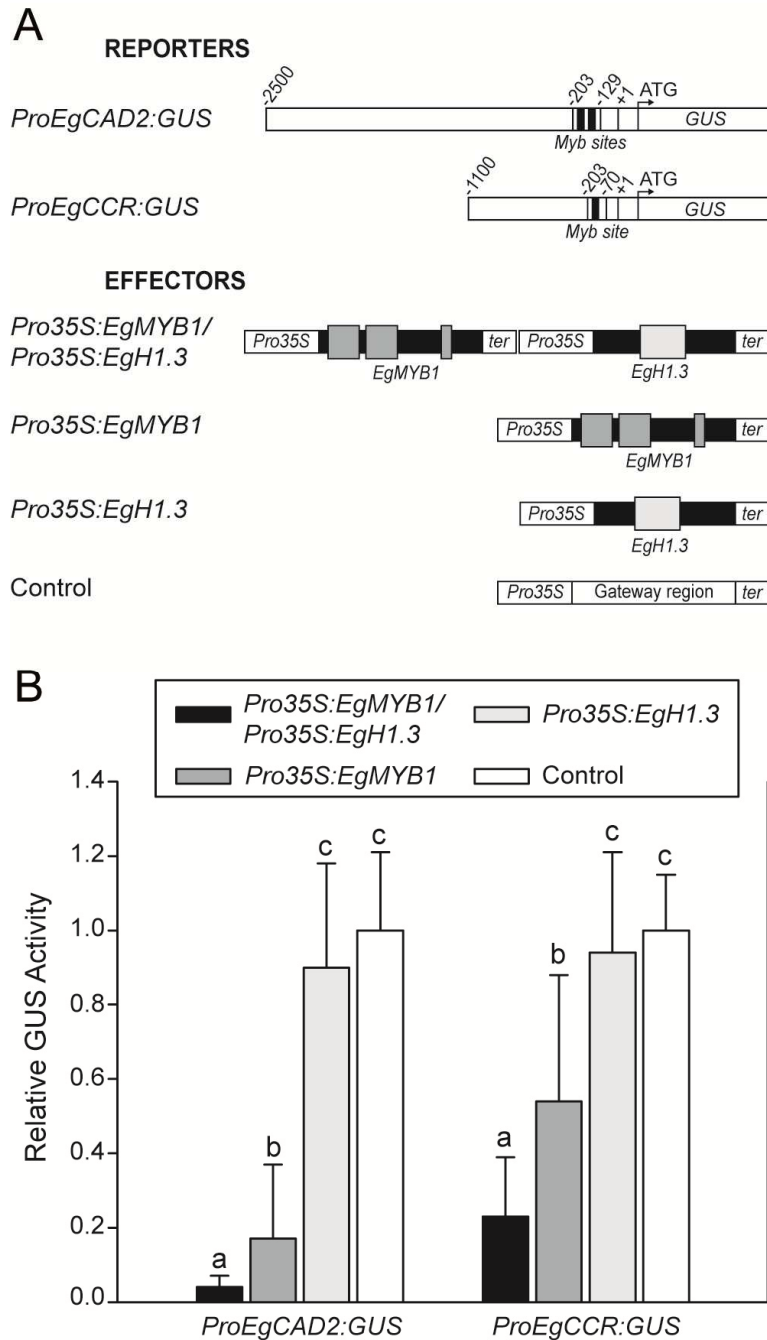
<sup>c</sup> FRET efficiency percentage ( $E=1-\tau_{DA}/\tau_D$ ) calculated as comparing the mean lifetime of the donor in the presence of the acceptor ( $\tau_{DA}$ ) against the mean lifetime of the donor in the absence of the acceptor ( $\tau_D$ ).

<sup>d</sup> Student's t test p value calculated comparing the CFP lifetime measures of the donor in the presence of the acceptor against the lifetime measures of the donor in the absence of the acceptor

#### 4.3.3 TRANSACTIVATION ASSAYS IN *N. BENTHAMIANA* LEAVES SHOWED A STRONG REPRESSION OF EGCCR AND EGCAD GENE PROMOTERS PRODUCED BY EGMYB1 TOGETHER WITH EGH1.3

Using gel-shift experiments and transactivation assays in tobacco leaves, we previously demonstrated that EgMYB1 was able to bind specifically to the promoters of two lignin biosynthetic genes, the Cinnamoyl CoA reductase (EgCCR) and the Cinnamyl Alcohol dehydrogenase (EgCAD2), and to further repress their transcriptional activities (Legay *et al.* 2007). Here, using transactivation assays with the same promoters fused to the GUS reporter gene, we tested if the repressive effect of EgMYB1 was modulated by the interaction with EgH1.3. When the effectors EgMYB1 and EgH1.3 were co-transformed simultaneously in tobacco leaf cells, the GUS activity driven either by the EgCAD2 or the EgCCR promoters was dramatically reduced as compared to controls (97% and 77% reduction, respectively; Figure 4.3). The repression effects mediated by the interaction between EgMYB1 and EgH1.3 were much stronger than those obtained by EgMYB1 alone (59% and 43% of reduction for the EgCAD2 and the EgCCR promoters, respectively). Overexpression of EgH1.3 alone did not show any significant transcriptional repressing activity. Similarly, neither EgMYB1Δ alone nor EgMYB1Δ co-transformed with EgH1.3 produced significant reduction of GUS activity driven by EgCAD2 promoter (Supplemental figure S4.3).





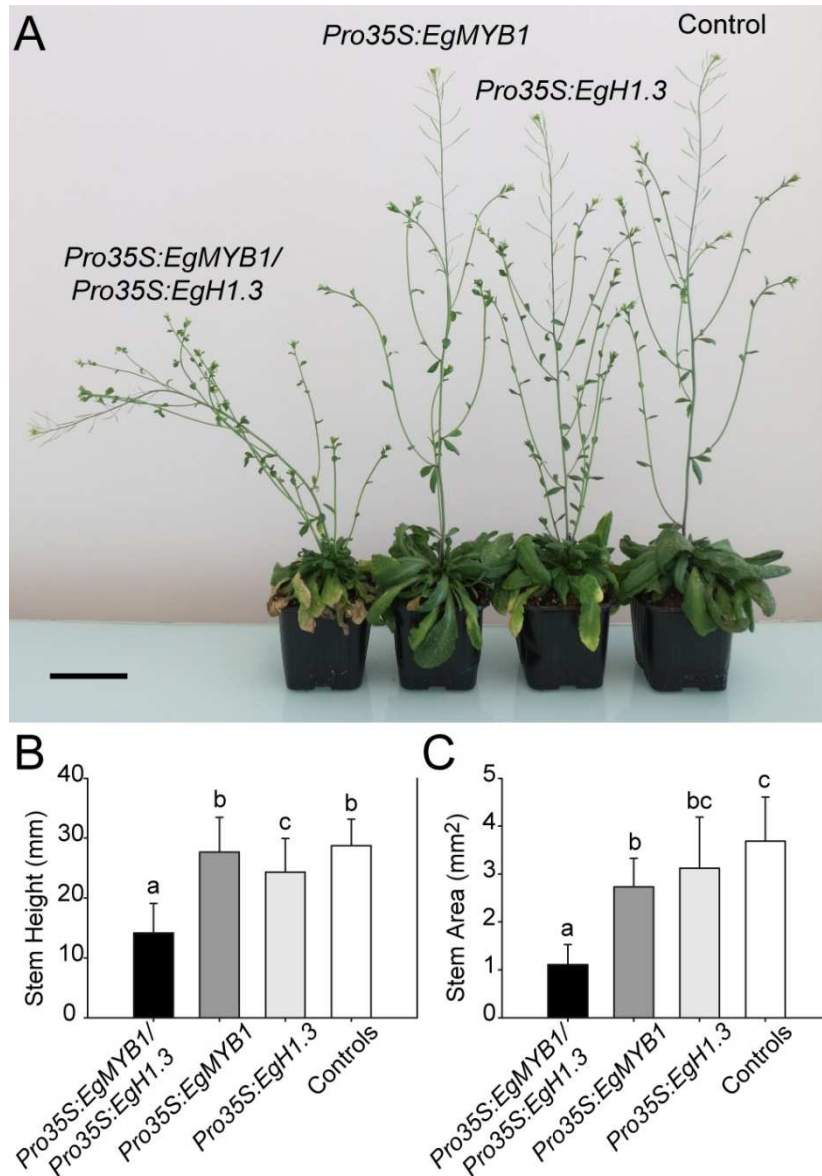
**Figure 4.3: Effects of EgMYB1 and EgH1.3 on the transcriptional activities of the EgCAD2 and EgCCR promoters using transactivation assays in *N. benthamiana* leaves.** Transactivation assays in *N. benthamiana* leaves were used to evaluate the effects of the EgMYB1-EgH1.3 interaction on the transcriptional activities of the EgCAD2 and EgCCR promoters fused to the GUS reporter gene. **A.** Schemes of the reporter and effector constructs. Pro35S, CaMV35S promoter; ter, 35S terminator; +1, transcription start site; GUS, uidA coding region. Black rectangles in effectors correspond to coding regions and grey boxes to protein domains or motifs. **B.** GUS activities measured after co-infiltration of reporters and effectors constructs in tobacco leaves. Data represent means  $\pm$  standard deviation resulting from nine biological replicates in at least three independent experiments. GUS activity is expressed as a ratio relative to control. Statistics were calculated with Student's t test, and significant differences with P value  $< 0.05$  are highlighted in lowercase letters.

#### 4.3.4 TRANSGENIC *ARABIDOPSIS* PLANTS OVEREXPRESSING TOGETHER EGMYB1 AND EGH1.3 SHOWED STRONG REDUCTION OF SECONDARY CELL WALL FORMATION

*Arabidopsis* plants overexpressing high levels of EgMYB1 transcripts under the control of 35S promoter exhibited severe phenotypes such as dwarfing, reduction in size, collapsed xylem vessels as well as a dramatic reduction of secondary cell wall deposition in interfascicular fibres (Legay *et al.*, 2010). Since the percentage of germination of these severely affected lines was very low, we used in the present study a transgenic line with a moderate phenotype to re-transform it with 35S:EgH1.3, in order to obtain double transformant plants expressing both proteins, EgMYB1 and EgH1.3 (referred as 35S:EgMYB1/35S:EgH1.3). To evaluate the function of EgH1.3 alone, we transformed Col-0 plants with 35S:EgH1.3 construct. As controls, we used wild-type (WT) plants as well as lines transformed with the empty vectors. Several independent lines of transgenic plants were obtained for each construction.

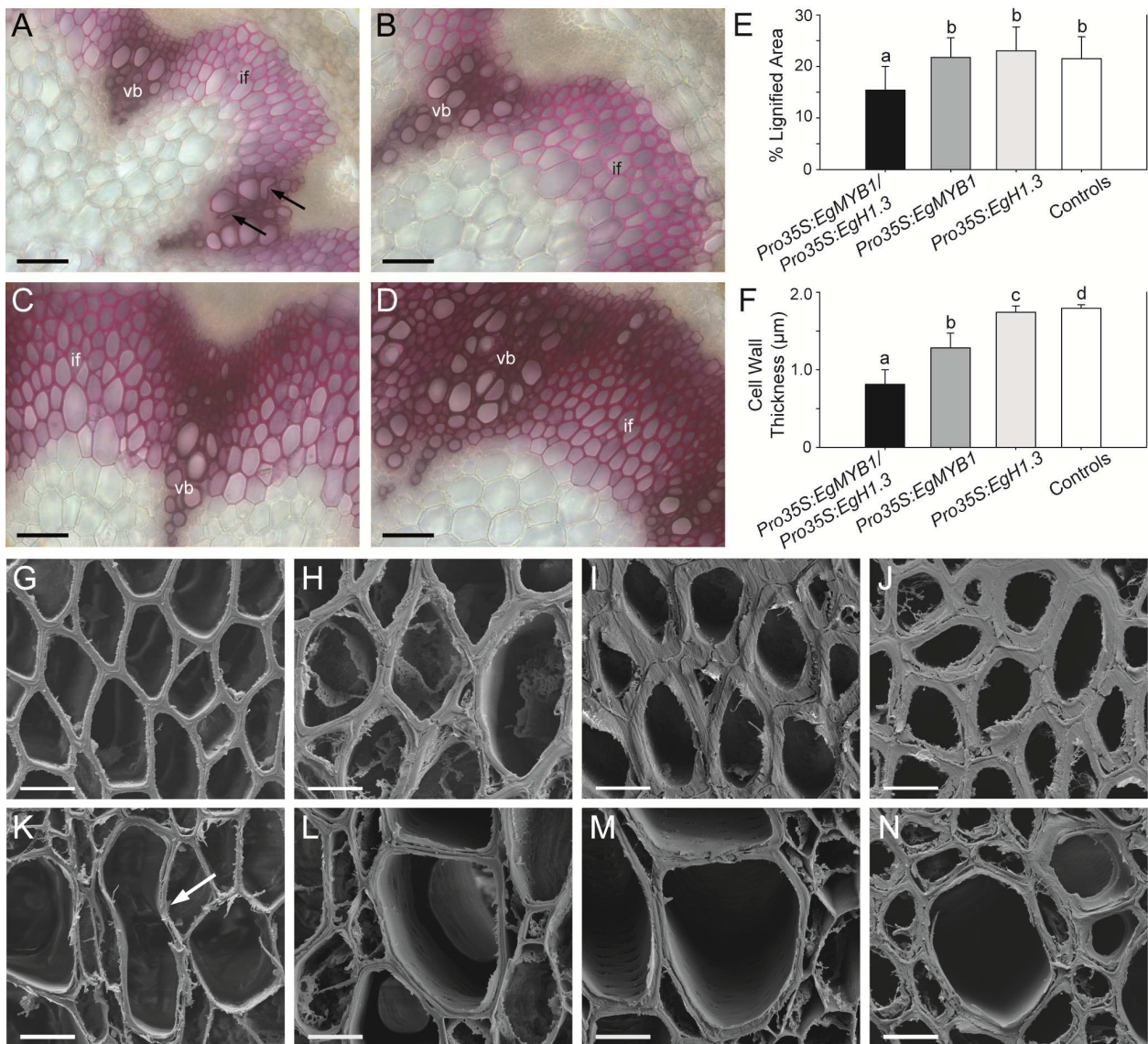
At macroscopic level, 35S:EgMYB1 plants exhibited a moderate phenotype, with no reduction in size as compared to controls but with some leaves showing a cup-shaped morphology and seeming less flexible and more fragile than controls (Figure 4.4A). 35S:EgH1.3 plants had a slightly shorter inflorescence stem than wild type (Figure 4.4B), and some leaves showed a yellowish colour that became more evident as the plant got older. The rosette of the double transformants (35S:EgMYB1/35S:EgH1.3) tended to be slightly smaller than controls and the inflorescence stem was shorter (Figure 4.4B), more flexible and unable to stand upright (Figure 4.4A). Similarly to 35S:EgH1.3 plants, some leaves showed a yellowish colour and they tend to senesce sooner (Figure 4.4A). There were no clear differences in the growing rates and in the flowering time among all these transgenic plants. As after visual inspection all plants showed similar macroscopic phenotypes, we selected three lines for each construction according to their transgene expression levels (Supplemental figure S4.5) to further characterize them in detail.

Microscopic observations at the base of the inflorescence revealed that the stem area was moderately smaller than controls in 35S:EgMYB1 lines, and substantially smaller in 35S:EgMYB1/35S:EgH1.3 plants (Figure 4.4C).



**Figure 4.4: Macroscopic phenotype of *Arabidopsis* transgenic plants.** **A.** Two-month-old *Arabidopsis* plants grown in short day conditions. Scale bar = 5 cm. **B-C.** Heights (B) and areas at the basis (C) of the inflorescence stems from *Arabidopsis* transgenic plants after 2 months of growth in short day conditions. Data represent the means  $\pm$  standard deviation of the measurements made on 15 plants (i.e. five plants for each of the three lines selected for each transformant type). Statistics were calculated with Student's t test, and significant differences with *P* value < 0.05 are highlighted in lowercase letters.

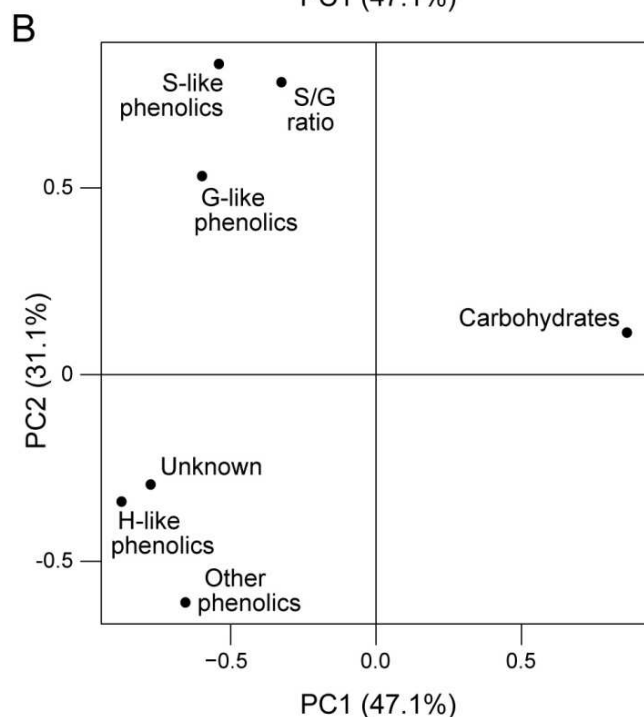
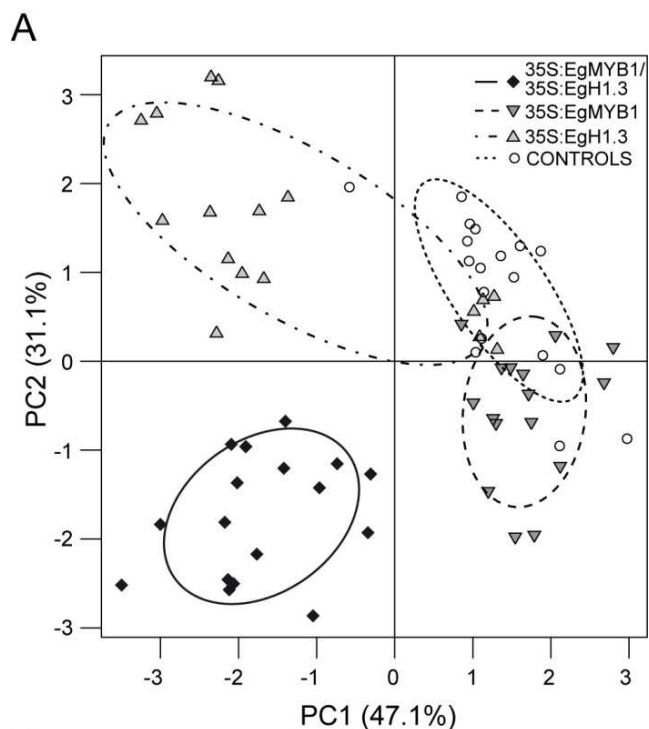
Using the lignin phloroglucinol-HCl staining, we found that the ratio of the lignified area over the total area of the stem was smaller in 35S:EgMYB1/EgH1.3 plants (Figure 4.5A-E). The secondary wall thickness of interfascicular xylem fibres was reduced in all transgenic lines when compared to controls. This reduction was small in 35S:EgH1.3 lines, but important in 35S:EgMYB1 and especially in 35S:EgMYB1/35S:EgH1.3 plants, which had less than half of the thickness found in control lines (Figure 4.5F). Reduction of cell wall thickness in interfascicular fibres was clearly seen under bright-field (Figure 4.5A-D) and scanning electron microscopy (Figure 4.5G-J) for 35S:EgMYB1 and 35S:EgMYB1/EgH1.3, but not such a clear reduction was observed in vessels (Figure 4.5K-N). However, in some cases, we observed irregular xylem (*irx*) vessels in 35S:EgMYB1/35S:EgH1.3 plants (Figure 4.5A, K).



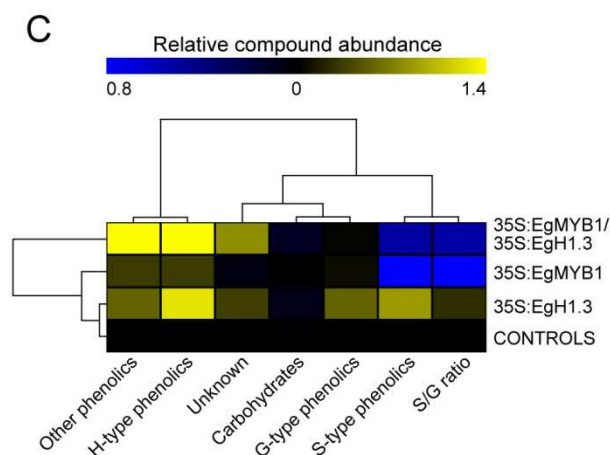
**Figure 4.5. Xylem structure and secondary cell wall thickness from the inflorescence stem of two-month-old *Arabidopsis* transgenic plants growing in short day conditions.** A-D. Transversal sections performed at the basis of inflorescence stems of *Pro35S:EgMYB1/Pro35S:EgH1.3* (A), *Pro35S:EgMYB1* (B), *Pro35S:EgH1.3* (C) and control plants (D) stained with phloroglucinol-HCl and observed under a bright-field microscope. Lignified secondary cell walls appear in red. Scale bars = 50 µm. E-F. Percentage of lignified area over the overall inflorescence stem area (E) and cell wall thickness of interfascicular fibres (F). Data represent the means ± standard deviation of the measurement done for 15 plants (i.e. five plants for each of the three lines selected by transformant type). Twenty measurements of cell wall thickness were made for each individual plant (F). Statistics were calculated with Student's t test, and significant differences with  $P$  value < 0.05 are highlighted in lowercase letters. G-N. Scanning electron microscopy images showing interfascicular fibres (G-J) and vessels (K-N) from stem sections of *Pro35S:EgMYB1/Pro35S:EgH1.3* (G, K), *Pro35S:EgMYB1* plants (H, L), *Pro35S:EgH1.3* (I, M) and control plants (J, N). Scale bars = 10 µm. vb, vascular bundles; if, interfascicular fibres. Arrows point to collapsed vessels.

#### 4.3.5 *ARABIDOPSIS* PLANTS OVEREXPRESSING BOTH EgMYB1 AND EgH1.3 SHOWED AN IMPORTANT REDUCTION OF THE TOTAL LIGNIN LEVEL AND AN ALTERATION OF THEIR LIGNIN MONOMERIC COMPOSITION

In order to get the most complete picture of the changes occurring at the secondary cell wall level, we first performed Pyrolysis-Gas Chromatography/Mass Spectrometry (Py-GC/MS) analyses of the *Arabidopsis* hypocotyls. Pyrogram spectra were collected and the area of each peak was measured and integrated into the following categories as described by Gerber *et al.* (2012): carbohydrates, G lignin, S lignin, H lignin, phenolics, non-identified phenolics, and unknown compounds (Supplemental Table 4.1). Py-GC/MS data was subjected to principal component analysis, which allowed the extraction of two main components, PC1 and PC2, responsible for 78.2% of system variation (47.1% and 31.1% respectively; Figure 4.6A). The first component mostly separated samples by their relative content in phenolic compounds versus carbohydrates, whereas the second component mostly separated samples rich in S and G units from those richer in H units as well as in other phenolics and unidentified compounds (Figure 4.6B). Therefore, 35S:EgH1.3 plants are separated from controls mostly because they contained a higher proportion of phenolic compounds, whereas 35S:EgMYB1 lines are separated from the controls because they showed a different composition of their S and G lignin units. Interestingly, 35S:EgMYB1/35S:EgH1.3 plants showed both the features of 35S:EgH1.3 and 35S:EgMYB1 lines. However, the huge separation of these plants as compared to the others is not just the result of an addition of the individual phenotypes of EgMYB1 and EgH1.3 lines, but rather highlight a synergistic effect of the interaction between these two proteins. Supporting that, when we hierarchically cluster the pyrolysis data, we found that the most distant lines corresponded to 35S:EgMYB1/35S:EgH1.3 plants.



**Figure 4.6: Principal components analysis and hierarchical clustering of the relative amounts of each compound category detected by Pyrolysis – Gas Chromatography / Mass Spectrometry. A.** Score plot of each plant replicate separated by the different transgenics screened. **B.** Loading plot of the different compound categories detected. **C.** Hierarchical clustering calculated using Pearson square as metric distance and the average linkage clustering method.



Pyrolysis results showed a good agreement when compared with those obtained using thioacidolysis. This more quantitative method is especially useful when limited amounts of material are available and gives rise to *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) monomers from H, G and S lignin units. Similarly to pyrolysis, thioacidolysis detects only monomers involved in labile ether  $\beta$ -O-4 linkages, generally accounting to the non-condensed

lignin fraction. Thioacidolysis yield expressed relative to cell wall dry weight gives an estimate of the amount of parent lignin structures, and in the case of *Pro35S:EgMYB1/Pro35S:EgH1.3* plants, this yield was reduced by 24% compared to controls (Table 4.2). This reduced thioacidolysis yield is indicative of a lower lignin content and/or of an altered lignin structure. The proportion of the different monomers was modified in the hypocotyls of all transgenic lines but with different tendencies. The overexpression of *EgH1.3* produced a strong increase in the frequency of H units at the expense of G, whereas the overexpression of *EgMYB1* led to a lower frequency of S and a higher frequency of G units. Interestingly, *Pro35S:EgMYB1/Pro35S:EgH1.3* plants showed a combination of both features (*i.e.* higher frequency of H and a lower frequency of S monomers) but more pronounced than in individual transformants, supporting a synergic effect of the interaction between EgMYB1 and EgH1.3 on lignin structure.

**Table 4.2. Thioacidolysis evaluation of lignin structure in the hypocotyls of *Arabidopsis* transgenic lines.**

Lines	Lignin derived H, G and S monomers measured with thioacidolysis: total yield and relative molar frequencies				
	Total yield ( $\mu\text{mol}$ s / g dry weight)	% H	%S	%G	S/G
<b>35S:EgH1.3</b>	99.0 $\pm$ 10.0	<b>13.8 <math>\pm</math> 3.1 *</b>	21.3 $\pm$ 0.8	<b>64.9 <math>\pm</math> 3.7 *</b>	0.33 $\pm$ 0.03
<b>35S:EgMYB1</b>	87.7 $\pm$ 10.7	8.3 $\pm$ 0.7	<b>17.3 <math>\pm</math> 1.3 **</b>	<b>74.4 <math>\pm</math> 0.7 *</b>	<b>0.23 <math>\pm</math> 0.02 **</b>
<b>35S:EgMYB1/EgH1.3</b>	<b>73.7 <math>\pm</math> 8.0 *</b>	<b>16.6 <math>\pm</math> 0.1 ***</b>	<b>16.5 <math>\pm</math> 2.1 *</b>	<b>67.0 <math>\pm</math> 2.2 *</b>	0.25 $\pm$ 0.04
<b>CONTROLS</b>	96.7 $\pm$ 7.5	7.0 $\pm$ 1.5	21.5 $\pm$ 0.6	71.4 $\pm$ 0.9	0.30 $\pm$ 0.01

Measures represent the mean  $\pm$  standard deviation from three biological replicates corresponding to each of the selected lines for each construct. Each biological replicate consisted of a pool of five plants. Statistics were calculated with Student's t test, \*\*\* p value < 0.001, \*\* p value < 0.01, \* p value < 0.05.

In the inflorescence stems of 35S:EgMYB1/35S:EgH1.3 plants, the total insoluble lignin content evaluated by the acetyl bromide method was importantly reduced, by 28%, compared to controls (Table 4.3). When expressed relative to the total lignin content, the thioacidolysis yield was reduced by about 20% in *Pro35S:EgMYB1/Pro35S:EgH1.3* plants (Table 4.3). This result reveals that lignin structure is also altered in these lines, with a lower frequency of lignin units only involved in  $\beta$ -O-4 labile lignin bonds. This structural alteration is accompanied with a significant reduction of the S/G ratio as a consequence of an increase in the proportion of G

units and a decrease in S units. No differences were observed in 35S:EgMYB1 or in 35S:EgH1.3 lines.

**Table 4.3. Evaluation of lignin content and structure in extractive-free inflorescence stems of *Arabidopsis* transgenic lines. Lignin content was measured as the acetyl bromide method (ABL) and expressed as percentage of dry weight. Lignin structure was evaluated by thioacidolysis.**

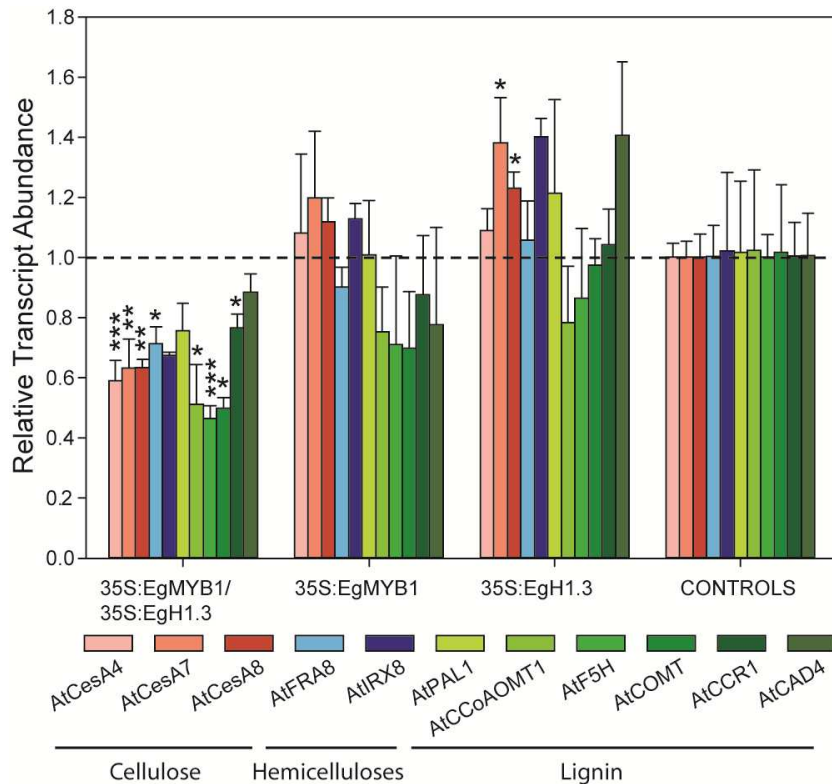
	Lignin derived H, G and S monomers measured with thioacidolysis: total yield and relative molar frequencies					
	% ABL	Total yield	% H	%S	%G	S/G
		( $\mu$ mol / g ABL)				
<b>35S:EgH1.3</b>	15.27 $\pm$ 1.06	1675 $\pm$ 41	0.7 $\pm$ 0.0	30.9 $\pm$ 0.9	68.4 $\pm$ 0.9	0.45 $\pm$ 0.02
<b>35S:EgMYB1</b>	13.34 $\pm$ 1.37	1499 $\pm$ 77	1.0 $\pm$ 0.3	29.3 $\pm$ 0.8	69.7 $\pm$ 0.6	0.42 $\pm$ 0.02
<b>35S:EgMYB1/EgH1.3</b>	<b>9.94 <math>\pm</math> 0.70 **</b>	<b>1418 <math>\pm</math> 73 *</b>	0.7 $\pm$ 0.1	<b>25.7 <math>\pm</math> 0.8 *</b>	<b>73.6 <math>\pm</math> 0.9 *</b>	<b>0.35 <math>\pm</math> 0.01 *</b>
<b>CONTROLS</b>	13.76 $\pm$ 0.52	1696 $\pm$ 85	0.8 $\pm$ 0.1	30.5 $\pm$ 1.7	68.7 $\pm$ 1.6	0.44 $\pm$ 0.04

Measures represent the mean  $\pm$  standard deviation from three biological replicates corresponding to each of the selected lines for each construct. Each biological replicate consisted of a pool of five plants. Statistics were calculated with Student's t test, \*\*\* p value < 0.001, \*\* p value < 0.01, \* p value < 0.05.

#### 4.3.6 TRANSCRIPTOMIC ANALYSIS REVEALED A REPRESSION OF LIGNIN, CELLULOSE AND HEMICELLULOSE GENES FOR 35S:EGMYB1/EGH1.3 ARABIDOPSIS PLANTS

Real-time qPCR analyses were performed to detect transcript variations in the levels of lignin, cellulose and hemicellulose biosynthetic genes, responsible to the formation of the secondary cell wall in xylem cells (Figure 4.7). Even if they were not statistically significant, nearly all of the lignin biosynthetic genes tested were repressed in 35S:EgMYB1 *Arabidopsis* lines exhibiting a mild phenotype, but not those involved in cellulose and hemicelluloses biosynthesis, suggesting that the primary targets of EgMYB1 are likely genes related to lignin and phenylpropanoid biosynthesis. In 35S:EgMYB1/35S:EgH1.3 lines, a clear repression of not only the lignin biosynthetic genes, but also those from cellulose and hemicellulose biosynthesis was found. In the 35S:EgH1.3 lines, no clear differences in the transcript abundance levels of these secondary cell wall biosynthetic genes were observed, with the exception of two cellulose biosynthetic genes Cesa7 and Cesa8, which were slightly but significantly increased.





**Figure 4.7: Relative transcript abundance of secondary cell wall biosynthetic genes measured by RT-qPCR.** Data represent the mean  $\pm$  standard deviation (SD) of three biological replicates corresponding to each of the three selected lines for each construct. Each biological replicate consisted of a pool of five plants. Statistics were calculated with Student's t test, \*\*\* p value < 0.001, \*\* p value < 0.01, \* p value < 0.05.

#### 4.3.7 METABOLITE PROFILING ANALYSIS IN STEMS REVEALED A STRONG INDUCTION OF INDOLE GLUCOSINOLATES IN 35S:EgMYB1/35S:EgH1.3 PLANTS

Because the Py-GC/MS pointed out important differences in phenolic compounds other than lignins between lines, we decided to perform a metabolic profiling in *Arabidopsis* stems. The most striking difference was an extremely abundant accumulation of some glucosinolates in 35S:EgMYB1/35S:EgH1.3 lines (Table 4.4). The great majority of these glucosinolates were indol-derived compounds originated from tryptophan, but two of them were aliphatic glucosinolates, even if their contents were lower. 35S:EgH1.3 plants showed also an increase in some glucosinolates, but with fewer compounds induced and in lower amounts as compared to 35S:EgMYB1/35S:EgH1.3 plants. In 35S:EgMYB1 lines, two glucosinolates compounds were reduced and one induced, but their amounts were not very different from control lines.

Other compounds were significantly induced in 35S:EgMYB1/35S:EgH1.3 plants, like the protocatechoyl xylose, the flavonoid Kaempferol-3O-(Rh-G)-7O-Rh and three yet unidentified compounds. Sinapoyl glucose was undetectable in 35S:EgMYB1/35S:EgH1.3 plants, being the only compound identified in these lines with statistically significant lower amounts than in

controls. In 35S:EgMYB1, a significant reduction of both sinapoyl malate and sinapoyl glucose was observed. Interestingly, the reduction of sinapoyl malate is in accordance with the higher accumulation of this compound in *Arabidopsis* AtMYB4 mutants (Jin *et al.*, 2000).

**Table 4.4: Soluble metabolites detected in the inflorescence stem of the *Arabidopsis* transgenic lines.** Measures represent the mean  $\pm$  SD from three biological replicates corresponding to each of the selected lines for each construct. Each biological replicate consisted of a pool of five plants. Statistics were calculated with Student's t test, \*\*\* p value < 0.001, \*\* p value < 0.01, \* p value < 0.05. Ratio of peak area of each compound was given in relation to control lines.

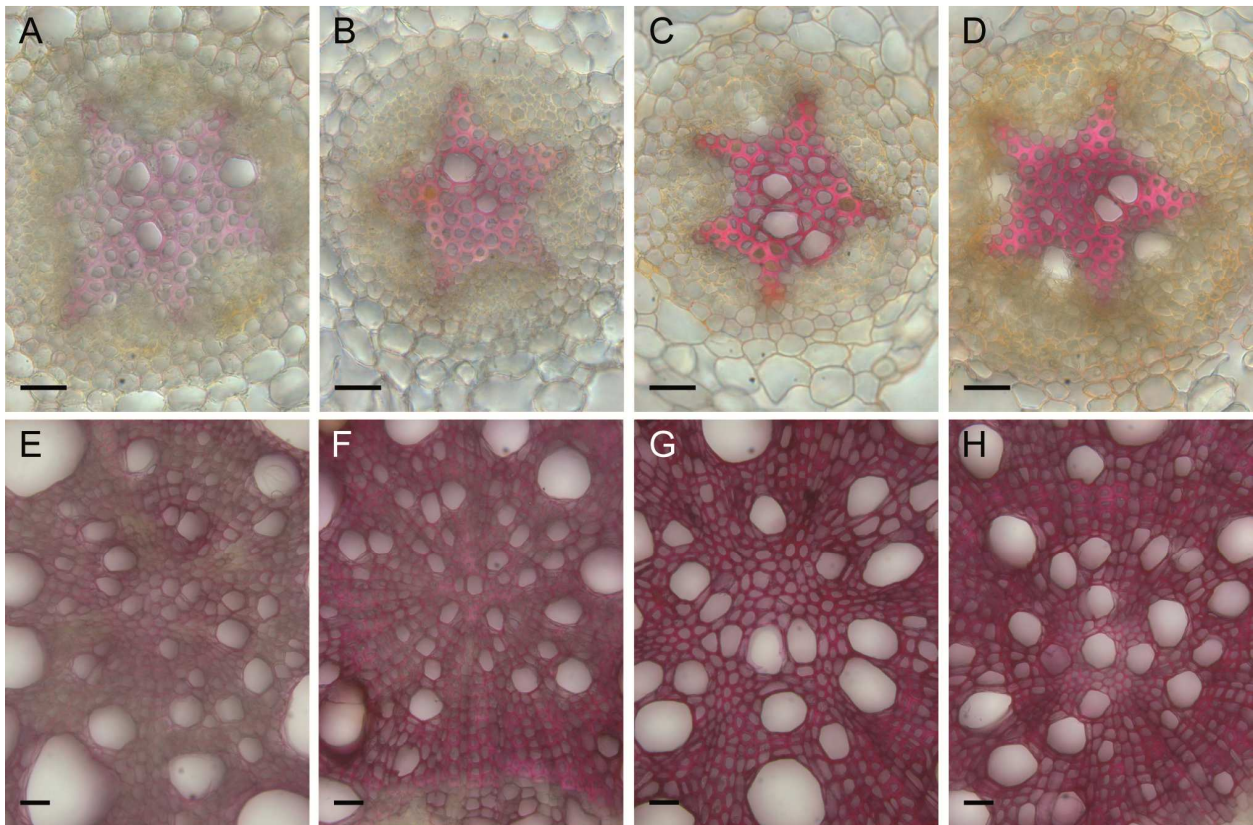
Peak Number	Retention Time (min)	m/z	Substance Name	35S:EgH1.3		35SEgMYB1/35S:EgH1.3		35S:EgMYB1		CONTROLS	
				Peak Area	Ratio	Peak Area	Ratio	Peak Area	Ratio	Peak Area	Ratio
1	1.379	-	unknown	85,001 $\pm$ 15,655		73,782 $\pm$ 21,630		79,687 $\pm$ 6,611		93,707 $\pm$ 10,668	
2	2.207	450.056	5-methylsulfinylpentyl glucosinolate (glucoalyslin, 5MSOP)	2,009 $\pm$ 626		6,348 $\pm$ 1,303 **		2,237 $\pm$ 240		2,166 $\pm$ 862	
3	2.448	463.047	1-Hydroxy-3-indolylmethyl glucosinolate (H13M)	17,052 $\pm$ 2,909		17,093 $\pm$ 16,188		14,516 $\pm$ 1,582 *		18,254 $\pm$ 1,804	
4	2.779	500.140	6-hydroxyindole-3-carboxylate dihexose (H13CDH)	ND		7,520 $\pm$ 1,509 ***		ND		ND	
5	3.203	338.068	6-hydroxyindole-3-carboxylate hexose (H13CH)	1,382 $\pm$ 370		29,069 $\pm$ 15,810 *		849 $\pm$ 20 *		1,008 $\pm$ 92	
6	3.628	447.053	3-indolylmethyl glucosinolate (glucobrassicin, 13M)	27,399 $\pm$ 9,622		97,628 $\pm$ 52,700 *		15,444 $\pm$ 2,418 ***		26,600 $\pm$ 809	
7	3.667	285.061	Protocatechyl xylose (PCX)	6,776 $\pm$ 3,091		16,874 $\pm$ 2,823 ***		4,706 $\pm$ 1,942		3,151 $\pm$ 649	
8	4.272	492.103	8-methylsulfinyl-n-octylglucosinolate (glucohisutin, 8MSOO)	1,042 $\pm$ 181 *		2,743 $\pm$ 348 ***		1,707 $\pm$ 620 *		685 $\pm$ 123	
9	4.564	477.064	4-methoxy-3-indolylmethyl glucosinolate (4MI3M)	4,526 $\pm$ 1,478 *		41,876 $\pm$ 10,740 ***		3,168 $\pm$ 339		2,589 $\pm$ 389	
10	4.96	385.113	Sinapoyl-glucose (SG)	9,314 $\pm$ 2,781		ND *		2,839 $\pm$ 2,067 *		11,194 $\pm$ 4,931	
11	5.314	477.064	1-methoxy-3-indolylmethyl glucosinolate (1MI3M)	3,166 $\pm$ 529 ***		7,973 $\pm$ 4,630 *		ND		ND	
12	5.669	755.201	Quercetin-3O-(Rh-G)-7O-Rh (Q3RG7R)	5,954 $\pm$ 3,062		9,174 $\pm$ 5,901		3,176 $\pm$ 5,223		3,263 $\pm$ 763	
13	5.659	-	unknown	3,069 $\pm$ 628		7,778 $\pm$ 3,625 *		2,596 $\pm$ 207 *		3,169 $\pm$ 321	
14	5.883	739.207	Kaempferol-3O-(Rh-G)-7O-Rh (K3RG7R)	99,944 $\pm$ 19,678		130,310 $\pm$ 27,048 *		65,054 $\pm$ 21,464		88,895 $\pm$ 11,351	
15	6.333	339.071	Sinapoyl malate (SM)	71,186 $\pm$ 4,183		83,694 $\pm$ 46,839		43,069 $\pm$ 12,486 *		75,774 $\pm$ 8,402	
16	6.921	593.150	Kaempferol rutinoside (K3R7G)	126,040 $\pm$ 46,256		97,614 $\pm$ 28,308		83,286 $\pm$ 37,555		123,159 $\pm$ 21,164	
17	7.504	577.155	Kaempferol-3O-Rh-7O-Rh (K3R7R)	213,907 $\pm$ 43,841		240,871 $\pm$ 50,668		137,081 $\pm$ 50,576		194,318 $\pm$ 24,960	
18	9.635	-	unknown	2,167 $\pm$ 526		3,248 $\pm$ 1,019 *		1,317 $\pm$ 154		1,590 $\pm$ 168	
19	10.526	-	unknown	6,273 $\pm$ 666		6,771 $\pm$ 649 *		4,757 $\pm$ 471		5,261 $\pm$ 395	



#### 4.3.8 SIMULTANEOUS OVEREXPRESSION OF EgMYB1 AND EgH1.3 IN A *EUCALYPTUS*

##### HOMOLOG TRANSFORMATION SYSTEM VALIDATES THE PHENOTYPE OBSERVED IN *ARABIDOPSIS*

We also investigated the effects of overexpressing either EgH1.3, EgMYB1, or both genes simultaneously in a homologous transformation system under the control of 35S promoter. To this end, we generated several independent *E. grandis* composite plants harbouring transgenic hairy roots using a protocol that we recently set up in the team (Plasencia *et al.*, 2015). Using RT-qPCR, we selected the eight independent composite plants for each construction that express the most the respective transgenes for further characterization (Supplemental figure S6). We did not observe any differences between the non-transformed aerial parts of all the composite *Eucalyptus* plants. Concerning transgenic roots, no clear differences at macroscopic level could be observed among lines. We made cross sections from young parts of the transgenic roots (sampled at 10cm from the root apex) mostly containing primary xylem, and from older parts of roots (at 40cm from root apex), exhibiting secondary growth with several layers of secondary xylem. The intensity of phloroglucinol-HCl staining, which is an indicator of the lignin content, was reduced in 75% of the young roots and in 89% of the older roots of the *Pro35S:EgH1.3/Pro35S:EgMYB1* composite *Eucalyptus* plants (Figure 4.8A, E). Similarly, 50% of the young and 43% of the older transgenic roots from *Pro35S:EgMYB1* *Eucalyptus* plants exhibited a fainter phloroglucinol-HCl staining, although to a lesser extent than in *Pro35S:EgH1.3/Pro35S:EgMYB1* roots (Figure 4.8B, F). In contrast, no clear difference in the staining intensity was observed between *Eucalyptus* hairy roots overexpressing *EgH1.3* and control roots (Figure 4.8C, D, G, H).

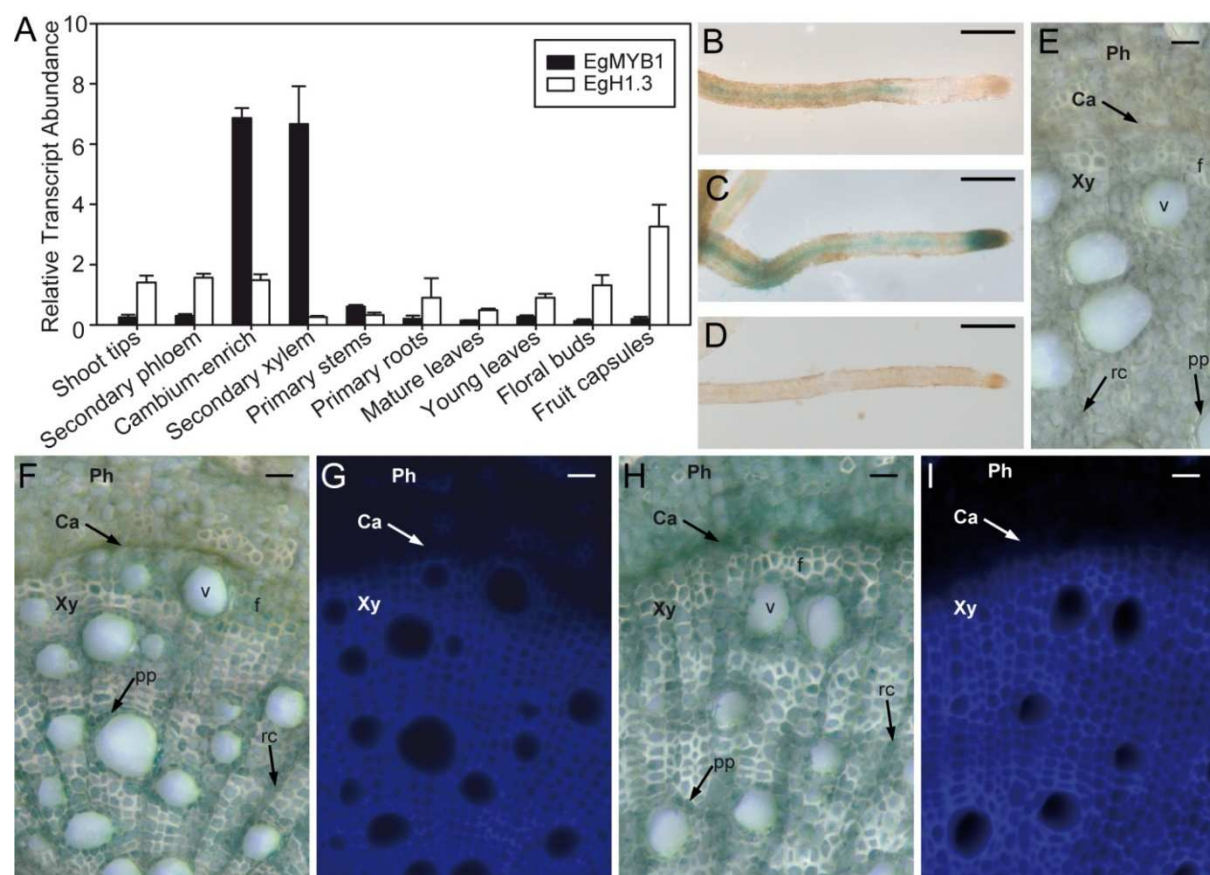


**Figure 4.8: Xylem structure of *Eucalyptus* hairy roots.** Hairy roots sections of *Pro35S:EgMYB1/Pro35S:EgH1.3* (A, E), *Pro35S:EgMYB1* (B, F), *Pro35S:EgH1.3* (C, G) and control constructs (D, H) after staining with phloroglucinol-HCl. Lignified cell walls are coloured in red. Young roots (A-D) were sampled at 10 cm from the root apex, whereas old roots with secondary growth (E-H) were harvested at 40 cm from the root apex. Scale bar = 25  $\mu$ m.

#### 4.3.9 PROMOTER EXPRESSION ANALYSIS AND QPCR REVEALED PARTIAL OVERLAPPING EXPRESSION PATTERNS BETWEEN EGMYB1 AND EGH1.3

RT-qPCR analyses on a large panel of *Eucalyptus* organs and tissues showed a preferential expression of *EgMYB1* in the developing secondary xylem and in the cambium-enriched fraction (Figure 9A), in agreement with recent results (Cassan-Wang et al., 2012; Soler et al., 2015). In contrast, *EgH1.3* was expressed at similar levels in all the tissues tested, although its transcript abundance was lower in developing secondary xylem as compared to other tissues or organs (Figure 9A). Thus, the highest transcript levels for both *EgMYB1* and *EgH1.3* were found in the cambium-enriched fraction, which includes cambial initials as well as xylem and phloem cells at early stages of their differentiation process.

To gain further insights into the expression profiles of *EgMYB1* and *EgH1.3*, we transformed *Eucalyptus* hairy roots with the corresponding promoters fused to the *GUS* reporter gene. At least five different transformation events were analysed for each construct, showing similar expression patterns. The activity of both *EgMYB1* and *EgH1.3* promoters was observed at the root vasculature (Figure 9B). The *EgH1.3* promoter activity was also found in root tips (Figure 9C). Sections from roots exhibiting secondary xylem showed activity of both promoters in the cambium and in the very first layers of developing xylem (Figure 9F-I). The *EgH1.3* promoter was also active in developing phloem (Figure 9H). In more mature xylem layers, GUS staining was observed for both promoters in ray cells and in paratracheal parenchyma cells surrounding vessels (Figure 9F, H). The absence of GUS activity detected in plants transformed with a control empty vector construct demonstrate the specificity of staining (Figure 9D, E).



**Figure 4.9. Expression pattern of *EgMYB1* and *EgH1.3* in *Eucalyptus*.** A. RT-qPCR transcript abundance patterns of *EgMYB1* and *EgH1.3* in the different organs and tissues from field-grown eucalypts. B-I. GUS activity profiles driven by the promoter of *EgMYB1* (B, F, G), and by the promoter of *EgH1.3*, (C, H, I) in transgenic *E. grandis* whole young hairy roots (B-D) and in cross sections from older hairy roots (E-I). Control constructs showed no GUS activity (D, E). UV fluorescence images G and I, which corresponded to F and H sections, respectively, allow the visualization of fluorescence in lignified xylem cells. Scale bar = 0.25 mm for B-D, 25  $\mu$ m for E-I. Ph, phloem; Ca, cambium; Xy, xylem; f, fibres; v, vessels; rc, ray cells; pp, paratracheal parenchyma cells.

These observations highlight that the overlapping expression of *EgMYB1* and *EgH1.3* is restricted to cambial cells and xylem cells at the beginning of their differentiation process, but also in mature xylem ray and parenchyma cells.

#### 4.4 DISCUSSION

By screening a *Eucalyptus* xylem yeast-two-hybrid library, we identified the linker histone protein EgH1.3 as a potential interactor of the transcription factor EgMYB1. This interaction was validated *in planta* using the FRET-FLIM technology. We further showed that the EgMYB1-EgH1.3 interaction enhances the transcriptional repression activity of EgMYB1 over the promoters of two lignin biosynthetic genes, *EgCAD2* and *EgCCR*, known to be direct targets of EgMYB1 (Legay et al., 2007). We then used two types of stable transformants (i.e. *Arabidopsis* plants and *Eucalyptus* hairy roots) overexpressing each transgene alone (*EgMYB1* or *EgH1.3*) or both simultaneously to analyse the effects of the EgMYB1-EgH1.3 interaction *in planta*. In both transformation systems, the main and common outcome of the interaction between EgMYB1 and EgH1.3 was a strong reduction of lignin deposition in xylem cell walls, a noticeably enhanced phenotype relative to transgenics overexpressing *EgMYB1* alone. It should be noted that transformants overexpressing only *EgH1.3* did not show any reduction of lignin deposition. In *Arabidopsis* lines overexpressing both *EgMYB1* and *EgH1.3*, a notable reduction of xylem secondary cell wall thickness was also observed, further supported by a clear repression of the transcript levels of genes involved in the biosyntheses of lignin, cellulose and hemicelluloses. We also observed a significant increase in secondary metabolites related to the phenylpropanoid metabolism concomitantly with the strong repression of lignin deposition in plants overexpressing together EgMYB1 and EgH1.3.

A growing number of studies performed essentially in animals revealed that, in addition to their well-known architectural role stabilizing chromatin structure, histone linkers can interact with many different proteins (reviewed in McBryant *et al.*, 2010). Some of these interactions have been demonstrated to affect post-translational modifications of core histones in order to modify gene expression at gene specific loci through the induction of changes in chromatin conformation (Lu *et al.*, 2009, 2013; Yang *et al.*, 2013). In plants, histone linker proteins are far

less studied and only two proteins interacting with histone linkers have been functionally characterized in *Arabidopsis*, and the two of them induce changes in DNA methylation to modify gene expression (Rea *et al.*, 2012; Zemach *et al.*, 2013). Another interaction involving a plant histone linker protein have been described in banana, this time with a transcription factor from the WRKY family, although no functional data of the interaction is provided (Wang *et al.*, 2012). Taking in consideration all these recent studies, one possible mechanism to explain the enhanced repressive effect of EgMYB1 when interacts with the histone linker is that EgMYB1 can provide the sequence specificity to EgH1.3. In turn, EgH1.3 could interact with other proteins to modify the degree of DNA methylation, or to post-translationally modify core histones to repress the transcription of EgMYB1 target genes.

EgMYB1 contains at the C-terminal region an EAR repressive domain, considered the most common mechanism of gene repression in plants, which works through the recruitment of chromatin-remodelling factors to mediate gene silencing by changing the chromatin conformation (Kagale and Rozwadowski, 2011). This EAR domain is indeed necessary for the repression of activity of EgMYB1, since the R2R3-MYB domain alone is not able to repress the EgCAD gene expression even when co-expressed with EgH1.3 in *N. benthamiana* transient assays. However, the R2R3-MYB domain of EgMYB1 is sufficient for interacting with EgH1.3, which suggests that the EAR domain is not essential for establishing this interaction but necessary for the biological function of the interaction. In any case, the mechanism by which EgMYB1 EAR-related repressing machinery could cooperate with the chromatin remodelling machinery of histone linkers remains to be elucidated.

It has been demonstrated that EgMYB1 can bind to the promoters of two lignin biosynthetic genes and repress their expression (Legay *et al.*, 2007), but when overexpressed in both poplar and *Arabidopsis* plants, not only genes involved in lignin biosynthesis showed reduced expression but also those from cellulose and hemicellulose biosynthesis (Legay *et al.*, 2010). In contrast, the *Arabidopsis* lines overexpressing EgMYB1 that we used in the present study exhibited a moderate phenotype, and we could only detect a tendency in the repression of lignin biosynthetic genes. However, when we overexpressed together EgMYB1 with EgH1.3, lignin, cellulose and hemicellulose biosynthesis genes were clearly repressed. In literature, orthologs of EgMYB1 in other species have been largely related to phenylpropanoid and lignin

biosynthesis, but they do not seem to have a clear role in controlling the whole secondary cell wall formation (Tamagnone *et al.*, 1998; Jin *et al.*, 2000; Fornalé *et al.*, 2006; Sonbol *et al.*, 2009; Fornalé *et al.*, 2010; Ma *et al.*, 2011; Shen *et al.*, 2012). Therefore, it is possible that the main role of EgMYB1 is to repress lignin biosynthesis, and when lignin repression is too strong such as the EgMYB1 overexpression lines used by Legay *et al.* (2010) or when overexpressed together with EgH1.3, plants significantly reduce all the secondary cell wall components including cellulose, hemicelluloses and lignin.

According to RT-qPCR results from field-grown *Eucalyptus* trees and promoter expression analyses performed in *Eucalyptus* hairy roots, expression profiles of *EgMYB1* and *EgH1.3* overlaps in cambial cells and in the early stages of xylem cells differentiation. Those tissues contain meristematic cells; and fibres, vessels and parenchymatic cells at the early stages of xylogenesis, respectively. According to the model proposed by Hertzberg *et al.* (2001), xylogenesis is a precise temporally and spatially fine-tuned process where cambial cells should maintain their meristematic nature, whereas xylem mother cells will just begin their differentiation process, being prevented from lignin deposition until the middle-later stages of cell differentiation. It is conceivable, therefore, that the expression of *EgH1.3* and *EgMYB1* overlapping in cambium and in cells starting their differentiation process would prevent them from premature deposition of secondary cell walls. Then, as differentiation progress and *EgH1.3* expression drops, the repressing activity of EgMYB1 would be weaker, allowing deposition of secondary cell walls. Other repressors and activators could also contribute to this fine regulation, like EgMYB2, a master activator of secondary cell wall formation highly abundant in xylem but not in cambium (Goicoechea *et al.*, 2005; Soler *et al.*, 2015) that is able to directly bind and activate the promoters of both *EgCAD2* and *EgCCR* genes competing with EgMYB1 (Goicoechea *et al.*, 2005).

The expression of *EgMYB1* and *EgH1.3* also overlaps in rays and in paratracheal cells surrounding vessels. These two kinds of xylem parenchymatic cells have thin secondary cell walls exhibiting different structure from that of tracheary elements (Chafe and Chauret, 1974; Donaldson *et al.*, 2015), being even unlignified in some species, and they can remain alive for several years until the formation of heartwood (Déjardin *et al.* 2010). Taking into account the characteristics of these cells, it seems that the repressing effect of the interaction between



EgMYB1 and EgH1.3 not only controls lignin deposition and secondary cell wall progression in differentiating fibres and vessels, but also integrates positional cell identity signals to limit the extent of secondary cell wall deposition in specific parenchyma cell types. In this sense, probably by working antagonistically to positive regulators, EgMYB1 and EgH1.3 could be part of a complex fine-tuned mechanism to control precisely the extent of secondary cell wall deposition depending on the cell types and on the cell differentiation stage.

Phylogenetic analyses indicate that EgH1.3 is the ortholog of a linker histone variant that is drought induced in several plant species (Ascenzi and Gantt, 1997, 1999b; Jerzmanowski *et al.*, 2000; Wei and O'Connell, 1996; Bray *et al.*, 1999; Scippa *et al.*, 2000, 2004; Trivedi *et al.*, 2012), even if its role was not always clearly related to drought resistance (Ascenzi and Gantt, 1999a; Przewloka *et al.*, 2002). Over and Michaels (2014) proposed that drought-induced linker histones may also have conserved functions in development, but this remains to be experimentally proved. Whether EgH1.3 is directly induced by drought or other biotic and abiotic stress is still to be established, but it is worth noting that ectopic expression of EgH1.3 in *Arabidopsis* results in an induction of some glucosinolates, like if plants were constitutively under stress conditions. The increase of glucosinolates, especially indole-derived, was much stronger in 35S:EgMYB1/EgH1.3 plants, suggesting that the interaction of these two proteins redirects the carbon flux and reinforces this stress simulating effect. Moreover, apart from glucosinolates, 35S:EgMYB1/EgH1.3 plants also accumulated other secondary metabolites, like flavonoids and phenolic glucosides, both derived from the phenylpropanoids metabolism. It has been recently shown that there is a crosstalk between indole glucosinolates and the phenylpropanoid metabolism (Kim *et al.*, 2015). As suggested by the dramatic increase in phenylpropanoid-related secondary metabolites in *Arabidopsis* overexpressing both genes, we hypothesize that EgMYB1 through its interaction with EgH1.3 could be involved in maintaining metabolic homeostasis playing a role at the cross-talk between the biosynthesis of lignin and the production of phenylpropanoid metabolites in response to stress. In the case of plants overexpressing EgMYB1, no important modifications in the abundance of glucosinolates were found, but they showed a strong repression of sinapoyl malate, implicated in UV-B resistance, in a similar way as described for others orthologs of EgMYB1 (Jin *et al.*, 2000), meaning that both lignin and sinapate esters biosynthesis are repressed in 35S:EgMYB1 plants. Thus, the big

increase of phenylpropanoid-derived secondary metabolites in 35S:EgMYB1/35S:EgH1.3 lines could be a result of a metabolic redirection after substantially blocking lignin biosynthesis.

## 4.5 MATERIAL AND METHODS

### 4.5.1 PHYLOGENY RECONSTRUCTION

Histone linker sequences were retrieved from different plant genomes using BlastX searches at Phytozome 10.3 database (<http://www.phytozome.net/>) using the three *Arabidopsis* histone linker proteins as query (Ascenzi and Gantt, 1997). Aminoacidic sequence alignment was performed using MAFFT with the L-INS-i method (Katoch *et al.*, 2002), and neighbor-joining phylogenetic tree (Saitou and Nei, 1987) with 1000 bootstrap replicates was constructed with Mega5 (Tamura *et al.*, 2011) using the pairwise deletion method. The evolutionary distances were computed using Poisson correction method with uniform variation among sites (Zuckerandl and Pauling, 1965).

### 4.5.2 VECTOR CONSTRUCTION

Coding sequences corresponding to EgMYB1, the R2R3-domain of EgMYB1, EgH1.3 and EgMYB137 were amplified from *E. gunnii* cDNA, and the promoter sequences from EgMYB1 and EgH1.3 were amplified using *E. grandis* genomic DNA (Specific primers are listed in Supplemental Table S4.2). PCR for cloning were performed using the PhusionTaq (Fisher Scientific) following manufacturer's instructions and amplicons were inserted into the pDONR207 vector using the BP clonase II (Invitrogen, Carlsbad, USA) or into pENTR/D-TOPO vector (Invitrogen) following manufacturer's instructions. Once inserted into pDONR207 or pENTR/D-TOPO, sequences were transferred into the different destination vectors using LR clonase II (Invitrogen) following specifications of the manual. Gateway destination vectors used are pGBG-BD-GTW and PGAD-AD-GTW (kindly provided by Laurent Deslandes) for the yeast-two-hybrid, pBin19-35S-GW-CFP and pBin19-35S-GW-YFPv (Froidure *et al.*, 2010) for sub-cellular localization and FRET-FLIM analysis, pFAST-GO2 (Shimada *et al.*, 2010) for over-expression in *Arabidopsis*, pGWAY-0 (Plasencia *et al.*, 2015) for *Eucalyptus* hairy roots transformation and *N. benthamiana* transactivation assays, pBGWFS7 (Karimi *et al.*, 2002) for promoter expression in *Arabidopsis*, and pGWAY-1 (Plasencia *et al.*, 2015) for promoter

expression in *Eucalyptus* hairy roots. Moreover, also for hairy roots transformation and transactivation assays, we also created a new vector to study the effects of overexpressing EgMYB1 together with other proteins, and we named it as pGWAY-MYB1. This vector was created by excising a 2.2 Kb fragment containing the 35S CaMV promoter, the EgMYB1 coding sequence and the 35S CaMV terminator, previously cloned in the pJR1 vector (Legay *et al.*, 2007), and then inserted into the pGWAY-0 vector (Plasencia *et al.*, 2015) by KpnI digestion and T4ligase ligation. This vector contains a cassette expressing EgMYB1 under the 35S promoter and a Gateway region to insert another gene, also under the control of another 35S promoter (Supplemental figure S4.7). We therefore inserted EgH1.3 in that Gateway region by LR cloning to overexpress both proteins together. We transformed *A. tumefaciens* strain GV3101:pMP90 or *A. rhizogenes* strain A4RS with the destination vectors using electroporation (Nagel *et al.*, 1990). We then incubated GV3101 in LB and A4RS in AG media (Franche *et al.*, 1997) with the appropriate antibiotics before transforming plants.

#### 4.5.3 YEAST-TWO-HYBRID LIBRARY CONSTRUCTION AND SCREENING

We constructed a yeast two-hybrid library from mixture of field-grown *Eucalyptus globulus* RNAs obtained by scrapping xylem samples at different development stages (juvenile and mature) and submitted or not to a mechanical stress (tension, opposite and straight xylems). Total xylem RNA from these different *Eucalyptus* xylem tissues was isolated using the method from Southerton *et al.* (1998). We took 10-20 µg of each RNA and we pooled them together until obtaining 75 µg. Pooled RNA was mRNA purified using “Dynabeads mRNA purification kit” (Ambion, Foster City, USA), obtaining 170 ng of purified mRNA. We then generated the cDNA for the yeast-two-hybrid library using the kit “Make Your Own Mate-&-Plate™ Library System” and (Clontech, Mountain View, USA) following manufacturer’s instructions. We generated two cDNA pools, one using random primers and the other using oligo d(T), we mixed them together with the SmaI linearised vector pGADT7-Rec, and we co-transformed yeast Y187 competent cells using the Library Scale protocol from the “Yeastmaker™ Yeast Transformation System” (Clontech). Library complexity was about  $1 \times 10^6$  clones mL<sup>-1</sup>, whereas density was around  $5 \times 10^8$  cells.

Screening was performed using the R2R3 domain of EgMYB1 as a bait to avoid possible repression of GAL4 activity due to the repression C-terminal domain of EgMYB1. We used the “Matchmaker™ Gold Yeast Two-Hybrid System” (Clontech) following the recommendation specified in the instructions, but using the more stringent media SD/-Trp-Leu-Ade-His. We identified 199 colonies growing in the stringent media, with a mating efficiency around 7.7%.

Sequencing of the colonies was performed using primers designed for the pGADT7 vector (Clontech). Plasmid isolation of the interesting colonies was done using the “Easy Yeast Plasmid Isolation Kit” (Clontech) following instructions. Targeted yeast two-hybrid assay was then repeated to test the repeatability and the specificity of the interactions, including the R2R3 domain of an unrelated MYB gene from *Arabidopsis* (AtMYB96, Canonne *et al.*, 2011) and the full-length sequence of an unrelated *Eucalyptus* MYB gene highly expressed in xylem (EgMYB137, Soler *et al.*, 2015).

#### 4.5.4 SUB-CELLULAR LOCALIZATION, CO-LOCALIZATION AND FRET-FLIM ANALYSIS

For the infiltration, transformed *A. tumefaciens* GV3101:pMP90 were grown at 28 °C in liquid LB media until a solution of OD<sub>600</sub>=0.8. Cells were then centrifuged at 2400g for 10 minutes and resuspended into infiltration medium (MES/KOH 10 mM, MgCl<sub>2</sub> 10 mM, acetosyringone 100µM, pH 5.6) at OD<sub>600</sub>=0.5 and incubated for 3 hours at room conditions in dark. Resuspended cells were then infiltrated in 4-5 weeks-old *N. benthamiana* plants using 1 mL needless syringe. 35S:EgMYB1-CFP and 35S:EgMYB1Δ-CFP *A. tumefaciens* cells were infiltrated alone and together with 35S:EgH1.3-YFP. As a negative control, 35S:EgMYB1-CFP was also infiltrated together with 35S:YFP, and for specificity controls, 35S:AtMYB96-CFP and 35S:EgrMYB137-CFP were also infiltrated alone and together with 35S:EgH1.3.

Once infiltrated, plants were grown in growth chambers in long day conditions (16h day length) and, two days after infiltration, leaves were observed under a Confocal Laser Scanning Microscope (TCS SP2-SE, Leica, Wetzlar, Germany) equipped with a 63X water-immersion objective lens (numerical aperture 1.20; PL APO). CFP was excited with the 458 nm ray line of the argon laser and the emitted fluorescence recorded in the 465– 520 nm emission range. YFP was excited with the 514 nm ray line of the argon laser and the fluorescence collected in the

520–575 nm range. Images were acquired in the sequential mode using Leica LCS software (version 2.61).

For FRET-FLIM analysis, fluorescence lifetime of the donor (CFP fusion protein) was measured in the presence and absence of the acceptor (YFP fusion protein). FRET efficiency ( $E$ ) was calculated by comparing the lifetime of the donor in the presence ( $\tau_{DA}$ ) or absence ( $\tau_D$ ) of the acceptor:  $E=1-\tau_{DA}/\tau_D$ . FRET-FLIM measurements were performed using FLIM system coupled to a streak camera. The light source was a 439 nm pulse diode laser (Hamamatsu Photonics, Hamamatsu City, Japan) delivering ultrafast, picosecond pulses with a frequency of 2 MHz. All images were acquired with a 60X oil-immersion lens (Plan Apo 1.4 Numerical Aperture) mounted on an inverted microscope (Eclipse TE2000E, Nikon, Tokyo, Japan) coupled to the FLIM system. The fluorescence emission was directed to the detection unit through a band pass filter (483/32 nm). The FLIM unit was composed of a streak camera (Streakscope C4334, Hamamatsu) coupled to a fast and highly sensitive CCD camera (model C8800-53C, Hamamatsu). For each nucleus, average fluorescence decay profiles were plotted and lifetimes were estimated by fitting data to a bi-exponential function using a non-linear least-squares estimation procedure ( $\tau$ topography-FLIM v3.1 software-APP CNRS).

#### 4.5.5 TRANSACTIVATION ASSAYS

*N. benthamiana* leaves were transformed with the effectors (35S:EgMYB1, 35S:EgH1.3 and 35S:EgMYB1/35S:EgH1.3) and the reporters (ProEgCAD2:GUS and ProEgCCR:GUS). Infiltration was done as described above and, two days after infiltration, 10 leaf discs were sampled and frozen with liquid nitrogen until no more than 48h. Quantitative GUS assay was performed as reported previously (Jefferson, 1987) using 4-methylumbelliferyl- $\beta$ -D-glucuronide as substrate and estimated as the mean of three independent assays, each containing at least three replicates. GUS activity results were normalized according to the total protein concentration or the DsRed fluorescent present in the effector's vector giving similar results (Supplemental figure S4.4), but we discussed only results normalized using total protein content as is the method used in literature about EgMYB1 (Legay *et al.*, 2007).

#### 4.5.6 PLANT MATERIAL

*A. thaliana* Col-0 plants were transformed under the 35S promoter with EgH1.3 inside the pFAST-G02 vector using the *A. tumefaciens* strain GV3101:pMP90 and the floral Dip method (Clough and Bent, 1998). *A. thaliana* plants harbouring EgMYB1 under the 35S promoter inside the pJR1 vector were already produced in the lab (Legay *et al.*, 2010), and one line with a moderate phenotype was subsequently transformed with EgH1.3 inside the pFAST-G02 vector to obtain lines overexpressing both genes. Primary transformants were selected using a fluorescent stereomicroscope with GFP long pass filters, as described in Shimada *et al.* (2010). More than 10 transformed independent lines were obtained for each construct with similar phenotypes, and we selected three lines for each construct according to transgene transcript abundance in leaves using RT-qPCR for detailed characterization of their phenotypes (Supplemental figure S5A, B). We further analysed the expression level of the transgene in *Arabidopsis* inflorescence stem (Supplemental figure S5C). *A. thaliana* col-0 plants were also transformed with the ProEgMYB1 and ProEgH1.3 sequences cloned inside the pBGWFS7 vector, and selected using Basta in soil-grown plantlets. For phenotype characterization, plants were grown in a growth chamber in short day conditions (9 hours light and 15h dark) to promote secondary growth. Sampling was done after eight weeks, when the inflorescence stem was fully developed and the first well-formed siliques were clearly observed. The base of the inflorescence stem was kept in ethanol 80% for histochemistry analysis. The rest of the stem after removing siliques and leaves was pooled from five plants of each transformation event/line, immediately frozen in liquid nitrogen and kept at -80 °C for biochemistry quantification and RNA extraction. Hypocotyls were also frozen in liquid nitrogen and kept at -80 °C or stored in ethanol 80%.

*Eucalyptus grandis* hairy root transformation was performed as described by Plasencia *et al.* (2015). EgMYB1 and EgH1.3 were overexpressed alone inside the pGWAY-0 vector, and for co-overexpression, EgH1.3 was inserted in the pGWAY-MYB1 vector. *E. grandis* hairy roots were also obtained for ProEgMYB1 and ProEgH1.3 sequences inserted in the pGWAY-1 vector. As detailed in Plasencia *et al.* (2015), transgenic hairy roots were detected using a fluorescent stereomicroscope equipped with DsRed filters. Composite plants were then transferred into substrate (Oil-Dri US-Special Type III / R; Damolin) and we let them hardening for 2 months

before harvesting. We obtained from nine to 20 independent transformation events for each construction. Sampling of 35S:EgMYB1, 35S:EgH1.3, 35S:EgMYB1/35S:EgH1.3 or the corresponding empty vector was done taking the first five cm from the tip from 4-5 roots, immediately freezing them in liquid nitrogen and then kept at -80 °C until RNA extraction. Five to eight composite plants of each type were selected for characterization according to their transgene expression evaluated by RT-qPCR (Supplemental figure S4.6). For biochemical analyses, samples of root segments between 15 and 40 cm from root apex were frozen in liquid nitrogen and kept at -80 °C until analysis. For histochemistry, samples of root segments at 10 (young) and 40 cm (old) from the tip were stored in ethanol (80%). For histochemical GUS assay (of the hairy roots transformed with ProEgMYB1:GFP-GUS and ProEgH1.3:GFP-GUS constructions), young and old root segments were harvested and immediately used to assay GUS activity.

#### 4.5.7 TRANSCRIPTOMIC ANALYSIS

*Arabidopsis* stems and hairy root samples were grinded to powder using a ball-mill (MM400, Retsch, Haan, Germany) and liquid nitrogen to keep them protected from degradation. *Arabidopsis* leaf samples were grinded using a plastic pestle with liquid nitrogen. RNA extraction from *Arabidopsis* leaves was done using the protocol for “vegetative tissues”, whereas RNA extraction from stems was done using the protocol for “seeds & siliques (Oñate-Sánchez and Vicente-Carbajosa, 2008). RNA extraction from *Eucalyptus* hairy roots was done using the protocol from Muoki *et al.* (2012) with the modification described by Plasencia *et al.* (2015). Once extracted, RNA was further digested using TurboDNase (Ambion) to remove residual traces of contaminating DNA. Quantity and quality of RNA was further analysed by spectrophotometry using the Nanodrop ND-1000 Spectrophotometer (Thermo Fisher Scientific, Waltham, USA). One µg RNA was retrotranscribed into cDNA using the High Capacity RNA to cDNA kit (Applied Biosystems, Foster City, USA) and diluted five-fold before used as template in RT-qPCR. The obtaining of cDNAs from field-grown *Eucalyptus* trees (including shoot tips, secondary phloem, cambium enriched fraction, secondary xylem, primary stems, primary roots, mature leaves, young leaves, floral buds and fruit capsules) is described in Cassan-Wang *et al.* (2012).

RT-qPCR was performed in technical triplicates using ABI 7900HT fast real-time PCR system (Applied Biosystems) with the Power SYBR Green PCR Master Mix (Applied Biosystems). PCR conditions were 95 °C for 10 minutes, followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. After amplification, a dissociation step was performed to confirm the presence of a simple amplicon. Relative Transcript Abundance was calculated using the  $2^{-\Delta\Delta Ct}$  method (Livak and Schmittgen, 2001), with the housekeeping gene AtACT2 (Legay *et al.*, 2010) for *Arabidopsis* or PP2A-1 and PP2A-3 (Cassan-Wang *et al.*, 2012) for *Eucalyptus* to normalize data, and the mean of the control samples to standardize results. Specific primer sequences used to amplify EgMYB1 and EgH1.3 are listed in Supplementary table S4.2, primer sequences to amplify cellulose, hemicelluloses and lignin biosynthetic genes from *Arabidopsis* are described in Legay *et al.* (2010). They included the following genes: *Cellulose synthase 4* (*CesA4*, At5g44030), *Cellulose synthase 7* (*CesA7*, At5g17420), *Cellulose synthase 8* (*CesA8*, At4g18780), *Fragile fibre 8* (*FRA8*, At2g28110), *Irregular xylem 8* (*IRX8*, At5g54690), *Phenylalanine ammonia lyase 1* (*PAL1*, At2g37040), *Caffeoyl-CoA 3-O-methyltransferase 1* (*CCoAOMT1*, At4g34050), *Ferulate 5-hydroxylase 1* (*F5H1*, At4g36220), *Caffeic acid O-methyltransferase* (*COMT*, At5g54160), *Cinnamoyl-CoA reductase 1* (*CCR1*, At1g15950), *Cinnamyl alcohol dehydrogenase 4* (*CAD4*, At3g19450).

#### 4.5.8 HISTOCHEMICAL ANALYSIS

Stems and hypocotyls from *Arabidopsis* plants were cut in 90 µm-thickness sections, whereas *Eucalyptus* hairy roots were cut in 50 µm-thickness sections after embedding them in agarose 5%, in both cases using a vibratome (VT 100S, Leica). Lignified secondary cell walls were stained using phloroglucinol-HCl reagent and observed under a bright-field inverted microscope (DM IRBE, Leica) equipped with a CDD colour camera (DFC300 FX, Leica). Stem area, lignified area and cell wall thickness were measured using imaging software (ImageJ).

Sections from *Arabidopsis* stems were also analysed by Scanning Electron Microscopy. Sections were dehydrated in an ethanol series, submitted to a critical point dry with CO<sub>2</sub> as a transitional fluid using a CPD300 unit (Leica) and, once dried, coated with nickel (5 nm) in a EM MED020 coating system (Leica) and analysed using a scanning electron microscope (ESEM Quanta 250 FEG, FEI) at 5KV.



Histochemical GUS assay was based on methods described by (Hawkins *et al.*, 1997). Briefly, transformed roots sections were incubated at 37 °C for 24 hours in 100 mM sodium phosphate (pH =7.0), 10 mM EDTA, 0.5 mM K<sub>3</sub>[Fe(CN)<sub>6</sub>], 0.5 mM K<sub>4</sub>[Fe(CN)<sub>6</sub>] and 1 mM 5-bromo-4-chloro-3-indolyl glucuronide (X-Gluc, Euromedex, Souffelweyersheim, France). Samples were further cleared with ethanol before observation under a bright-field inverted microscope (DM IRBE, Leica) equipped with a CDD colour camera (DFC300 FX, Leica).

#### 4.5.9 BIOCHEMICAL ANALYSIS

*Arabidopsis* stem and hypocotyl samples were grinded to powder, as explained before for the transcriptomic analysis, and subsequently lyophilised.

Pyrolysis – Gas Chromatography / Mass Spectrometry (Py-GC/MS) analysis were performed using three technical replicates in three independent runs with 80 µg of milled hypocotyls each. Powder was applied to a pyrolyzer equipped with an auto sampler (PY-2020iD and AS-1020E, Frontier Lab, Fukushima, Japan) connected to a Gas Chromatography/Mass Spectrometry (7890A/5975C; Agilent Technologies, Santa Clara, USA). Pyrolysate separation and analysis, including peak detection, integration, normalization and identification was done according to Gerber *et al.* (2012). Pyrolysis data were subjected to a principal component analysis using R package (<https://www.r-project.org/>) and to a hierarchical clustering using MultiExperiment Viewer (MeV; <http://www.tm4.org/mev.html>)

For measuring the quantity and composition of lignin linked by  $\beta$ -O-4 bounds in *Arabidopsis* hypocotyl, samples were analysed by thioacidolysis as previously described (Méchin *et al.*, 2014). In that case, measures of uncondensed lignin were given as µmols  $\beta$ -O-4 lignin / mg of cell wall. For *Arabidopsis* stems, since the quantity of material was sensibly higher, extractive compounds were removed by exhaustive water, then ethanol extraction in a Soxhlet apparatus. The recovered extract-free samples were dried at 40°C overnight before analysis. In *Arabidopsis* stems, the total lignin was measured by the acetyl bromide method as described by Pitre *et al.* (2007) and the quantity and composition of lignin linked by  $\beta$ -O-4 bounds was measured by thioacidolysis (Méchin *et al.*, 2014). For stems, measures of acetyl bromide-measured lignin were referred as lignin percentage in extract-free cell wall residue, whereas uncondensed lignin was given as µmols  $\beta$ -O-4 lignin / g of acetyl bromide-measured lignin.

Soluble phenolics analysis was performed using 30 mg of freeze-dried *Arabidopsis* stem powder. Samples were extracted with 1 ml of MeOH 70%, blended for 1 min with an Ultra-Thurrax and centrifuged for 10 min at 10000g. The supernatant was transferred to a new tube. One additional millilitre was added to the pellet, then vortexed and left 2 hours at room temperature. The mixture was centrifuged 10 min at 10000g and the supernatant was recovered and pooled with the first one. The 2ml solution was evaporated in a speed vacuum until dryness. The dried pellet was then resuspended in 500 $\mu$ l MeOH 70% and filtrated on 2 $\mu$ m filter before U-HPLC analysis. One microliter of extract was analysed on a U-HPLC system (Shimadzu) equipped with a PDA and a MS detector. The sample was separated on a C18 Kinetex (100 mm  $\times$  2.1 mm) column (Phenomenex). The mobile phase consisted in 0.1% formic acid in ultra-pure water (solvent A) and 0.1% formic acid in methanol (solvent B). The molecules of the sample were eluted through a gradient elution from 1 to 50% B for 7 min, then to 99% B for 1 min with a flow rate of 400  $\mu$ L min<sup>-1</sup>. The column was rinsed for 4 min with 99% B and re-equilibrated to 1% B prior to the next run. Mass analysis was carried out in ESI negative mode. Quantification was realized by measuring the area under peak at 280nm, 320nm or 350nm depending on the  $\lambda$  max of each molecule.

Experimental exact masses and MS2 fragments were compared to metabolomics data banks (Respect: <http://spectra.psc.riken.jp/>; Mass Bank: <http://massbank.jp/>, DNP: <http://dnp.chemnetbase.com/>) and data available in the literature in order to identify the nature of the metabolites.

## 4.6 SUPPLEMENTARY TABLES AND FIGURES

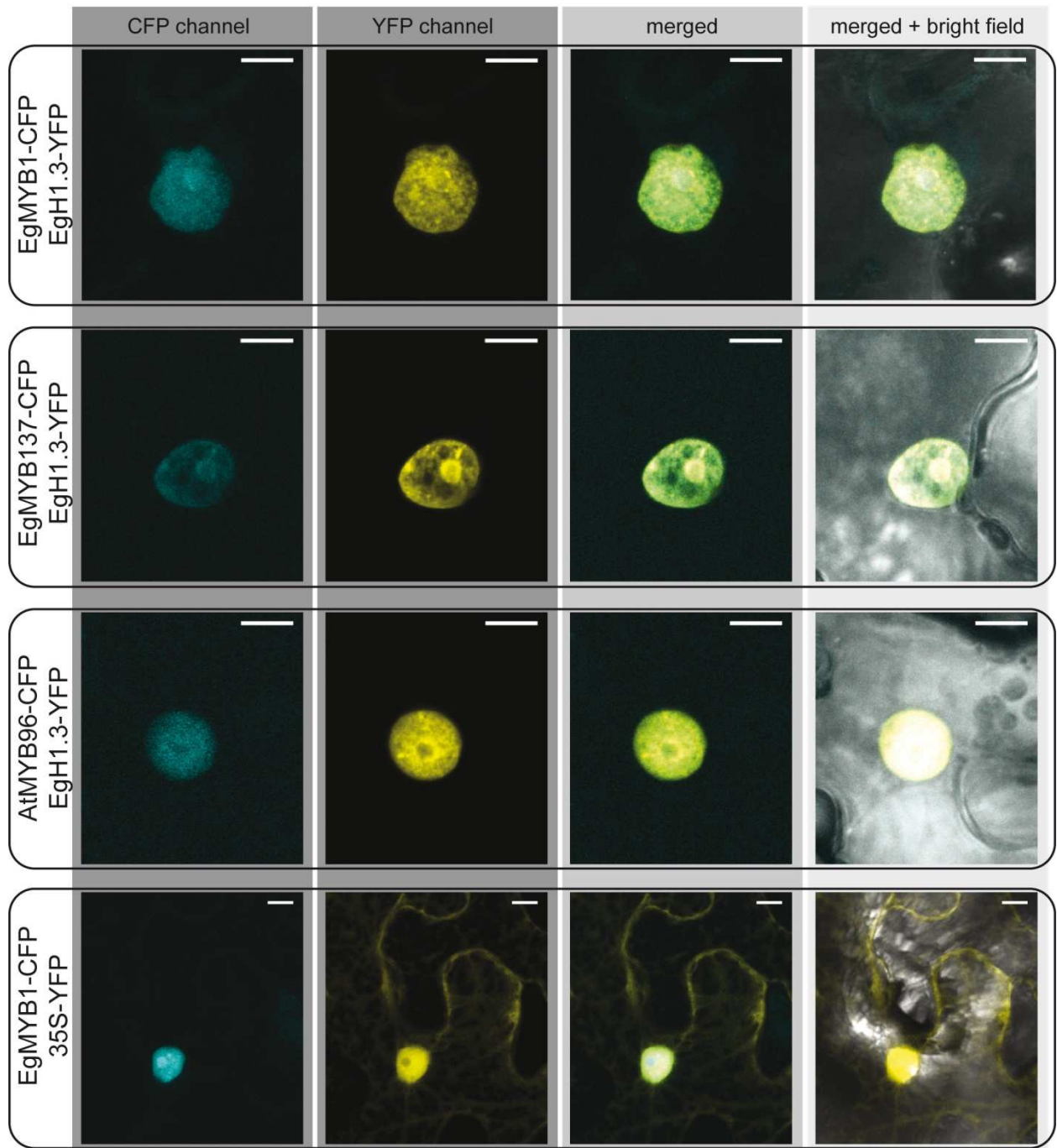
**Table S4.1.** Relative amounts of the compounds categories detected in Pyr-GC/MS.

	Total		Total phenolics		H-type		G-type		S-type		Other		S/G
	carbohydrates	phenolics	phenolics	phenolics	phenolics	phenolics	phenolics	phenolics	phenolics	phenolics	phenolics	phenolics	
<b>35S:EgH1.3</b>	0.98 ± 0.01 ***	1.19 ± 0.19 ***	1.34 ± 0.28 ***	1.13 ± 0.19 **	1.22 ± 0.31 *	1.19 ± 0.24 **	1.06 ± 0.15						
<b>35S:EgMYB1</b>	1.00 ± 0.01	1.03 ± 0.06	1.10 ± 0.09 ***	1.02 ± 0.09	0.79 ± 0.09 ***	1.10 ± 0.12 **	0.78 ± 0.08 ***						
<b>35S:EgMYB1/EgH1.3</b>	0.97 ± 0.02 ***	1.20 ± 0.08 ***	1.66 ± 0.15 ***	1.01 ± 0.09	0.88 ± 0.08 **	1.64 ± 0.26 ***	0.87 ± 0.08 ***						
<b>CNTL</b>	1.00 ± 0.01	1.00 ± 0.06	1.00 ± 0.05	1.00 ± 0.07	1.00 ± 0.15	1.00 ± 0.06	1.00 ± 0.13						

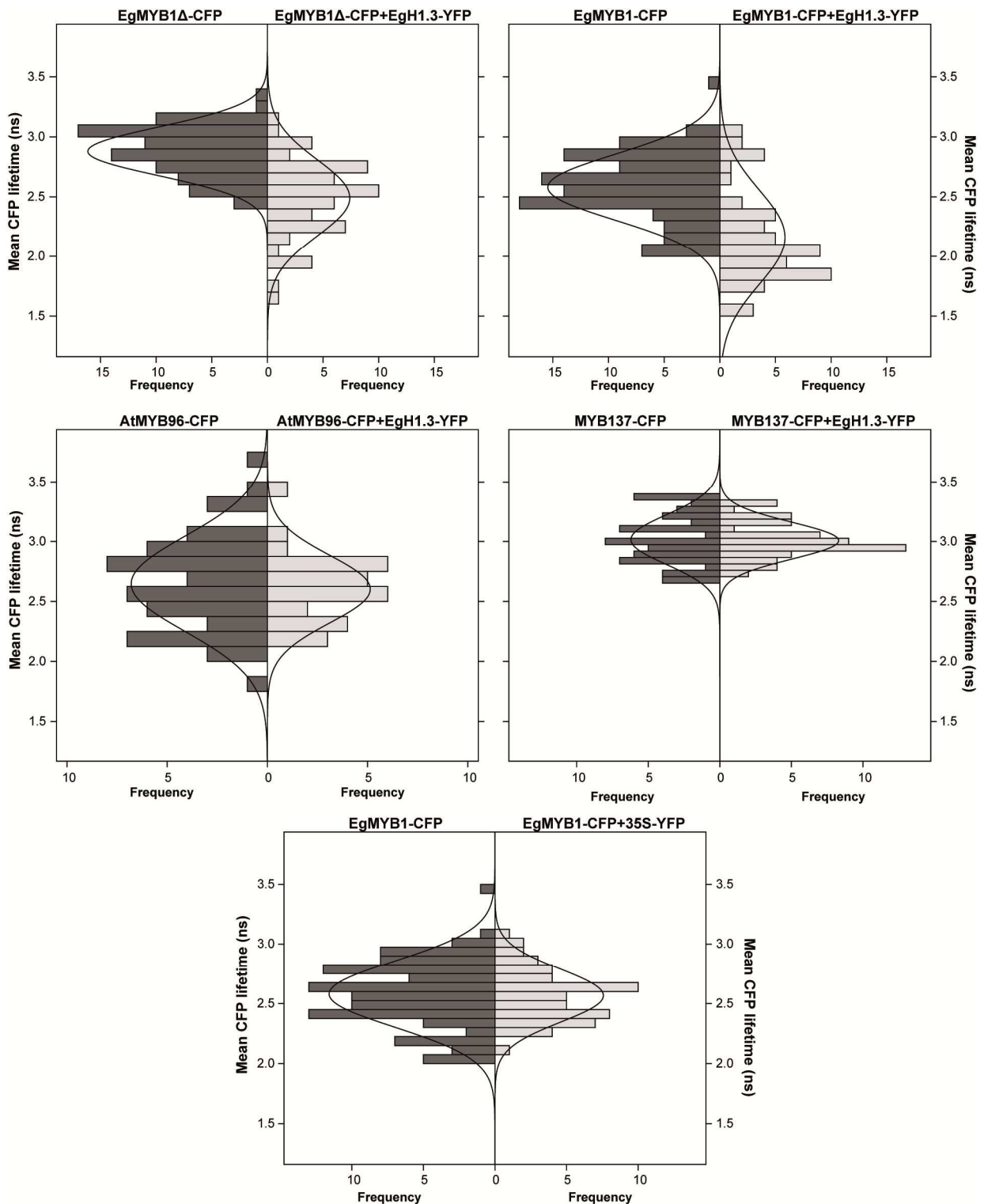
Measures represent the mean ± SD from six biological replicates corresponding to each of the selected lines for each construct growing in two different batches. Each biological replicate consisted of a pool of five plants. Statistics were calculated with Student's t test, \*\*\* p value < 0.001, \*\* p value < 0.01, \* p value < 0.05.

**Table S4.2.** Primers designed in this study.

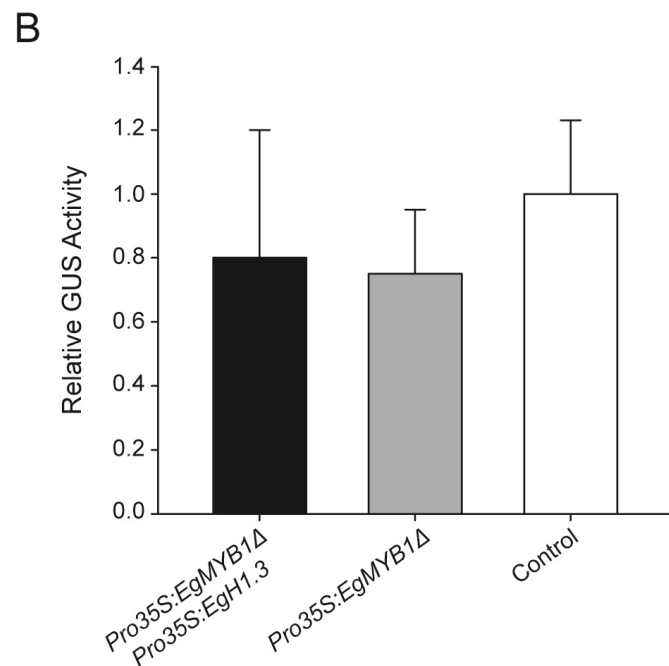
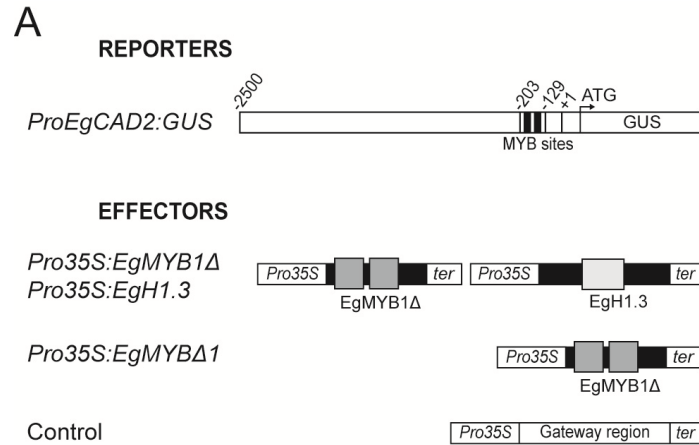
Name	Forward primer 5'-3'	Reverse primer 5'-3'
<b>Cloning primers</b>		
EgMYB1 (AJ576024.1)	GGAGATAGAACCATGGGAAGGTCTCTTGTGCG	CAAGAAAGCTGGGTCGCTCTATAATCCAAAACACTTGCC
EgMYB1Δ(Eucgr.G01774)	CACCTACCACATGGGAAGGTCTCC	GTTCAAAGCTTCTCTTATGTG
EgH1.3 (Eucgr.I02364)	GGGACAAGTTTGTACAAAAAGCAGGCTTCGCCAT GTCGACCACTGTAGA	GGGACCACTTTGTACAAGAAAGCTGGTTAGCAGCAGC TACCTTCTGG
EgrMYB137	CACCTTCTCGCGAGTTCTCAAT	GAAGATGTTGAAATGTCCTGTA
EgrProMYB1 (Eucgr.G01774)	CACCTGTTGTGCCAATTTAGTTCTCA	GTGGTTGATGTTGTTGTCGTTG
EgrProH1.3 (Eucgr.I02364)	CACCCCATGCTTGGCATTGATGA	GGCGTAGTTAGCAGCGGCAGA
<b>RT-qPCR primers</b>		
EgMYB1(Eucgr.G01774)	ACCATGACGAGCCACCATTTTC	TCAGGTCAGGACACCTTTCTCG
EgH1.3(Eucgr.I02364)	AGCAGCCCAAGTCTATCAAGTCG	AAGCAGCAGCTACCTTCTTGGC



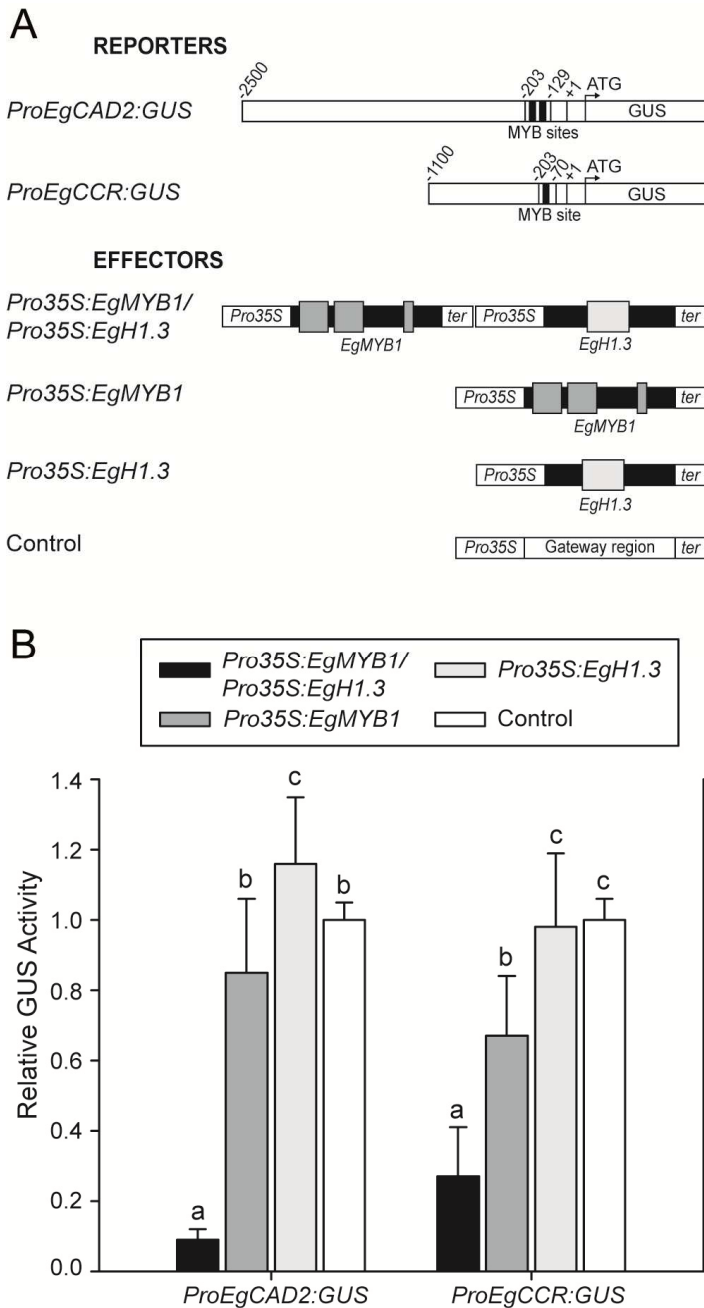
**Figure S4.1.** Subcellular localization of EgMYB1-CFP, AtMYB96-CFP and EgMYB137-CFP infiltrated together with EgH1.3-YFP in *N. benthamiana* leaves. EgMYB1-CFP was also co-infiltrated with YFP.



**Figure S4.2.** Comparison of the FRET-FLIM measures of the donor alone and in presence of the acceptor, expressed in frequency histograms according to the CFP mean lifetime ( $\tau$ ) intervals. Curves show the normal distribution inferred to each histogram.

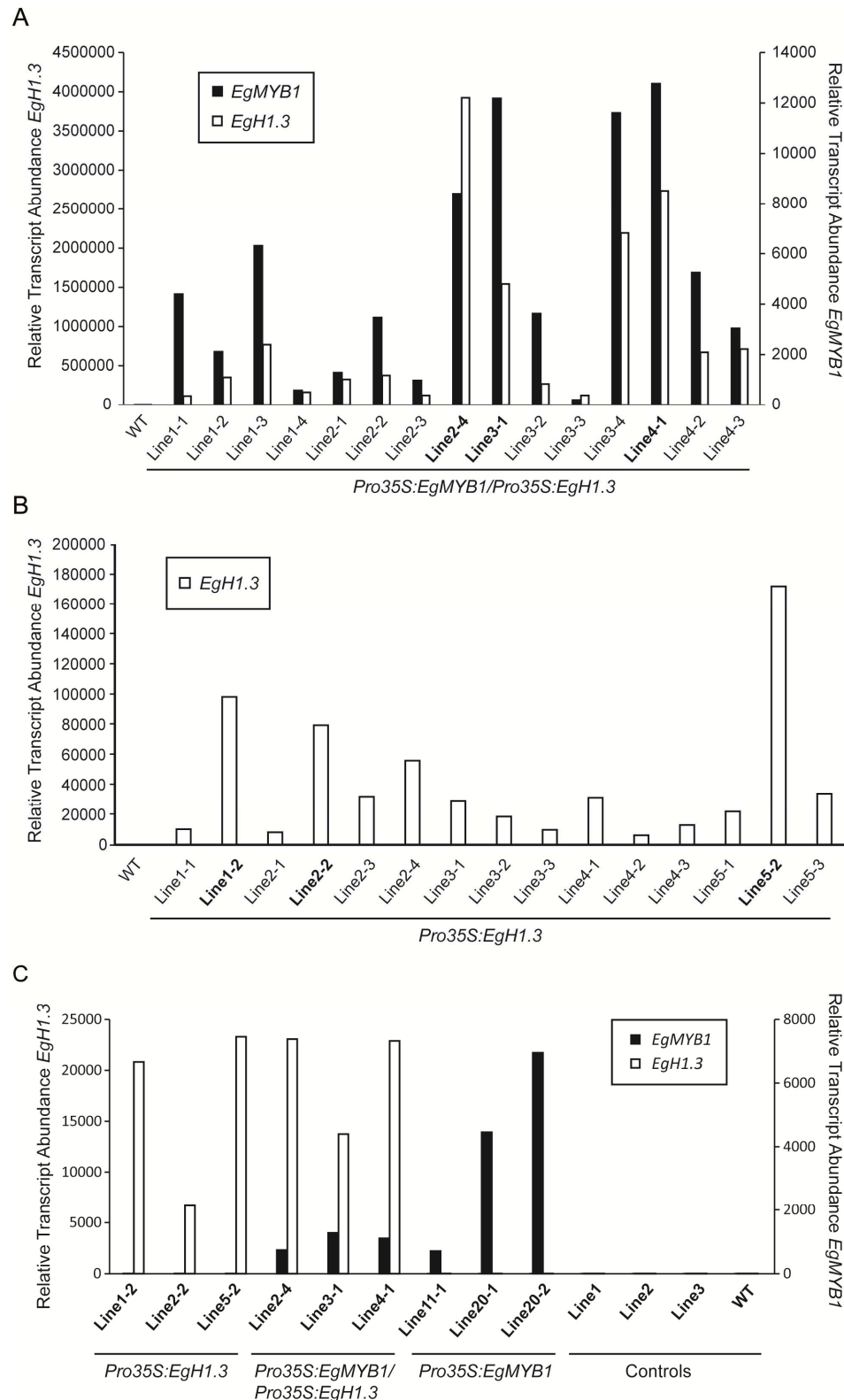


**Figure S4.3. Effect of EgMYB1Δ expressed alone or together with EgH1.3 on the transcriptional activation of the *EgCAD2* promoter using transactivation assays in *N. benthamiana* leaves. **A.** Schemes of the reporter and effector constructs. *Pro35S*, CaMV35S promoter; *ter*, 35S terminator; +1, transcription start site; *GUS*, *uidA* coding region. Black rectangles in effectors correspond to coding regions and grey boxes to protein domains and motifs. **B.** GUS activities measured after co-infiltration of reporters and effectors constructs in tobacco leaves. Data represent means ± standard deviation (SD) resulting from nine biological replicates in at least three independent experiments. GUS activity is expressed as a ratio relative to control. No significant statistical differences were found with Student's *t* test, *P* value < 0.05.**

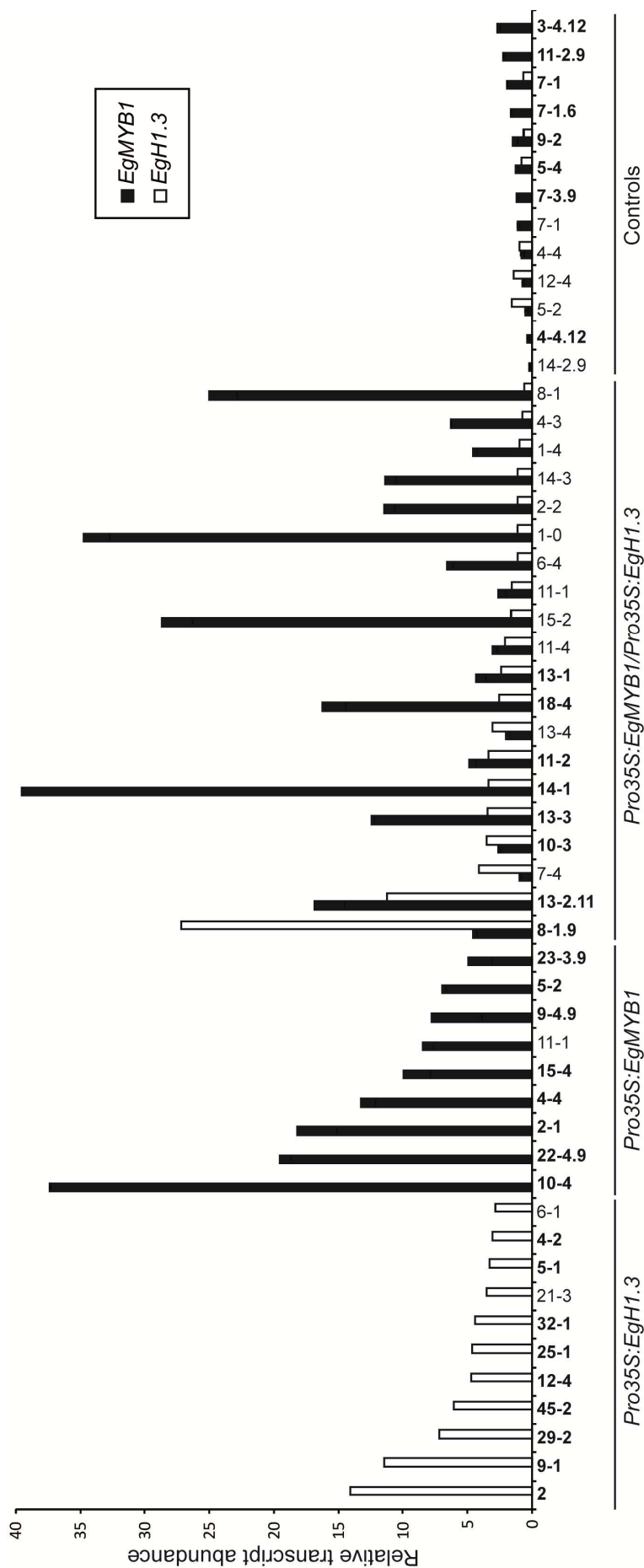


**Figure S4.4: Effect of EgMYB1 and EgH1.3 on the transcriptional activities of the *EgCAD2* and *EgCCR* promoters using transactivation assays in *N. benthamiana* leaves normalized according to DsRED fluorescence. **A.** Schemes of the reporter and effector constructs. *Pro35S*, CaMV35S promoter; *ter*, 35S terminator; +1, transcription start site; GUS, *uidA* coding region. Black rectangles in effectors correspond to coding regions and grey boxes to protein domains and motifs. **B.** GUS activities measured after co-infiltration of reporters and effectors constructs in tobacco leaves. Data represent means  $\pm$  standard deviation (SD) resulting from nine biological replicates in at least three independent experiments. GUS activity is expressed as a ratio relative to control. Statistics were calculated with Student's t test, and significant differences with  $P$  value  $< 0.05$  are highlighted in lowercase letters.**



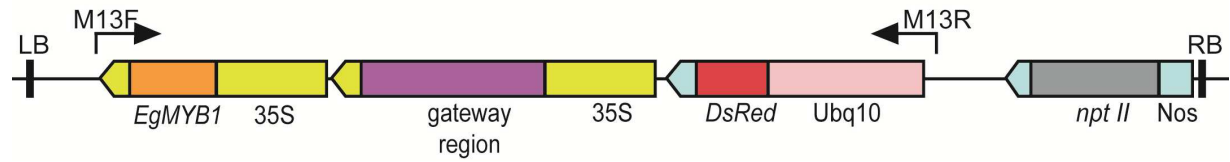


**Figure S4.5: Relative transcript abundance of transgenes in *Arabidopsis* transgenic plants. A.** Relative transcript level abundance in 35S:EgMYB1/35S:EgH1.3 measured in leaves. Lines selected for detailed characterization are highlighted in bold. **B.** Relative transcript level abundance in 35S:EgH1.3 measured in leaves. Lines selected for detailed characterization are highlighted in bold. **C.** Relative transcript level abundance measured in inflorescence stems in all the *Arabidopsis* transgenic lines characterized in detail in the present study.



**Figure S4.6. Relative transcript abundance of EgMYB1 and EgH1.3 transgenes in Eucalyptus transgenic hairy roots.** Selected plants for detailed characterization according to RT-qPCR results are highlighted in bold.

### pGWAY-MYB1



**Figure S4.7. Map of pGWAY-MYB1 destination vector.** LB: T-DNA left border; M13F and M13R: Hybridization sites for primers M13 forward and M13 reverse respectively; 35S: *CAMV 35S* promoter; gateway region: GATEWAY™ cassette (including attR1, CmR, ccdB, attR2 features); *DsRed*: *DsRed* fluorescent protein; Ubq10: *Arabidopsis thaliana* Ubiquitin-10 promoter; *npt II*: *neomycin phosphotransferase II*; Nos: *Nopaline synthase* promoter; RB: Right border. 35S and Nos terminators are shown as yellow and light blue arrows respectively.







CHAPTER 5:

---

## GENERAL DISCUSSION AND PERSPECTIVES





The hierarchical multi-level regulation network that regulates xylogenesis has been intensively studied during last years in the model plant *Arabidopsis*. *Arabidopsis* can undergo secondary growth in hypocotyls or under specific growth conditions (Chaffey *et al.*, 2002), and allowed many insights in our understanding of wood formation. However, some limitations and differences compared to trees demand proper investigations in woody perennial angiosperms. For instance, ray cells are missing in *Arabidopsis* hypocotyl xylem and differences between developmental stages (juvenile versus mature wood) and due to seasonal cycle can obviously not be addressed in *Arabidopsis*. Moreover, recent findings from our group pointed out other major differences between *Arabidopsis* and woody plants. Indeed, some members of the R2R3-MYB TF family highly expressed in the vascular cambium in *Eucalyptus* are not found in the *Arabidopsis* genome (Soler *et al.*, 2015).

The objective of my PhD was to get a better understanding of the regulation of secondary xylem in *Eucalyptus*, the most planted forest tree on a worldwide basis. Eucalypts wood is used to produce pulp, paper, timber or firewood, and its use as feedstock to bioethanol production, is promising. A better understanding of the molecular basis of wood formation in eucalypts should, for instance, highlight candidates to overcome SCW recalcitrance to degradation, a major obstacle both for pulp and bioethanol production, and consequently have a direct economic impact.

The sequence of the genome of *E. grandis* (Myburg *et al.*, 2014) together with the genome-wide characterization of gene families involved in lignin, cellulose and xylan biosynthesis (Carocha *et al.*, 2015; Myburg *et al.*, 2014) as well as transcription factor families such as MYB, NAC, ARF and AUX/IAA (Soler *et al.*, 2015; Hussey *et al.*, 2015; Yu *et al.*, 2014, 2015), pointed out many candidate genes putatively implicated in the regulation of SCW biosynthesis.

My PhD project focused on the MYB R2R3 family, which together with the NAC TF family contains key members shown to be involved in xylem differentiation and SCW biosynthesis in *Arabidopsis*.

In Chapter 2, we report the functional characterization of EgMYB137, a new and promising candidate for wood regulation (Soler *et al.*, 2015). We also report deeper insights in the roles of EgMYB1 and EgMYB2, which were previously studied in pioneer works in our team (Goicoechea *et al.*, 2005; Legay *et al.*, 2007; Hussey *et al.*, 2011; Yu *et al.*, 2015). Since the genetic transformation of *Eucalyptus* is long and tedious, we performed the functional characterization of these genes in another woody angiosperm, poplar. We overexpressed the dominant repression chimeric proteins of the three MYBs (EgMYB1-DR, EgMYB2-DR, and EgMYB137-DR). The dominant repression is a useful strategy that allows obtaining phenotypes by overcoming the functional redundancy of TFs (Hiratsu *et al.*, 2003). For each construct, we characterized 2 to 3 independent lines, which gave consistent results. The biochemical characterization of xylem SCW was performed using pyrolysis and wet chemistry for lignin (Klason and thioacidolysis), and we also analyzed the composition in phenylpropanoid-derived molecules such as phenolic glucosides and oligolignols.

Our results showed that the three MYB dominant repressors induced a reduction in lignin content as well as of the levels of oligolignols. EgMYB2-DR plants were the less affected and EgMYB137 exhibited the highest reduction. In contrast to studies where lignin biosynthetic genes were down-regulated (Meyermans *et al.*, 2000; Morreel *et al.*, 2004), we did not observe any accumulation of glucosylated phenolics, thought to act as storage and/or detoxification mechanism for specific phenylpropanoid intermediates accumulation. Thus, the dominant repression constructs of EgMYB1, EgMYB2 and EgMYB137 reduce the carbon flux directed to the whole lignin biosynthesis process and do not induce the accumulation of intermediate compounds.

The composition of lignin of EgMYB1-DR plants was similar to that of the EgMYB1-OE poplar plants, with no changes in the S/G ratio (Legay *et al.*, 2010) but showing a significant increase in H-lignin units, which was already found in the bark of poplar plants overexpressing EgMYB1 but not in their xylem (Legay, 2008). Accumulation of H-units has been proposed to be related with defence reactions against pathogens or in response to wounding (Sattler *et al.*, 2013 and references therein), suggesting that the overexpression of EgMYB1-DR might induce a constitutive stress response in these transgenic lines. This putative constitutive activation of the stress response seems more important than that occurring when overexpressing the native form of EgMYB1, which is already a strong repressor.

The saccharification yields obtained without pre-treatment were inversely correlated with the lignin content, in agreement with a key role of lignin as a major determinant of lignocellulosic recalcitrance, impairing cellulose accessibility to degrading enzymes (Chen and Dixon, 2007; Sykes *et al.*, 2015; Van Acker *et al.*, 2013). With an increase of two fold of the saccharification yield, EgMYB137-DR poplar lines showed higher yields than EgMYB1-DR lines, which exhibited also a notorious 65% yield increase. Together with the decrease in lignin content, EgMYB137-DR lines showed an increased S/G ratio, which can also be a favourable change to improve cellulose accessibility.

The saccharification yield without pre-treatment of EgMYB2-DR lines was not improved as compared to controls and was even lower when measured after alkali pre-treatment. These results support the role of EgMYB2, as its *Arabidopsis* orthologs (AtMYB46/AtMYB83), as a key regulator of the biosynthesis of the three main polymers of the SCW. Thus, not only lignin content, but also cellulose is probably decreased in EgMYB2-DR lines. The dosage of polysaccharides is in progress to confirm this assumption.

Since EgMYB137 was never characterized before, we had no *a priori* idea of its potential targets and we could not test its activity as transactivator or repressor using transient transactivation assays, as we did for EgMYB1 and 2. Moreover, because poplar plants overexpressing EgMYB137 died during the rooting and elongating process, we could only characterize EgMYB137-DR lines and, therefore, we cannot know if EgMYB137 is an activator turned into a repressor by the EAR motif or if it is already repressor. However, it is worth noting that overexpression of strong activators of lignin biosynthesis such as EgMYB2 (Legay *et al.*, 2008)

and PtMYB8 (Bomal *et al.*, 2008) could not be obtained in poplar and spruce, respectively. For EgMYB2, Legay *et al.* (2008) reported that after a series of propagation cycles during the *in vitro* multiplication phase, the poplar cuttings became gradually unable to develop their root system. The only poplars that made it through the *in vitro* stages were phenotypically identical to controls and did not express the transgene (Legay, 2008). Reasons for this inability to root had not been investigated, but overexpression of EgMYB2 was suggested to increase prematurely the thickness and lignin content of the walls of cells that should differentiate into a root system, preventing their proper normal differentiation. Another possibility is that an excess of phenolics in the cuttings might have interfered with the hormone signalling involved in adventitious rooting during the multiplication phase. In addition, it is worth noting that only half of the poplar lines obtained along this thesis harbouring the EgMYB2-DR and the EgMYB137-DR expressed the transgene, suggesting that silencing phenomena may have occurred for these genes.

Although it does not appear as a clear EgMYB137's ortholog, PtMYB8 and its orthologs in *Arabidopsis* (AtMYB61 and AtMYB26) are however phylogenetically related to EgMYB137. All are activators of the lignin pathway, suggesting that EgMYB137 could also be an activator of the lignin pathway. More investigations are needed to decipher the function of EgMYB137 as well as its position in the hierarchical network controlling SCW deposition, especially because preliminary phenotyping data of recently generated *Arabidopsis* plants either overexpressing the native form of EgMYB137 or the dominant repression chimeric protein suggest unexpectedly that EgMYB137 could act a repressor of SCW deposition.

Transcriptome profiling of EgMYB137-DR poplar and *Arabidopsis* EgMYB137-DR and EgMYB137-OE lines is in progress and should help deciphering direct and indirect target genes. In parallel, the use of an inducible promoter that would be triggered after the *in vitro* multiplication phase, or of a tissue-specific promoter preventing gene expression in cells that do not normally express the gene should also help with the identification of the target genes of the TF in a woody plant, using for instance chromatin immunoprecipitation. A similar strategy could be useful for EgMYB2.

However, these approaches would be more accurate if they could be performed in a homologous transformation system. *Eucalyptus* and poplar are relatively distant species, with only very fragmented microsynchronic regions (Gion and Jorge, personal communication). In addition, because the relatively recent whole genome duplication of the poplar genome, many genes appear as pairs whereas they are single genes in eucalypts. The main handicap is that transformation of *Eucalyptus* mediated by *A. tumefaciens* is long, tedious and with low efficiency. The need for a suitable, rapid and easy transformation system for eucalypts become also urgent since the genome sequence of *E. grandis* is available and many genes of unknown function deserve functional characterization. Overcoming this last bottleneck would be a breakthrough making *Eucalyptus* a model tree to study wood formation, and the discoveries could also have a direct impact on industrial processing of wood.

With this objective, we set up and optimized an *A. rhizogenes*-mediated transformation system for *E. grandis* (the specie whose genome has been sequenced), which is suitable for mid-throughput functional characterization of candidate genes including those involved in SCW formation. In chapter 3, in a publication accepted in “Plant biotechnology journal”, we describe protocols (i) to obtain composite plants (containing wild type aerial parts and co-transformed hairy roots), (ii) to harden these composite plants and also (iii) the conditions to cultivate co-transformed roots *in vitro*. We further show the usefulness of hairy roots to study pattern of promoters’ expression and subcellular localisation of proteins, but also to perform reverse genetic approaches. Since we are mostly interested in genes involved in wood formation, we examined as a proof of concept the expression pattern of the promoters of two monolignol biosynthetic genes (EgCAD2 and EgCCR1) and the functional characterisation of silencing the EgCCR1 gene. The results we obtained were comparable to previous experiments performed in heterologous systems, thus validating the usefulness of the method. The system is rapid, easy and functional, and has been already used to characterise EgMYB1 and the effect of its interaction with EgH1.3 (see chapter 4). However, some improvements would be useful, for instance to extract good quality RNA from the oldest parts of the roots. In addition, as co-transformed hairy roots can be maintained *in vitro*, it would be interesting to set up liquid hairy root cultures to obtain large amounts of biomass.

Finally, in chapter 4, we describe the role a protein-protein interaction involving EgMYB1 and a linker histone linker (EgH1.3) in xylem formation. To achieve this, we have used *Arabidopsis* as a heterologous transformation system, but we also took advantage of the homologous system consisting of *E. grandis* hairy roots described in chapter 3.

As in many other plant processes, protein-protein interactions are important in the regulation of xylogenesis, probably through modulating the activity of transcription factors to provide a tighter and more precise control of xylem cell differentiation. However, despite its importance, examples of interactions in xylogenesis are scarce (see chapter 1). With the aim of understanding how the activity of EgMYB1 is regulated by protein-protein interaction, a yeast-two hybrid library was constructed from *Eucalypts* xylem and screened with EgMYB1 as bait. This was performed in the team just before I started my PhD project.

In chapter 4, we demonstrate that EgMYB1 interacts specifically with EgH1.3 *in planta* and that this interaction increases the repressing effect of EgMYB1 over lignin deposition, both in *Arabidopsis* and in *E. grandis* hairy roots, while increasing the soluble phenolic metabolites. Taking into account the overlap between the expression profiles of EgMYB1 and EgH1.3, we propose a model in which the interaction between EgMYB1 and EgH1.3 provides a fine tuned mechanism that increases the repression activity of EgMYB1. By that mechanism, EgMYB1 and EgH1.3 would prevent a premature deposition of secondary cell walls in the cambial cells and in xylem cells at the first stages of xylogenesis, thus ensuring a metabolic homeostasis.

The mechanism by which EgH1.3 reinforces EgMYB1 repression effect is still to be determined. From studies performed in animals, we know that histone linkers can modify chromatin

condensation by methylating their surrounding DNA and/or by methylating or acetylating the Neighbor core histones. In plants, histones linkers have been shown to interact with demethylases and methylases. Taking into account this information, one possible mechanism to explain the enhanced repressive effect of EgMYB1 when interacting with the histone linker is that EgMYB1 can provide the sequence specificity to EgH1.3. In turn, EgH1.3 could interact with other proteins to modify the degree of DNA methylation, or to post-translationally modify core histones to further repress the transcription of EgMYB1 target genes. The analyses of the DNA methylation of the double transformants as well as the post-translational modification of histones would be helpful to verify this hypothesis. However, how the interaction between EgMYB1 and EgH1.3 is integrated with the repressing effect of EgMYB1 mediated by the EAR motif, which may modify the chromatin conformation by interacting with the co-repressor Topless, is an important question that would need to be addressed in the future

Another interesting aspect is that the interaction of EgMYB1 with EgH1.3 increases the amount of secondary metabolites related to biotic stress response. Since lignin itself has a role in defence against pathogens and its content was reduced, we have two possible scenarios. One is that the interaction causes an activation of the biosynthesis of defence-related compounds in detriment to lignin biosynthesis and thus, plants could be more resistant to pathogens although they have less lignin. Alternatively, plants can be weaker due to the strong reduction of lignin and thus more susceptible to pathogens. In this last case, the increase in defence-related secondary metabolites would be a side effect of the reduction of lignin. To test these hypotheses, the susceptibility of *Arabidopsis* plants to vascular and non-vascular pathogens is currently under investigation.

To better characterise the effects of the interaction, additional experiments using the homolog transformation systems of *Eucalyptus* hairy roots are required, notably at thicker regions of hairy roots with important production of secondary growth. Additionally, as EgH1.3 has been described as drought responsive, *E. grandis* composite plants with roots co-transformed with EgMYB1 or EgH1.3 promoters will be submitted to drought stress and the expression patterns of both promoters will be compared with control plants.

In conclusion, we have made significant progress in understanding the transcriptional regulation of wood formation in *Eucalyptus*, notably with the identification of a new R2R3-MYB regulator of lignin biosynthesis (EgMYB137) and the characterization of the protein-protein interaction between EgMYB1 and EgH1.3. Moreover, this is the first time that an interaction with a linker histone and a transcription factor is characterized in plants.

On an applied perspective, we have identified promising candidates to improve bioethanol production from eucalypts wood, with no significant biomass penalties.

Last but not least, we have set up a fast and efficient protocol for homologous transformation of *Eucalyptus* whose utility has already been demonstrated in our hands for functional

characterization of genes but also of protein-protein interactions. This homologous transformation system will be useful for the whole scientific community working on eucalypts.







---

## REFERENCES



- Agusti,J., Herold,S., Schwarz,M., Sanchez,P., Ljung,K., Dun,E.A., Brewer,P.B., Beveridge,C.A., Sieberer,T., Sehr,E.M., and Greb,T. (2011) Strigolactone signaling is required for auxin-dependent stimulation of secondary growth in plants. *Proc. Natl. Acad. Sci. U. S. A.*, **108**, 20242–7.
- Alejandro,S., Lee,Y., Tohge,T., Sudre,D., Osorio,S., Park,J., Bovet,L., Lee,Y., Geldner,N., Fernie,A.R., and Martinoia,E. (2012) AtABCG29 Is a Monolignol Transporter Involved in Lignin Biosynthesis. *Curr. Biol.*, **22**, 1207–1212.
- Aloni,R. (2007) Phytohormonal mechanisms that control wood quality formation in young and mature trees. *Compromised Wood Work. 2007*, 1–22.
- Alpizar, E., Dechamp, E., Espeout, S., Royer, M., Lecouls, A. C., Nicole, M., Bertrand, B., Lashermes, P. and Etienne, H. (2006) Efficient production of *Agrobacterium rhizogenes*-transformed roots and composite plants for studying gene expression in coffee roots. *Plant Cell Rep.* **25**, 959–967.
- Alves,A., Schwanninger,M., Pereira,H., and Rodrigues,J. (2006) Analytical pyrolysis as a direct method to determine the lignin content in wood: Part 1: Comparison of pyrolysis lignin with Klason lignin. *J. Anal. Appl. Pyrolysis*, **76**, 209–213.
- Arabidopsis* Interactome Mapping Consortium (2011) Evidence for network evolution in an *Arabidopsis* interactome map. *Science*, **333**, 601–607.
- Ascenzi,R. and Gantt,J.S. (1997) A drought-stress-inducible histone gene in *Arabidopsis thaliana* is a member of a distinct class of plant linker histone variants. *Plant Mol. Biol.*, **34**, 629–41.
- Ascenzi,R. and Gantt,J.S. (1999a) Molecular genetic analysis of the drought-inducible linker histone variant in *Arabidopsis thaliana*. *Plant Mol. Biol.*, **41**, 159–69.
- Ascenzi,R. and Gantt,J.S. (1999b) Subnuclear distribution of the entire complement of linker histone variants in *Arabidopsis thaliana*. *Chromosoma*, **108**, 345–55.
- Arvidsson, S., Kwasniewski, M., Riaño-Pachón, D.M. and Mueller-Roeber, B. (2008) QuantPrime—a flexible tool for reliable high-throughput primer design for quantitative PCR. *BMC Bioinformatics* **9**, 465.
- Baghdady, A., Blervacq, A.-S., Jouanin, L., Grima-Pettenati, J., Sivadon, P. and Hawkins, S. (2006) *Eucalyptus gunnii* CCR and CAD2 promoters are active in lignifying cells during primary and secondary xylem formation in *Arabidopsis thaliana*. *Plant Physiol. Biochem*, **44**, 674–683.
- Balasubramanian, A., Venkatachalam, R., Selvakesavan, K.R., Mary, A., Gherbi, H., Svistoonoff, S., Franche, C., Bogusz, D., Krishna Kumar, N. and Nambiar-Veetil, M. (2011) Optimisation of methods for *Agrobacterium rhizogenes* mediated generation of composite plants in *Eucalyptus camaldulensis*. *BMC Proc.* **5**, O45.

- Barnett, J.R. and Bonham, V. a (2004) Cellulose microfibril angle in the cell wall of wood fibres. *Biol. Rev. Camb. Philos. Soc.*, **79**, 461–472.
- Barros, J., Serk, H., Granlund, I., and Pesquet, E. (2015) The cell biology of lignification in higher plants. *Ann. Bot.*, **115**, 1053–1074.
- Baucher, M., Halpin, C., Petit-Conil, M., and Boerjan, W. (2003) Lignin: genetic engineering and impact on pulping.
- Baucher, M., El Jaziri, M., and Vandeputte, O. (2007) From primary to secondary growth: origin and development of the vascular system. *J. Exp. Bot.*, **58**, 3485–3501.
- Bécard, G. and Fortin, J.A. (1988) Early events of vesicular-arbuscular mycorrhiza formation on Ri T-DNA transformed roots. *New Phytol.* **108**, 211–218.
- Bedon, F., Grima-Pettenati, J., and Mackay, J. (2007) Conifer R2R3-MYB transcription factors: sequence analyses and gene expression in wood-forming tissues of white spruce (*Picea glauca*). *BMC Plant Biol.*, **7**, 17.
- Bennett, T., van den Toorn, a., Sanchez-Perez, G.F., Campilho, a., Willemsen, V., Snel, B., and Scheres, B. (2010) SOMBRERO, BEARSKIN1, and BEARSKIN2 Regulate Root Cap Maturation in *Arabidopsis*. *Plant Cell*, **22**, 640–654.
- Berthet, S., Demont-Caulet, N., Pollet, B., Bidzinski, P., Cezard, L., Le Bris, P., Borrega, N., Herve, J., Blondet, E., Balzergue, S., Lapierre, C., and Jouanin, L. (2011) Disruption of LACCASE4 and 17 Results in Tissue-Specific Alterations to Lignification of *Arabidopsis thaliana* Stems. *Plant Cell*, **23**, 1124–1137.
- Bevan, M., Lane, M. and Cb, C. (2010) Binary Agrobacterium vectors for plant transformation. *Nucleic Acids Res.*, **38** 8711–8721.
- Bhalerao, R.P. and Fischer, U. (2014) Auxin gradients across wood-instructive or incidental? *Physiol. Plant.*, **151**, 43–51.
- Bhargava, a., Mansfield, S.D., Hall, H.C., Douglas, C.J., and Ellis, B.E. (2010) MYB75 Functions in Regulation of Secondary Cell Wall Formation in the *Arabidopsis* Inflorescence Stem. *Plant Physiol.*, **154**, 1428–1438.
- Bhargava, A., Ahad, A., Wang, S., Mansfield, S.D., Haughn, G.W., Douglas, C.J., and Ellis, B.E. (2013) The interacting MYB75 and KNAT7 transcription factors modulate secondary cell wall deposition both in stems and seed coat in *Arabidopsis*. *Planta*, **237**, 1199–1211.
- Boerjan, W., Ralph, J., and Baucher, M. (2003) Lignin biosynthesis. *Annu. Rev. Plant Biol.*, **54**, 519–46.

- Boisson-Dernier, A, Chabaud, M., Garcia, F., Bécard, G., Rosenberg, C. and Barker, D.G. (2001) *Agrobacterium rhizogenes*-transformed roots of *Medicago truncatula* for the study of nitrogen-fixing and endomycorrhizal symbiotic associations. *Mol. Plant. Microbe. Interact.* **14**, 695–700.
- Bollhoner, B., Prestele, J., and Tuominen, H. (2012) Xylem cell death: emerging understanding of regulation and function. *J. Exp. Bot.*, **63**, 1081–1094.
- Bomal, C., Bedon, F., Caron, S., Mansfield, S.D., Levasseur, C., Cooke, J.E.K., Blais, S., Tremblay, L., Morency, M.-J., Pavy, N., Grima-Pettenati, J., Seguin, A., and MacKay, J. (2008) Involvement of *Pinus taeda* MYB1 and MYB8 in phenylpropanoid metabolism and secondary cell wall biogenesis: a comparative *in planta* analysis. *J. Exp. Bot.*, **59**, 3925–3939.
- 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.
- Bosselut, N., Van Ghelder, C., Claverie, M., Voisin, R., Onesto, J.P., Rosso, M.N. and Esmenjaud, D. (2011) *Agrobacterium rhizogenes*-mediated transformation of *Prunus* as an alternative for gene functional analysis in hairy-roots and composite plants. *Plant Cell Rep.* **30**, 1313–1326.
- Boudet, A.M., Kajita, S., Grima-Pettenati, J., and Goffner, D. (2003) Lignins and lignocellulosics: A better control of synthesis for new and improved uses. *Trends Plant Sci.*, **8**, 576–581.
- Bragatto, J., Segato, F., Cota, J., Mello, D.B., Oliveira, M.M., Buckeridge, M.S., Squina, F.M., and Driemeier, C. (2012) Insights on how the activity of an endoglucanase is affected by physical properties of insoluble celluloses. *J. Phys. Chem. B*, **116**, 6128–6136.
- Bray, E.A., Shih, T., Moses, M.S., Cohen, A., Imai, R., and Plant, L. (1999) Water-deficit induction of a tomato H1 histone requires abscisic acid. *Plant Growth Regul.*, **29**, 35–46.
- Brown, D.M., Zeef, L. a H., Ellis, J., Goodacre, R., and Turner, S.R. (2005) Identification of Novel Genes in *Arabidopsis* Involved in Secondary Cell Wall Formation Using Expression Profiling and Reverse Genetics. *Plant Cell*, **17**, 2281–2295.
- Brunner, A.M., Busov, V.B., and Strauss, S.H. (2004) Poplar genome sequence: functional genomics in an ecologically dominant plant species. *Trends Plant Sci.*, **9**, 49–56.
- Burton, R. a, Gidley, M.J., and Fincher, G.B. (2010) Heterogeneity in the chemistry, structure and function of plant cell walls. *Nat. Chem. Biol.*, **6**, 724–732.

- Camargo,E.L.O., Nascimento,L.C., Soler,M., Salazar,M.M., Lepikson-Neto,J., Marques,W.L., Alves,A., Teixeira,P.J.P.L., Mieczkowski,P., Carazzolle,M.F., Martinez,Y., Deckmann,A.C., Rodrigues,J.C., Grima-Pettenati,J., and Pereira,G.A.G. (2014) Contrasting nitrogen fertilization treatments impact xylem gene expression and secondary cell wall lignification in *Eucalyptus*. *BMC Plant Biol.*, **14**, 256.
- Canonne, J., Marino, D., Jauneau, A., Pouzet, C., Brière, C., Roby, D., and Rivas, S. (2011). The Xanthomonas type III effector XopD targets the *Arabidopsis* transcription factor MYB30 to suppress plant defense. *Plant Cell*. **23**: 3498-511.
- Carocha,V., Soler,M., Hefer,C., Cassan-Wang,H., Fevereiro,P., Myburg,A.A., Paiva,J.A.P., and Grima-Pettenati,J. (2015) Genome-wide analysis of the lignin toolbox of *Eucalyptus grandis*. *New Phytol.*, **206**, 1297–1313.
- Cassan-Wang,H., Goué,N., Saidi,M.N., Legay,S., Sivadon,P., Goffner,D., and Grima-Pettenati,J. (2013) Identification of novel transcription factors regulating secondary cell wall formation in *Arabidopsis*. *Front. Plant Sci.*, **4**, 1–14.
- Cassan-Wang,H., Soler,M., Yu,H., Camargo,E.L.O., Carocha,V., Ladouce,N., Savelli,B., Paiva,J. a P., Leplé,J.-C., and Grima-Pettenati,J. (2012) Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of *Eucalyptus* grown in various environmental conditions. *Plant Cell Physiol.*, **53**, 2101–16.
- Causier,B., Ashworth,M., Guo,W., and Davies,B. (2012) The TOPLESS Interactome: A Framework for Gene Repression in *Arabidopsis*. *PLANT Physiol.*, **158**, 423–438.
- Chabaud, M., Boisson-dernier, A., Zhang, J., Taylor, C.G., Yu, O. and Barker, D.G. (2006) *Agrobacterium rhizogenes*-mediated root transformation. In, Mathesius, U., Journet, EP., Sumner, LW. (eds), *The Medicago truncatula handbook*. The Samuel Roberts Noble Foundation, Ardmore, PA, USA, pp. 1–8.
- Chafe, S.C, and Chauret, G. (1974). Cell wall structure in the xylem parenchyma of trembling aspen. *Protoplasma*. **80**: 129-147.
- Chaffey,N., Cholewa,E., Regan,S., and Sundberg,B. (2002) Secondary xylem development in *Arabidopsis*: a model for wood formation. *Physiol. Plant.*, **114**, 594–600.
- Chandra, S. (2012) Natural plant genetic engineer *Agrobacterium rhizogenes*: Role of T-DNA in plant secondary metabolism. *Biotechnol. Lett.* **34**, 407–415.
- Chen,F. and Dixon,R. a (2007) Lignin modification improves fermentable sugar yields for biofuel production. *Nat. Biotechnol.*, **25**, 759–761.

- Chilton, M.-D., Tepfer, D. a., Petit, A., David, C., Casse-Delbart, F. and Tempé, J. (1982) *Agrobacterium rhizogenes* inserts T-DNA into the genomes of the host plant root cells. *Nature*, **295**, 432–434.
- Christey, M.C. (2001) Use of ri-mediated transformation for production of transgenic plants. *Vitr. Cell. Dev. Biol. - Plant*, **37**, 687–700.
- Clough,S.J. and Bent, a F. (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.*, **16**, 735–43.
- Cosgrove,D.J. (2005) Growth of the plant cell wall. *Nat. Rev. Mol. Cell Biol.*, **6**, 850–61.
- Creux, N.M., Bossinger, G., Myburg, A. a. and Spokevicius, A. V. (2013) Induced somatic sector analysis of cellulose synthase (CesA) promoter regions in woody stem tissues. *Planta*, **237**, 799–812.
- Creux, N.M., Ranik, M., Berger, D.K. and Myburg, A.A. (2008) Comparative analysis of orthologous cellulose synthase promoters from *Arabidopsis*, *Populus* and *Eucalyptus*: Evidence of conserved regulatory elements in angiosperms. *New Phytol.* **179**, 722–737.
- Damiani,I., Morreel,K., Danoun,S., Goeminne,G., Yahiaoui,N., Marque,C., Kopka,J., Messens,E., Goffner,D., Boerjan,W., Boudet,A.M., and Rochange,S. (2005) Metabolite profiling reveals a role for atypical cinnamyl alcohol dehydrogenase CAD1 in the synthesis of coniferyl alcohol in tobacco xylem. *Plant Mol. Biol.*, **59**, 753–769.
- Dauwe, R., Morreel, K., Goeminne, G., Gielen, B., Rohde, A., Van Beeumen, J., Ralph, J., Boudet, A.M., Kopka, J., Rochange, S.F., Halpin, C., Messens, E. and Boerjan, W. (2007) Molecular phenotyping of lignin-modified tobacco reveals associated changes in cell-wall metabolism, primary metabolism, stress metabolism and photorespiration. *Plant J.* **52**, 263–285.
- De la Torre, F., Rodríguez, R., Jorge, G., Villar, B., Álvarez-Otero, R., Grima-Pettenati, J. and Gallego, P.P. (2014) Genetic transformation of *Eucalyptus globulus* using the vascular-specific *EgCCR* as an alternative to the constitutive CaMV35S promoter. *Plant Cell. Tissue Organ Cult.* **117**, 77–84.
- Déjardin,A., Laurans,F., Arnaud,D., Breton,C., Pilate,G., and Leplé,J.-C. (2010) Wood formation in Angiosperms. *C. R. Biol.*, **333**, 325–334.
- Didi,V., Jackson,P., and Hejatko,J. (2015) Hormonal regulation of secondary cell wall formation. *J. Exp. Bot.*, **66**, 5015–5027.

- Diouf, D., Gherbi, H., Prin, Y., Franche, C., Duhoux, E. and Bogusz, D. (1995) Hairy root nodulation of *Casuarina glauca*: a system for the study of symbiotic gene expression in an actinorhizal tree. *Mol. Plant. Microbe. Interact.* **8**, 532–7.
- Doblin, M.S., Pettolino, F., and Bacic, A. (2010) Evans Review: Plant cell walls: The skeleton of the plant world. *Funct. Plant Biol.*, **37**, 357–381.
- Donaldson, L.A., Nanayakkara, B., Radotić, K., Djikanovic-Golubović, D., Mitrović, A., Bogdanović Pristov, J., Simonović Radosavljević, J., and Kalauzi, A. (2015). Xylem parenchyma cell walls lack a gravitropic response in conifer compression wood. *Planta*. **242**: 1413-1424.
- Donner, T.J., Sherr, I., and Scarpella, E. (2010) Auxin signal transduction in *Arabidopsis* vein formation. *Plant Signal. Behav.*, **5**, 70–72.
- Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C., and Lepiniec, L. (2010) MYB transcription factors in *Arabidopsis*. *Trends Plant Sci.*, **15**, 573–81.
- Endo, H., Yamaguchi, M., Tamura, T., Nakano, Y., Nishikubo, N., Yoneda, a., Kato, K., Kubo, M., Kajita, S., Katayama, Y., Ohtani, M., and Demura, T. (2015) Multiple Classes of Transcription Factors Regulate the Expression of VASCULAR-RELATED NAC-DOMAIN7, a Master Switch of Xylem Vessel Differentiation. *Plant Cell Physiol.*, **56**, 242–254.
- Esau, K. (1965) *Plant anatomy* Second edi. John Wiley and Sons Inc, New York, NY.
- Estrada-Navarrete, G., Alvarado-Affantranger, X., Olivares, J.-E., Guillén, G., Díaz-Camino, C., Campos, F., Quinto, C., Gresshoff, P.M. and Sanchez, F. (2007) Fast, efficient and reproducible genetic transformation of *Phaseolus* spp. by *Agrobacterium rhizogenes*. *Nat. Protoc.* **2**, 1819–1824.
- Feuillet, C., Lauvergeat, V., Deswarte, C., Pilate, G., Boudet, a, and Grima-Pettenati, J. (1995) Tissue- and cell-specific expression of a cinnamyl alcohol dehydrogenase promoter in transgenic poplar plants. *Plant Mol. Biol.*, **27**, 651–67.
- Fliegmann, J., Canova, S., Lachaud, C., Uhlenbroich, S., Gascioli, V., Pichereaux, C., Rossignol, M., Rosenberg, C., Cumener, M., Pitorre, D., Lefebvre, B., Gough, C., Samain, E., Fort, S., Driguez, H., Vauzeilles, B., Beau, J.M., Nurisso, A., Imbert, A., *et al.* (2013) Lipochitooligosaccharidic symbiotic signals are recognized by LysM receptor-like kinase LYR3 in the legume *Medicago truncatula*. *ACS Chem. Biol.* **8**, 1900–1906.
- Fornalé, S., Shi, X., Chai, C., Encina, A., Irar, S., Capellades, M., Fuguet, E., Torres, J.L., Rovira, P., Puigdomènech, P., Rigau, J., Grotewold, E., Gray, J., and Caparrós-Ruiz, D. (2010) ZmMYB31 directly represses maize lignin genes and redirects the phenylpropanoid metabolic flux. *Plant J.*, **64**, 633–644.



- Fornalé,S., Sonbol,F.-M., Maes,T., Capellades,M., Puigdomènech,P., Rigau,J., and 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.
- Foucart, C., Jauneau, A., Gion, J.-M., Amelot, N., Martinez, Y., Panegos, P., Grima-Pettenati, J. and Sivadon, P. (2009) Overexpression of *EgROP1*, a *Eucalyptus* vascular-expressed Rac-like small GTPase, affects secondary xylem formation in *Arabidopsis thaliana*. *New Phytol.* **183**, 1014–29.
- Franche,C., Diouf,D., Le,Q.V., Bogusz,D., N’Diaye,A., Gherbi,H., Gobe,C., and Duhoux,E. (1997) Genetic transformation of the actinorhizal tree *Allocasuarina verticillata* by *Agrobacterium tumefaciens*. *Plant J.*, **11**, 897–904.
- Froidure,S., Canonne,J., Daniel,X., Jauneau,A., Brière,C., Roby,D., and Rivas,S. (2010) AtsPLA2- $\alpha$  nuclear relocalization by the *Arabidopsis* transcription factor AtMYB30 leads to repression of the plant defense response. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 15281–6.
- Fukuda,H. (2004) Signals that control plant vascular cell differentiation. *Nat. Rev. Mol. Cell Biol.*, **5**, 379–391.
- Funk,V., Kositsup,B., Zhao,C., and 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.
- Gago,J., Grima-Pettenati,J., and Gallego,P.P. (2011) Vascular-specific expression of GUS and GFP reporter genes in transgenic grapevine (*Vitis vinifera* L. cv. Albariño) conferred by the *EgCCR* promoter of *Eucalyptus gunnii*. *Plant Physiol. Biochem.*, **49**, 413–419.
- Gallardo,F., Fu,J., Canton,F., Garcia-Gutierrez,A., Canovas,F., and Kirby,E. (1999) Expression of a conifer glutamine synthetase gene in transgenic poplar. *Planta*, **210**, 19–26.
- Georgiev, M.I., Agostini, E., Ludwig-Müller, J. and Xu, J. (2012) Genetically transformed roots: From plant disease to biotechnological resource. *Trends Biotechnol.* **30**, 528–537.
- Gerber, L., Eliasson, M., Trygg, J., Moritz, T., and Sundberg, B. (2012) Multivariate curve resolution provides a high-throughput data processing pipeline for pyrolysis-gas chromatography/mass spectrometry. *J. Anal. Appl. Pyrol.* **95**, 95–100.
- Gherbi, H., Nambiar-Veetil, M., Zhong, C., Félix, J., Autran, D., Girardin, R., Vaissayre, V., Auguy, F., Bogusz, D. and Franche, C. (2008) Post-transcriptional gene silencing in the root system of the actinorhizal tree *Allocasuarina verticillata*. *Mol. Plant. Microbe. Interact.* **21**, 518–524.

- Girijashankar, V. (2011) Genetic transformation of *Eucalyptus*. *Physiol. Mol. Biol. Plants*, **17**, 9–23.
- Goicoechea, M., Lacombe, E., Legay, S., Mihaljevic, S., Rech, P., Jauneau, A., Lapierre, C., Pollet, B., Verhaegen, D., Chaubet-Gigot, N., and Grima-Pettenati, J. (2005) EgMYB2, a new transcriptional activator from *Eucalyptus* xylem, regulates secondary cell wall formation and lignin biosynthesis. *Plant J.*, **43**, 553–67.
- Gonzalez, A., Zhao, M., Leavitt, J.M., and Lloyd, A.M. (2008) Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in *Arabidopsis* seedlings. *Plant J.*, **53**, 814–827.
- Goujon, T., Ferret, V., Mila, I., Pollet, B., Ruel, K., Burlat, V., Joseleau, J.-P., Barrière, Y., Lapierre, C. and Jouanin, L. (2003) Down-regulation of the AtCCR1 gene in *Arabidopsis thaliana*: effects on phenotype, lignins and cell wall degradability. *Planta*, **217**, 218–228.
- Grima-Pettenati, J., Feuillet, C., Goffner, D., Borderies, G., and Boudet, A. (1993) Molecular cloning and expression of a *Eucalyptus gunnii* cDNA clone encoding cinnamyl alcohol dehydrogenase. *Plant Mol. Biol.*, **21**, 1085–1095.
- Grima-Pettenati, J., Soler, M., Camargo, E.L.O., and Wang, H. (2012) Transcriptional Regulation of the Lignin Biosynthetic Pathway Revisited: New Players and Insights. In: Jacquot, J.-P. and Gadgil, P. (eds), *Lignins: Biosynthesis, Biodegradation and Bioengineering*. Elsevier, pp. 173–218.
- Guillon, S., Trémouillaux-Guiller, J., Pati, P.K., Rideau, M. and Gantet, P. (2006) Hairy root research: recent scenario and exciting prospects. *Curr. Opin. Plant Biol.* **9**, 341–346.
- Hansen, J., Jørgensen, J.E., Stougaard, J. and Marcker, K.A. (1989) Hairy roots - a short cut to transgenic root nodules. *Plant Cell Rep.* **8**, 12–15.
- 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.
- Hertzberg, M., Aspeborg, H., Schrader, J., Andersson, a, Erlandsson, R., Blomqvist, K., Bhalerao, R., Uhlén, M., Teeri, T.T., Lundberg, J., Sundberg, B., Nilsson, P., and Sandberg, G. (2001) A transcriptional roadmap to wood formation. *Proc. Natl. Acad. Sci. U. S. A.*, **98**, 14732–7.
- Hiratsu, K., Matsui, K., Koyama, T., and Ohme-Takagi, M. (2003) Dominant repression of target genes by chimeric repressors that include the EAR motif, a repression domain, in *Arabidopsis*. *Plant J.*, **34**, 733–9.

- Holladay, J.E., White, J.F., Bozell, J.J., and Johnson, D. (2007) Top Value-Added Chemicals from Biomass. Volume II-Results of Screening for Potential Candidates from Biorefinery Lignin Richland, WA.
- Huis, R., Morreel, K., Fliniaux, O., Lucau-Danila, a., Fenart, S., Grec, S., Neutelings, G., Chabbert, B., Mesnard, F., Boerjan, W., and Hawkins, S. (2012) Natural Hypolignification Is Associated with Extensive Oligolignol Accumulation in Flax Stems. *Plant Physiol.*, **158**, 1893–1915.
- Hussey, S.G., Mizrachi, E., Creux, N.M., and Myburg, A.A. (2013) Navigating the transcriptional roadmap regulating plant secondary cell wall deposition. *Front. Plant Sci.*, **4**, 325.
- Hussey, S.G., Mizrachi, E., Spokevicius, A. V, Bossinger, G., Berger, D.K., and Myburg, A.A. (2011) SND2, a NAC transcription factor gene, regulates genes involved in secondary cell wall development in *Arabidopsis* fibres and increases fibre cell area in *Eucalyptus*. *BMC Plant Biol.*, **11**, 173.
- Hussey, S.G., Saïdi, M.N., Hefer, C.A., Myburg, A.A., and Grima-Pettenati, J. (2015) Structural, evolutionary and functional analysis of the NAC domain protein family in *Eucalyptus*. *New Phytol.*, **206**, 1337–1350.
- leamkhang, S. and Chatchawankanphanich, O. (2005) Augmentin<sup>®</sup> as an alternative antibiotic for growth suppression of *Agrobacterium* for tomato (*Lycopersicon esculentum*) transformation. *Plant Cell. Tissue Organ Cult.* **82**, 213–220.
- Imanishi, L., Vayssières, A., Franche, C., Bogusz, D., Wall, L. and Svistoonoff, S. (2011) Transformed Hairy Roots of the actinorhizal shrub *Discaria trinervis*: a valuable tool for studying actinorhizal symbiosis in the context of intercellular infection. *BMC Proc.* **5**, P85.
- Jefferson, R.A. (1987) Assaying chimeric genes in plants: The GUS gene fusion system. *Plant Mol. Biol. Report.*, **5**, 387–405.
- Jerzmanowski, A., Przewłoka, M., and Grasser, K.D. (2000) Linker Histones and HMG1 Proteins of Higher Plants. *Plant Biol.*, **2**, 586–597.
- Jin, H., 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.
- Jones, L., Ennos, a R. and Turner, S.R. (2001) Cloning and characterization of irregular xylem4 (irx4): A severely lignin-deficient mutant of *Arabidopsis*. *Plant J.* **26**, 205–216.
- Jouanin, L., Tourneur, J., Tourneur, C. and Casse-Delbart, F. (1986) Restriction maps and homologies of the three plasmids of *Agrobacterium rhizogenes* strain A4. *Plasmid*, **16**, 124–134.

- Kagale,S. and Rozwadowski,K. (2011) EAR motif-mediated transcriptional repression in plants: An underlying mechanism for epigenetic regulation of gene expression. *Epigenetics*, **6**, 141–146.
- Karimi,M., Inzé,D., and Depicker,A. (2002) GATEWAY™ vectors for Agrobacterium-mediated plant transformation. *Trends Plant Sci.*, **7**, 193–195.
- Katoh,K., Misawa,K., Kuma,K., and Miyata,T. (2002) MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.*, **30**, 3059–3066.
- Kazan,K. (2006) Negative regulation of defence and stress genes by EAR-motif-containing repressors. *Trends Plant Sci.*, **11**, 109–112.
- Kim,J.I., Dolan,W.L., Anderson,N. a., and Chapple,C. (2015) Indole Glucosinolate Biosynthesis Limits Phenylpropanoid Accumulation in *Arabidopsis thaliana*. *Plant Cell*, **27**, 1529–46.
- Ko,J.-H., Jeon,H.-W., Kim,W.-C., Kim,J.-Y., and Han,K.-H. (2014) The MYB46/MYB83-mediated transcriptional regulatory programme is a gatekeeper of secondary wall biosynthesis. *Ann. Bot.*, **114**, 1099–1107.
- 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. *M G G Mol. Gen. Genet.*, **204**, 383–396.
- Kubo,M., Udagawa,M., Nishikubo,N., Horiguchi,G., Yamaguchi,M., Ito,J., Mimura,T., Fukuda,H., and Demura,T. (2005) Transcription switches for protoxylem and metaxylem vessel formation. *Genes Dev.*, **19**, 1855–60.
- Kumagai, H. and Kouchi, H. (2003) Gene silencing by expression of hairpin RNA in *Lotus japonicus* roots and root nodules. *Mol. Plant. Microbe. Interact.* **16**, 663–8.
- Lacombe,E., Van Doorselaere,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**, 663–76.
- Lacombe,E., Hawkins,S., Van Doorselaere,J., Piquemal,J., Goffner,D., Poeydomenge,O., Boudet, a M., and Grima-Pettenati,J. (1997) Cinnamoyl CoA reductase, the first committed enzyme of the lignin branch biosynthetic pathway: cloning, expression and phylogenetic relationships. *Plant J.*, **11**, 429–441.
- Lakhal,W. (2013) Etude fonctionnelle de trois facteurs de transcription impliqués dans la formation de la paroi secondaire chez le peuplier.
- Lapierre,C. (1993) Application of new methods for the investigation of lignin structure. In, ASA-CSSA-SSSA (ed), *Forage cell wall structure and digestibility*. Madison (USA), pp. 133–166.

- Lauvergeat,V., Lacomme,C., Lacombe,E., Lasserre,E., Roby,D., and Grima-Pettenati,J. (2001) Two cinnamoyl-CoA reductase (CCR) genes from *Arabidopsis thaliana* are differentially expressed during development and in response to infection with pathogenic bacteria. *Phytochemistry*, **57**, 1187–1195.
- 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.
- Legay,S. (2008) Caractérisation fonctionnelle de deux facteurs de transcription MYB R2R3: Rôle dans la formation du bois chez les Angiospermes.
- Legay,S., Lacombe,E., Goicoechea,M., Brière,C., Séguin,A., Mackay,J., and Grima-Pettenati,J. (2007) Molecular characterization of EgMYB1, a putative transcriptional repressor of the lignin biosynthetic pathway. *Plant Sci.*, **173**, 542–549.
- Legay,S., Sivadon,P., Blervacq,A.-S., Pavy,N., Baghdady,A., Tremblay,L., Levasseur,C., Ladouce,N., Lapierre,C., Séguin,A., Hawkins,S., Mackay,J., and Grima-Pettenati,J. (2010) EgMYB1, an R2R3 MYB transcription factor from *Eucalyptus* negatively regulates secondary cell wall formation in *Arabidopsis* and poplar. *New Phytol.*, **188**, 774–86.
- Lens,F., Smets,E., and Melzer,S. (2012) Stem anatomy supports *Arabidopsis thaliana* as a model for insular woodiness. *New Phytol.*, **193**, 12–17.
- Lepikson-Neto,J., Nascimento,L.C., Salazar,M.M., Camargo,E. Lo, Cairo,J.P., Teixeira,P.J., Marques,W.L., Squina,F.M., Mieczkowski,P., Deckmann,A.C., and Pereira,G.A. (2014) Flavonoid supplementation affects the expression of genes involved in cell wall formation and lignification metabolism and increases sugar content and saccharification in the fast-growing *Eucalyptus* hybrid *E. urophylla* x *E. grandis*. *BMC Plant Biol.*, **14**, 301.
- Léplé, J.-C., Dauwe, R., Morreel, K., Storme, V., Lapierre, C., Pollet, B., Naumann, A., Kang, K.-Y., Kim, H., Ruel, K., Lefèbvre, A., Joseleau, J.-P., Grima-Pettenati, J., De Rycke, R., Andersson-Gunnerås, S., Erban, A., Fehrle, I., Petit-Conil, M., Kopka, J., *et al.* (2007) Downregulation of *cinnamoyl-coenzyme A reductase* in poplar: multiple-level phenotyping reveals effects on cell wall polymer metabolism and structure. *Plant Cell*, **19**, 3669–3691.
- Léplé,J.-C., Grima-Pettenati,J., Van Montagu,M., and Boerjan,W. (1998) A cDNA Encoding Cinnamoyl-CoA Reductase from *Populus trichocarpa* (Accession no. AJ224986) (PGR 98-121). *Plant Physiol.*, **117**, 1126.
- Lev-Yadun (1994) Induction of sclereid differentiation in the pith of *Arabidopsis thaliana*. *J. Exp. Bot.*, **45**, 1845–1849.

- Li,E., Bhargava,A., Qiang,W., Friedmann,M.C., Forneris,N., Savidge,R. a, Johnson,L. a, Mansfield,S.D., Ellis,B.E., and Douglas,C.J. (2012) The Class II KNOX gene KNAT7 negatively regulates secondary wall formation in *Arabidopsis* and is functionally conserved in *Populus*. *New Phytol.*, **194**, 102–115.
- Li,E., Wang,S., Liu,Y., Chen,J.-G., and Douglas,C.J. (2011) OVATE FAMILY PROTEIN4 (OFP4) interaction with KNAT7 regulates secondary cell wall formation in *Arabidopsis thaliana*. *Plant J.*, **67**, 328–41.
- Limpens, E., Ramos, J., Franken, C., Raz, V., Compaan, B., Franssen, H., Bisseling, T. and Geurts, R. (2004) RNA interference in *Agrobacterium rhizogenes*-transformed roots of *Arabidopsis* and *Medicago truncatula*. *J. Exp. Bot.* **55**, 983–992.
- Liu,Y. and Douglas,C.J. (2015) A role for OVATE FAMILY PROTEIN1 (OFP1) and OFP4 in a BLH6-KNAT7 multi-protein complex regulating secondary cell wall formation in *Arabidopsis thaliana*. *Plant Signal. Behav.*, **10**, e1033126.
- Liu,C.J., Miao,Y.C., and Zhang,K.W. (2011) Sequestration and transport of lignin monomeric precursors. *Molecules*, **16**, 710–727.
- Liu,Y., You,S., Taylor-Teeples,M., Li,W.L., Schuetz,M., Brady,S.M., and Douglas,C.J. (2014) BEL1-LIKE HOMEODOMAIN6 and KNOTTED *ARABIDOPSIS THALIANA*7 interact and regulate secondary cell wall formation via repression of REVOLUTA. *Plant Cell*, **26**, 4843–61.
- Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>(-Delta Delta C(T))</sup> Method. *Methods*, **25**, 402–8.
- Lu,X., Wontakal,S.N., Emelyanov,A. V., Morcillo,P., Konev,A.Y., Fyodorov,D. V., and Skoultchi,A.I. (2009) Linker histone H1 is essential for *Drosophila* development, the establishment of pericentric heterochromatin, and a normal polytene chromosome structure. *Genes Dev.*, **23**, 452–465.
- Lu,X., Wontakal,S.N., Kavi,H., Kim,B.J., Guzzardo,P.M., Emelyanov,A. V, Xu,N., Hannon,G.J., Zavadil,J., Fyodorov,D. V, and Skoultchi,A.I. (2013) *Drosophila* H1 Regulates the Genetic Activity of Heterochromatin by Recruitment of Su(var)3-9. *Science (80-. )*, **340**, 78–81.
- Ma,Q.-H., Wang,C., and Zhu,H.-H. (2011) TaMYB4 cloned from wheat regulates lignin biosynthesis through negatively controlling the transcripts of both cinnamyl alcohol dehydrogenase and cinnamoyl-CoA reductase genes. *Biochimie*, **93**, 1179–86.
- Matsui,K., Umemura,Y., and Ohme-Takagi,M. (2008) AtMYBL2, a protein with a single MYB domain, acts as a negative regulator of anthocyanin biosynthesis in *Arabidopsis*. *Plant J.*, **55**, 954–967.

- Mauriat, M., Le, G., Philipp, R., Delzon, S., Bréda, N., Clair, B., Coutand, C., Domec, J.-C., Thierry, F., Grima-Pettenati, J., Herrera, R., Leplé, J.-C., Richet, N., Trontin, J.-F., and Plomion, C. (2014) Wood Formation in Trees. In, *Tree Biotechnology*. CRC Press, pp. 1513–1523.
- McBryant, S.J., Lu, X., and Hansen, J.C. (2010). Multifunctionality of the linker histones: an emerging role for protein-protein interactions. *Cell Res.* **20**: 519-5128.
- McCarthy, R.L., Zhong, R., Fowler, S., Lyskowski, D., Piyasena, H., Carleton, K., Spicer, C., and Ye, Z.H. (2010) The poplar MYB transcription factors, PtrMYB3 and PtrMYB20, are involved in the regulation of secondary wall biosynthesis. *Plant Cell Physiol.*, **51**, 1084–1090.
- McCarthy, R.L., Zhong, R., and Ye, Z.H. (2009) MYB83 is a direct target of SND1 and acts redundantly with MYB46 in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. *Plant Cell Physiol.*, **50**, 1950–1964.
- MacRae, S. and Van Staden, J. (1993) *Agrobacterium rhizogenes*-mediated transformation to improve rooting ability of eucalypts. *Tree Physiol.* **12**, 411–8.
- Méchin, V., Laluc, A., Legée, F., Cézard, L., Denoue, D., Barrière, Y., and Lapierre, C. (2014) Impact of the Brown-Midrib bm5 Mutation on Maize Lignins. *J. Agric. Food Chem.*, **62**, 5102–7.
- Melzer, S., Lens, F., Gennen, J., Vanneste, S., Rohde, A., and Beeckman, T. (2008) Flowering-time genes modulate meristem determinacy and growth form in *Arabidopsis thaliana*. *Nat. Genet.*, **40**, 1489–1492.
- Meyermans, H., Morreel, K., Lapierre, C., Pollet, B., De Bruyn, A., Busson, R., Herdewijn, P., Devreese, B., Van Beeumen, J., Marita, J.M., Ralph, J., Chen, C., Burggraeve, B., Van Montagu, M., Messens, E., and Boerjan, W. (2000) Modifications in lignin and accumulation of phenolic glucosides in poplar xylem upon down-regulation of caffeoyl-coenzyme A O-methyltransferase, an enzyme involved in lignin biosynthesis. *J. Biol. Chem.*, **275**, 36899–36909.
- Mir Derikvand, M., Sierra, J.B., Ruel, K., Pollet, B., Do, C.T., Thévenin, J., Buffard, D., Jouanin, L. and Lapierre, C. (2008) Redirection of the phenylpropanoid pathway to feruloyl malate in *Arabidopsis* mutants deficient for *cinnamoyl-CoA reductase 1*. *Planta*, **227**, 943–956.
- Mitsuda, N., Iwase, A., Yamamoto, H., Yoshida, M., Seki, M., Shinozaki, K., and 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*, **19**, 270–80.
- Mitsuda, N., Seki, M., Shinozaki, K., and Ohme-Takagi, M. (2005) The NAC transcription factors NST1 and NST2 of *Arabidopsis* regulate secondary wall thickenings and are required for anther dehiscence. *Plant Cell*, **17**, 2993–3006.

- Miyashima,S., Sebastian,J., Lee,J.-Y., and Helariutta,Y. (2012) Stem cell function during plant vascular development. *EMBO J.*, **32**, 178–193.
- Moore, L., Warren, G. and Strobel, G. (1979) Involvement of a plasmid in the hairy root disease of plants caused by *Agrobacterium rhizogenes*. *Plasmid*, **2**, 617–626.
- Morreel,K., Dima,O., Kim,H., Lu,F., Niculaes,C., Vanholme,R., Dauwe,R., Goeminne,G., Inzé,D., Messens,E., Ralph,J., and Boerjan,W. (2010) Mass spectrometry-based sequencing of lignin oligomers. *Plant Physiol.*, **153**, 1464–1478.
- Morreel,K., Ralph,J., Kim,H., Lu,F., Goeminne,G., Ralph,S., Messens,E., and Boerjan,W. (2004) Profiling of oligolignols reveals monolignol coupling conditions in lignifying poplar xylem. *Plant Physiol.*, **136**, 3537–3549.
- Muoki, R.C., Paul, A., Kumari, A., Singh, K. and Kumar, S. (2012) An improved protocol for the isolation of RNA from roots of tea (*Camellia sinensis* (L.) O. Kuntze). *Mol. Biotechnol.* **52**, 82–88.
- Myburg,A.A., Grattapaglia,D., Tuskan,G.A., Hellsten,U., Hayes,R.D., Grimwood,J., Jenkins,J., Lindquist,E., Tice,H., Bauer,D., Goodstein,D.M., Dubchak,I., Poliakov,A., Mizrachi,E., Kullam,A.R.K., Hussey,S.G., Pinard,D., van der Merwe,K., Singh,P., *et al.* (2014) The genome of *Eucalyptus grandis*. *Nature*, **509**, 356–62.
- Myburg,A.A., Potts,B.M., Marques,C., Kirst,M., Gion,J.-M., Grattapaglia,D., and Grima-Pettenati,J. (2007) Eucalypts. In, Kole,C. (ed), *Genome Mapping and Molecular Breeding in Plants*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 115–160.
- Nagel,R., Elliott,A., Masel,A., Birch,R.G., and Manners,J.M. (1990) Electroporation of binary Ti plasmid vector into *Agrobacterium tumefaciens* and *Agrobacterium rhizogenes*. *FEMS Microbiol. Lett.*, **67**, 325–328.
- Nakano,Y., Yamaguchi,M., Endo,H., Rejab,N.A., and Ohtani,M. (2015) NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants. *Front. Plant Sci.*, **6**, 1–18.
- Newman,L.J., Perazza,D.E., Juda,L., and Campbell,M.M. (2004) Involvement of the R2R3-MYB, AtMYB61, in the ectopic lignification and dark-photomorphogenic components of the det3 mutant phenotype. *Plant J.*, **37**, 239–250.
- Newman,R.H., Hill,S.J., and Harris,P.J. (2013) Wide-Angle X-Ray Scattering and Solid-State Nuclear Magnetic Resonance Data Combined to Test Models for Cellulose Microfibrils in Mung Bean Cell Walls. *Plant Physiol.*, **163**, 1558–1567.



- Niculaes,C., Morreel,K., Kim,H., Lu,F., McKee,L.S., Ivens,B., Haustraete,J., Vanholme,B., Rycke,R. De, Hertzberg,M., Fromm,J., Bulone,V., Polle,A., Ralph,J., and Boerjan,W. (2014) Phenylcoumaran Benzylic Ether Reductase Prevents Accumulation of Compounds Formed under Oxidative Conditions in Poplar Xylem. *Plant Cell*, **26**, 3775–91.
- Nieminen,K., Blomster,T., Helariutta,Y., and Mähönen,A.P. (2015) Vascular Cambium Development. *Arab. B.*, **13**, e0177.
- Nieminen,K.M., Kauppinen,L., and Helariutta,Y. (2004) A weed for wood? *Arabidopsis* as a genetic model for xylem development. *Plant Physiol.*, **135**, 653–659.
- Nilsson,J., Karlberg,A., Antti,H., Lopez-Vernaza,M., Mellerowicz,E., Perrot-Rechenmann,C., Sandberg,G., and Bhalerao,R.P. (2008) Dissecting the molecular basis of the regulation of wood formation by auxin in hybrid aspen. *Plant Cell*, **20**, 843–855.
- Ohashi-Ito,K., Kubo,M., Demura,T., and Fukuda,H. (2005) Class III homeodomain leucine-zipper proteins regulate xylem cell differentiation. *Plant Cell Physiol.*, **46**, 1646–56.
- Ohashi-Ito,K., Oda,Y., and Fukuda,H. (2010) *Arabidopsis* VASCULAR-RELATED NAC-DOMAIN6 Directly Regulates the Genes That Govern Programmed Cell Death and Secondary Wall Formation during Xylem Differentiation. *Plant Cell*, **22**, 3461–3473.
- Ohdaira,Y., Kakegawa,K., Amino,S., Sugiyama,M., and Fukuda,H. (2002) Activity of cell-wall degradation associated with differentiation of isolated mesophyll cells of *Zinnia elegans* into tracheary elements. *Planta*, **215**, 177–184.
- Ohman,D., Demedts,B., Kumar,M., Gerber,L., Gorzsas,A., Goeminne,G., Hedenström,M., Ellis,B., Boerjan,W., and Sundberg,B. (2013) MYB103 is required for FERULATE-5-HYDROXYLASE expression and syringyl lignin biosynthesis in *Arabidopsis* stems. *Plant J.*, **73**, 63–76.
- Ohta,M., Matsui,K., Hiratsu,K., Shinshi,H., and Ohme-Takagi,M. (2001) Repression domains of class II ERF transcriptional repressors share an essential motif for active repression. *Plant Cell*, **13**, 1959–1968.
- Ohtani,M., Nishikubo,N., Xu,B., Yamaguchi,M., Mitsuda,N., Goué,N., Shi,F., Ohme-Takagi,M., and Demura,T. (2011) A NAC domain protein family contributing to the regulation of wood formation in poplar. *Plant J.*, **67**, 499–512.
- Oñate-Sánchez,L. and Vicente-Carbajosa,J. (2008) DNA-free RNA isolation protocols for *Arabidopsis thaliana*, including seeds and siliques. *BMC Res. Notes*, **1**, 93.
- Over, R.S., and Michaels, S.D. (2014). Open and closed: the roles of linker histones in plants and animals. *Mol. Plant* **7**: 481-491.
- Paiva,J.A., Rodrigues,J., Fevereiro,P., Neves,L., Araújo,C., Marques,C., Freitas,A., Bergès,H., and Grima-Pettenati,J. (2011) Building up resources and knowledge to unravel transcriptomics dynamics underlying *Eucalyptus globulus* xylogenesis. *BMC Proc.*, **5**, O52.

- Patzlaff, A., McInnis, S., Courtenay, A., Surman, C., Newman, L.J., Smith, C., Bevan, M.W., Mansfield, S., Whetten, R.W., Sederoff, R.R., and Campbell, M.M. (2003) Characterisation of a pine MYB that regulates lignification. *Plant J.*, **36**, 743–754.
- Paux, E., Tamasloukht, M., Ladouce, N., Sivadon, P., and Grima-Pettenati, J. (2004) Identification of genes preferentially expressed during wood formation in *Eucalyptus*. *Plant Mol. Biol.*, **55**, 263–80.
- Persak, H. and Pitzschke, A. (2014) Dominant repression by *Arabidopsis* transcription factor MYB44 causes oxidative damage and hypersensitivity to abiotic stress. *Int. J. Mol. Sci.*, **15**, 2517–2537.
- Pesquet, E., Zhang, B., Gorzsás, A., Puhakainen, T., Serk, H., Escamez, S., Barbier, O., Gerber, L., Courtois-Moreau, C., Alatalo, E., Paulin, L., Kangasjärvi, J., Sundberg, B., Goffner, D. and Tuominen, H. (2013) Non-cell-autonomous postmortem lignification of tracheary elements in *Zinnia elegans*. *Plant Cell*, **25**, 1314–28.
- Phelep, M., Petit, A., Martin, L., Duhoux, E. and Tempé, J. (1991) Transformation and Regeneration of a Nitrogen-Fixing Tree, *Allocasuarina Verticillata* Lam. *Nat. Biotechnol.* **9**, 461–466.
- Pichon, M., Courbou, I., Beckert, M., Boudet, A.M., and Grima-Pettenati, J. (1998) Cloning and characterization of two maize cDNAs encoding cinnamoyl-CoA reductase (CCR) and differential expression of the corresponding genes. *Plant Mol. Biol.*, **38**, 671–676.
- Piquemal, J., Lapierre, C., Myton, K., O’Connell, A., Schuch, W., Grima-Pettenati, J. and 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.
- Pitre, F.E., Pollet, B., Lafarguette, F., Cooke, J.E., MacKay, J.J., and Lapierre, C. (2007). Effects of increased nitrogen supply on the lignification of poplar wood. *J Agric. Food Chem.* **55**: 10306-10314.
- Plasencia, A., Soler, M., Dupas, A., Ladouce, N., Silva-Martins, G., Martinez, Y., Lapierre, C., Franche, C., Truchet, C. and Grima-Pettenati, J. (2015). *Eucalyptus* hairy roots, a fast, efficient and versatile tool to explore function and expression of genes involved in wood formation. *Plant Biotech. J.*, in press.
- Plomion, C., Leprovost, G., and Stokes, A. (2001) Wood formation in trees. *Plant Physiol.*, **127**, 1513–1523.
- Preston, J., Wheeler, J., Heazlewood, J., Li, S.F., and Parish, R.W. (2004) AtMYB32 is required for normal pollen development in *Arabidopsis thaliana*. *Plant J.*, **40**, 979–995.

- Przewloka, M.R., Wierzbicki, A.T., Slusarczyk, J., Kuras, M., Grasser, K.D., Stemmer, C., and Jerzmanowski, A. (2002) The 'drought-inducible' histone H1s of tobacco play no role in male sterility linked to alterations in H1 variants. *Planta*, **215**, 371–379.
- Quandt, H.J., Puehler, A. and Broer, I. (1993) Transgenic root nodules of *Vicia hirsuta*: a fast and efficient system for the study of gene expression in indeterminate-type nodules. *MPMI-Molecular Plant Microbe*, **6**, 699–703.
- Raes, J., Rohde, A., Christensen, J.H., Peer, Y. Van De, and Boerjan, W. (2014) Genome-Wide Characterization of the Lignification Toolbox in *Arabidopsis* 1 [ w ]. **133**, 1051–1071.
- Rahantamalala, A., Rech, P., Martinez, Y., Chaubet-Gigot, N., Grima-Pettenati, J., and Pacquit, V. (2010) Coordinated transcriptional regulation of two key genes in the lignin branch pathway - CAD and CCR - is mediated through MYB- binding sites. *BMC Plant Biol.*, **10**, 130.
- Ralph, J., Kim, H., Lu, F., Grabber, J.H., Leplé, J.C., Berrío-Sierra, J., Derikvand, M.M., Jouanin, L., Boerjan, W. and Lapierre, C. (2008) Identification of the structure and origin of a thioacidolysis marker compound for ferulic acid incorporation into angiosperm lignins (and an indicator for *cinnamoyl CoA reductase* deficiency). *Plant J.* **53**, 368–379.
- Ranik, M. and Myburg, a a (2006) Six new cellulose synthase genes from *Eucalyptus* are associated with primary and secondary cell wall biosynthesis. *Tree Physiol.*, **26**, 545–556.
- Rea, M., Zheng, W., Chen, M., Braud, C., Bhangu, D., Rognan, T.N., and Xiao, W. (2012) Histone H1 affects gene imprinting and DNA methylation in *Arabidopsis*. *Plant J.*, **71**, 776–86.
- Rennie, E. a. and Scheller, H.V. (2014) Xylan biosynthesis. *Curr. Opin. Biotechnol.*, **26**, 100–107.
- Riker, A.J., Banfield, W.M., Wright, W.H., Keitt, G.W. and Sagen, H.E. (1930) Studies on Infectious Hairy Root of Nursery Apple Trees. *J. Agric. Res.* **41**, 507–540.
- Růžička, K., Ursache, R., Hejátko, J., and Helariutta, Y. (2015) Xylem development - from the cradle to the grave. *New Phytol.*, **207**, 519–35.
- Saitou, N. and Nei, M. (1987) The Neighbor-joining Method: A New Method for Reconstructing Phylogenetic Trees. *Mol. Biol. Evol.*, **4**, 406–425.
- Šamaj, J., Hawkins, S., Lauvergeat, V., Grima-Pettenati, J. and Boudet, A. (1998) Immunolocalization of *cinnamyl alcohol dehydrogenase 2* (CAD 2) indicates a good correlation with cell-specific activity of CAD 2 promoter in transgenic poplar shoots. *Planta*, **204**, 437–443.
- Sattler, S.E. and Funnell-Harris, D.L. (2013) Modifying lignin to improve bioenergy feedstocks: strengthening the barrier against pathogens?†. *Front. Plant Sci.*, **4**, 1–8.

- Sawaki, Y., Kihara-Doi, T., Kobayashi, Y., Nishikubo, N., Kawazu, T., Kobayashi, Y., Koyama, H. and Sato, S. (2013) Characterization of Al-responsive citrate excretion and citrate-transporting MATEs in *Eucalyptus camaldulensis*. *Planta*, **237**, 979–989.
- Sawaki, Y., Kobayashi, Y., Kihara-Doi, T., Nishikubo, N., Kawazu, T., Kobayashi, M., Kobayashi, Y., Iuchi, S., Koyama, H. and Sato, S. (2014) Identification of a STOP1-like protein in *Eucalyptus* that regulates transcription of Al tolerance genes. *Plant Sci.* **223**, 8–15.
- Schuetz, M., Smith, R., and Ellis, B. (2013) Xylem tissue specification, patterning, and differentiation mechanisms. *J. Exp. Bot.*, **64**, 11–31.
- Scippa, G.S., Griffiths, A., Chiatante, D., and Bray, E. A. (2000) The H1 histone variant of tomato, H1-S, is targeted to the nucleus and accumulates in chromatin in response to water-deficit stress. *Planta*, **211**, 173–81.
- Scippa, G.S., Di Michele, M., Onelli, E., Patrignani, G., Chiatante, D., and Bray, E.A. (2004) The histone-like protein H1-S and the response of tomato leaves to water deficit. *J. Exp. Bot.*, **55**, 99–109.
- Shen, H., He, X., Poovaiah, C.R., Wuddineh, W. A., Ma, J., Mann, D.G.J., Wang, H., Jackson, L., Tang, Y., Neal Stewart, J., Chen, F., and Dixon, R. A. (2012) Functional characterization of the switchgrass (*Panicum virgatum*) R2R3-MYB transcription factor PvMYB4 for improvement of lignocellulosic feedstocks. *New Phytol.*, **193**, 121–136.
- Shi, R., Sun, Y.H., Li, Q., Heber, S., Sederoff, R., and Chiang, V.L. (2010) Towards a systems approach for lignin biosynthesis in populus trichocarpa: Transcript abundance and specificity of the monolignol biosynthetic genes. *Plant Cell Physiol.*, **51**, 144–163.
- Shimada, T.L.T., Shimada, T.L.T., and Hara-Nishimura, I. (2010) A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of *Arabidopsis thaliana*. *Plant J.*, **61**, 519–528.
- Sibout, R. and Höfte, H. (2012) Plant cell biology: The ABC of monolignol transport. *Curr. Biol.*, **22**, 533–535.
- Soler, M., Camargo, E.L.O., Carocha, V., Cassan-Wang, H., San Clemente, H., Savelli, B., Hefer, C.A., Paiva, J.A.P., Myburg, A.A., and Grima-Pettenati, J. (2015) The *Eucalyptus grandis* R2R3-MYB transcription factor family: evidence for woody growth-related evolution and function. *New Phytol.*, **206**, 1364–77.
- Sonbol, F.-M., Fornalé, S., Capellades, M., Encina, A., Touriño, S., Torres, J.-L., Rovira, P., Ruel, K., Puigdomènech, P., Rigau, J., and Caparrós-Ruiz, D. (2009) The maize ZmMYB42 represses the phenylpropanoid pathway and affects the cell wall structure, composition and degradability in *Arabidopsis thaliana*. *Plant Mol. Biol.*, **70**, 283–296.

- Song, D., Shen, J., and Li, L. (2010) Characterization of cellulose synthase complexes in *Populus* xylem differentiation. *New Phytol.* **187**, 777-790.
- Sorce, C., Giovannelli, A., Sebastiani, L., and Anfodillo, T. (2013) Hormonal signals involved in the regulation of cambial activity, xylogenesis and vessel patterning in trees. *Plant Cell Rep.*, **32**, 885–898.
- Southerton, S.G., Marshall, H., Mouradov, a, and Teasdale, R.D. (1998) Eucalypt MADS-box genes expressed in developing flowers. *Plant Physiol.* **118**, 365–372.
- Spokevicius, A. V., Van Beveren, K., Leitch, M. a. and Bossinger, G. (2005) *Agrobacterium*-mediated in vitro transformation of wood-producing stem segments in eucalypts. *Plant Cell Rep.* **23**, 617–624.
- Studer, M.H., DeMartini, J.D., Davis, M.F., Sykes, R.W., Davison, B., Keller, M., Tuskan, G. a, and Wyman, C.E. (2011) Lignin content in natural *Populus* variants affects sugar release. *Proc. Natl. Acad. Sci. U. S. A.*, **108**, 6300–6305.
- Sykes, R.W., Gjersing, E.L., Foutz, K., Rottmann, W.H., Kuhn, S. a., Foster, C.E., Ziebell, A., Turner, G.B., Decker, S.R., Hinchee, M. a. W., and Davis, M.F. (2015) Down-regulation of *p*-coumaroyl quinate/shikimate 3'-hydroxylase (C3'H) and cinnamate 4-hydroxylase (C4H) genes in the lignin biosynthetic pathway of *Eucalyptus urophylla* × *E. grandis* leads to improved sugar release. *Biotechnol. Biofuels*, **8**, 128.
- Tamagnone, L., Merida, a, Parr, a, Mackay, S., Culianez-Macia, F., 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**, 135–54.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolution distance, and maximum parsimony methods. *Mol. Biol. Evol.*, **28**, 2731–2739.
- Taylor-Teeple, M., Lin, L., de Lucas, M., Turco, G., Toal, T.W., Gaudinier, A., Young, N.F., Trabucco, G.M., Veling, M.T., Lamothe, R., Handakumbura, P.P., Xiong, G., Wang, C., Corwin, J., Tsoukalas, A., Zhang, L., Ware, D., Pauly, M., Kliebenstein, D.J., *et al.* (2014) An *Arabidopsis* gene regulatory network for secondary cell wall synthesis. *Nature*, **517**, 571–575.
- Timell, T.E. (1967) Recent progress in the chemistry of wood hemicelluloses. *Wood Sci. Technol.* **1**, 45-70.
- Tournier, V., Grat, S., Marque, C., El Kayal, W., Penchel, R., de Andrade, G., Boudet, A.-M. and Teulières, C. (2003) An efficient procedure to stably introduce genes into an economically important pulp tree (*Eucalyptus grandis* × *Eucalyptus urophylla*). *Transgenic Res.* **12**, 403–11.

- Trivedi,I., Ranjan,A., Sharma,Y.K., and Sawant,S. (2012) The histone H1 variant accumulates in response to water stress in the drought tolerant genotype of *Gossypium herbaceum* L. *Protein J.*, **31**, 477–86.
- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M. and Rozen, S.G. (2012) Primer3-new capabilities and interfaces. *Nucleic Acids Res.* **40**, 1–12.
- Umezawa,T. (2003) Diversity in lignan biosynthesis. *Phytochem. Rev.*, **2**, 371–390.
- Ursache,R., Nieminen,K., and Helariutta,Y. (2013) Genetic and hormonal regulation of cambial development. *Physiol. Plant.*, **147**, 36–45.
- Van Acker,R., Vanholme,R., Storme,V., Mortimer,J.C., Dupree,P., and Boerjan,W. (2013) Lignin biosynthesis perturbations affect secondary cell wall composition and saccharification yield in *Arabidopsis thaliana*. *Biotechnol. Biofuels*, **6**, 46.
- Van Beveren, K.S., Spokevicius, A. V., Tibbits, J., Wang, Q. and Bossinger, G. (2006) Transformation of cambial tissue in vivo provides an efficient means for induced somatic sector analysis and gene testing in stems of woody plant species. *Funct. Plant Biol.*, **33** 629.
- Vanholme,R., Cesarino,I., Rataj,K., Xiao,Y., Sundin,L., Goeminne,G., Kim,H., Cross,J., Morreel,K., Araujo,P., Welsh,L., Haustraete,J., McClellan,C., Vanholme,B., Ralph,J., Simpson,G.G., Halpin,C., and Boerjan,W. (2013) Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in *Arabidopsis*. *Science*, **341**, 1103–6.
- Vanholme,R., Demedts,B., Morreel,K., Ralph,J., and Boerjan,W. (2010) Lignin biosynthesis and structure. *Plant Physiol.*, **153**, 895–905.
- Vanholme,R., Morreel,K., Darrah,C., Oyarce,P., Grabber,J.H., Ralph,J., and Boerjan,W. (2012) Metabolic engineering of novel lignin in biomass crops. *New Phytol.*, **196**, 978–1000.
- Wang,J., Kuang,J., Shan,W., Chen,J., Xie,H., Lu,W., Chen,J., and Chen,J. (2012) Expression profiles of a banana fruit linker histone H1 gene MaHIS1 and its interaction with a WRKY transcription factor. *Plant Cell Rep.*, **31**, 1485–94.
- Wang,Y., Chantreau,M., Sibout,R., and Hawkins,S. (2013) Plant cell wall lignification and monolignol metabolism. *Front. Plant Sci.*, **4**, 220.
- Wei,T. and O’Connell,M. a (1996) Structure and characterization of a putative drought-inducible H1 histone gene. *Plant Mol. Biol.*, **30**, 255–68.
- Willemsen,V., Bauch,M., Bennett,T., Campilho,A., Wolkenfelt,H., Xu,J., Haseloff,J., and Scheres,B. (2008) The NAC Domain Transcription Factors FEZ and SOMBRERO Control the Orientation of Cell Division Plane in *Arabidopsis* Root Stem Cells. *Dev. Cell*, **15**, 913–922.

- Yamaguchi,M., Kubo,M., Fukuda,H., and Demura,T. (2008) VASCULAR-RELATED NAC-DOMAIN7 is involved in the differentiation of all types of xylem vessels in *Arabidopsis* roots and shoots. *Plant J.*, **55**, 652–664.
- Yamaguchi,M., Mitsuda,N., Ohtani,M., Ohme-Takagi,M., Kato,K., and Demura,T. (2011) VASCULAR-RELATED NAC-DOMAIN 7 directly regulates the expression of a broad range of genes for xylem vessel formation. *Plant J.*, **66**, 579–590.
- Yamaguchi,M., Ohtani,M., Mitsuda,N., Kubo,M., Ohme-Takagi,M., Fukuda,H., and Demura,T. (2010) VND-INTERACTING2, a NAC Domain Transcription Factor, Negatively Regulates Xylem Vessel Formation in *Arabidopsis*. *Plant Cell*, **22**, 1249–1263.
- Yang,S.-M., Kim,B.J., Norwood Toro,L., and Skoultchi,A.I. (2013) H1 linker histone promotes epigenetic silencing by regulating both DNA methylation and histone H3 methylation. *Proc. Natl. Acad. Sci. U. S. A.*, **110**, 1708–13.
- Yu,H., Soler,M., Mila,I., San Clemente,H., Savelli,B., Dunand,C., Paiva,J. a. P., Myburg,A.A., Bouzayen,M., Grima-Pettenati,J., and Cassan-Wang,H. (2014) Genome-Wide Characterization and Expression Profiling of the AUXIN RESPONSE FACTOR (ARF) Gene Family in *Eucalyptus grandis*. *PLoS One*, **9**, e108906.
- Yu,H., Soler,M., San Clemente,H., Mila,I., Paiva,J. a. P., Myburg,A.A., Bouzayen,M., Grima-Pettenati,J., and Cassan-Wang,H. (2015) Comprehensive Genome-Wide Analysis of the Aux/IAA Gene Family in *Eucalyptus*: Evidence for the Role of EgrIAA4 in Wood Formation. *Plant Cell Physiol.*, **56**, 700–714.
- Zemach,A., Kim,M.Y., Hsieh,P.-H., Coleman-Derr,D., Eshed-Williams,L., Thao,K., Harmer,S.L., and Zilberman,D. (2013) The *Arabidopsis* Nucleosome Remodeler DDM1 Allows DNA Methyltransferases to Access H1-Containing Heterochromatin. *Cell*, **153**, 193–205.
- Zhang,J., Elo,A., and Helariutta,Y. (2011) *Arabidopsis* as a model for wood formation. *Curr. Opin. Biotechnol.*, **22**, 293–9.
- Zhang,J., Nieminen,K., Serra,J.A.A., and Helariutta,Y. (2014) The formation of wood and its control. *Curr. Opin. Plant Biol.*, **17**, 56–63.
- Zhao,J., Zhang,W., Zhao,Y., Gong,X., Guo,L., Zhu,G., Wang,X., Gong,Z., Schumaker,K.S., and Guo,Y. (2007) SAD2, an importin -like protein, is required for UV-B response in *Arabidopsis* by mediating MYB4 nuclear trafficking. *Plant Cell*, **19**, 3805–18.
- Zhong,R., Demura,T., and Ye,Z.-H. (2006) SND1, a NAC Domain Transcription Factor, Is a Key Regulator of Secondary Wall Synthesis in Fibers of *Arabidopsis*. *Plant Cell Online*, **18**, 3158–3170.

- Zhong,R., Lee,C., McCarthy,R.L., Reeves,C.K., Jones,E.G., and Ye,Z.-H. (2011) Transcriptional Activation of Secondary Wall Biosynthesis by Rice and Maize NAC and MYB Transcription Factors. *Plant Cell Physiol.*, **52**, 1856–1871.
- Zhong,R., Lee,C., and Ye,Z.-H. (2010a) Evolutionary conservation of the transcriptional network regulating secondary cell wall biosynthesis. *Trends Plant Sci.*, **15**, 625–32.
- Zhong,R., Lee,C., and Ye,Z.-H. (2010b) Functional Characterization of Poplar Wood-Associated NAC Domain Transcription Factors. *Plant Physiol.*, **152**, 1044–1055.
- Zhong,R., Lee,C., and Ye,Z.-H. (2010c) Global Analysis of Direct Targets of Secondary Wall NAC Master Switches in *Arabidopsis*. *Mol. Plant*, **3**, 1087–1103.
- Zhong,R., Lee,C., Zhou,J., McCarthy,R.L., and Ye,Z.-H. (2008) A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in *Arabidopsis*. *Plant Cell*, **20**, 2763–82.
- Zhong,R., McCarthy,R.L., Haghghat,M., and Ye,Z.H. (2013) The Poplar MYB Master Switches Bind to the SMRE Site and Activate the Secondary Wall Biosynthetic Program during Wood Formation. *PLoS One*, **8**, 18–21.
- Zhong,R., Richardson,E. a, and Ye,Z.-H. (2007) The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in *Arabidopsis*. *Plant Cell*, **19**, 2776–2792.
- Zhong,R. and Ye,Z.-H. (2012) MYB46 and MYB83 Bind to the SMRE Sites and Directly Activate a Suite of Transcription Factors and Secondary Wall Biosynthetic Genes. *Plant Cell Physiol.*, **53**, 368–380.
- Zhong,R. and Ye,Z.-H. (2007) Regulation of cell wall biosynthesis. *Curr. Opin. Plant Biol.*, **10**, 564–572.
- Zhong,R. and Ye,Z.-H. (2015) Secondary Cell Walls: Biosynthesis, Patterned Deposition and Transcriptional Regulation. *Plant Cell Physiol.*, **56**, 195–214.
- Zhong,R. and Ye,Z.-H. (2009) Transcriptional regulation of lignin biosynthesis. *Plant Signal. Behav.*, **4**, 1028–1034.
- Zhou,J., Lee,C., Zhong,R., and Ye,Z.-H. (2009) MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in *Arabidopsis*. *Plant Cell*, **21**, 248–66.
- Zhou,J., Zhong,R., and Ye,Z.-H. (2014) *Arabidopsis* NAC Domain Proteins, VND1 to VND5, Are Transcriptional Regulators of Secondary Wall Biosynthesis in Vessels. *PLoS One*, **9**, e105726.



- Zhu,T., Nevo,E., Sun,D., and Peng,J. (2012) Phylogenetic analyses unravel the evolutionary history of nac proteins in plants. *Evolution (N. Y.)*, **66**, 1833–1848.
- 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.
- Zuckerkindl,E. and Pauling,L. (1965) Evolutionary Divergence and Convergence in Proteins. In, *Evolving Genes and Proteins*. Elsevier, pp. 97–166.



**Author** - Anna PLASENCIA CASADEVALL

**Title** - Transcriptional regulation of wood formation in *Eucalyptus*. Role of MYB transcription factors and protein-protein interactions.

**PhD co-supervisors** - Jacqueline GRIMA-PETTENATI et Isabelle TRUCHET

**Abstract**

Our objective was to better understand the regulation of the biosynthesis of the lignified secondary cell walls during wood formation in *Eucalyptus*, the most planted hardwood tree, and the second whose genome has been sequenced. We functionally characterized three *Eucalyptus* transcription factors of the R2R3-MYB family and identified EgMYB137 as a new regulator of secondary cell wall deposition. We also showed that the transcriptional activity of EgMYB1, a repressor of lignin biosynthesis was modulated by protein-protein interactions involving a linker histone (EgH1.3). Finally, we set up a homologous transformation system for *Eucalyptus* using *Agrobacterium rhizogenes*. The transgenic hairy roots are suitable for high throughput functional characterization of cell wall-related genes. Our findings not only allowed getting new insights into the complexity of the network regulating secondary cell walls but also open new avenues to improve wood quality for industrial applications such as second-generation bioethanol.

**Key words** – *Eucalyptus*, Wood formation, Xylem, Secondary cell wall, Lignin, MYB transcription factor, Linker histone, Hairy roots, Secondary metabolites, second-generation bioethanol.

**Autrice** - Anna PLASENCIA CASADEVALL

**Titre** - Régulation transcriptionnelle de la formation du bois chez l'*Eucalyptus*: rôle des facteurs de transcription MYB et des interactions protéines-protéines

**Directrices de thèse** - Jacqueline GRIMA-PETTENATI et Isabelle TRUCHET

**Lieu et date de soutenance** - 15 décembre 2015, Salle Marc Ridet, Campus INRA-Auzeville

### **Résumé**

Notre objectif était de mieux comprendre la régulation de la biosynthèse des parois secondaires lors de la formation du bois chez l'*Eucalyptus*, le feuillu le plus planté au monde et le deuxième dont le génome est séquencé. Nous avons caractérisé trois facteurs de transcription de la famille MYB-R2R3 et montré que EgMYB137 était un nouveau régulateur de la biosynthèse des parois secondaires. Nous avons aussi démontré que l'activité transcriptionnelle de EgMYB1, un répresseur de la biosynthèse des lignines, était régulée par une interaction protéine-protéine impliquant une histone linker (EgH1.3). Enfin, nous avons mis au point une méthode de transformation homologue chez l'*Eucalyptus* via *A. rhizogenes*. Les « hairy roots » transgéniques sont adaptées à la caractérisation fonctionnelle de gènes reliés à la formation du xylème. Nos résultats ont permis de découvrir de nouveaux acteurs impliqués dans la régulation des parois secondaires, mettant en lumière la complexité de ce processus mais aussi offrant de nouvelles perspectives pour l'amélioration du bois pour des applications industrielles comme la production de bioéthanol de deuxième génération.

**Mots clés** - *Eucalyptus*, Formation du bois, Xylème, Paroi secondaire, Lignine, Facteurs de transcription MYB, Histone linker, Hairy roots, Métabolites secondaires, Bioéthanol de deuxième génération.

**Discipline administrative** - Développement des plantes

### **Intitulés et adresses des laboratoires**

UMR5546 UPS/CNRS, Laboratoire de Recherche en Sciences Végétales (LRSV), Pôle de Biotechnologies Végétales. 24, chemin de Borde Rouge, B.P. 42617, Auzeville, 31326 Castanet-Tolosan cédex, France.