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 :
Hong YU

Le mardi 30 septembre 2014

Titre :

Caractérisation des familles de facteurs de transcription ARF et Aux/IAA chez l'Eucalyptus, rôles dans la formation du bois

---

École doctorale et discipline ou spécialité :

ED SEVAB : Développement des plantes

Unité de recherche :

Laboratoire de Recherche en Sciences Végétales (LRSV), UMR 5546, UT3 Paul Sabatier/CNRS

Directeur(s) de Thèse :

Dr Jacqueline GRIMA-PETTENATI, Directrice de Recherches CNRS, Toulouse
Dr Hua CASSAN-WANG, Maître de conférences, Université Toulouse 3

Rapporteurs :

Dr Marie BAUCHER, Maître de Recherches, Université Libre de Bruxelles, Belgique
Pr Catherine BELLINI, Professeure, Université de Umeå, Suède

Autre(s) membre(s) du jury :

Pr Vincent BURLAT, Professeur, Université Toulouse 3
Pr Zhengguo LI, Professeur, Université de Chongqing (Chine)

Acknowledgments

Foremost, I would like to express my sincere gratitude to my supervisors Dr. Jacqueline Grima-Pettenati and Dr. Hua Cassan-Wang for allowing me to be part of their team, and thanks for their guidance, advice, involvement and support in my work. They helped me in all the time of research and writing of this thesis.

I am very grateful to the committee members: Dr. Marie Baucher, Dr. Catherine Bellini and Dr. Laurent Laplaze for having accepted to judge my research work.

I would like also to acknowledge the help and advice of Prof. Mondher Bouzayen and Isabelle Mila in GBF during my PhD. Thanks Dr. Edouard Pesquet and Junko Takahashi Schmidt for hosting me in Umeå Plant Science Centre and for their precious help and advice with the Py-GC/MS analysis. Thanks to Dr. Françoise Monéger and Dr. Didier Aldon for their precious advice during my PhD. Thanks Yves Martinez for his help with the microscope analysis in Plateforme Imagerie-TRI.

Thanks to Nathalie Ladouce for her help with my experiments especially for Fluidigm. Thanks to Marçal Soler, for his stimulating discussions and sharing his knowledge with me. Thanks to Hélène San-clemente, Dunand Christophe and Qiang Li, for their support of the bioinformatics.

I also would like to thank all members of the GFE team and all the friends studying and living around me during these four years. Because of their support, advice and friendship, they all contributed to the accomplishment of this thesis. And I would like to thank to all my colleagues from Chongqing University for their friendship and support especially Prof. Zhengguo Li, Zhongkang Wang and Youping Yin.

My scholarship was granted by China Scholarship Council in Beijing. Thank you for making my staying in France possible.

And last but not the least, I would like to thank to my whole family for their loving support and encouragement.

Résumé

L'auxine est une hormone centrale du développement chez les végétaux, impliquée dans un large éventail de processus tels que l'organogenèse, la croissance racinaire, le développement des fruits et le développement vasculaire. Chez les plantes ligneuses, l'auxine joue un rôle clé dans la régulation de la formation du bois en stimulant l'activité du cambium vasculaire et le développement du xylème. Les ARFs (Auxin-response factors) et les Aux/IAAs (Auxin/Indole acetic acid) sont des régulateurs importants de la voie de perception et de signalisation de l'auxine. La disponibilité récente de la séquence du génome d'*Eucalyptus grandis* nous a permis d'étudier les caractéristiques et l'histoire évolutive de ces deux familles de gènes chez cette plante ligneuse de grande importance économique. Dans ce travail, nous avons identifié et caractérisé 17 *EgrARFs* et 24 *EgrIAAs* que nous avons nommés en fonction de leurs orthologues putatifs chez *Arabidopsis*. Chez *E. grandis*, ces deux familles ont un peu moins de gènes que chez la plupart des angiospermes étudiés jusqu'ici. L'analyse de phylogénie comparative de génomes appartenant à des lignées taxonomiques pertinentes a révélé, chez la famille *ARF*, la présence d'une clade trouvée préférentiellement chez les plantes ligneuses pérennes. Des analyses d'expression à haut débit sur un large panel d'organes et de tissus et en réponse à des signaux environnementaux ont mis en évidence des gènes exprimés préférentiellement dans le cambium vasculaire et/ou le xylème en différentiation, certains présentant des réponses à des stimuli abiotiques. Sur la base de leurs profils d'expression, nous avons sélectionné certains gènes candidats prometteurs et effectué leur caractérisation fonctionnelle afin d'évaluer leur rôle potentiel dans la formation du bois. Nous avons utilisé *Arabidopsis*, qui peut dans certaines conditions présenter une croissance secondaire, pour surexprimer des versions mutées et stabilisées de *EgrIAA4, 9 et 20*. Les lignées transgéniques présentent des phénotypes aberrants d'insensibilité à l'auxine. Parmi ceux-ci, des tailles de plantes réduites ou naines, des racines qui croissent de façon agravitropique, la réduction ou même l'absence de racines latérales et la stérilité partielle. Les analyses histochimiques complétées par des analyses en pyrolyse ont montré des modifications évidentes dans

la différentiation des cellules du xylème et/ou des fibres interfasciculaires ainsi que composition chimique de leurs parois secondaires. En conclusion, cette étude offre une caractérisation complète de deux familles de régulateurs importants de l'auxine chez l'*Eucalyptus* et met en évidence l'implication de certains membres dans la régulation de la formation du bois.

Abstract

Auxin is a central hormone involved in a wide range of developmental processes including organogenesis, tropic movement, root growth, fruit development, tissue and organ patterning and vascular development. In woody plants, auxin has been proposed to play a key role in regulating wood formation through its effects on cambial activity and xylem development. Auxin Response Factors (ARF) and Auxin/Indole-3-Acetic Acid (Aux/IAA) are important regulators of auxin responses in plants. The recent availability of the *Eucalyptus grandis* genome allowed us to investigate the characteristics and evolutionary history of these two gene families in a woody plant of high economic importance. In this work, we identified and characterized 17 *EgrARF* and 24 *EgrIAA* gene members and named them according to their putative orthologs in *Arabidopsis*. Both of these two gene families are slightly contracted, as compared to those of most angiosperms studied hitherto. Comparative phylogenetic analyses with genomes of relevant taxonomic lineages revealed the presence of a new ARF clade found preferentially in woody and/or perennial plants. High-throughput expression profiling among different organs and tissues and in response to environmental cues highlighted genes expressed in vascular cambium and/or developing xylem, in addition to dynamic modifications in response to environmental stimuli. Based on their expression profiles, we selected some promising candidates and carried out their functional characterization to get insights into their potential roles in wood formation. We used *Arabidopsis*, which was shown to undergo secondary growth in some conditions, to overexpress stabilized versions of *EgrIAA4m*, *9Am* and *20*. The transgenic lines exhibited auxin-related aberrant phenotypes, such as smaller size, impaired root growth and development, and partial sterility. Histochemistry and Py-GC/MS analysis revealed that the transgenic plants showed obvious modifications in xylem and/or fiber cells differentiation and secondary cell wall composition. Altogether, the present study provides a comprehensive characterization of the *Eucalyptus ARF* and *Aux/IAA* gene families and highlights the involvement of some members in the regulation of wood development.

摘 要

生长素参与植物生长和发育诸多过程，包括器官的形成、植物的向地和向光反应、根的发育生长、果实的发育，以及维管束组织的形成和分化。生长素在木材的形成过程中也发挥着重要作用，主要通过对维管形成层及木质部细胞发育的影响来调控其形成。生长素响应因子（ARF）和生长素/吲哚乙酸蛋白（Aux/IAA）是植物生长素响应的主要调节蛋白。桉树作为木材在人类的生产生活中具有重要的经济价值，其全基因组测序工作已于近期完成，这使得我们可以对桉树 *ARF* 和 *Aux/IAA* 这两个家族进行全面的分析。在本论文中，我们一共分离出了 17 个 *ARF* 与 24 个 *Aux/IAA* 基因，并且根据其与拟南芥的同源性对它们进行了重命名。目前为止，这两个基因家族的数量在桉树中较其它研究过的大多数的被子植物有所减少。与其它物种的系统发育进化树分析表明在 ARF 家族中存在一个木本植物植物中占有优势的分支。对不同组织器官、环境因素及植物激素处理实验的高通量表达分析显示有些基因在维管形成层和/或木质部中有优势表达，而且大部分基因会对不同的环境因素及激素处理作出响应。为了深入了解桉树 *ARF* 与 *Aux/IAA* 在木材发育过程中的作用，我们选择了一些可能调控木质部发育的 *ARF* 与 *Aux/IAA* 基因构建了转基因拟南芥。其中 *EgrIAA4m*, *EgrIAA9Am* 与 *EgrIAA20* 转基因拟南芥表现出了一些与生长素缺陷相关的表型，比如矮小株型，根与侧根非正常的生长发育，根向地性的丧失以及育性降低。组织化学及高温分解气相色谱质朴分析表明转基因植株中木质部细胞的发育以及次生细胞壁的组成成分都发生了明显的改变。总之，该研究全面分析了桉树 ARF 与 Aux/IAA 两个家族，并通过反义基因功能分析揭示了有些成员在木质部的生长发育中起着重要的调控作用。

Contents

List of Figures

List of Tables

Abbreviations

2, 4-D: 2,4-dichlorophenoxyacetic acid

4CL: 4-Hydroxycinnamoyl-CoA ligase

4-Cl-IAA: 4-Chloroindole-3-acetic

ABC: ATP-binding cassette

ABCB: ATP Binding Cassette subfamily B

ABP1: AUXIN BINDING PROTEIN 1

ACC: 1-aminocyclopropane-1-carboxylic-acid

AD: Activation domain

AFB: Auxin-signaling F-box proteins

ARF: Auxin response factors

AuxREs: Auxin response elements

Aux/IAA: Auxin/Indole-3-Acetic Acid

AUX1/LAX: AUXIN RESISTANT 1/LIKE AUX1

BD: Binding domain

C_2O: Carbon dioxide

C4H: Cinnamate 4-hydroxylase

C3H: *p*-coumarate 3-hydroxylase

CAD: Cinnamyl alcohol dehydrogenase

CaMV: Cauliflower Mosaic Virus

CCoAOMT: Caffeoyl-CoA O-methyltransferase

CCR: Cinnamoyl-CoA reductase

cDNA: Complementary deoxyribonucleic acid

CDS: Coding sequence

CesA: Cellulose synthase

CGs: Candidate genes

COMT: Caffeic acid O-methyltransferase

CSC: Cellulose synthase complexes

CTAB: Cetyltrimethylammonium bromide

CTD: C-terminal dimerization domain

DBD: DNA binding domains

EAR: Ethylene-responsive element binding factor-associated Amphiphilic Repression

EST: Expression Sequence Tag

F5H: Ferulate 5-hydroxylase

G: Guaiacyl

GA: Gibberellic acid or gibberellin

GFP: Green Fluorescent Protein

GH3: Gretchen Hagen3

GTs: Glycosyltransferases

H: *p*-hydroxyphenyl

HD-ZIP III: Class III homeodomain-leucine zipper

IAA: Indole-3-Acetic Acid

IAM: Indole-3-acetamide

IAOX: Indole-3-acetaldoxime

IBA: Indole-3-butyric

IPA: Indole-3-pyruvate acid

MP: ARF5/MONOPTEROS

MR: Middle region

MS: Murashige and Skoog

NAA: 1-Naphthaleneacetic acid

NLS: Nuclear localization signal

OFP: Ovate Family Protein

PAA: Phenylacetic acid

PAL: Phenylalanine ammonia lyas

PAT: Auxin polar transport

PCD: Programed cell death

PCR: Polymerase chain reaction

PIN: PIN-FORMED

PILS: PIN-LIKES

Py-GC/MS: Pyrolysis–gas chromatography/mass spectrometry

qRT-PCR: Quantitative reverse transcription Polymerase chain reaction

RD: Repression domain

S: Syringyl

SAUR: Small Auxin-up RNA

SCF: SKP1-Cullin-F-box

SCW: Secondary cell wall

SD: Segmental duplication

TAM: Tryptamine

TD: Tandem duplication

TDIF: Tracheary element differentiation inhibitory factor

TDR/PXY: TDIF RECEPTOR/ PHLOEM INTERCALATED WITH XYLEM

TF: Transcription factor

TIR1: Transport inhibitor response 1

TPL: TOPLESS

Y2H: Yeast two-hybrid

General introduction

General introduction

Scientific context

The plant hormone auxin plays a prominent role in regulation of plant growth in response to diverse developmental and environmental cues such as organogenesis, tropic movement, root growth, fruit development, tissue and organ patterning (Friml 2003). In woody plants, auxin has been proposed to play a key role in the development of secondary xylem cells, a differentiation process involving cell division, expansion, secondary cell wall formation and cell death (Miyashima et al. 2013; Sundberg et al. 2000). Auxin is believed to exert its function in wood formation through its perception/signalling pathway, of which Auxin Response Factors (ARFs) and Aux/IAAs are two well-known components regulating auxin responsive gene expression (Guilfoyle and Hagen 2007; Tiwari et al. 2003). These two families have been studied in several annual plants and more deeply in *Arabidopsis*, but remain largely under investigated in tree species. The recent availability of *Eucalyptus grandis* genome (Myburg et al. 2014), the second hardwood forest tree genome fully sequenced, offers new opportunities to get insights into the regulation of secondary growth and cambial activity by *ARF* and *IAA*. *Eucalyptus* is indeed the most planted hardwood in the world, mainly for pulp and paper production but is also foreseen as a dedicated energy crop for lignocellulosic biofuel production. My PhD work was supported by the European Plant KBBE "Tree for Joules" project aiming at identifying candidate transcription factors involved in the regulation of secondary cell wall formation for improving their recalcitrance to degradation during the saccharification process.

Objectives of the thesis

The overall objective of my PhD was to better understand how auxin is involved in the regulation of secondary xylem formation in *Eucalyptus* through the analysis of two main components of the auxin-signalling pathway.

The first overarching objective was to survey, identify and characterize all members of the *ARF* and *Aux/IAA* families in *E. grandis* and to compare their evolutionary histories to that of other genomes. The second aim was to analyse the expression patterns in a wide range of tissues and organs and in response to environmental cues in order to select the best candidates potentially involved in wood formation. The third objective was to functionally characterize using reverse genetics some promising candidates to get insights on their potential roles in wood formation.

Organization of the manuscript

This thesis manuscript comprises four main chapters. The first chapter consists in a bibliographic review on wood formation and auxin. First, it presents actual knowledge regarding the development of secondary xylem and its regulation at molecular level in this chapter. It also describes auxin homeostasis, auxin signalling and the main roles of auxin on regulation different plant development process and wood formation. At last, it also explained why *Eucalyptus* was chosen as a model tree to study auxin action on wood formation.

Chapter II is presented in the form of an article (accepted in Plos one) and presents a genome-wide identification and characterization of the *ARF* family in *E. grandis*. With 17 members, the *E. grandis ARF* gene family is slightly contracted, as compared to those of most angiosperms studied hitherto, lacking traces of duplication events. Alternative splicing seemed to be a preeminent mechanism in shaping the functional diversity of the *ARF* family in *Eucalyptus*. We identified a new ARF clade found preferentially in woody plants. Finally, this study allowed identification of three *ARF* candidates (*ARF4, 10, 19A*) potentially involved in the auxin-regulated transcriptional program underlying wood formation

Chapter III is also presented in the form of an article (will be submitted to PCP) and presents a comprehensive genome-wide analysis of *Aux/IAA* gene family in *Eucalyptus* with evidence for *EgrIAA4*‘s role in wood formation. In this work, 24 functional

EgrIAAs were identified. High-throughput expression profiling highlighted *Aux/IAAs* expressed in vascular cambium and/or developing xylem, some showing differential expression in response to developmental (juvenile *versus* mature) and/or to environmental (tension stress) cues. We selected *EgrIAA4* as the most promising candidate gene. Overexpressing of *EgrIAA4m* in *Arabidopsis* strongly and negatively affected both xylary and interfascicular fibers development, and lignified secondary cell wall formation.

Chapter IV focuses on wood-associated *ARFs* and *IAAs* candidate gene selection and their functional characterization by reverse genetic in *Arabidopsis*. Thirteen promising candidate genes were selected and six of them have been overexpressed into *Arabidopsis* hitherto. By now, *EgrIAA4*, *EgrIAA9A* and *EgrIAA20* transgenic lines were analyzed both exhibiting alternated lignification pattern in stems and hypocotyls.

Chapter I

Bibliographic review

Chapter I: Bibliographic review

From the numerous adaptations that land plants have developed during evolution, the acquisition of the vascular system some 400 million years ago has been a crucial event ensuring their successful earth colonization. Plant vascular systems are composed of xylem and phloem. They provide physical strength to plant bodies and transport water, nutrients and other substances required for growth and defense. They interconnect all the plant body parts by their conductive function, from the root tip to the various organs in the shoot. Xylem is the main tissue for transporting water and solute minerals, whereas phloem is the route for distributing photosynthetic products and various signaling molecules. These two conductive tissues consist of highly specialized cell types that arise from undifferentiated stem cells located in lateral meristems via asymmetric periclinal cell division (Eames and MacDaniels, 1947; Esau, 1965). In the first part of the introduction, we will focus on secondary xylem formation and in a second part on auxin and the roles it plays in the different steps of xylem ontogeny.

Part I Wood or secondary xylem

1 Wood plays crucial roles for trees and mankind

1.1 Wood role in trees

The major functions of wood are to conduct water from roots to the crown, to support an ever-increasing mass of the growing tree, whilst adjusting to various environmental cues (wind, snow, slope, light) and to contribute to tree growth over more than one year by storing temporary reserves (Déjardin et al, 2010). In angiosperm trees, different cell types fulfill these three functions. Vessels and fibres are involved, respectively, in water conduction and mechanical support, while parenchyma cells, organized in rays, are involved in the radial transfer of assimilates between phloem and xylem, their temporary storage as starch or lipids, and their remobilization at the new season. As fibres and vessels account for the major part of the xylem cell population, wood is

mostly made of the secondary cell walls of dead cells.

1.2 Wood is composed of lignified secondary cell walls of dead fibres

The majority of plant cells have only primary cell wall, while wood cells are characterized by thick lignified secondary cell wall (SCW). The cell starts producing the SCW after the primary cell wall is complete and the cell has stopped expanding. The SCW is located between the primary cell wall and the plasma membrane and contains three different layers S1, S2 and S3 (Fig. I-1), each composed of a network of long bundles of cellulose microfibrils, oriented at a fixed angle, and cemented in an amorphous matrix of hemicelluloses – xylans mainly –, and lignins. In smaller quantities, proteins and pectins are also present (Déjardin et al, 2010). From a

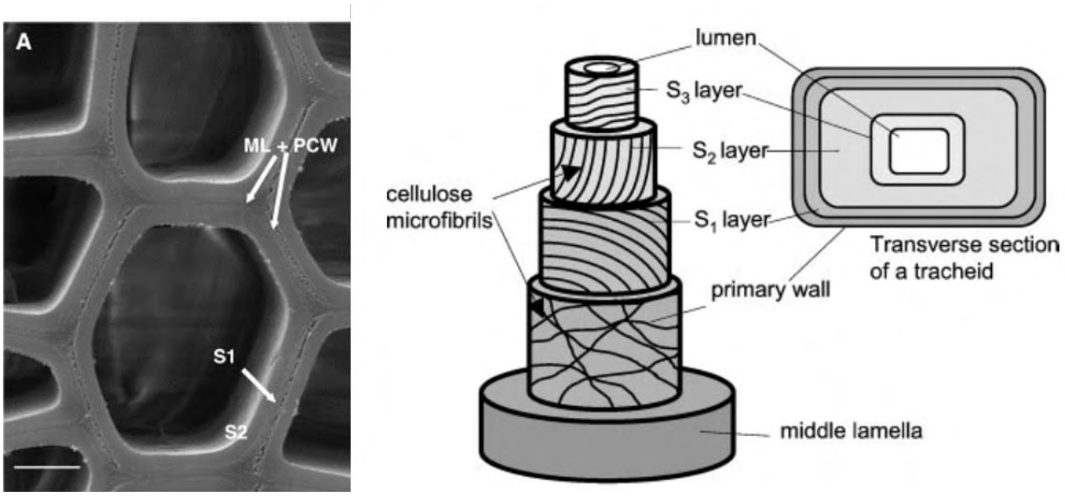


Fig. I-1 Three-dimensional structure of the secondary cell wall of a tracheid (xylem cell). The cell wall is divided into different layers, each layer having its own particular arrangement of cellulose MFs, which determine the mechanical and physical properties of the wood in that cell. These MFs may be aligned irregularly (as in the primary cell wall), or at a particular angle to the cell axis (as in layer S1, S2, and S3). The middle lamella ensures the adhesion between cells. (Adapted to Plomion *et al.* 2001)

mechanical point of view, the SCW can indeed be regarded as a composite material, like reinforced concrete. The composition of the SCW typically comprises 40-50% of cellulose, 25% of hemicellulose, 25-35% of lignin, pectin and cell wall proteins. The

three layers differ from one another with respect to thickness and the orientation of their cellulose microfibrils (Plomion et al. 2001). The microfibril angle (MFA) in the S2 layer, which is the thickest one, is a parameter widely used in wood technology. The biosynthesis of the main polymers will be presented below.

1.3 Economic importance and end-uses of wood

Wood represents the most important natural and endlessly renewable resource for humans providing timber (e.g. for house building, furniture, packaging), fibres (for pulp, paper, plywood) and energy (firewood). It is also an important sink for excess atmospheric CO_2, one of the major causes of global warming due to the greenhouse effect (Demura and Fukuda 2007). Wood is expected to play a significant role in the future as a renewable and environmentally cost effective alternative to fossil fuels. As the world population is predicted to reach over 9 billion in 2050, the global demand for wood for renewable energy, building and pulp and paper will grow rapidly (Mauriat et al, 2014).

The major obstacle to the use of lignocellulosic biomass for the pulp industry and for bioethanol production resides in its recalcitrance to degradation, due to the structure and composition of the lignified secondary cell walls (SCW) (Séguin 2011). For the plant, the presence of lignin confers rigidity and also protects the cell wall polysaccharides from pathogens and microbial degradation (Vanholme et al. 2010). This high resistance to degradation is also one of the most important industrial limitations, where lignin impairs the accessibility of cellulose during kraft pulping as well as during saccharification, reducing in the latter the yield of fermentable sugars and rendering the whole process costly. The economic importance of wood for the pulp industry and more recently for the production of bioethanol have driven many researches worldwide aiming at improving SCW degradation (Pauly and Keegstra 2008; Burton and Fincher 2014; Carpita 2012).

2. 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 show large differences not only between hardwoods and softwood and between different tree species, but individual trees also exhibit huge variations in their wood properties. 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). Anatomical, chemical and physical differences in wood characteristics are not only common from tree to tree within the same area, but are also present within a single tree (Plomion et al., 2001). This difference in wood properties is important for the end-uses of wood as a raw material. Wood plasticity occurs during development and in responses to environmental cues.

Developmental wood plasticity:

The age of the cambium has an important effect on the type of wood produced. Indeed, juvenile wood is formed during the rapid early growth of a tree (usually first 10–12 years of the tree life but depending on the species). The young cambium located in the crown of adult trees also produces juvenile wood. Once this period is finished another type of wood, called mature wood, is produced by the mature cambium (Mauriat et al, 2014). Mature wood differs from juvenile wood by having thicker cell walls, narrower cell lumens, larger cellulose microfibril angles and a higher specific density. In terms of chemical composition, mature wood shows higher cellulose and lower lignin contents (Mauriat et al, 2014; Zobel and Sprague 1998).

Environmental plasticity:

Environmental cues trigger the formation of different types of wood, which can be found within a single tree genotype.

In reaction to gravitropic (wind, slope) or light stimuli, trees develop a reaction wood at the upper side of the inclined axes in order to reorient their axes, trunks and branches, and to allow their adaptation and harmonious development in their environment (Pilate et al, 2004). In angiosperm trees, it is called tension wood (TW), because it is formed in zones of the tree held in tension – the upper side of a leaning stem. At the macroscopic level, a high eccentricity is often observed in the transverse section of stem, suggesting that cell divisions are activated on tension wood side. Tension wood generally has fewer, smaller vessels whereas fibres are significantly longer. The most obvious and striking feature of TW is the presence in fibres of an inner gelatinous cell-wall layer. This so-called G-layer consists of almost pure crystalline cellulose in parallel microfibrils to the long cell axis (Jourez et al. 2001). The lignin content of the TW cell walls is generally decreased (Pilate et al. 2004; Furuy et al. 1970; Habrant et al. 2003). The changes in the biophysical properties and morphology of reaction wood cell walls have a negative impact on the quality of wood products and their potential industrial uses. For instance, TW is a problem for the solid wood industry as it increases longitudinal, radial and tangential shrinkage during the drying process (Plomion et al., 2001).

Nitrogen availability has been described to influence growth and development as well as xylogenesis. Fiber morphology, SCW structure and composition were modified in response to high N supply, including lignification pattern (Pitre et al. 2007). Further mRNA profiles analysis showed that nitrogen fertilization had overlapping effects with tension wood formation (Pitre et al. 2010). Moreover, using pedigree of pseudo-backcrossed hybrid poplar (*Populus trichocarpa* × *Populus deltoides*), Novaes et al. (2009) have shown that N fertilization significantly increased all growth traits as well as the amount of cellulose and hemicelluloses in the wood whereas a decrease of the lignin content was observed.

Other environmental cues such as temperature, drought have been shown to impact wood properties (Gindl et al. 2014; Searson et al. 2004; Moura et al. 2010).

3. Wood develops during secondary growth

Wood, also called secondary xylem, consists of three main cell types: xylem tracheary (vessel) elements, xylary fibres, and xylem parenchyma cells (Fig. I-2). Tracheary elements, which facilitate water and solute transport between organs, and fibres, which provide structural support for the plant, both possess thick secondary cell walls. Xylem parenchyma cells lack well-defined secondary cell walls and are implicated in a variety of biological processes, including aiding the lignification of secondary cell walls in neighbouring vessel elements and fibres (Pesquet et al, 2013; McCann et al. 2001).

Wood derives from a secondary meristem, the vascular cambium, and is manufactured

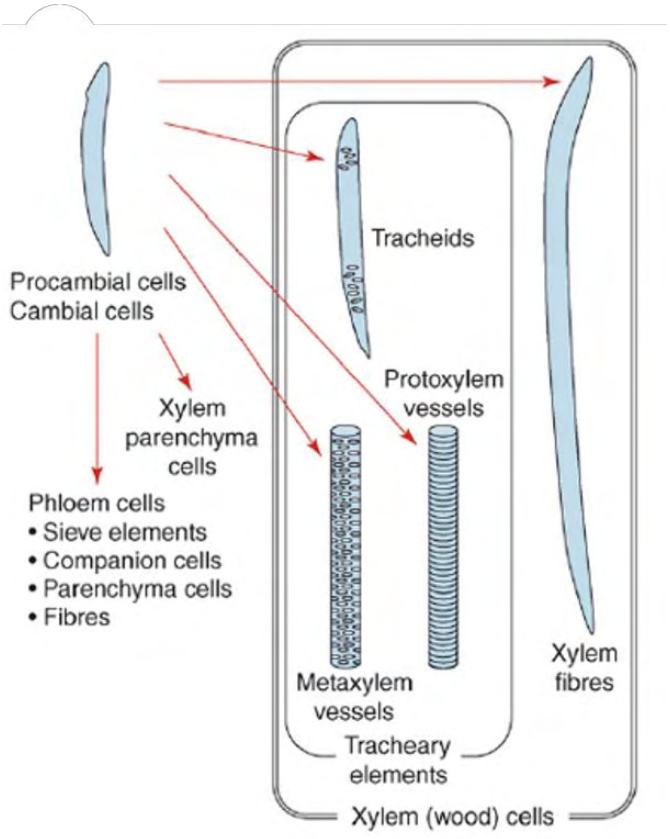


Fig. I-2 Schematic model of xylem (wood) formation. Procambial cells and daughter cells produced by cambial initials differentiate into phloem cells and xylem (wood) cells. Xylem (wood) cells include tracheary elements and fibres. Tracheids and vessels are constituents of tracheary elements. Two types of vessels are observed in angiosperms: protoxylem vessels that commonly have annular and spiral secondary wall thickenings and metaxylem vessels that usually have reticulate and pitted thickenings. (Adapted to Demura and Fukuda 2005)

by a succession of four major steps, including cell division, cell expansion (elongation and radial enlargement), secondary cell wall thickening (involving cellulose, hemicellulose, cell wall proteins, and lignin biosynthesis and deposition) (Plomion et al. 2001). In addition, vessels undergo programmed cell death (PCD), resulting in a continuous system of adjoining hollow cells that function in water/solute transport. During fibre development, PCD is delayed, allowing for more extensive thickening and lignification of secondary cell walls, consistent with a primary role for this cell type in providing structural support.

The vascular cambium is derived from the procambium, and parenchyma cells that regain the capacity to divide and form a layer of meristematic cells located between the primary xylem and phloem in the vascular bundles. Whereas the procambium provides a source of vascular stem cells during primary growth, the vascular cambium ensures the same role during secondary growth, a prominent phenomena in woody plants where it is responsible for the diametral growth of the plant. It is especially important for the development of trees, ensuring their perennial life through the regular renewal of functional xylem and phloem. An increase in the amount of vascular tissues mediated by cambium is one of the characteristics that distinguish dicotyledons and gymnosperms from monocots.

In woody species, the vascular cambium presents unique characteristics since it is composed of two distinct cell types, the fusiform initials and the ray initials, which differentiate into several particular cell types closely interconnected to constitute the secondary xylem or wood, a complex tridimensional tissue (Chaffey, 2002). The fusiform initials divide length-wise and produce secondary vascular tissues through periclinal divisions in a position-dependent manner: on the inner side, wood elements (mostly tracheids in gymnosperms, but also vessel elements, vessel-associated cells, axial parenchyma and fibers in dicotyledons) and, on the outer side, phloem cells (sieve tubes, and, in dicotyledons, companion cells, axial parenchyma, and fibers) (Fig. I-3). Compared to phloem mother cells, xylem mother cells having a greater rate of division,

so there is more secondary xylem than phloem in the tree. Anticlinal (radial) divisions of the fusiform initials also produce daughter cells similar to mother cells and ensure the harmonious increase in circumference of the cambium. Radial initials give rise to rays that are essential to the translocation of nutrients between phloem and xylem (Du and Groover 2010; Plomion et al. 2001).

The economic importance of secondary growth and wood formation has focused considerable research attention on the function of the vascular cambium in tree species, but this question has also been actively studied in *Arabidopsis thaliana*. With several advantages such as genomic resources, *Arabidopsis* has emerged as a useful model for investigating the secondary growth.

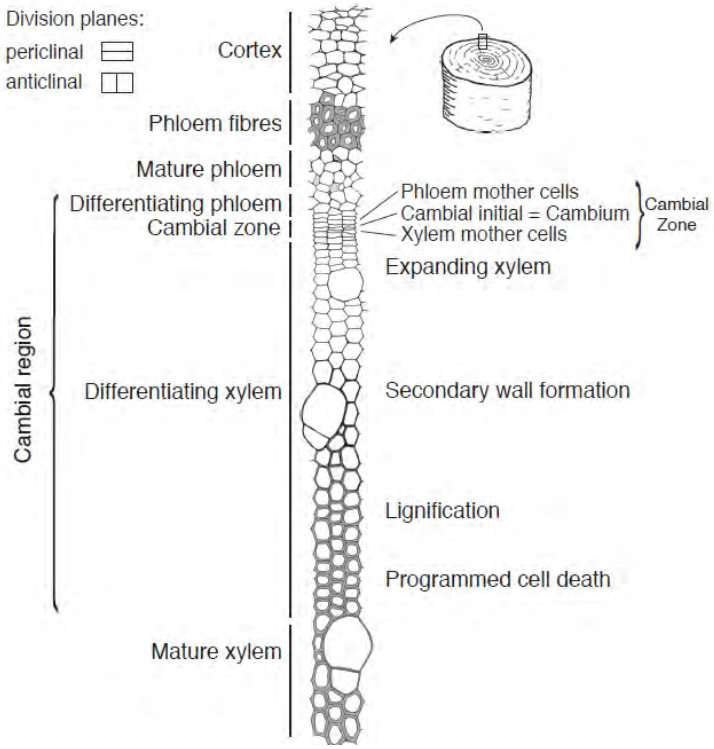


Fig. I-3 Terminology of wood-forming tissues. (Schrader et al. 2003)

4. *Arabidopsis* as a model to study secondary growth

Although this herbaceous plant does not normally undergo extensive secondary growth, two regions of vascular cambium—zones of fascicular cambium and the neighbouring zones of interfascicular cambium—are found within the *Arabidopsis* inflorescence stem (Fig. I-4). Despite their proximity to each other and their apparent similarity, these two cambial niches have different developmental origins. The fascicular cambium is derived from the procambium that developed within the original vascular tissue as it was formed during the primary growth of the stem. It is not surprising, therefore, that radial differentiation of fascicular cambial cells gives rise to the full range of both xylem and phloem cell types. The inflorescence stem fascicular cambium can thus be considered functionally analogous to the circumferential vascular cambium of woody plants.

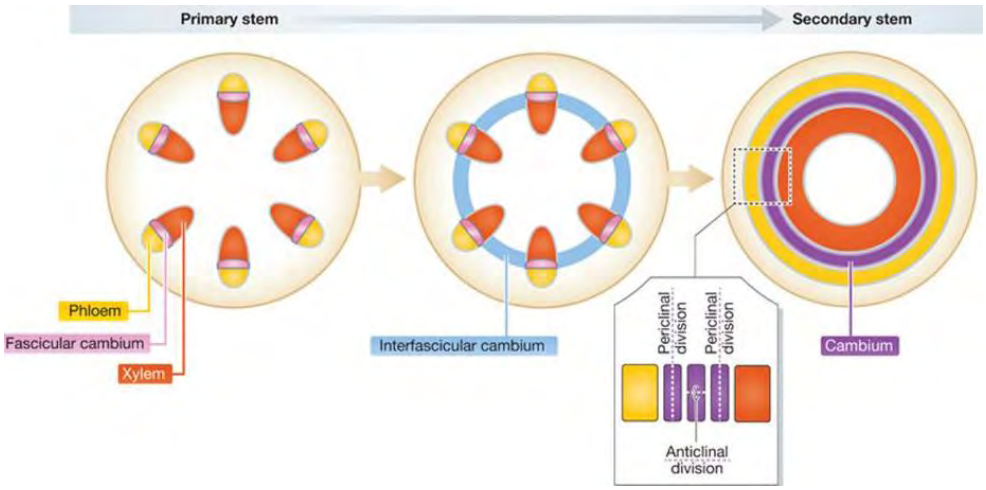


Fig. I-4 Schematic illustration of the primary and secondary stem anatomy in *Arabidopsis*. The primary stem exhibits disconnected vascular bundles with procambium. In the secondary developmental phase, this procambium turns into a fascicular cambium and the cells between bundles become an interfascicular cambium. Fascicular and interfascicular cambia interconnect to each other and establish a cambium in a circular form. (Miyashima et al. 2013)

The interfascicular cambium, on the other hand, is thought to arise through the *de novo* recruitment of interfascicular parenchyma cells as primary growth in the stem slows. It represents a specialized vascular meristem that gives rise exclusively to the structurally important interfascicular fibres.

In contrast to the bundled arrangement of vascular tissues and intervening cambial cells in *Arabidopsis* inflorescence stems, the true stem (hypocotyl) of *Arabidopsis* forms a circumferential vascular cambium that can also undergo substantial secondary growth under certain conditions to form so-called '*Arabidopsis* wood' (Chaffey *et al*., 2002). Examination of cross-sections of this 'secondary growth' hypocotyl reveals an axial tissue arrangement of large metaxylem tracheary elements dispersed among thick xylem fibres, similar to that observed in woody stems of trees (Fig. I-5) (Chaffey *et al*., 2002).

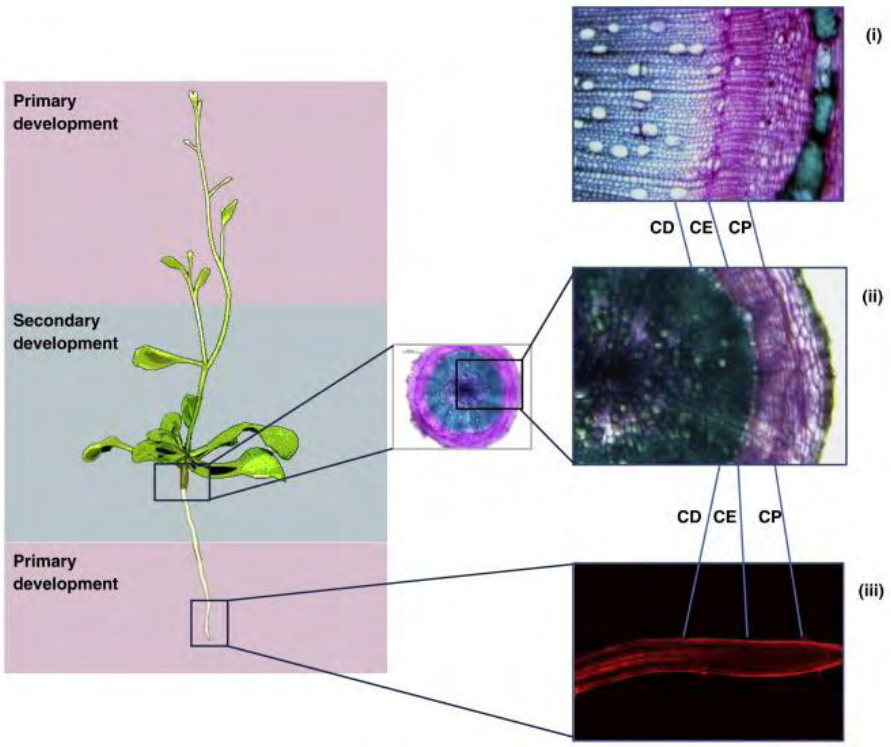


Fig. I-5 Anatomy of primary and secondary vascular tissues in *Arabidopsis* and in comparison with *Populus* wood-forming region. The same developmental phases can be identified in the cross section of **(I)***Populus* stem and **(II)** 6-week-old *Arabidopsis* hypocotyl, both representing secondary vascular development, as well as **(III)** in the confocal microscopy image of *Arabidopsis* root tip, representing primary vascular development. CP: cell proliferation zone; CE: cell expansion (elongation) zone; CD: cell differentiation zone. (Zhang et al. 2010)

The most prominent difference between the secondary xylem in true woody plant stems and in these *Arabidopsis* 'secondary' stems is the lack of well-defined xylem rays and the determinate growth habit in the latter. Despite these differences, *Arabidopsis*

provides a useful model for understanding the process of secondary growth in woody plants.

5. Regulation of cambium activity and wood specification

Plant hormones play important roles to regulate wood formation. For example, ethylene and cytokinin acts as positive regulators of wood formation, promoting cambial cell proliferation (Love et al. 2009; Nieminen et al. 2008); gibberellin promotes both cell division and xylem fibers elongation (Mauriat and Moritz 2009); Auxin is involved in different steps: formation and maintenance of the pro(cambium), cell division and cell specification. Its role will be examined in more details in a specific paragraph. It is worth noting that the different hormones also show crosstalking in control cambial growth and differentiation (Sorce et al. 2013).

5.1 Regulation of cambium identity and activity

The procambium is formed and maintained by directed auxin flow in the precursors of procambium cells (Donner et al. 2009; Ohashi-Ito and Fukuda 2010). In the leaf veins, auxin flow is directed by polar localization of an auxin efflux carrier, PINFORMED 1 (PIN1) and the expression of PIN1 precedes differentiation of the procambium precursor cells (Donner et al. 2009). *ATHB8*, a class III HD-ZIP transcription factor, is required for preprocambial development and procambium differentiation, and its expression depends on the activity of the auxin response factor *MONOPTEROS* (*MP*) (Donner et al. 2009). There are five *HD-ZIP III* genes in *Arabidopsis*, *PHB/ATHB14*, *PHV/ATHB9*, *REV/IFL1*, *ATHB8* and *CNA/ATHB15*. They have been shown to regulate the number of procambium cells by promoting xylem differentiation during vascular development (Baima et al. 2001; Carlsbecker et al. 2010; Ilegems et al. 2010).

The maintenance of the pluripotent identity of the cambium is crucial for continuous meristem activity. Current evidence indicates that a similar molecular mechanism regulating shoot apical meristem (SAM) and root apical meristem (RAM) is likely

applicable in cambial meristems. Recent findings have identified a peptide-receptor-transcription factor signaling pathway, TDIF/CLE41/CLE44-TDR/PXY-WOX4, that plays a crucial role in the maintenance of the vascular meristem organization during secondary growth (Hirakawa et al. 2010), controls cambium maintenance (Fig. I-6) (Hirakawa et al. 2010; Eitchells and Turner 2010; Elo et al. 2009). The small peptide TDIF (Tracheary element Differentiation Inhibitory factor), that is processed from the translated products of *CLE (CLAVATA3/EMBRYO SURROUNDING REGION-related) 41/CLE44* in *Arabidopsis*, is produced in the phloem; it interacts with its receptor, the receptor-like kinase TDR/PXY (TDIF RECEPTOR/PHLOEM INTERCALATED WITH XYLEM), which is expressed in (pro)cambium. This TDIF-TDR/PXY signal promotes proliferation of procambial cells and suppresses their xylem differentiation. PXY-CLE41/CLE44 loop suppresses the differentiation of vascular stem cells into xylem cells and thus maintains the undifferentiated pluripotent status in the cambium. *WOX4*, a *WUSCHEL HOMEOBOX RELATED* gene, mediates this ligand-receptor signaling to regulate the maintenance of (pro)cambium cells. Cambium activity is reduced in the hypocotyl and inflorescence stem of the *wox4* mutant, indicating that

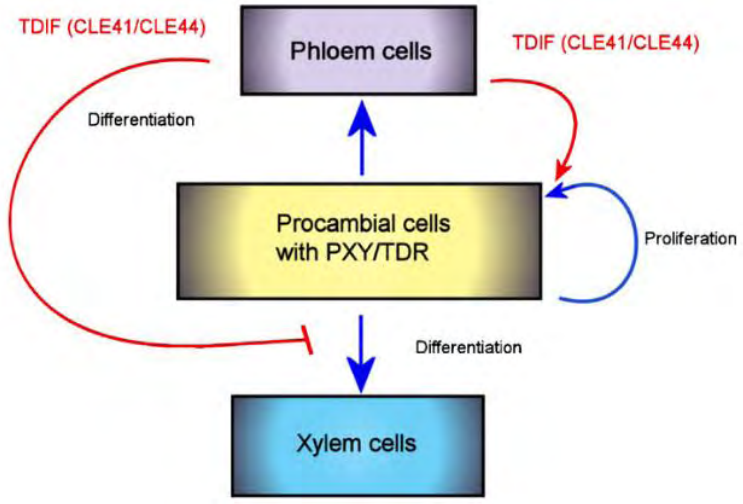


Fig. I-6 A current model suggests that the procambial cell population is maintained by the autoregulation loop of CLE41/44–PXY signalling system in a non-cell-autonomous manner. According to this model a signal molecule, TDIF (CLE41/CLE44), is produced and secreted by phloem and their neighboring cells. Procambial cells perceive the TDIF signal through the PXY (TDR) receptor. TDIF signal promotes maintenance and proliferation of cambial (stem) cells and inhibits their differentiation into xylem cells. (Elo et al. 2009)

WOX4 regulates, but is not required to establish, the meristem (Hirakawa et al. 2010; Suer et al. 2011). Another WOX-family gene, *WOX14*, acts redundantly with *WOX4* to regulate cambial cell proliferation (Etchells et al. 2013). The TDIF/CLE41/CLE44-TDR/PXY-WOX4 signaling pathway appears to act downstream of auxin signaling in regulating cambial cell proliferation and seems to be evolutionarily conserved between both woody and herbaceous species, as it has been described in both *Arabidopsis* and *Populus* (Schrader et al. 2004). Two other receptor-like kinases, REDUCED IN LATERAL GROWTH1 (RUL1) and MORE LATERAL GROWTH1 (MOL1), were also found to be activator and repressor of cambium activity, respectively (Agusti et al. 2011b).

In *Arabidopsis*, class I KNOX transcription factors are major regulators of SAM activity. Plants overexpressing a class I KNOX gene, *BREVIPEDICELLUS (BP)/KNAT1*, exhibited impaired lignin deposition while *bp* mutants show ectopic lignification in the inflorescence stem, indicating that BP regulates xylem cell differentiation during vascular development (Mele et al. 2003). In *Populus*, an ortholog of *BP*, *ARBORKNOX2 (ARK2)*, is expressed in both SAM and cambium region. Overexpression of *ARK2* in *Populus* results in the expansion of the cambium region and inhibition of differentiation of tracheary elements and fibres in secondary xylem as well as phloem fibres. Knock-down of *ARK2* results in early appearance of lignified secondary xylem and thicker SCWs (Du et al. 2009).

KANADI genes (*KAN1-4*) belong to one clade of the GARP transcription factor family. Gain-of-function *KANADI* alleles result in a loss of cambium activity, while loss-of-function result in increased cambium activity. For example, ectopic *KAN1* expression results in a complete loss of vascular tissue development. *KANADI* and *Class III HD-ZIP* show function mutual antagonism through effects on the auxin flow (Ilegems et al. 2010).

5.2 Regulation of xylem specification

Several *Arabidopsis* mutants defective in the radial patterning of vascular tissues also have defects in the radial (abaxial–adaxial) patterning of leaves, suggesting the involvement of HD-ZIP III and KANADI transcription factors in the xylem cell type patterning in root and stem (McConnell et al. 2001; Kerstetter et al. 2001; Emery et al. 2003; Cano-Delgado et al. 2010). Class III homeodomain-leucine zipper (HD-ZIP III) transcription factors, including *AtHB8*, *CORONA/AtHB15* (*CNA*), *PHABULOSA* (*PHB*), *PHAVOLUTA* (*PHV*), and *REVOLUTA* (*REV*), are expressed in procambium, cambium and developing xylem, and play important roles in cambium and xylem differentiation in *Arabidopsis*. Loss-of-function mutants *rev* plants display disruption of the differentiation of interfascicular fibers and secondary xylem, while *AtTHB8* and *AtHB15* appear to have functions antagonistic to the *REV* in vascular formation (Prigge et al. 2005). *HD-ZIP III* gene expression is positively regulated by auxin, and the expression of *AtHB8* is directly activated by the auxin-response transcription factor *MONOPTEROS (MP)* at *Arabidopsis* preprocambial stages (Donner et al. 2009). The stability of transcripts are regulated by microRNA165/166, and all the gain-of-function mutations disrupt the miR165/166 target sequence (Demura and Fukuda 2007). In additional, brassinosteroids can also activate *HD-ZIP III* expression and thus promote xylem differentiation (Ohashi-Ito and Fukuda 2003). In *Populus*, knock down of *POPCORONA* results in abnormal lignification, while overexpression of a miRNA-resistant *POPCORONA* results in delayed lignification of xylem and phloem fibres during secondary growth (Du et al. 2011). When the microRNA resistant form of *popREVOLUTA* (*PRE*) (orthologous to *AtREV*) is overexpressed, it leads to abnormal cambium formation in the cortex (Robischon et al. 2011).

Other transcription factors such as class I KNOX and bHLH TFs are also involved in secondary growth. *Populus ARK2*, the ortholog of *Arabidopsis BP/KNAT1*, is expressed in both SAM and cambium region, and *ARK2* overexpression *Populus* displays expansion of the cambium region and inhibited xylem differentiation. Knock-down of

ARK2 results in early appearance of lignified secondary xylem and thicker SCWs (Du et al. 2009). *TMO* quadruple mutants completely lost vascular tissue in roots, and *TMO5* and *LHW* co-expression plants induced dramatic periclinal divisions within the vasculature of roots (De Rybel et al. 2013).

6. Secondary cell wall formation and its transcriptional regulation

6.1 Biosynthesis of the three main polymers

Cellulose

Cellulose, as a major structural of the cell wall, is the most abundant biopolymer synthesized on land. It is consists of several linear polymer chains of ß-1, 4 linked glucose residues. The fundamental structure units of cellulose are the microfibrils (MFs), which are formed through intra- and inter-chain hydrogen bonds between the different parallel linear glucan chains. In higher plants, the synthesis of cellulose is believed to be catalyzed by cellulose synthase (CesA), organizing into cellulose synthase complexes (CSCs) localized on the plasma membrane (Brett 2000; Saxena and Brown 2005). CSCs exist as rosettes, containing six subunits arranged in a hexagonal structure, and it has been postulated that each of the six rosette subunits contains six cellulose synthase (CesA) proteins. It means that each rosette therefore has a total of 36 CesA proteins (Doblin et al. 2002; Li et al. 2014).

CesAs are int*Egr*al plasma membrane proteins with multiple transmembrane domains and a central catalytic domain (Sethaphong et al. 2013; Slabaugh et al. 2014). The *Arabidopsis* genome encodes ten *CesA* genes, which are roughly classified into two groups: *AtCesA1*, *AtCesA3*, *AtCesA6*, *AtCesA2*, *AtCesA5* and *AtCesA9* are involved in primary cell walls cellulose synthesis (Li et al. 2014); while *AtCesA4*, *AtCesA7* and *AtCesA8* are required for cellulose synthesis in secondary cell walls (Taylor et al. 2003). Among primary *CesAs, AtCesA1* and *AtCesA3* are essential, while At*CesA6* is redundant with *AtCesA2, AtCesA5* and *AtCesA9*. For example, the temperature sensitive mutant, *radial swelling root 1* (*rsw1*) caused the rosette to disintegrate,

revealing the importance of *AtCesA1* in the formation of an intact CSC during the synthesis of primary cell walls (Arioli et al. 1998). Secondary cell wall cellulose synthesis gene mutation results in collapsed or irregular xylem cells, and loss of function of any single secondary *CesA* of *AtCesA4 (irx5)*, *AtCesA7 (irx3)* or *AtCesA8 (irx1)* causes a complete xylem morphology defect, indicating the CSCs require at least three subunits to function (Taylor et al. 2000). Furthermore, 18 *CesA* gene loci have been identified in *Populus*, and based on sequence analysis, *PtiCesA4*, *PtiCesA7-A* and *-B*, and *PtiCesA8-A* and *-B* were found to be homologous to *AtCesA4*, *AtCesA7* and *AtCesA8*, respectively (Song et al. 2010).

In additional, many non-CESA encoding genes have also been identified to be involved in cellulose synthesis, such as the korrigan gene and sucrose synthase gene (*SuSy*). Mutations in *korrigan* gene (*KOR1*), which encodes a putative membrane bound β-1, 4 endoglucanase, resulted in defects in cellulose synthesis in both primary and secondary cell walls (Liebminger et al. 2013; Paredez et al. 2008; Szyjanowicz et al. 2004). UDP-glucose, a substrate of CesA in the synthesis of cellulose glucan in plants, can be made from sucrose through a reaction catalyzed by sucrose synthase (SUSY) (Amor et al. 1995; Haigler et al. 2001).

Hemicellulose

Hemicelluloses, which present along with cellulose in almost all plant cell walls, are polysaccharides containing many sugar monomers including glucose, xylose, mannose, galactose, rhamnose, and arabinose. They are a heterogeneous group of polysaccharides, consisting of xyloglucan, xylans, mannans and glucomannans, and β-(1→3, 1→4)-glucans. Xylans is a predominant hemicellulose and accounts for ~20% of the total dry weight of wood in angiosperms woody plants, while mannans such as galactoglucomannans are the major hemicellulose in the SCW of gymnosperms (Capek et al. 2002; Mellerowicz et al. 2001).

Hemicelluloses are synthesized by a variety of glycosyltransferases (GTs) located in

the Golgi membranes. In *Arabidopsis*, many GT genes have been identified involved in xylan biosynthesis through the characterization of collapsed xylem *Arabidopsis* mutants, such as *AtFRA8/IRX7*, *AtIRX8*, *AtIRX14*, *AtIRX14-L*, *AtF8H*, *AtPARVUS*, *AtIRX10* and *AtIRX10-Like* (*IRX10-L*), *AtIRX9*, *AtIRX9-L*. *AtIRX9*, *AtIRX14*, *AtIRX10* and *AtIRX10-L* are involved in synthesis of the β-D-(1→4)-xylan backbone elongation, while the *AtFRA8/IRX7*, *AtIRX8*, *AtF8H*, and *PARVUS* are thought to play a role in forming this oligosaccharide (Brown et al. 2009; Lee et al. 2007a; Lee et al. 2007b; Pauly et al. 2013; Persson et al. 2007; Wu et al. 2009). In poplar, a few *GT* genes involved in xylan biosynthesis have also been studied. *PoGT47C*, *PoGT8D* and *PoGT43B* show high sequence similarity to *AtFRA8*, *AtIRX8* and *AtIRX9*, respectively, and it has also been shown that *PoGT43B* and *PoGT47C* can rescue the defects of *Arabidopsis Atirx9* and *Atfra8* mutants, respectively. Moreover, poplar *GT8E*, *GT8F/PdGATL1.1* and *PdGATL1.2* are identified as functional orthologs of *Arabidopsis PARVUS* since they are able to rescue the *Arabidopsis* irregular xylem phenotype (Kong et al. 2009; Lee et al. 2009a, b).

Lignin

Lignin is phenolic biopolymer produced by the dehydrogenative polymerization of essentially three different hydroxycinnamyl alcohols, the *p*-coumaryl, coniferyl, and sinapyl alcohols, that differ in the d*egr*ee of methoxylation at the C_3 and C_5 positions of the aromatic ring giving raise to *p*-hydroxyphenyl (H), guaiacyl (G), syringyl (S), and units, respectively (Fig. I-7) (Baucher et al. 2003).

Lignin embeds the polysaccharide matrix of the SCW to make them rigid and impervious. Lignin content and monomeric composition vary widely among different species, individuals, cell types, and cell wall layers and are influenced by developmental and environmental cues. In general, lignins from gymnosperms and related species are rich in G units and contains low amounts of H units, whereas dicots lignins are mainly composed of G and S units (Weng and Chapple 2010). H-units are elevated in softwood compression wood and may be slightly higher in grasses (Boerjan

et al. 2003; Weng and Chapple 2010). Vessel element walls are rich in G lignin whereas fibre walls contain S–G lignin. In addition to developmentally programmed deposition of lignin, some biotic and abiotic stress such as wounding and pathogen infection can also induce the lignin biosynthesis (Vanholme et al. 2010).

The monolignols are synthesized from phenylalanine through the general phenylpropanoid and monolignol-specific pathways via a series of enzymes (Fig. I-7). Phenylalanine ammonia lyase (PAL) converts phenylalanine to *trans*-cinnamic acid at the entry point of the phenylpropanoid pathway. Cinnamate 4-hydroxylase (C4H) and *p*-coumarate 3-hydroxylase (C3H), two cytochrome P450 enzymes, catalyze the first two aromatic hydroxylation reactions. 4-Hydroxycinnamoyl-CoA ligase (4CL)

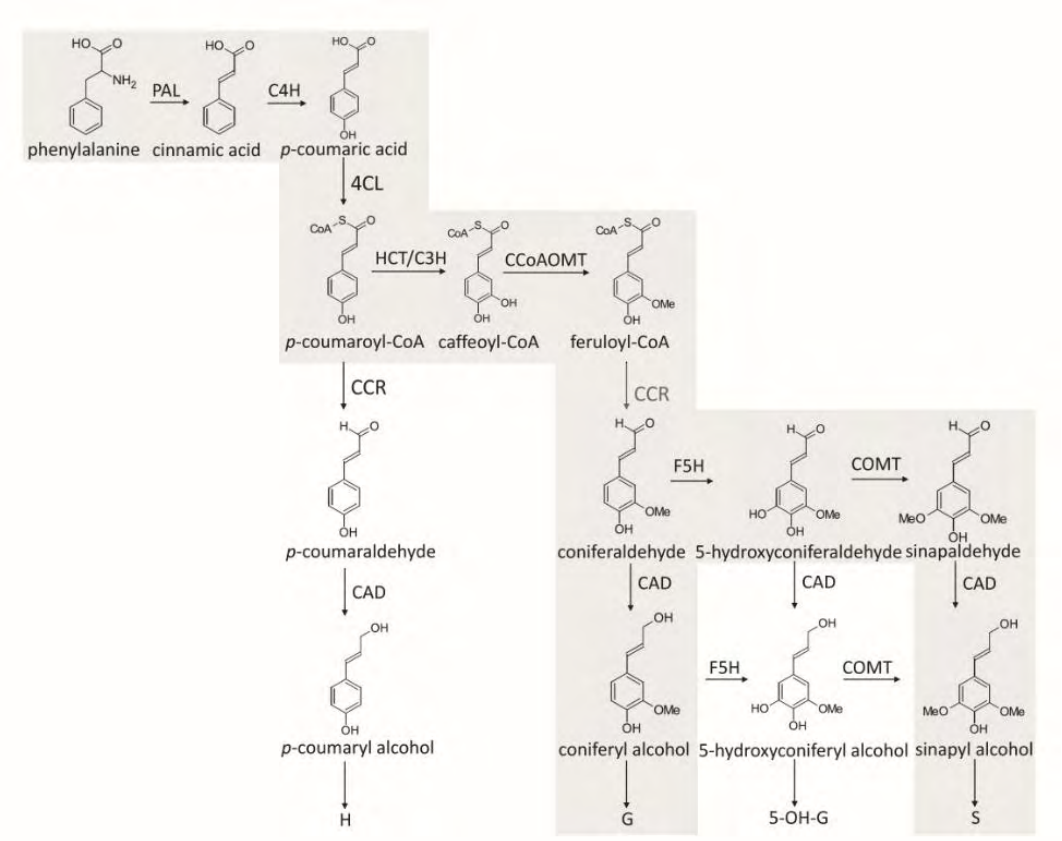


Fig. I-7 Phenylpropanoid and monolignol biosynthetic pathways. The grey box represents the phenylpropanoid and monolignol biosynthetic pathways that are generally accepted for angiosperms. The general phenylpropanoid pathway starts with *PAL* and ends with *CCoAOMT*, whereas the monolignol-specific biosynthesis starts with *CCR*. (Adapted to Van Acker et al. 2013)

activates *p*-coumaric acid to the activated thioester form *p*-coumaroyl CoA, which represents the most important branchpoint within the central phenylpropanoid biosynthesis in plants (Vogt 2010). Hydroxycinnamoyl-CoA: shikimate hydroxycinnamoyl transferase (HCT), which belongs to the BAHD acyltransferase superfamily, catalyzes the formation of *p*-coumarate esters using shikimate or quinate as an acyl acceptor. Caffeoyl-CoA O-methyltransferase (CCoAOMT) catalyzes the first transmethylation reaction in the phenylpropanoid pathway, synthesizing feruloyl-CoA from caffeoyl-CoA (Fig. I-7). The first step of the monolignol-specific pathway starts with cinnamoyl-CoA reductase (CCR) (Van Acker et al. 2014). CCR and Cinnamyl alcohol dehydrogenase (CAD), two oxidoreductases, convert the hydroxycinnamoyl-CoA esters to their corresponding alcohols (Fig. I-7). Ferulate 5-hydroxylase (F5H) and caffeic acid O-methyltransferase (COMT) are two S lignin biosynthesis specific enzymes.

The different monolignol subunits are synthesized in the cytoplasm and then transported to the cell wall for subsequent polymerization (Wang et al. 2013). Recently, it has been reported that AtABCG29, an ATP-binding cassette transporter, acts as a *p*-coumaryl alcohol transporter (Alejandro et al. 2012). However, as a large transporter gene family in plant, ABC transporters for coniferyl alcohol and sinapyl alcohol remain to be identified and characterized (Wang et al. 2013). Following transport to the secondary cell wall, monolignols are oxidized and subsequently polymerized by peroxidases and laccases to form lignin (Baucher et al. 2003).

6.2 The SCW transcriptional network

The SCW formation is a critical step of wood formation. Through the recent studies in *Arabidopsis* as well as other species such as *Populus*, *Eucalyptus* and pine, several key transcriptional switches have been identified to regulate the entire differentiation process. A simplified SCW regulatory network is shown in Fig. I-8.

NAC transcription factors

NAC (NAM, ATAF1/2 and CUC2) transcription factors is a large gene family in plant and they have been extensively studied in several plants especially in model plant *Arabidopsis* and poplar. Some of them have been shown to play pivotal roles in wood formation. Recent studies show that four of them (VND6, VND7, NST1 and SND1/NST3) are the top of the master switches regulating several downstream

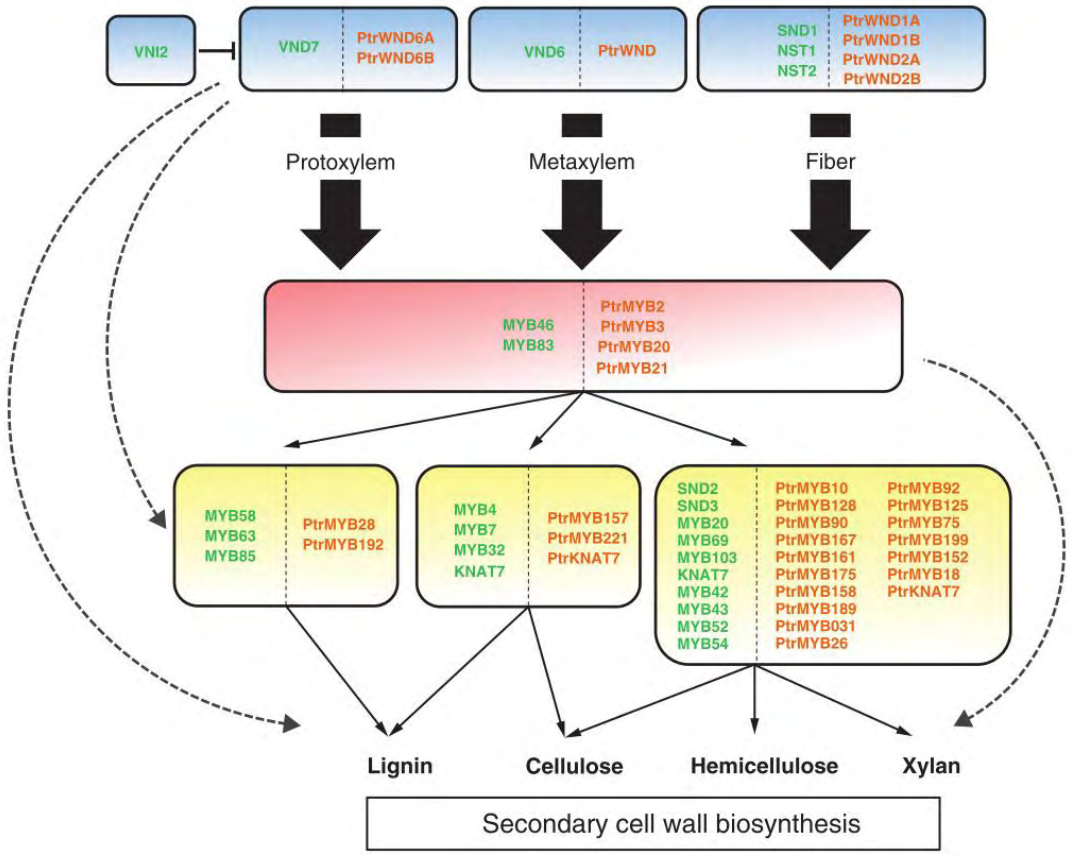


Fig. I-8 Transcriptional regulatory network controlling secondary cell wall biosynthesis in *Arabidopsis* and *Populus*. *Arabidopsis* genes are presented in green and their *Populus* orthologs in orange. The NAC genes (blue boxes) function as first-level master switches; they induce expression of the second-level master switches, MYB46 and MYB83 (red box), which in turn activates a plethora of downstream TFs (yellow boxes), as well as many genes directly involved in secondary wall biosynthesis. The MYB target TFs promote the biosynthesis of lignin, cellulose, hemicellulose and xylan biosynthesis. A multilevel feed-forward loop structure is int*Egr*ated in the transcriptional network: both NAC and MYB master switches directly induce expression of many of the same genes (dashed arrows). (Zhang *et al.* 2014a)

transcription factors which lead to the secondary cell wall formation. VASCULAR-RELATED NAC-DOMAIN6 (VND6) and VND7 are transcription switches for plant metaxylem and protoxylem vessel formation, respectively (Kubo et al. 2005), while NAC SECONDARY WALL THICKENING PROMOTING FACTOR1 (NST1) and SECONDARY WALL-ASSOCIATED NAC DOMAIN PROTEIN1 (SND1/NST3) are two mater regulators promoting fiber differentiation (Mitsuda et al. 2007; Zhong et al. 2007). The other two NAC transcription factors SND2 and SND3, which are indirect and direct targets of SND1 respectively, positively regulate the secondary wall thickness of both xylary fibers and interfascicular fibers (Hussey et al. 2011; Zhong et al. 2008). In *Populus*, *PtrWND* (*PtVNS*) genes, homologs of the four *Arabidopsis* NAC transcription factors, which are highly expressed in developing secondary xylem tissue, has been shown to be master switches activating the secondary cell wall biosynthesis (Ohtani et al. 2011; Zhong and Ye 2010).

MYB transcription factors

Several MYB transcription factors have also been shown to be important regulators of secondary cell wall formation. First, *Eucalyptus Eg*MYB2 was found to bind the *EgCCR* and *EgCAD2* gene regulatory regions and act as a positive regulator of secondary cell wall formation and lignin biosynthesis (Goicoechea et al. 2005). Then in *Arabidopsis*, the MYB proteins have been identified as the direct target of secondary wall NACs master switches regulating secondary wall biosynthesis. For example, *MYB46* and *MYB83*, both of which are SND1 direct targets, were found to function redundantly as second-level master switches regulating secondary wall biosynthesis. Over-expression of *MYB46* or *83* in *Arabidopsis* induces activation of secondary wall biosynthetic genes for cellulose, xylan and lignin and results in ectopic deposition of secondary walls in cells that are generally parenchymatous, while mutations of *MYB46* and *83* results in lack of secondary wall thickening (McCarthy et al., 2009; Zhong and Ye, 2012). More recently, four *Populus* orthologs *PtrMYB2*, *PtrMYB3*, *PtrMYB20* and *PtrMYB21* were also proved to be the direct target of the master regulators PtrWND

and function as second-level master switches of wood formation (McCarthy et al. 2010; Zhong et al. 2013). In addition, other MYB transcription factors such as MYB58, MYB63 and MYB85, were shown to be regulated by the SND1 close homologs NST1, NST2, VND6, and VND7 and their downstream target MYB46 and/or MYB83 to control biosynthesis of cellulose, hemicellulose, xylan and lignin (Zhong et al. 2008; Zhou et al. 2009).

The Class II KNAT7 transcription factor

KNAT7, which has been shown to be one of the direct target of both SND1 and MYB46 (Ko et al. 2009; Zhong et al. 2008), is a regulator of secondary cell wall biosynthesis. Loss-of-function *knat7* mutants displayed exhibit both *irx* and enhanced fiber cell wall thickness, while overexpression plants displayed opposite phenotype with decreased secondary wall thickening in interfascicular fibres (Li et al. 2012). So KNAT7 acts as a transcriptional repressor rather than an activator in regulating secondary wall biosynthesis. KNAT7 can also interact with members of the Ovate Family Protein (OFP) transcription co-regulators, such as OFP1 and OFP4. The *ofp4* mutant exhibited similar *irx* and fibre cell wall phenotypes as *knat7*, so it has been proposed that KNAT7 forms a functional complex with OFP proteins to regulate aspects of secondary cell wall formation (Li et al. 2011).

Part II Auxin-key regulators of plant growth and development

1 Auxin homeostasis

The plant hormone auxin has a crucial role in plant development, and this depends on the graded distribution of auxin. The auxin homeostasis maintains the active auxin at an optimal level for growth and development. This homeostasis depends on the combined effects of all the process: biosynthesis, transport, conjugation and degradation.

1.1 Auxin biosynthesis

IAA is mainly produced in the shoot apical meristems and young leaves of plant. In 2001, Ljung and colleagues proved that even all parts of the *Arabidopsis* seedling can synthesize low levels of IAA, but young leaves have the highest biosynthetic capacity (Ljung et al. 2001). Auxin biosynthesis in plants is fairly complex. Multiple pathways that contribute to *de novo* auxin biosynthesis have been postulated. There are two major widely accepted pathways to synthesize IAA: the Trp-independent and Trp-dependent pathways (Mano and Nemoto 2012). In Trp-independent pathway, indole-3-glycerol phosphate or indole is the likely precursor, but little is known about the biochemical

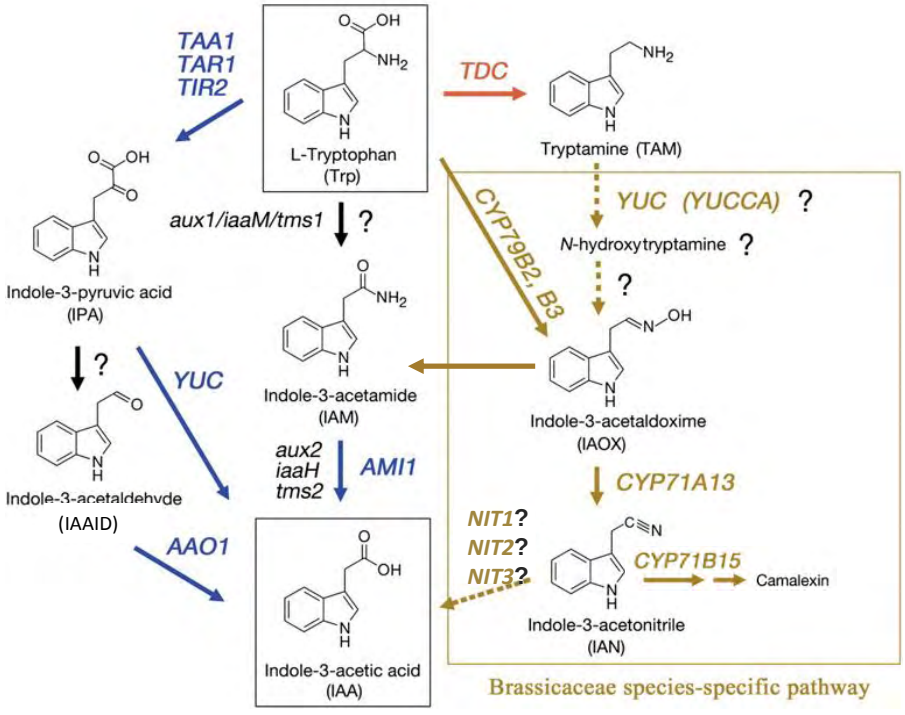


Fig. I-9 Presumptive pathways for IAA biosynthesis in plants. Blue arrows indicate steps for which the gene and enzymatic function are known in the tryptophan-dependent IAA biosynthetic pathway. Red arrows indicate the indole alkaloid and serotonin biosynthetic pathway. Mustard-coloured arrows indicate the Brassicaceae species-specific pathway. Black arrows indicate steps for which the gene(s) and enzymatic function(s) are unknown. Dashed mustard-coloured arrows indicate steps for which the gene and enzymatic function(s) remain poorly understood. Letters in italics show genes involved in the conversion process. Lower case letters in italics indicate bacterial genes. (Adapted to Mano and Nemoto, 2012).

pathway to IAA. In Trp-dependent pathway, there are four proposed pathways for biosynthesis of IAA from tryptophan in plants (Fig. I-9): i) the IAM (indole-3-acetamide) pathway; ii) the IPA (indole-3-pyruvate acid) pathway; iii) the TAM (tryptamine) pathway; and iv) the IAOx (indole-3-acetaldoxime) pathway.

1.2 Auxin conjugation and degradation

Although free IAA is the biologically active form of the hormone, the majority of auxin in plants is found in a conjugated state. These Auxin conjugations have been identified in both higher and lower plants and are considered hormonally inactive. There are three main types of conjugation: 1) ester conjugates with sugar moieties; 2) amide conjugates with amino acids; 3) high molecular weight conjugates with peptides or proteins also via amide bonds (Ludwig-Muller 2011). These conjugates are involved in auxin storage, transport and degradation to maintain auxin homeostasis. Auxin ester conjugates, including IAA-glucose, IAA-myo-inositol and molecular weight IAA-glycan, have been identified from a variety of plant species (Normanly 2010).

1.3 Auxin transport

Auxin synthesis occurs in some specific tissues, and the distribution of auxin is generally realized by two distinct pathways. Throughout the plant, IAA is transported over long distances (from the source tissues to the roots) through the phloem by mass flow. In addition, there is also another auxin transport system called auxin polar transport (PAT), which is a slower, regulated and carrier-mediated cell-to-cell directional transport system (Zazimalova et al. 2010). This PAT within plant tissues appears to be unique to auxin, as it has not been detected for any other signaling molecule. PAT provides essential directional and positional information for developmental processes, such as vascular differentiation, apical dominance, organ development and tropic growth (Blakeslee et al. 2005). The PAT is based on chemiosmotic hypothesis. Because IAA is a weak acid, it exists in a charged anionic form (IAA^-) in the cytoplasm (pH ~ 7). In the more acidic cell wall environment (pH ~

5.5), ~ 15% of the molecules are in the uncharged form (IAAH), which can enter the cell by lipophilic diffusion across the plasma membrane (Zazimalova et al. 2010). This pH differential between the cytoplasm and wall means that auxin can move into but not out of the plant cells. So the auxin ion requires facilitators or transporters to exit cells (Fig. I-10). The majority (85%) of IAA remains in its dissociated form (IAA^-) and would require a carrier for its active uptake across the cell. Three types of auxin transporters have been studied. The AUX1/LAX (AUXIN RESISTANT 1/LIKE AUX1) are influx auxin carriers (Bennett et al. 1996), while PIN (PIN-FORMED), and ABCB (ATP Binding Cassette subfamily B) are efflux auxin carriers (Friml et al. 2002; Yang and Murphy 2009). Recently, a novel PIN like family of auxin transport facilitators termed PILS (PIN-LIKES) has been discovered by *in silico* studies and appears to be involved in the regulation of auxin homeostasis in *Arabidopsis* (Barbez et al. 2012). More recently, WALLS ARE THIN1 (WAT1) has also been demonstrated to be a vacuolar auxin transport facilitator mediate auxin homoeostasis (Ranocha et al. 2013).

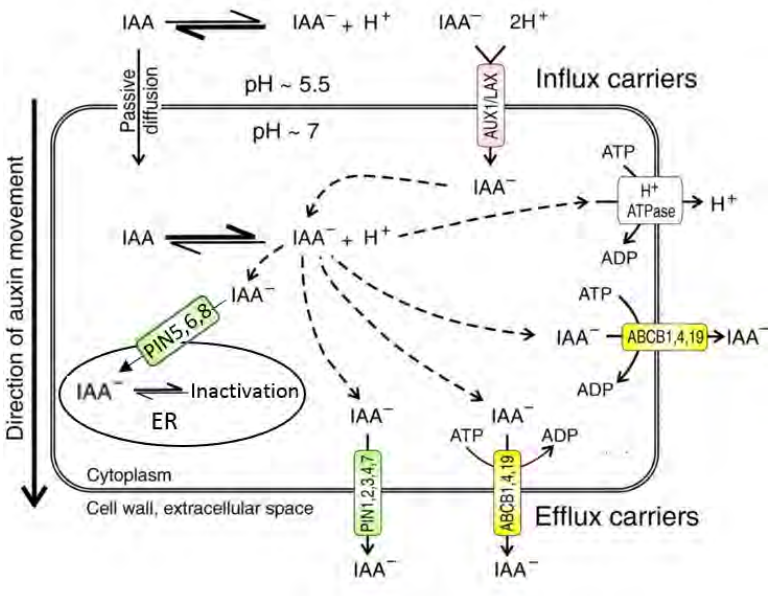


Fig. I-10 Model of intercellular auxin transport. (Adapted to Friml, 2010)

2 Auxin signaling

In different plant species, auxin can rapidly regulate gene expression in a few minutes

and control many plant developmental processes. Auxin acts by regulating transcription through the action of at least three protein families called the TIR1/AFB F-box proteins, the Aux/IAA transcriptional repressors, and the ARF transcription factors. The best characterized pathway model is the $SCF^{TIR1/AFB}$-proteasome complex mediated the degradation of Aux/IAA which releases auxin response factor (ARF) to induce early auxin-responsive gene expression (Chapman and Estelle 2009). Auxin signaling is also involved in non-transcriptional regulation, but the mechanism is not as well characterized. AUXIN BINDING PROTEIN 1 (ABP1) is the best characterized auxin binding proteins and implicated in non-transcriptional auxin signaling.

2.1 Auxin receptors

By now, two distinct classes of auxin receptors have been recognized: TIR1 (transport inhibitor response 1) and AFB (auxin-signaling F-box proteins) that control transcriptional responses to auxin, and ABP1 (auxin binding protein 1), that controls a wide variety of growth and developmental processes.

TIR1 (transport inhibitor response 1)/AFB (auxin-signaling F-box proteins)

TIR1 and additional five AFB proteins, AFB1 to AFB5, are six F-box proteins, and function as auxin receptors (Dharmasiri et al. 2005a; Dharmasiri et al. 2005b). These six proteins can be grouped into three subclusters according to the phylogenetic analysis, TIR1/AFB1, AFB2/AFB3 and AFB4/AFB5. They form SCF (SKP1-Cullin-F-box) protein complex with SUPPRESSOR OF KINETOCHORE PROTEIN 1 (SKP1 or ASK1), CULLIN1 (CUL1) and RING BOX1 (RBX1) (Cardozo and Pagano 2004). SCF complex is the largest family of E3 ubiquitin – protein ligases and catalyze the ubiquitination of diverse regulatory and signalling protein destined for proteasomal degradation.

Previous study showed TIR1/AFB1-3 function as auxin receptors and interact with Aux/IAA in an auxin dependent manner. Genetic experiments indicate that reducing the number of functional TIR1/AFB1-3 proteins in plant results in increasing resistance to

exogenous auxin. Comparison of single and multiple mutants in the *TIR1*, *AFB1*, *AFB2*, and *AFB3* genes and protein interaction analysis with Aux/IAA showed that TIR1 appears to make the largest contribution to auxin response followed by AFB2. Double, triple and quadruple mutants showed a gradual decrease in auxin response and defects in hypocotyl, cotyledon, root growth, leaf morphology and flow development. Although *TIR1/AFB1-3* show close relationships, their biochemically function are distinct. For example, AFB1 and AFB2 do not rescue the *tir1* mutant even regulated by the TIR1 promoters (Parry et al. 2009).

Members of the TIR1 and AFB2 groups act as positive regulators of auxin signaling, while both AFB4 and AFB5 function as auxin receptors based on *in vitro* assays (Greenham et al. 2011; Hu et al. 2012). AFB4 and AFB5 are the major targets of the picolinate herbicides in *Arabidopsis*. Genetic experiments showed that *afb4-2* single and *afb4-2 afb5-5* double mutants displayed opposite effects compare to *tir1 afb2* mutant, such as elongated petioles and longer hypocotyls. In addition, the *afb4-2* mutant had shorter roots and produced more lateral roots/primary root length, so it suggests that AFB4 has a role in anchor or adventitious root production (Hu et al. 2012).

ABP1 (auxin binding protein 1)

ABP1 was discovered from maize coleoptiles 40 years ago (Hertel et al. 1972) and it was a soluble 22kDa glycoprotein that specifically binds to auxin. The majority of ABP1 protein localizes to the endoplasmic reticulum (ER), while some ABP1 (~22%) is also found on the plasma membrane (Timpte 2001) and works as a functional receptor for auxin perception. ABP1 is essential for early embryogenesis, root development, leaf expansion, cell morphogenesis, and subcellular distribution of PIN auxin transporters. In addition, auxin binding by ABP1 coordinates the cytoskeleton structure by regulating the activity of two GTPases ROP2 and ROP6 (Sauer and Kleine-Vehn 2011). However, it is not known how ABP1 transmits the auxin signal to regulate these cytoplasmic responses, while the mechanism of the auxin recognition site is quite different from TIR/AFB receptors. The crystal structure analysis of ABP1 suggested that ABP1 forms

a dimer through the ß-barrel structure, and auxin binds within a predominantly hydrophobic pocket with a zinc ion coordinated with three histidines and a glutamate in each subunit (Fig. I-11A) (Woo et al. 2002). Until recently, it was proved that the plasma membrane–localized TMK1 receptor–like kinase is functionally and physically associated with ABP1 at the cell surface to regulate auxin- and ABP1-mediated activation of ROP GTPase signaling in non-transcriptional signaling way (Fig. I-11B) (Xu et al. 2014). It was also demonstrated that ABP1 is involved in the control of Aux/IAA homeostasis (Fig. 11B) (Tromas et al. 2013).

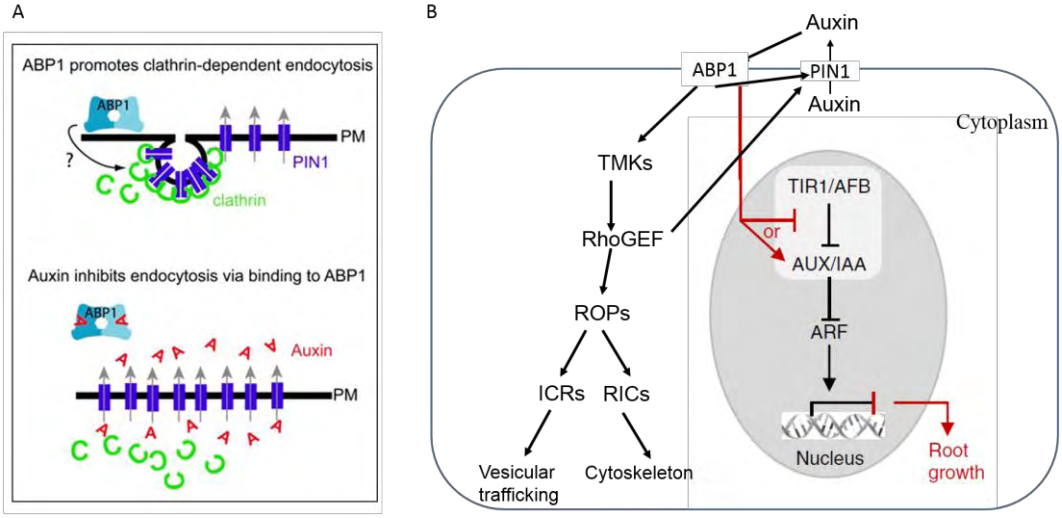


Fig. I-11 ABP1 auxin signaling pathway. (A) ABP1 Regulates Clathrin-Dependent Endocytosis. Auxin binding to ABP1 reduces clathrin-dependent endocytosis, leading to higher retention of PIN proteins and, subsequently, higher auxin efflux rates. (Sauer and Kleine-Vehn 2011) (B) Model for ABP1 action on TIR1/AFB–AUX/IAA pathway and for coordination of root growth. ABP1 promotes stabilization of AUX/IAA repressors and maintenance or restoration of transcriptional repression by acting negatively on the TIR1/AFB pathway (Adapted to Xu et al 2014 and Tromas et al. 2013).

2.2 Aux/IAAs family

Several classes of auxin-responsive gene families have been identified and characterized in different plants, such as *Aux/IAA*, *GH3* (*Gretchen Hagen3*) and *SAUR* (*Small Auxin-up RNA*) (Guilfoyle et al. 1998). Most of their transcription level can be directly altered without novo protein synthesis, so these gene families are termed

primary/early auxin-responsive genes (Abel and Theologis 1996). *Aux/IAA* gene as a primary response to auxin was firstly identified in soybean and pea (Ainley et al. 1988; Theologis et al. 1985; Walker and Key 1982). Then *Aux/IAA* genes have been found in many plant species, including *Arabidopsis* (Overvoorde et al. 2005), poplar (Kalluri et al. 2007), tomato (Audran-Delalande et al. 2012), maize (Wang et al. 2010) and rice (Jain et al. 2006).

Canonical Aux/IAA proteins contain four highly conserved amino acid sequence motif designated domain I, II, III and IV (Fig. I-12), although several proteins lacking one or

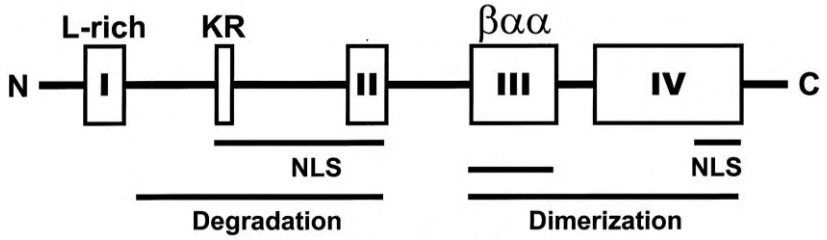


Fig. I-12 Structural and functional domains of Aux/IAA proteins. (Colón-Carmona A et al. 2000)

more of these domains are also included in the family. Domain I which contains a conserved Leu repeat motif (LxLxLx) similarity within the EAR motif is responsible for the repressing activity of the proteins (Tiwari et al. 2004), and it can also form co-repressor with TOPLESS through this domain (Szemenyei et al. 2008). The instability of Aux/IAA proteins involves conserved domain II composed by a hydrophobic motif GWPPV (Worley et al. 2000), which mediates interaction with SCF^{TIR1} ubiquitin-ligase complex for targeting of Aux/IAA to the 26S proteasome (Gray et al. 2001; Kepinski and Leyser 2004). An amino acid substitution in the d*Egr*on sequence (mutation of G or P in GWPP motif) results in a stabilization of the affected protein and a decrease in auxin response (Ouellet et al. 2001; Reed 2001; Tian et al. 2003). Domain III and IV which are located in the carboxyl part of the protein, can mediate homo- and hererо-dimerization with other Aux/IAA members, as well as dimerization with ARFs which contain similar domains (Kim et al. 1997; Ouellet et al. 2001; Ulmasov et al. 1997). The Aux/IAA proteins do not contain DNA binding domain, so they cannot bind to the

auxin-responsive elements (*AuxREs*) directly and regulate auxin-mediated gene expression by interaction with ARF proteins, which have a B3-like DNA binding domain (Guilfoyle and Hagen 2007). In addition to conserved domains, Aux/IAA proteins display putative nuclear localization signals (NLS). NLS are composed by both the bipartite structure of conserved basic doublet KR between domains I and II and basic amino acids in domain II, and the SV40-type NLS located in domain IV.

Aux/IAA function has been mostly studied in *Arabidopsis* through the characterization of gain-of-function mutants (mutation in domain II). *axr5/iaa1 Arabidopsis* mutant showed defects in root and shoot tropisms, leaf morphology and inflorescence (Yang et al. 2004). Both gain-of-function and loss-of-function mutant of *AtIAA3* is characterized. Gain-of-function shy2/iaa3 mutant showed a reduced lateral root formation and increased root hair and slow root re-orientation, while loss-of-function mutant displayed the opposite phenotype, indicating that *IAA3* plays a central role in auxin regulation of root growth and gravitropism. *Shy2/iaa3* mutant also causes short hypocotyls, leaf formation in dark and short inflorescence stems (Tian and Reed 1999). *axr2/iaa7* mutant causes agravitropic root and shoot growth, a short hypocotyl and stem, and auxin-resistant root growth (Nagpal et al. 2000). The *bdl/iaa12 Arabidopsis* mutant fails in initiating root meristem during early embryogenesis showing a specific role of *AtIAA12* during this process (Hamann et al. 2002). *iaa28-1*, *slr/iaa14* and *msg2-1/iaa19* suppress lateral root development (Fukaki et al. 2002; Rogg et al. 2001; Tatematsu et al. 2004). The *axr3/iaa17 Arabidopsis* mutant presents shorter primary roots and formation of adventitious roots with altered root gravitropism. Moreover, leaves are darker and smaller than wild type and plants are dwarf showing altered apical dominance (Leyser et al. 1996). 18 of 29 *Aux/IAA* T-DNA insertion mutants in *Arabidopsis* have been characterized, but none of them displays visible developmental defects, even for the double or triple mutants of closely homologs, suggesting that there is an extensive functional redundancy among *Aux/IAA* gene family members (Overvoorde et al. 2005).

2.3 ARFs (auxin response factors) family

Auxin response factors are transcription factors that regulate the expression of auxin response genes. *Arabidopsis ARF* family is composed of 23 members (Okushima et al. 2005b). Most ARFs contains a conserved N-terminal DNA-binding domain (DBD), a variable middle transcriptional regulatory region (MR) and a carboxy-terminal dimerization domain (CTD) (Fig. I-13) (Guilfoyle and Hagen 2007). The DBD of ARFs bind with specificity to TGTCTC auxin response elements (AuxRE) in promoter regions of auxin response genes. The middle region (MR) is very variable. The QSL-

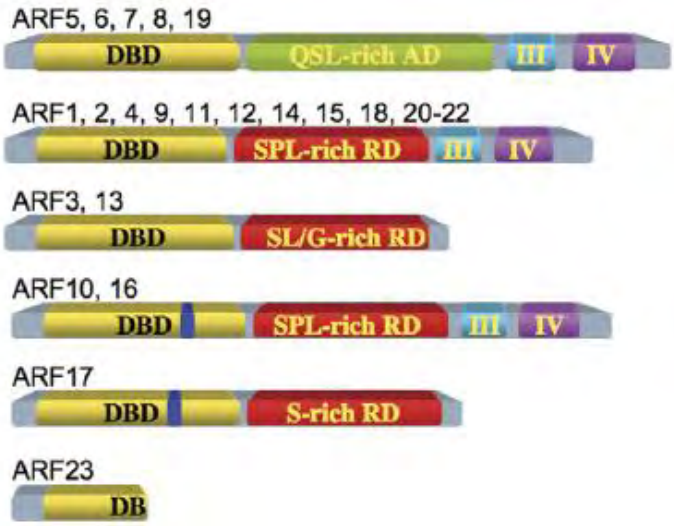


Fig. 1-13 The classification and structures of the ARF protein family in *Arabidopsis* (Guifoyle and Hagen, 2007).

rich (glutamine, serine, leucine) middle region ARFs function as an activatior whereas the S-rich (serine), SPL-rich (serine, proline, leucine) and SL/G-rich (serine, leucine and/or glycine) middle region ARFs are repressor. Five *Arabidopsis* ARFs, ARF5-8 and 19, function as transcriptional activators when tested on auxin response genes in transfected protoplasts, while the remainder of the ARFs function as transcriptional repressors. The carboxy-terminal dimerization domain is similar to those found in Aux/IAA proteins designated to domains III and IV, makes it can function as a homo- and hetero- dimerization among the ARFs or with several Aux/IAA proteins. ARF TF family also has been identified and characterized in a number of species, such as maize

(Xing et al. 2011), rice (Wang et al. 2007), tomato (Kumar et al. 2011, Zouine et al. 2014), poplar (Kalluri et al. 2007), etc.

To understand their function in plant growth and development, the expression pattern of *ARF* genes have been studied. *ARF* genes display specific expression pattern, *ARF1* and *ARF6* mainly expressed in developing flowers, *ARF2* in floral organs and seedlings, *ARF3* and *ARF4* in reproductive and vegetative tissues, *ARF5* in embryos and vascular tissues, *ARF7* in roots, seedlings and embryos, *ARF8* in seedlings and flowers, *ARF12* in seeds, *ARF16* in roots, leaves, vascular tissues and embryos, *ARF19* in roots and seedlings (Guilfoyle and Hagen 2007). A number of closely related *ARFs* (*ARF12*, *-14*, *-15*, *-17*, *-20-23*) which located near the centromere of chromosome 1 were expressed exclusively in the embryo (Rademacher et al. 2011). Moreover, lots of *Arabidopsis arf* mutants have been studied. *arf2* exhibits pleiotropic developmental phenotypes including delayed flowing, leaf senescence and floral abscission, thick and long inflorescence as well as defective apical hook formation, and *arf1* mutations enhanced many *arf2* phenotypes (Ellis et al. 2005; Okushima et al. 2005a). *arf3/ettin* displays abnormal development of flowers (Sessions et al. 1997). *arf4* mutant shows no phenotype whereas *ett arf4* double mutants presents transformation of abaxial tissues into adaxial ones in all aerial parts. It has been proven that *ARF3* and *ARF4* have overlapping functions in leaf and floral organ patterning and cooperate with *KANADI* genes to specify abaxial cell identity (Pekker et al. 2005). *ARF5*/ *MONOPTEROS (MP)* gene is involved in embryo development and vascular tissue formation (Hardtke and Berleth 1998). The *AtARF7*/*NPH4* gene regulates hypocotyls tropism and the sensitivity of auxin and ethylene (Harper et al. 2000). *arf7 arf19* double mutant displays strong auxin-related phenotype not observed in single mutant, including severely impaired lateral roots formation and abnormal gravitropism in both hypocotyl and root (Okushima et al. 2005b). *AtARF6* and *AtARF8* both regulate flowering process by coordinating stamen development, petal expansion, anther dehiscence and gynoecium maturation. *arf6 arf8* double mutant has complete arrested flower development before bud opening and are completely infertile, while *arf6* and *arf8* single mutants display

delayed flower maturation and reduced fertility (Wu et al. 2006). *arf10* and *arf16* single mutant present no phenotype whereas the *arf10 arf16* double mutant has root cap defects and abnormal root gravitropism (Wang et al. 2005b). Screens for T-DNA insertion mutations have identified mutations in at least 18 of the *Arabidopsis ARF* genes, but few of these have distinctive growth or developmental phenotypes beyond those genes identified by forward genetic approaches, suggesting that there are functional redundancies among the ARF proteins (Okushima et al. 2005b).

The expression of *ARF* genes have been shown to respond to environmental and hormonal signals. Such as *ARF1* and *8* response to light, while *ARF4, 5, 16, 1*9 response to auxin. Studies in *Arabidopsis* also have found that several *ARF* genes are post-transcriptionally regulated by micro-RNA (miR). *ARF10, 16* and *ARF17* display site-recognition of miR160 (Axtell and Bartel 2005), *ARF6* and *ARF8* regulated by miR167 (Wu et al. 2006), while *ARF2, 3* and *ARF4* are post-transcriptionally regulated by a trans-acting-small interfering RNAs (ta-siRNA) called TAS3 ta-siRNA (Fahlgren et al. 2006).

2.4 Protein-protein interactions in auxin signaling

Aux/IAA and ARF are two groups of well-studied transcription factors mediate auxin response, and auxin signaling is mainly regulated by the interactions of Aux/IAA and ARF proteins. The similar conserved domains III and IV present in these proteins allow the formation of Aux/IAA-ARF heterodimers. The effects of auxin depend on its concentration, with high and low doses eliciting different responses. At basal auxin levels, Aux/IAAs are relatively stable, so the (activating) ARFs are bound by Aux/IAAs (Tiwari et al. 2001), which recruit the TOPLESS co-repressor and associated chromatin-modification machinery, thereby inhibiting transcription (Szemenyei et al. 2008). When auxin levels are high, auxin promotes the formation of $SCF^{TIR1/AFB}$ – Aux/IAA complex which in turn facilitates ubiquitination and degradation of the Aux/IAAs (Chapman and Estelle 2009). The degradation of Aux/IAAs results in the release of ARFs which can then activate the transcription of target genes via binding to

the Auxin Responsive Elements (*AuxRE*) present in the promoter regions of auxin-regulated genes (Fig. 1-14) (Guilfoyle and Hagen 2007).

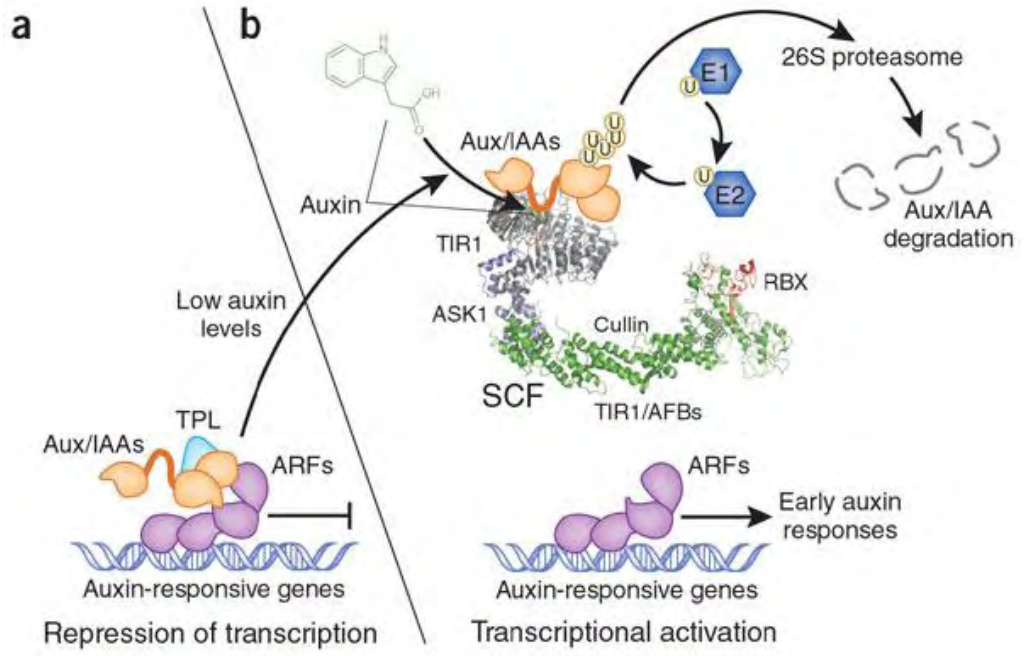


Fig. I-14 Auxin signaling in *Arabidopsis*. SCFTIR1 and related SCF complexes bind auxin and target Aux/IAA proteins for degradation. (a) At low cellular levels of auxin, transcription of auxin response genes is repressed by the Aux/IAAs. (b) When auxin cellular levels increase, auxin binds to TIR1, enhances its affinity for the Aux/IAAs, and promotes their ubiquitination and subsequent degradation, thus permitting the ARF proteins to promote transcription. (Santner, et al. 2009)

29 Aux/IAAs and 23 ARFs protein-protein interactions in *Arabidopsis* have been deeply studied by high-throughput yeast two-hybrid approach (Vernoux et al. 2011). This analysis revealed that 433 interactions among the 1,225 tested combinations between Aux/IAAs and ARFs and themselves. Among the 433 interactions, the majority of Aux/IAA proteins interact with themselves and ARF activators, while ARF repressors have no or very limited interactions with other proteins in the network (Fig. I-15). So it is suggesting that ARF repressors may more act as a competition for *AuxRE* element binding with ARF activators rather than be directly involved in auxin signaling through Aux/IAA regulation (Vernoux et al. 2011; Weijers et al. 2005). Moreover, the in planta analysis also proved that auxin signals are converted into specific responses

by matching pairs of co-expressed ARF and Aux/IAA proteins. AtIAA12 paralog IAA13 could interact with ARF5, regulating embryonic root formation (Weijers et al. 2005), and the specific interaction between AtIAA14 and AtARF7 or AtARF19 is essential for the inactivation of lateral root formation (Fukaki et al. 2006).

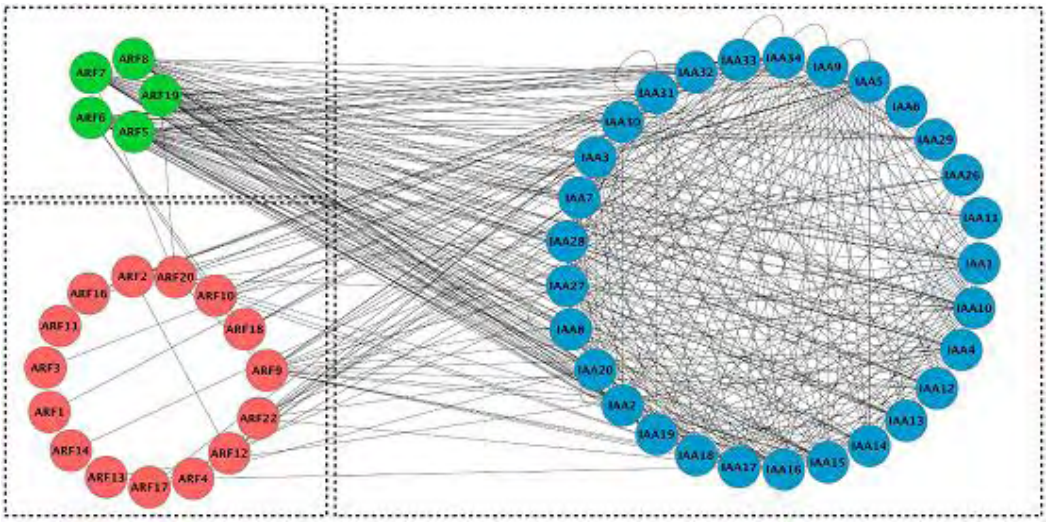


Fig. I-15 The *Arabidopsis* ARF-Aux/IAA interaction map. Most of Aux/IAAs are able to interact with ARF activators while ARF repressors display very few interactions in *Arabidopsis*. (Vernoux *et al*. 2011)

3 Roles of auxin in plant development

Auxin plays critical roles in plant cell growth and in a wide variety of developmental processes. The effects of auxin was first documented by Darwin in 1880, then 45 years later, by Went in 1926. The term generic term “auxin” in fact represents a family of related compounds: indole-3-acetic acid (IAA) is considered to be the major biologically active auxin, while a number of related compounds with auxin activity like indole-3-butyric (IBA), 4-Chloroindole-3-acetic (4-Cl-IAA) and phenylacetic acid (PAA) have also been identified in various plant species (Simon and Petrášek 2011).

3.1 Auxin promotes cell division, and cell expansion

On the cellular level, auxin influences aspects of cell division, cell elongation and cell differentiation according to auxin concentration and cell/tissue-specific responses.

Auxin starvation of suspension cells causes cell division arrest while addition of auxin to the arrest cells will lead to restoration of cell division (Perrot-Rechenmann 2010). Auxin can induce the main gene expression promoted the cell division, such as Cyclin-dependent kinase (CDKA) and D-type cyclins which plays a critical role in G1 to S and G2 to M progression (Koroleva et al. 2004), while reduce the expression of cell division inhibitors such as KRP (Himanen et al. 2002). One of the auxin receptors ABP1 is also critical for auxin regulation of the cell cycle and expansion. Inactivation of ABP1 in tobacco cells leads to cell-cycle arrest in G1 phase even in the presence of auxin (David et al. 2007). Overexpression of ABP1 leads to an increased expansion of some leaf tissues and the null *Arabidopsis ABP1* mutant displays altered cell expansion (Chen et al. 2001a; Chen et al. 2001b). Plant cell extension requires uptake of water and irreversible extension of the cell wall, which includes wall loosening (short time frame) and deposition of new wall material (long time frame). Auxin is one of the major stimuli affecting these mechanisms, and this auxin-dependent cell expansion follows a dose-response curve in which high concentrations are inhibitory (Barbier-Brygoo et al. 1991; Evans et al. 1994).

3.2 Auxin roles in organ patterning

Auxin plays critical role in regulation of organ patterning, such as embryo and vascular pattern, root initiation, apical dominance, tropism and fruit development. The role of auxin in vascular patterning will be reviewed in the next paragraph.

Vascular patterning

Vascular tissues are generated during embryogenesis and organogenesis and response for water and nutrients conduction. Auxin plays a key role in embryonic vascular pattern, and disrupt auxin polar transport result in defects in leaf veins (reviewed in Berleth and Mattsson 2000). In *Arabidopsis, lax2* knockouts exhibited vascular breaks in their cotyledons. Mutations in auxin response related genes such as *MONOPTEROS* (*MP*)/*AUXIN RESPONSE FACTOR 5* (*AFR5*), *AUXIN RESISTANT 6* (*AXR6*), and

BODENLOS (*BDL*)/*IAA12* result in incomplete vascular systems and defects in embryo axis formation and embryonic root (Berleth T 1993; Hamann T 1999; Hobbie L 2000; Przemeck et al. 1996). Auxin can induce xylem tracheary element differentiation in suspension cells under special conditions (Fukuda 1997) and local auxin sources can induce the formation of new vascular strands from parenchymatic cells (Sachs 1986).

Lateral root formation

In most eudicot plants, lateral roots (LRs) initiate from anticlinal cell divisions of pericycle cell layers adjacent to the xylem pole (Bellini et al. 2014; Barlow et al. 2004; Dubrovsky et al. 2001). Recent studies in several model plants have shown that LR formation is dependent on auxin (Peret et al. 2009; Fukaki et al. 2007, Fukaki and Tasaka 2009; Lavenus et al. 2013; Lucas et al. 2008). In plant cells, many auxin-responsive genes are regulated by auxin through Aux/IAA–ARF auxin-signalling modules as described in 2.2.4. Molecular genetic studies with mutants defective in LR formation have shown that several Aux/IAA–ARF modules play important roles in the developmental steps during LR formation. Gain-of-function *slr-1* mutation in the *SLR/IAA14* blocks pericycle cell division for LR initiation, resulting solitary-root phenotype (Fukaki et al. 2002; Vanneste et al. 2005). The *arf7 arf19* loss-of-function double mutant also has a few LRs, but the *arf7* and *arf19* single mutants do produce LRs, indicating that ARF7 and ARF19 have redundant functions for LR formation (Okushima et al. 2005). In addition, the *SLR/IAA14*, *ARF7* and *ARF19* genes are co-expressed in root tissues and SLR/IAA14 interacts with ARF7 and ARF19 in yeast yeast two-hybrid system, which strongly suggested that the stabilized mutant IAA14 constitutively inhibits the activity of ARF7 and 19, thereby repressing the downstream genes for LR initiation (Okushima et al. 2005; Fukaki et al. 2005). Therefore, auxin was proposed to promote the degradation of SLR/IAA14, resulting in the activation of ARF7/19-dependent transcription of the target genes (such as *LBD16/ASL18*) involved in LR initiation (Wilmoth et al. 2005; Okushima et al. 2005; Okushima et al. 2007; Fukaki et al. 2005). At present, several gain-of-functions *Aux/IAA* members, including

IAA28, AXR5/IAA1, SHY3/IAA3, CRANE/IAA18 and MSG2/IAA19 also showed decreased number of LRs, indicating that auxin signalling dependent on these *Aux/IAAs* is necessary for LR formation (Overvoorde et al 2010). More recently, a multiple AUX/IAA–ARF modules regulate lateral root formation has been demonstrated by Goh et al. (2012) (Fig. I-16): IAA28 module regulates LR founder cell specification (De Rybel et al. 2010), then SLR/IAA14-ARF7-ARF19 module regulates nuclear migration and asymmetric cell divisions of the LR founder cells (Goh et al. 2012). Moreover, this

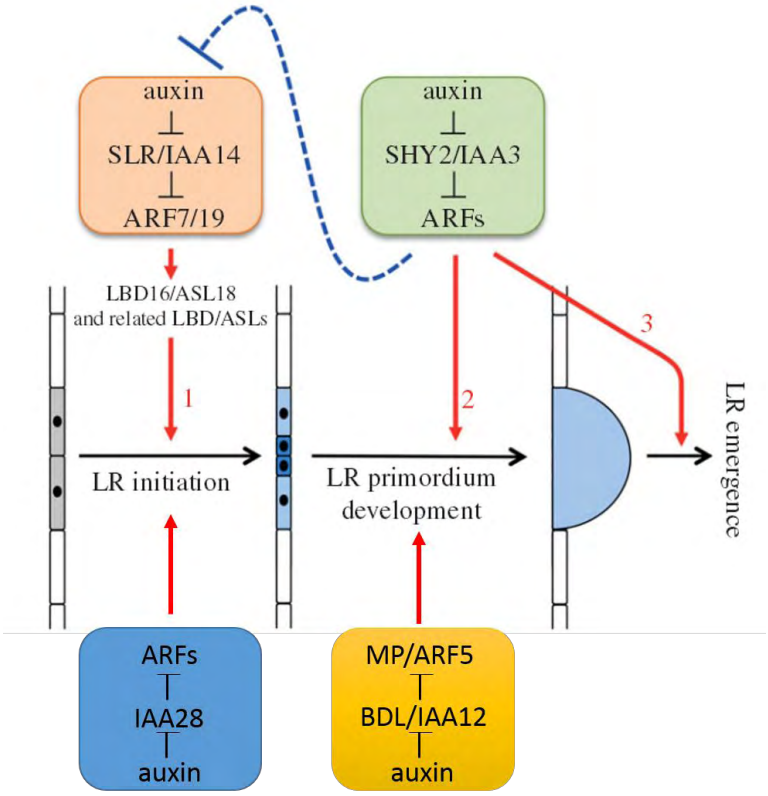


Fig. I-16 Schematic of lateral root (LR) formation regulated by SLR/IAA14–ARF7–ARF19 and SHY2/IAA3–ARFs auxin-signalling modules. LR initiation is controlled by the SLR/IAA14–ARF7–ARF19 auxin-signalling module (orange box) by the activation of LBD16/ASL18 and its related LBD/ASL proteins (red arrow 1). After initiation, the SHY2/IAA3–ARFs signalling module (green box) plays a role not only for LR primordium development and LR emergence after the SLR/IAA14–ARF7–ARF19 module (red arrows 2 and 3) but also for inhibition of SLR/IAA14–ARF7–ARF19-dependent LR initiation in the xylem pericycle cells by affecting auxin homeostasis (blue dotted line). (Goh et al. 2012)

process is also regulated by BDL/IAA12-MP/ARF5 which represses ectopic pericycle cell divisions (Goh et al. 2012). SHY2/IAA3-ARFs positively regulate LR primordium development and in parallel inhibit LR formation by affecting auxin homeostasis (Goh et al. 2012).

Flower and fruit development

As known, auxin plays a critical role in regulating flower and subsequent fruit development (Dharmasiri et al., 2005b; Liscum and Reed, 2002). Auxin accumulation controls the floral primordium initiation (Reinhardt et al., 2000); it is also essential for anther development (Cecchetti et al. 2008) and involved in stamen and pollen development. Several mutants disrupted in either auxin biosynthesis (*yuc2yuc6*), transport (*tir1afb*), or signalling (*arf6arf8, ettin/arf3, mp/arf5*) exhibit flowering defects that are variable but typically involve alterations in organ numbers, organ spacing, and gynoecium morphology (Nagpal et al. 2005; Cecchetti et al. 2008).

Auxin plays a critical role in fruit development, beginning with flower formation, through fruit set, fruit growth and ripening (de Jong et al. 2009b, Sundberg and Ostergaard 2009, Ruan et al. 2012). In *Arabidopsis* alteration of *AtARF8* expression leads to the formation of parthenocarpic fruits (Goetz et al. 2007). In the loss-of-function tomato *iaa9* mutant ovary develops prior to pollination leading to precocious fruit set and marker fruit parthenocarpy (Wang et al. 2005a). Moreover application of auxin results in a delayed fruit ripening (Aharoni et al. 2002; Cohen 1996). Recent data are also uncovering new roles for other Aux/IAA and ARF proteins in fruit growth after fruit set. It was recently demonstrated that tomato *IAA27* and *ARF10* are involved in the regulation of seed development and the size and shape of the fruit (Bassa et al. 2012, Hendelman et al. 2012). Silencing of *ARF4*, which is expressed strongly in tomato pericarp resulted in enhanced accumulation of starch and chlorophyll content, which suggests the involvement of auxin signaling in the control of chloroplastic activity and sugar metabolism in the fruit (Jones et al. 2002, Sagar et al. 2013).

Apical dominance

In intact plants, the shoot apex grows predominantly and inhibits outgrowth of axillary buds. After decapitation of the shoot apex, outgrowth of axillary buds begins. This phenomenon is called an apical dominance, and it is regulated by the plant hormones auxin and cytokinin (CK). Auxin, derived from the shoot apex, inhibits the growth of axillary buds, while CK, thought to be derived from the roots, promotes the growth of axillary buds (Leyser 2003; Shimizu-Sato et al. 2001). Tanaka et al. (2006) proposed that in intact plants, auxin is basipetally transported and represses expression of the *IPT* gene (a key enzyme in CK biosynthesis) in the stem, then consequently, axillary buds lack the ability to grow out. On the other hand, once the shoot apex is decapitated, the auxin level in the stem decreases, repression of *IPT* gene expression is released, CK levels increase and axillary buds grow out. After axillary buds grow out, *de novo* synthesized IAA derived from a new shoot apex flows to the stem and again represses *IPT* gene expression.

3.3 Typical auxin insensitive responses

Auxin regulates a host of plant developmental and physiological processes through the $SCF^{TIR1/AFB}$-mediated proteolysis of the Aux/IAA family of transcriptional regulators. The domain II of the Aux/IAA is responsible for ubiquitin-mediated proteolysis, so when the mutation happened in this domain, the Aux/IAA protein will be stabilized and further resulted in altered auxin sensitivity and response. Most of the insight into the biological function of Aux/IAA proteins gathered to date comes from this characterization of dominant, gain-of-function mutations. Gain-of-function *AtIAA12* mutant *bdl* is insensitive to auxin and displayed several auxin related phenotypes. The strong lines displayed lacked hypocotyl, root and primary root meristem and the vasculature of the cotyledons was reduced (Hamann et al. 1999). The *bdl* homozygous mutant plants were short with rolled-up leaves and the apical dominance was reduced. Semidominant *AtIAA3* mutant *shy2* showed less sensitive to auxin and adult plants were extremely dwarfed with curled-up leaves. The seedlings displayed larger and curled up

cotyledons and shorter hypocotyls, and it could make leaves in dark. The lateral roots were largely reduced and primary roots were shorter than wild-type, and the root gravitropic response were altered (Tian and Reed 1999). Other gain-of-function Aux/IAA mutants also displayed these similar auxin insensitive responses, such as *slr/IAA14*, *msg2/IAA19*, *IAA28*, etc.

4 Auxin signal and wood formation

Several hormonal signals have been shown to be involved in cambial activity and xylem development (reviewed in Sorce et al. 2013), while the auxin indole acetic acid (IAA) is considered to be the primary hormonal signal among those (Fig. I-17).

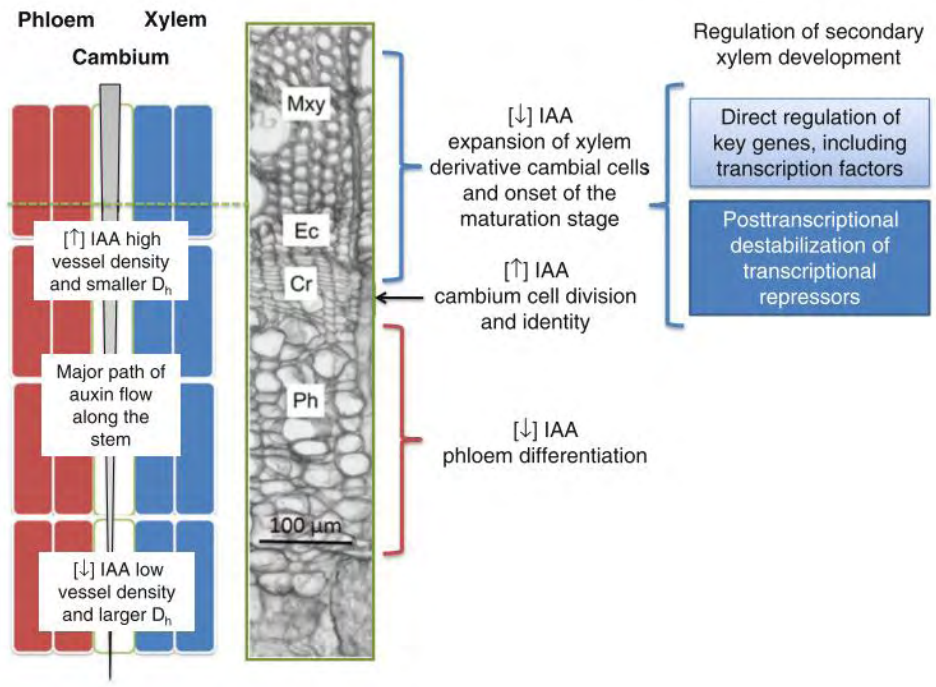


Fig. I-17 Summary of the main biological effects of IAA on cambium development and xylem and phloem differentiation processes. The diagram on the left shows the hypothesized relationship between the longitudinal auxin gradient concentration (decreasing from the top to the base of the stem) and the pattern of vessel development. Light micrograph shows a transverse section of *Populus* stem. Cr cambial region, Ec enlarging xylem cells, MXy maturing xylem, Ph phloem (Sorce et al. 2013)

The phenomenon of recruiting existing parenchyma cells for differentiation into files of procambial precursors is observed both during normal leaf vein development and in response to rupture of existing vascular strands (Sachs 1986). Early experiments established that ectopic application of the hormone auxin [indole-3-acetic acid (IAA)] was sufficient to trigger the specification of vascular tissue, including proliferation of cambial cells and final differentiation of vascular cell types (reviewed in Berleth and Mattsson 2000). Auxin (supported by cytokinin) can induce xylem tracheary element differentiation in suspension culture cells of *Zinnia* and *Arabidopsis* (Fukuda 1997). In 1981, Sachs proposed the 'canalization of auxin flow hypothesis' as a model for the auxin-mediated formation of vascular tissues (Sachs, 1981). In this model, channels of preferential auxin flow are created when a series of cells gradually become specialized for directional auxin transport. Once established, these channels effectively drain auxin from surrounding cells, resulting in localized concentration of auxin within distinct cell files, and this shift in auxin distribution was hypothesized to subsequently induce vascular tissue formation. The 'canalization of auxin' model thus provided a single mechanism that could account for both the initial specification and the physical contiguity of developing vascular tissues. The auxin-induced vascular differentiation response is: (1) local, as vascular strand formation is initiated at the specific site of auxin application; (2) polar, as vascular strand formation progresses from the auxin source toward the basal side of the plant; and (3) continuous, as it generates uninterrupted files of vascular cells (Scarpella et al. 2010). In normal plants, auxin is predominantly produced in apical regions, such as young leaves, from which it is transported basally to roots, such a steady polar flow of auxin induces the continuity of vascular tissues along the plant.

The proposed canalization of auxin was supported by the fact that disruption of auxin polar transport result in defects in leaf veins (Berleth and Mattsson 2000) and by the identification of genes involved in auxin signalling and transport. Mutations in auxin response related genes such as *MONOPTEROS* (*MP*)/*AUXIN RESPONSE FACTOR 5* (*AFR5*), *AUXIN RESISTANT 6* (*AXR6*), and *BODENLOS* (*BDL*)/*IAA12* result in

incomplete vascular systems and defects in embryo axis formation and embryonic root (Berleth 1993; Hamann 1999; Hobbie 2000; Przemeck et al. 1996). *mp* mutants also have severely attenuated levels of expression of the *PINFORMED1* (*PIN1*) gene (Wenzel et al. 2007). Asymmetric localization of PIN1 in plant cells is thought to establish a directional auxin flow. In the *Arabidopsis* leaf, *PIN1* and *MP* are co-expressed during very early stages of procambial cell specification, and their spatial pattern of expression gradually changes from initially broader domains to a single file of cells, as predicted for positive feedback onto the auxin canalization process (Scarpella et al, 2006; Wenzel et al. 2007).

Recent findings have identified auxin-mediated basic helix-loop-helix (bHLH) TF dimers as important factors regulating early vascular development, including TARGET OF MONOPTEROS5 (TMO5), LONESOME HIGHWAY (LHW) and their closest homologs (reviewed in Zhang et al. 2014a). Vascular tissue differentiation was totally blocked in the roots of *TMO* quadruple mutants. By contrast, co-overexpression of *TMO5* and *LHW* induced dramatic periclinal divisions within the vasculature of roots. Whether these factors interact to regulate secondary growth in different species is still an open question.

The major path of auxin flow along the stems of pine and poplar trees is in the vascular cambium, then auxin is distributed in a radial concentration gradient, with highest levels in vascular cambium, where cell division takes place, then decline rapidly toward the xylem and phloem (Fig. I-17, 18) (Uggla et al. 1998; Uggla et al. 1996). Thus, it has been suggested that auxin acts as in a manner similar to a morphogen to provide positional information for the division of wood-forming tissue into distinct differentiation zones. High auxin concentrations may be interpreted as a signal for cell division, intermediate levels may promote cell expansion and low levels may be read out as a signal inducing the deposition of secondary cell walls. So auxin concentration could define the fate of differentiating cells. High levels of auxin in cambium initial mother cells would maintain cambial cells meristematic, whereas low levels would lead

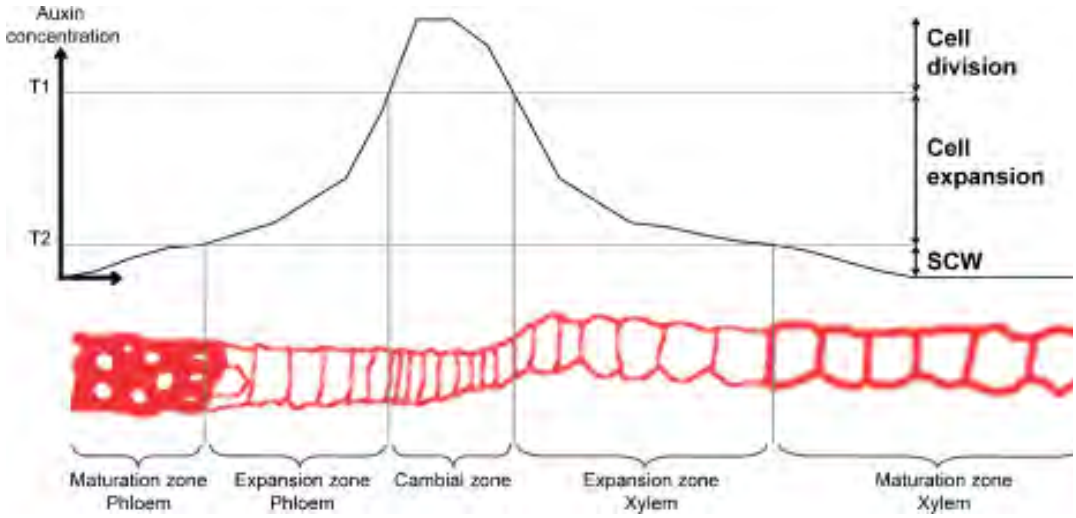


Fig. I-18 Auxin concentration gradient across wood-forming tissue. Auxin concentrations higher than threshold 1 (T1), correspond to the cell division zone, concentrations between T1 and T2 with cell expansion and below T2 with secondary cell wall formation (SCW). (Bhalerao and Fischer, 2014)

to differentiation into either xylem or phloem. In *pinus* explants lacking an apical auxin source, cambial derivatives are not able to maintain their fusiform shape and differentiate into parenchymatic cells instead of xylem tracheids (Savidge 1983). This suggest that auxin is required for maintaining the meristematic identity of the cambium, while it seems factors other than auxin likely to contribute to subsequent cell fate decision. So the role of auxin in promoting procambial cells to differentiate into xylem is still unclear. Moreover, Nilsson et al. (2008) did auxin-responsive transcriptome analysis in hybrid aspen wood forming tissues through depleting stem segments form endogenous auxin and then adding exogenous auxin to monitor global gene expression. Surprisingly, expression patterns of only a few auxin-responsive genes correlated well with the auxin concentration gradient, including transcription factors such as *PttHB8* and those of the *Aux/IAA* gene family (Nilsson et al. 2008). In addition, the transcripts of auxin-responsive genes responded dynamically to the changes in auxin levels, rather than being dependent on the steady state of auxin concentration. Moreover, although the auxin levels in cambium undergo only subtle seasonal changes (Uggla et al. 2001), the stability of PttIAA3 protein is quite different between the active growth and

dormancy stage (Baba et al. 2011). These findings argue against auxin displaying morphogen-like activity, and support the idea that auxin signaling controls cambial activity by modulation of auxin responsiveness (Bhalerao and Fischer 2014). Nilsson et al. (2008) hypothesized that auxin could regulate secondary xylem development through two mechanisms: (1) direct regulation of the expression of a few key genes, including some transcription factors; (2) posttranscriptional destabilization of transcriptional repressors.

During the auxin concentration gradient formation, PIN-FORMED 1 (PIN1)-like auxin efflux carries has been suggested to play an important role in trees (Hellgren et al. 2004; Schrader et al. 2003). In order to redistribute polarly transported auxin from the cambium into phloem and xylem, efflux carriers would need to be localized to the lateral membrane of cambial cells and their derivatives. Although PIN proteins have not yet been systematically localized in cambia, PIN1, which is strongly expressed in *Arabidopsis* inflorescence stems, has been shown to localize not only to basal plasma membranes but also in the lateral plasma membrane at the basal end of parenchymatic xylem cells (Galweiler et al. 1998; Sauer et al. 2006). Similarly, PIN3, which is involved in redirecting the auxin flux upon gravistimulation in the root tip, localizes to the inner lateral membrane of the starch sheath of the *Arabidopsis* inflorescence stem (Friml et al. 2002). In both a *pin1* and *pin3* mutant the radial extension of derivatives of the interfascicular cambium is reduced (Agusti et al. 2011a). In stems of *revolute* (*rev*) mutants, which are defective in the formation of interfascicular fibres, *PIN3* and *PIN4* is dramatically reduced (Zhong and Ye 2001). Another auxin transporter which be involved in this auxin gradient is the ATP-binding cassette (ABC) transporter ABCB14. The *abcb14* mutant displays reduced polar auxin transport and a mild radial cell expansion phenotype in the xylem (Kaneda et al. 2011). More recently, WAT1 has been demonstrated as a vacuolar auxin transporter. *WAT1* mutants revealed two cell wall-related phenotypes in *Arabidopsis* stems: a defect in cell elongation, resulting in a dwarfed habit and little to no secondary cell walls in fibres. Secondary walls of vessel elements were unaffected by the mutation (Ranocha et al. 2010). In summary,

phenotypes of auxin carrier mutants related to wood formation have either not been reported or are difficult to interpret due to an absence of auxin distribution data in those mutants (Agusti et al. 2011a, Kaneda et al. 2011).

The control of vessel size is an important trait for ensuring the ascent of water and minerals from roots to leaves and adaptation of plants to the environment. Auxin concentration has also been proposed to be involved in the regulation of vessel size and density (Aloni and Zimmermann 1983). The hypothesis is that high auxin concentration near the leaves stimulate rapid cell differentiation resulting in narrow vessels, while low auxin levels further down result in wide vessels because of the slow differentiation, which permits more cell expansion before secondary wall deposition. Thus, along the tree axis, auxin induces gradual gradients of increasing vessel diameter and decreasing vessel density from leaves to roots, which ensuring the ascent of water and minerals from roots to leaves and the adaptation of plants to the environment (Aloni 2001). Although the detailed mechanisms of how IAA would regulate growth and differentiation are not fully understood, the aforesaid hypothesis is substantiated by several observations: (1) the major sources of IAA are developing buds and young shoots (Uggla et al. 1998); (2) IAA is transported basipetally from the leaves towards the roots (Aloni 2001); (3) the major path of the IAA flow in tree stems is in the vascular cambium (Sundberg et al. 2000); (4) there is a general decrease in the IAA concentration towards the stem base (Aloni 2001).

As trees are long living organisms with sessile lifestyle, they have to adapt to changing environmental conditions throughout their lifetimes which may span decades and centuries in some cases. In particular, vascular stem cell activity shows plasticity in response to mechanical stress which affects wood formation and quality. In angiosperm woody species, a local increase in cambial cell division induces the formation of tension wood in the upper side of the leaning tree stems (Paux et al. 2004; Hellgren et al. 2004). Auxin has been proposed to be implicated in the tension response, and many different experiments have indicated this (Little and Savidge 1987). For example, application of

either exogenous auxin or auxin transport inhibitors was shown to induce the gelatinous G-fibres characteristics of tension wood (Pilate et al 2004). Although measurements of endogenous auxin failed to reveal significant changes in auxin balance in the cambial region tissues (Hellgren et al. 2004), a rather large set of auxin-related genes were found to be differentially expressed in developing poplar tension wood (Andersson-Gunneras et al. 2006).

Despite the fact that auxin has long been proposed as primary regulator of cambial activity and wood formation (Uggla et al. 1996, Sundberg et al. 2000), the auxin-regulated transcriptional programs underlying wood formation remain largely under investigated.

Part III The *Eucalyptus*, an evergreen angiosperm tree for wood formation

Eucalyptus species, commonly referred to as eucalypts, are native to Australia and the islands to its north. It is a large genus of evergreen hardwood trees and comprises more than 700 species, belonging to the angiosperm family Myrtaceae (Brooker 2000). *Eucalyptus* species occur naturally from sea level to the alpine tree line, from high rainfall to semi-arid zones and from the tropics to latitudes as high as 43 °south. Species of *Eucalyptus* are cultivated widely in the tropical and temperate world, though most species do not tolerate frost. A mature *Eucalyptus* may take the form from a low shrubby mallees (less than 10 m) to a straight-trunked forest up to 90 m tall.

Due to its wide adaptability, extremely fast growth rate, good form, and excellent wood and fibres properties, *Eucalyptus* species have become the most valuable and most widely planted forest trees on a world-wide basis. They are cultivated mainly in tropical and temperate zones. Today *Eucalyptus* can be found in Brazil, India, Portugal, South Africa, China, Uruguay, Ethiopia, California and more than 100 countries across six continents about more than 20 million hectares. The rotation cycles have relatively

shorter (5-7 years) compared to teak (20-40 years). The rotation time also depends on the trunk diameter desired, for example 4-5 years for heating wood, 7-10 years for paper pulp and 15 years or more for furniture wood. The productivity of managed *Eucalyptus* forests can reach to 20-50 m^3 ha^{-1} yr^{-1} due to improved genetic breeding and silviculture, against about 15 m^3 ha^{-1} yr^{-1} for the classical coniferous forests.

Eucalyptus is the most common short fibres source for pulpwood to make pulp. *Eucalyptus globulus* (in temperate climates) and the hybrid of *Eucalyptus urophylla x Eucalyptus grandis* (in tropical climates) are the most used varieties in papermaking. They are also utilized for a diverse array of products including sawn timber, mine props, poles, firewood, essential oils, honey and tannin as well as for shade, shelter and soil reclamation. In rural communities of many developing countries, eucalypt wood is an important source of fuel and building material. In today's "new carbon economy", eucalypts are receiving attention as fast-growing, short-rotation, renewable biomass crops for energy production. It is for instance listed as one of the U.S. Department of Energy's candidate biomass energy crops.

The genome of *Eucalyptus grandis* BRASUZ1 has been fully sequenced and annotated (Myburg et al. 2014), offering new opportunities to get insights into the regulation of secondary growth and cambial activity by auxin-signalling mediators. Understanding the mechanisms that underlie auxin regulation in *Eucalyptus* wood formation is of interest both in the context of plant development and as a path to improve lignocellulosic biomass production and quality as a dedicated energy crop for lignocellulosic biofuel production. It is worth noting that *Eucalyptus* is an evergreen tree that do not present dormancy in contrast to poplar a deciduous tree, allowing further comparison of the regulation of cambium activity and xylem differentiation between the two only forest genus sequenced hitherto.

Chapter II

Genome-wide Characterization and Expression Profiling of the AUXIN RESPONSE FACTOR (ARF) Gene Family in Eucalyptus grandis

OPEN ACCESS Freely available online

PLOS ONE

Chapter II: Genome-Wide Characterization and Expression Profiling of the *AUXIN RESPONSE FACTOR* (*ARF*) Gene Family in *Eucalyptus grandis*

Hong Yu[1], Marçal Soler[1], Isabelle Mila[2], Hélène San Clemente[1], Bruno Savelli[1], Christophe Dunand[1], Jorge A. P. Paiva[3,4], Alexander A. Myburg[5], Mondher Bouzayen[2], Jacqueline Grima-Pettenati[1]*, Hua Cassan-Wang[1]

1 LRSV Laboratoire de Recherche en Sciences Végétales, UMR5546, Université Toulouse III/CNRS, Castanet Tolosan, France, **2** Université de Toulouse, Institut National Polytechnique-Ecole Nationale Supérieure Agronomique de Toulouse, Laboratoire de Génomique et Biotechnologie des Fruits, Castanet-Tolosan, France, **3** Instituto de Investigação Científica e Tropical (IICT/MCTES), Lisboa, Portugal, **4** Instituto de Biologia Experimental e Tecnológica (IBET), Oeiras, Portugal, **5** Department of Genetics, Forestry and Agricultural Biotechnology Institute (FABI), Genomics Research Institute (GRI), University of Pretoria, Pretoria, South Africa

Abstract

Auxin is a central hormone involved in a wide range of developmental processes including the specification of vascular stem cells. Auxin Response Factors (ARF) are important actors of the auxin signalling pathway, regulating the transcription of auxin-responsive genes through direct binding to their promoters. The recent availability of the *Eucalyptus grandis* genome sequence allowed us to examine the characteristics and evolutionary history of this gene family in a woody plant of high economic importance. With 17 members, the *E. grandis ARF* gene family is slightly contracted, as compared to those of most angiosperms studied hitherto, lacking traces of duplication events. *In silico* analysis of alternative transcripts and gene truncation suggested that these two mechanisms were preeminent in shaping the functional diversity of the *ARF* family in *Eucalyptus*. Comparative phylogenetic analyses with genomes of other taxonomic lineages revealed the presence of a new ARF clade found preferentially in woody and/or perennial plants. High-throughput expression profiling among different organs and tissues and in response to environmental cues highlighted genes expressed in vascular cambium and/or developing xylem, responding dynamically to various environmental stimuli. Finally, this study allowed identification of three *ARF* candidates potentially involved in the auxin-regulated transcriptional program underlying wood formation.

Citation: Yu H, Soler M, Mila I, San Clemente H, Savelli B, et al. (2014) Genome-Wide Characterization and Expression Profiling of the *AUXIN RESPONSE FACTOR* (*ARF*) Gene Family in *Eucalyptus grandis*. PLoS ONE 9(9): e108906. doi:10.1371/journal.pone.0108906

Editor: Marcel Quint, Leibniz Institute of Plant Biochemistry, Germany

Received July 2, 2014; **Accepted** August 27, 2014; **Published** September 30, 2014

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All Eucalyptus ARF sequences used in this paper are available in phytozome http://www.phytozome.net/. E. globulus RNA Seq Illumina reads are provided in File S1.

Funding: This work was supported by the Centre National pour la Recherche Scientifique (CNRS), the University Paul Sabatier Toulouse III (UPS), the French Laboratory of Excellence project ''TULIP'' (ANR-10-LABX-41; ANR-11-IDEX-0002-02) and the Plant KBBE TreeForJoules project (ANR-2010-KBBE-007-01 (FR) and P-KBBE/AGR_GPL/0001/2010 (FCT, PT)). HY was supported by PhD grants from the China Scholarship council. MS was supported by a Postdoc fellowship ''Beatriu de Pinós'' thanks to the Departament d'Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya. JAPP acknowledges Fundação para a Ciência e a Tecnologia (FCT, Portugal) and his research contract funded by Programa Ciência 2008 and POPH (QREN). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors declare that no competing interests exist.

* Email: yu@lrsv.ups-tlse.fr

1 Introduction

The plant hormone auxin plays a prominent role in the regulation of plant growth in response to diverse developmental and environmental cues such as organogenesis, tropic movement, root growth, fruit development, tissue and organ patterning and vascular development [1]. Auxin plays a crucial role in the specification of vascular stem cells (procambium) and in cambial activity [2]. Analysis of auxin distribution across the cambial region in hybrid aspen trees showed a radial auxin gradient reaching a peak level in the cambial zone or at the border between the cambial zone and the expansion zone towards developing wood cells [3,4]. The auxin gradient was indeed shown to overlap with the sequential and numerous auxin-regulated genes responding dynamically to the change in auxin levels in wood forming cells [5].

As trees are long living organisms with sessile lifestyle, they have to adapt to changing environmental conditions throughout their lifetimes which may span decades and centuries in some cases. In particular, vascular stem cell activity shows plasticity in response to mechanical stress which affects wood formation and quality. In angiosperm woody species, a local increase in cambial cell division induces the formation of tension wood in the upper side of the leaning tree stems [6,7]. Auxin has been proposed to be implicated in the tension response, and application of either exogenous auxin or auxin transport inhibitors was shown to induce the gelatinous G-fibres characteristics of tension wood [8]. Although measurements of endogenous auxin failed to reveal significant changes in auxin balance in the cambial region tissues, a rather large set of auxin-related genes were found to be differentially expressed in developing poplar tension wood [9]. A recent study indicated that the auxin signalling pathway is significantly disrupted during

cambial dormancy in hybrid aspen [10]. Despite the fact that auxin has long been proposed as primary regulator of cambial activity and wood formation [4,11], the auxin-regulated transcriptional programs underlying wood formation remain largely under investigated.

Auxin exerts its function through modulating the expression of numerous genes among which is a set of transcriptional regulators. Auxin Response Factors (ARFs) and Aux/IAAs are two well-known mediators which regulate auxin responsive gene expression [12,13]. Most ARF proteins contain a highly conserved N-terminal B3-like DNA binding domain that recognizes an auxin-response element (AuxRE: TGTCTC) present in the promoters of auxin-responsive genes. The C-terminal domain contains two motifs, called III and IV, also found in Aux/IAA proteins and shown to enable the formation of homo- and heterodimers among ARFs and Aux/IAAs [14,15]. The middle region whose sequence is less conserved confers transcription activation or repression depending on its amino acid composition [13]. Biochemical and genetic studies in *Arabidopsis* and other plants have led to a working model of the mediation of auxin response by ARF proteins [14,16]. In the absence of auxin, Aux/IAAs bind to ARFs and recruit co-repressors of the TOPLESS (TPL) family, preventing the ARFs from regulating target genes [17]. The presence of auxin induces Aux/IAA protein degradation via the 26S proteasome through SCF-TIR1 ubiquitin ligase complex; thus liberating the trapped ARF proteins, allowing them to modulate the transcription of target auxin-responsive genes (for review, see Guilfoyle and Hagen) [12]. This model based on limited ARF-Aux/IAA interaction studies which provides a framework for understanding how members of these families may function. More recently, a large-scale analysis of the Aux/IAA-ARF interactions in the shoot apex of *Arabidopsis* showed that the vast majority of Aux/IAAs interact with all ARF activators, suggesting that most Aux/IAAs may repress the transcriptional activity of ARF activators [18]. In contrast, Aux/IAAs have limited interactions with ARF repressors suggesting that the role of the latter is essentially auxin-independent and that they might simply compete with the ARF activators for binding to the promoter of auxin-inducible genes [18]. This finding is particularly important taking into account that auxin predominantly activates transcription [19–21] and that a large complement of the ARF family acts as transcriptional repressors [12]. Whereas the above proposed scenario applies to the shoot apical meristem, it is likely that specific interactions between Aux/IAAs and ARFs might also affect the dynamics of the ARF-Aux/IAA signalling pathway in other developmental processes such as cambial development.

The *ARF* gene family has been most extensively studied in *Arabidopsis* where phenotyping of mutants revealed involvement of specific *ARF* genes in various plant growth and development processes [20,22–30]. For instance, ARF5/MONOPTEROS (MP) is a transcriptional activator known to play a critical role in the specification of vascular stem cells [27,31].

The *ARF* family has also been characterized in several annual herbaceous plants including monocots (rice, maize) [32,33] and dicots (*Arabidopsis*, tomato, soybean, *Brassica rapa*) [24,34–37] and in only two woody perennial genera, *Populus* [38] and *Vitis* [39]. However, so far, no *ARF* candidate has been identified as specifically involved in vascular cambium activity and xylem differentiation.

The recent availability of *Eucalyptus grandis* genome [40], the second hardwood forest tree genome fully sequenced, offers new opportunities to get insights into the regulation of secondary growth and cambial activity by ARFs, especially because *Eucalyptus* belongs to evergreen trees that do not present dormancy in their cambial activity in sharp contrast with deciduous trees like *Populus*. *Eucalyptus* is also the most planted hardwood in the world, mainly for pulp and paper production but is also foreseen as a dedicated energy crop for lignocellulosic biofuel production. Thus, understanding the mechanisms that underlying auxin regulation in *Eucalyptus* wood formation is of interest both in the context of plant development and as a path to improve lignocellulosic biomass production and quality.

In the present paper, we report a genome-wide identification and characterization of the *ARF* family in *Eucalyptus grandis*. We analyzed gene structure, protein motif architecture, and chromosomal location of the members of the *E. grandis ARF* family. We also performed comparative phylogenetic relationships and large scale transcript profiling with a special focus on vascular tissues to get insights in their evolution, expression characteristics and possible functions.

2 Materials and Methods

2.1 Identification of *ARF* gene family in *Eucalyptus grandis* and chromosomal location

The identification procedure is illustrated in Fig. S1. Firstly we used *Arabidopsis* ARF proteins as queries in BLASTP searches for predicted protein in *Eucalyptus* genome (JGI assembly v1.0, annotation v1.1, http://www.phytozome.net/eucalyptus). A total of 64 *Eucalyptus* proteins identified in this initial search were examined by manual curation of protein motif scan using Pfam domain IDs (http://pfam.wustl.edu) and NCBI conserved domain database (http://www.ncbi.nlm.nih.gov/cdd). Redundant and invalid gene models were eliminated based on gene structure, intactness of conserved motifs and EST support. Three incomplete gene models were identified and completed by FGeneSH (http://linux1.softberry.com). To complete partial sequence of Eucgr.K02197.1, we cloned the corresponding genomic fragment using forward primer: 5′-AATTGACCGCGGTTGGATA-3′ and reverse primer 5′-GAGCAGGCCAACATCCTCA-3′, which located up-stream and down-stream respectively of the non-determined sequence (N). According to sequencing result we complete the missing part (1156 bp), corresponding to a part of promoter region and a part of 5′end CDS of the Eucgr.K02197.1 (submitted to GenBank data library under the accession number KC480258). All these manual curations enabled us to obtain 17 complete *Eucalyptus* ARF proteins sequences. We then used them as query in two subsequent additional searches: 1) BLASTP against *Eucalyptus* proteome for exhaustive identification of divergent *Eucalyptus* gene family members, and 2) tBLASTn searches against *Eucalyptus* genome for seeking any possible non-predicted genes. For validation, we also used poplar ARF proteins as queries to do the search procedure described above, and we obtained exactly the same result.

In the course of the above identification process we completed and expertly re-annotated three partial sequences (accession numbers Eucgr.F02090.1, Eucgr.F04380.1, and Eucgr.K03433.1 in the Phytozome database) initially annotated in the *Eucalyptus* genome-sequencing project (Table 1). In addition, we found one gene (accession number Eucgr.K02197.1) that corresponded to a partial sequence for which the 5′ end was not determined (1240 N as sequencing results). Information on chromosomal location was retrieved from the *Eucalyptus* genome browser (http://www.phytozome.net/eucalyptus). *EgrARF* genes were mapped to their loci using MapChart 2.2 [41].

Table 1. *ARF* gene family in *Eucalyptus*.

Accession no.[a]	Arabidopsis best hit (blastp)[b]	Amino acid identity %[c]	Short name[d]	Number of predicted alternative transcript[e]	Chromosome[f]	Genome location[g]	ORF (bp)[h]	Deduced polypeptide[i]			Exon No.
								Length (aa)	MW (kDa)	PI	
Eucgr.G00076.1	*AtARF1*	62.2	*EgrARF1*	4	7	848,248..860,087	2028	675	75.3	6.06	14
Eucgr.K02197.1	*AtARF2*	59.3	*EgrARF2A***	5	11	29,232,770..29,239,103	2508	835	93.13	6.55	14
Eucgr.B03551.1	*AtARF2*	49.5	*EgrARF2B*	2	2	60,136,697..60,143,120	2364	787	88.09	6.62	14
Eucgr.D00588.1	*AtARF3*	46.2	*EgrARF3*	2	4	10,835,264..10,841,977	2172	723	79.19	6.36	10
Eucgr.B02480.1	*AtARF4*	57.8	*EgrARF4*	3	2	46,950,723..46,956,747	2394	797	88.53	6.31	12
Eucgr.F02090.1	*AtARF5*	59.3	*EgrARF5**	1	6	28,226,789..28,233,884	2835	944	103.53	5.43	14
Eucgr.D00264.1	*AtARF6*	69.1	*EgrARF6A*	2	4	4,224,849..4,232,295	2694	897	99.04	6.04	14
Eucgr.A02065.1	*AtARF6*	65.6	*EgrARF6B*	3	1	31,463,716..31,473,722	2613	870	96.64	6.06	14
Eucgr.D01764.1	*AtARF9*	58.6	*EgrARF9A*	1	4	31,589,333..31,593,948	1971	656	73.53	6.15	14
Eucgr.E00888.1	*AtARF9*	58.4	*EgrARF9B*	2	5	9,303,573..9,308,209	2064	687	76.46	6.09	14
Eucgr.J00923.1	*AtARF10*	60.1	*EgrARF10*	1	10	10,070,677..10,074,840	2139	712	77.68	8.32	4
Eucgr.G02838.1	*AtARF16*	58	*EgrARF16A*	1	7	46,569,603..46,573,031	2079	692	76.33	6.80	3
Eucgr.K01240.1	*AtARF16*	52.5	*EgrARF16B*	1	11	15,584,770..15,588,716	2124	707	77.91	6.91	3
Eucgr.F04380.1	*AtARF17*	42.1	*EgrARF17**	1	6	52,730,940..52,733,856	1881	626	67.75	7.54	3
Eucgr.C03293.1	*AtARF19*	60.4	*EgrARF19A*	2	3	62,460,408..62,469,132	3360	1119	124.93	6.30	14
Eucgr.C02178.1	*AtARF19*	44.8	*EgrARF19B*	2	3	39,505,080..39,513,066	3360	1119	123.24	5.98	14
Eucgr.K03433.1	-	-	*EgrARF24**	1	11	43,352,151..43,357,873	1836	611	68.44	7.24	14

*Using FGeneSH to complete the complete sequence.
**using specific primers to amplify the genomic DNA to complete the sequence.
[a]Gene model of Eucalyptus (version 1.1) in phytozome V8.0.
[b]The best hit of EgrARF in Arabidopsis by using blastp.
[c]The amino acid identity percentage between EgrARF and corresponding AtARF.
[d]Designation related to Arabidopsis best hit.
[e]The number of predicted alternative transcripts of EgrARF in phytozome.
[f,g]Location of the EgrARF genes in the Chromosome.
[h]Length of open reading frame in base pairs.
[i]The number of amino acids, molecular weight (kilodaltons), and isoelectric point (pI) of the deduced polypeptides.
doi:10.1371/journal.pone.0108906.t001

2.2 Sequence, phylogenetic, gene structure analysis

Conserved protein motifs were determined by Pfam [42]. Multiple protein sequences alignment was performed using Clustal X program (Version 2.0.11). Using full length sequences of all predicted protein, phylogenetic trees were constructed with MEGA5 program by neighbor-joining method with 1000 bootstrap replicates. Their exon-intron structures were extracted from Phytozome (http://www.phytozome.net/eucalyptus) and visualized in Fancy Gene V1.4 (http://bio.ieo.eu/fancygene/). The prediction of small RNA target sites in *EgrARF* genes was performed through the web application psRNATarget (http://plantgrn.noble.org/psRNATarget/). The stem-loop structures were predicted using RNAfold web server (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) and visualized by RNAstructure 5.3 program.

2.3 Plant material

The plant materials provenance and preparation are described in Cassan-Wang et al. [43]. Hormone treatments were performed in an *in vitro* culture system. 10 µM NAA (1-Naphthaleneacetic acid, for auxin), or 20 µM gibberellic acid or 100 µM ACC (1-aminocyclopropane-1-carboxylic-acid, for ethylene) were added to the medium of 65-d-old young trees, and trees were sampled 14 days after treatments.

2.4 Total RNA extraction, cDNA synthesis, quality controls and high throughput quantitative RT-PCR

All the procedures used for the qRT-PCR, from the RNA extraction to the calculation of transcript abundance are described in Cassan-Wang et al. [43]. Only samples with a RNA integrity number >7 (assessed by Agilent 2100 Bioanalyzer) were retained for reverse transcription. cDNA quality was assessed as described by Udvardi et al. [44] using housekeeping genes *IDH* and *PP2A3* (primers see Table S1). Primer pairs were designed using the software QuantPrime (http://www.quantprime.de) [45], showing in Table S1. qRT-PCR was performed by the Genotoul service in Toulouse (http://genomique.genotoul.fr/) using the BioMark 96:96 Dynamic Array integrated fluidic circuits (Fluidigm Corporation, San Francisco, USA) described in Cassan-Wang et al. [43]. The specificity of the PCR products was confirmed by analysing melting curves. Only primers that produced a linear amplification and qPCR products with a single-peak melting curves were used for further analysis. The efficiency of each pair of primers was determined from the data of amplification Ct value plot with a serial dilution of mixture cDNA and the equation $E = 10^{(-1/slope)} - 1$. $E^{-\Delta\Delta Ct}$ method was used to calculate relative mRNA fold change compared to control sample using formula $(E_{target})^{\Delta Ct\_target\ (control-sample)} / (E_{reference})^{\Delta Ct\_reference\ (control-sample)}$ [46] and five reference genes (*IDH, PP2A1, PP2A3, EF-1a* and *SAND*, Table S1) were used for data normalization. We chose *in vitro* plantlets as control sample, because it contains the main organs and tissues of our studies such as stem, leaves, shoot tips, xylem, phloem and cambium, and it is a relative stable and less variable sample as being grown under the same *in vitro* culture condition from one experiment to another.

2.5 Transactivation analysis in single cell system

For testing the ability of ARF transcription factors to up or down regulate the expression of auxin responsive promoter DR5, the full-length cDNAs of the ARF transcription factors were cloned in pGreen vector under 35SCaMV promoter to create the effector constructs. The reporter constructs use a synthetic auxin-responsive promoter DR5 fused with the GFP reporter gene. Tobacco BY-2 protoplasts were co-transfected with the reporter and effector constructs as described in Audran-Delalande et al. [47]. After 16 h incubation, GFP expression was quantified by flow cytometry (LSR Fortessa, BD Biosciences). Data were analysed using BD FacsDiva software. Transfection assays were performed in three independent replicates and 3000–4000 protoplasts were gated for each sample. GFP fluorescence corresponds to the average fluorescence intensity of the protoplasts population after subtraction of auto-fluorescence determined with non-transformed protoplasts. 50 µM 2, 4-D was used for auxin treatment. We tested two independent protoplast preparations and for each of them, we performed in three independent transformation replicates. Similar results were obtained with the independent protoplast preparations and the data were represented by one of the preparations. For normalization, protoplasts were transformed with the reporter vector and the effector plasmid lacking the *ARF* gene.

3 Results and Discussion

3.1 Identification and chromosomal distribution of *Eucalyptus ARF* genes

The procedure to identify all members of the *ARF* family in the *E. grandis* genome (JGI assembly v1.0, annotation v1.1 (http://www.phytozome.net/cgi-bin/gbrowse/eucalyptus/), included expert manual curation as illustrated in Fig. S1. It allowed the identification of 17 genes encoding full length *Eucalyptus* ARF proteins (henceforth referred to as *EgrARF* genes). We named these genes according to their potential orthologs in *Arabidopsis* (Table 1). Where two *EgrARFs* matched the same potential *Arabidopsis* ortholog *AtARFx*, they were named as *EgrARFxA* and *xB*, with *xA* being the closest to the *Arabidopsis* ortholog; e.g. *EgrARF2A* and *EgrARF2B*. The percentage of identity between the *Arabidopsis* and the *Eucalyptus* predicted ARF protein sequences, and among the *Eucalyptus* ARFs themselves are given as Table S2 and S3, respectively. Eight *Arabidopsis* genes have no corresponding *Eucalyptus* orthologs (*AtARF12* to *15* & *20* to *23*), while only one *EgrARF* gene, *EgrARF24*, has no ortholog in *Arabidopsis* (Table 1). *In silico* chromosomal mapping of the gene loci revealed that the 17 *EgrARF* genes are scattered on nine of the eleven chromosomes, with one to three *EgrARF* genes per chromosome and with chromosomes 8 and 9 being devoid of *ARF* genes (Fig. S2).

The predicted proteins encoded by the *EgrARF* genes ranged from 593–1119 amino acid residues (Table 1), with PIs in the range of 5.43–8.32, suggesting that they can work in very different subcellular environments. Sequence analyses of the predicted proteins and Pfam protein motif analysis showed that most of them (14 of the 17 predicted proteins) harbour the typical ARF protein structure comprising a highly conserved DNA-binding domain (DBD) in the N-terminal region composed of a plant specific B3-type subdomain and an ARF subdomain, a variable middle region (MR) that functions as an activation or repression domain, and a carboxy-terminal dimerization (CTD) domain consisting of two highly conserved dimerization subdomains III and IV, similar to those found in Aux/IAAs (Fig. 1). We analysed and aligned the predicted amino acid sequences of the *Egr*ARFs (Fig. 1 and Fig. S3). Four out of the 17 *Egr*ARFs (*10, 16A, 16B* and *17*) exhibited an additional short segment of amino acids (between 15 to 43 amino-acids) in their DBD, between the B3 and ARF subdomains (Fig. 1 and Fig. S3). Such a feature has already been reported in *Arabidopsis* and soybean [36]. At the end of the DBD domain, all of the *Egr*ARFs excepty *Egr*ARF6A, 6B and 19A contain a conserved putative mono-partite nuclear localization signal (NLS) (Fig. S3) shown to direct the proteins into the nucleus [36,48].

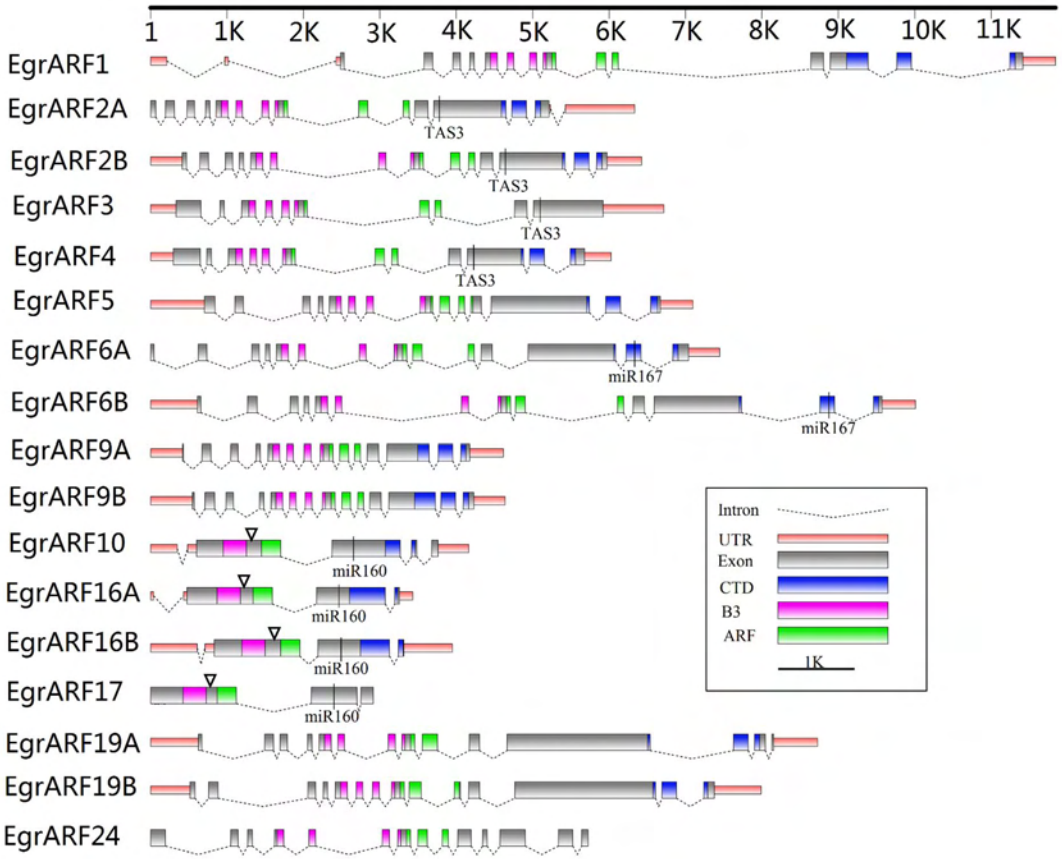


Figure 1. Gene structure of the *EgrARF* family. The information on exon–intron structure was extracted from the Phytozome database and visualized by using the FancyGene software (http://bio.ieo.eu/fancygene/). The sizes of exons and introns are indicated by the scale at the top. The domains of *EgrARF* gene were predicted by Pfam (http://pfam.xfam.org/) and are indicated by different colours. The B3 together with ARF subdomains constitute the DNA binding domain (DBD). The CTD contains two sub-domains III and IV. The TAS3 and microRNA target sites are marked on the corresponding target genes. The triangles underline the insertion sites of additional short amino-acids segments between the B3 and ARF subdomains.
doi:10.1371/journal.pone.0108906.g001

Five *Egr*ARFs (*Egr*ARF5, 6A, 6B, 19A and 19B) harbour a glutamine (Q)-rich middle region (Fig. S3) implying that these proteins are likely transcriptional activators since glutamine enrichment seems to be a distinctive feature of ARF activators in all plant lineages [15,49]. The other 12 *Egr*ARFs may function as repressors based on their middle regions enriched either in S (serine), SPL (Serine, Proline, Leucine) or SGL (Serine, Glycine, Leucine) [36] (Fig. S3).

The predicted protein structures of *Egr*ARF3 and *Egr*ARF17 are lacking dimerization domains III and IV like their potential orthologs in *Arabidopsis* (Fig. 1 and Fig. S3). *Egr*ARF24, which has no ortholog in *Arabidopsis*, has a truncated CTD since only Aux/IAA subdomain III is present. The percentage of *Egr*ARFs displaying a truncated CTD (17.6%) is similar to that in *Arabidopsis* (17.4%), but lower than in rapeseed (22.6%) or tomato (28.6%) [24,37,50]. These truncated *Egr*ARFs are predicted to be unable to interact with Aux/IAA, a sequestration mechanism which may regulate their activity, and hence, they are likely to be insensitive to auxin. However, ARF repressors seem to display very limited interactions with Aux/IAA proteins [18], therefore the lack of domains III and/or IV could also have consequences for the interaction of ARFs with other transcriptional regulators [49].

Compared to *Arabidopsis*, the ARF family in *Eucalyptus* is slightly contracted with 17 *versus* 23 members. It is worth noting that we found the exact same number of *ARF* genes in another *Eucalyptus* species, *E. camaldulensis* (http://www.kazusa.or.jp/eucaly/). Indeed when comparing to other species, in which the *ARF* family has been characterized (Table 2), *Eucalyptus* and grapevine appeared to have the smallest families with 17 and 19 members respectively, whereas poplar and soybean had the largest families with 39 and 51 members, respectively. We did not find evidence that any of the 17 *EgrARF* genes arose by tandem, segmental, or whole genome duplication, or even the more ancient hexaploidization in the *E. grandis* genome [40] and it appears that any such duplicates have been lost in *Eucalyptus* as is the case for 95% of whole-genome duplicates. This is sharply contrasting with the intensive tandem duplication events found for *Arabidopsis* ARF members [14,51], the segmental duplication found in *Populus* [38], and the whole-genome duplication events in soybean [36].

As duplication and alternative splicing are the two main mechanisms involved in diversification of function within gene families, sometimes viewed as opposite trends in gene family evolution, we performed an *in silico* survey of the alternative transcripts predicted in the *E. grandis* genome JGI assembly v1.0, annotation v1.1 (http://www.phytozome.net/eucalyptus), and compared them to those in *Arabidopsis* (Table 1 and Fig. S4). More than half of the *Eucalyptus* ARF family members (10 out of 17) have evidence of alternative splicing (Fig. S4). Taking into

Table 2. Summary of *ARF* gene content in angiosperm species.

Species	ARFs content	Reference
Eucalyptus grandis	17	This study
Vitis vinifera	19	[39] Wan *et al.* (2014)
Solanum lycopersicon	22	[35] Zouine *et al.* (2014)
Arabidopsis thaliana	23	[24] Okushima *et al.* (2005)
Oryza sativa	25	[33] Wang *et al.* (2007)
Zea mays	31	[32] Xing *et al.* (2011)
Brassica rapa	31	[37] Mun *et al.* (2012)
Populus trichocarpa	39	[38] Kalluri *et al.* (2007)
Glycine max	51	[36] Ha *et al.* (2013)

doi:10.1371/journal.pone.0108906.t002

account the number of possible alternative transcripts in *Eucalyptus* (17) and in *Arabidopsis* (15), the total number of possible transcripts in both species becomes very similar, 34 and 38, respectively. Some of the transcripts resulted in truncated versions of the genes like *EgrARF1.4, 4.3* and *9B.2* lacking the Aux/IAA interaction domain and *EgrARF2B.2* lacking the B3/DBD domain. We further compared the *in silico* predicted *ARF* alternative transcripts from *E. grandis* to those expressed in a dataset of in-house RNA-Seq data from *E. globulus* (Table S4, Fig. S5, File S1). Remarkably, the vast majority of the alternative transcripts predicted in *E. grandis* were found expressed in *E. globulus* providing strong experimental support to their occurrence and conservation in the two *Eucalyptus* species. The importance of alternative splicing in the *ARF* family, has been highlighted recently by Finet et al. [49], who have shown that two *Arabidopsis* alternative transcripts of *AtARF4* have very different functions in flower development, and by Zouine et al. [35] who have shown that in tomato, one third of the *ARF* members displays alternative splicing as a mode of regulation during the flower to fruit transition. In *Arabidopsis* and in many other species, not only domain rearrangement through alternative splicing but also extensive gene duplication played a significant role in *ARF* functional diversification [49], whereas in *Eucalyptus* the first mechanism appeared to be preeminent.

3.2 Comparative Phylogenetic analysis of the *ARF* family

To assess the relationship of *Eucalyptus ARF* family members to their potential orthologs in other landmark genomes, we constructed a comparative phylogenetic tree using all predicted ARF protein sequences from genomes of relevant taxonomic lineages. The core rosids were represented by *Arabidopsis* and *Populus* (Malvids) while the Myrtales, the Vitales and the Asterides were represented by *Eucalyptus grandis*, *Vitis vinifera* and *Solanum lycopersicum*, respectively. The monocots were represented by the O*ryza sativa* genome (Fig. S6). A simplified tree with only *Arabidopsis*, *Populus* and *Eucalyptus* (Fig. 2A) showed that ARFs are distributed into four major groups I, II, III, and IV. *Eucalyptus* (and also grapevine) which harbour the smallest number of *ARF* genes as compared to all other species (Table 2), have the fewest number of ARF proteins in each of the four groups. The positions and phases of the introns were well conserved within each group (Fig. 1 and Fig. 2), whereas their sizes were poorly conserved even within the same group. All five predicted *Eucalyptus* ARF transcriptional activators fell within group II as their potential orthologs from *Arabidopsis* and other species; the remaining *Egr*ARFs were distributed among the three other groups.

Some lineage-specific clades were found in the *Solanaceae ARF* family [35] as well as in *Arabidopsis ARF* family [24]. In *Arabidopsis*, group I was substantially expanded with a subgroup containing seven tandem duplicated genes (encoding proteins AtARF12 to 15 and AtARF20, to 22), and the sister pair of AtARF11–AtARF18, for which orthologs were found only in Brassicaceae.

In group I, an isolated clade (highlighted in green in Fig. 2) contained *Egr*ARF24 clustering with *Ptr*ARF2.5 and *Ptr*ARF2.6 and did not contain any obvious *Arabidopsis* ortholog. This clade was absent from the herbaceous annual plants (*Arabidopsis*, tomato and rice), but present in woody perennial plants (*Eucalyptus*, *Populus* and *Vitis*; Fig. S6). To verify if this clade could be more specific to woody perennial plants, we performed a BLAST similarity search in 33 plant genomes available in Phytozome and found potential orthologs of *EgrARF24* in 13 plant species out of 33 (Table S5) which are presented in a phylogenetic tree (Fig. 2B). Among these 13 plant species, 11 are trees such as *M. domestica, C. sinensis, C. clementina, P. persica*, or tree-like plants and shrubs such as *C. papaya, T. cacao*, *G. raimondii*, although the latter is often grown as an annual plant. *A. coerulea and F. vesca* are perennial herbaceous plants. The two notable exceptions are two members of the Fabaceae family (*G. max*, and *P. vulgaris*) which are annual herbaceous plants. We thus considered this clade as woody-preferential. Regarding Group III, there was no evidence of large expansion of *ARF3* and *ARF4* genes in any of the three species, with only *ARF3* duplicated in *Populus*. Group IV contained four members from *Eucalyptus*, i.e. one more than in *Arabidopsis*. All of the *Egr*ARFs belonging to this group have in common an additional fragment (between 15 to 43 amino-acids residues) within their DBD (Fig. 1 and Fig. 2) and, noteworthy, alternative splicing was not detected for any of these genes in *Eucalyptus* and *Arabidopsis* (Fig. S4).

3.3 Prediction of small regulatory RNAs and their potential ARF targets

In *Arabidopsis*, several *ARF* genes are targets of microRNAs miR160 and miR167, or of a trans-acting short interfering RNA (tasiRNA) *TAS3* [29,52–54]. Since these small RNAs and their targets are very often conserved across plant species [32,55,56], we searched for their potential orthologs in the *Eucalyptus* genome. Their chromosomal locations, genomic sequences and the sequences of their mature forms are presented in Table S6. We identified three potential *Eucalyptus miR160* loci and three

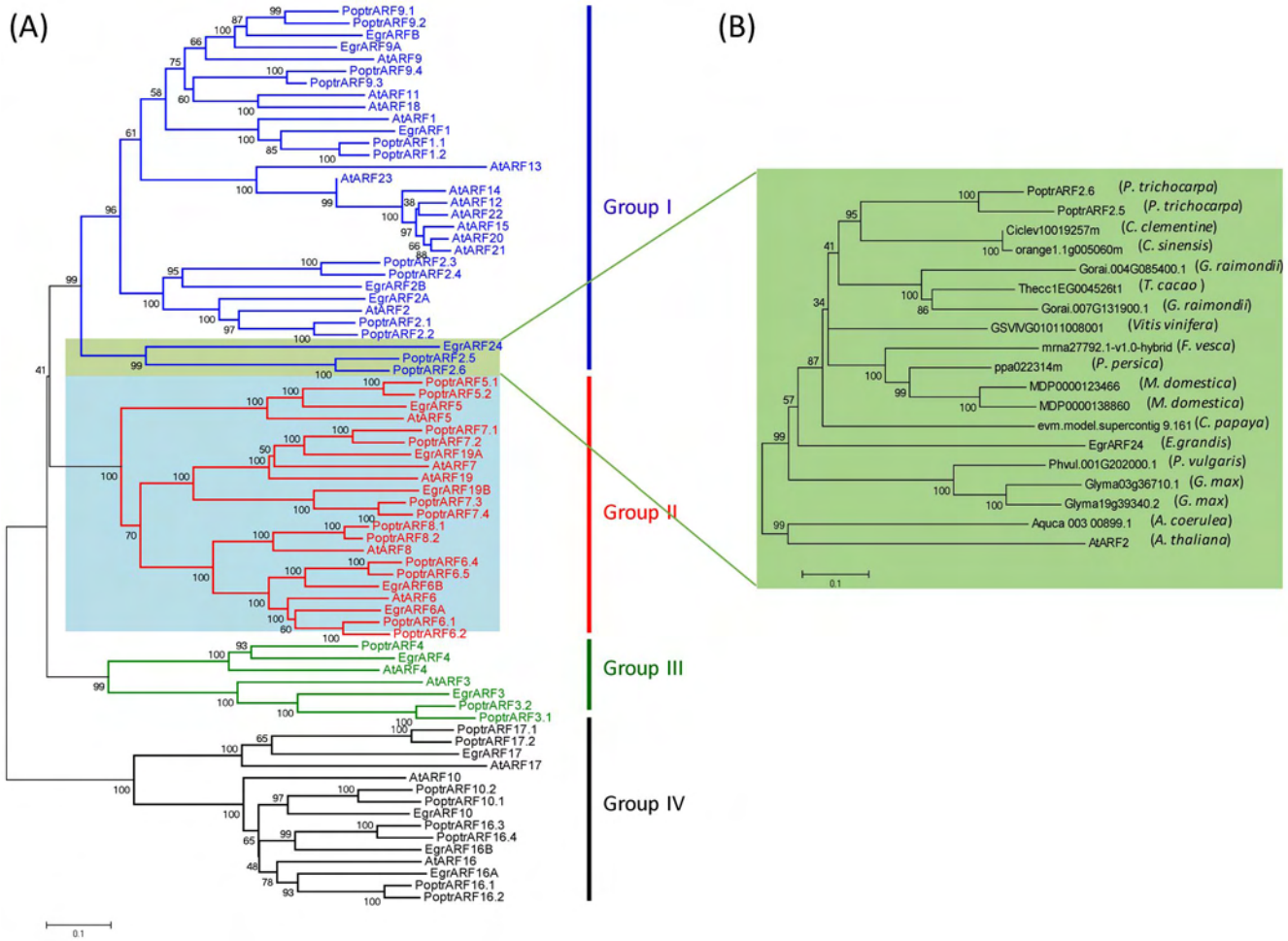


Figure 2. Phylogenetic relationships of ARF proteins between *Eucalyptus* and other species. (A) Phylogenetic relationships between ARF proteins from *Arabidopsis*, *Populus* and *Eucalyptus*. Full-length protein sequences were aligned by using the Clustal_X program. The phylogenetic tree was constructed by using the MEGA5 program and the neighbor-joining method with predicted ARF proteins. Bootstrap support is indicated at each node. The blue shade highlights the activators, and the green shade indicates the distinct likely woody preferential clade containing *EgrARF24*. (B) Phylogenetic relationships between the orthologs of *EgrARF24* in other species. *EgrARF24* proteins were used to blast 33 species genomes in Phytozome. An E-value of 1.0E-50 as used as a cut off to select the ARF potential orthologs from each species. A phylogenetic tree was constructed used the procedure as in (A) and using *AtARF2* was used as an outgroup. The species containing putative orthogs of *EgrARF24* were the followings: 1 *Aquilegia coerulea*, 2 *Glycine max*, 1 *Phaseolus vulgaris*, 1 *Carica papaya*, 2 *Malus domestica*, 1 *Prunus persica*, 1 *Fragaria vesca*, 1 *Vitis vinifera*, 2 *Populus trichocarpa*, 1 *Citrus sinensis*, 1 *Citrus clementine*, 2 *Gossypium raimondii*, 1 *Theobroma cacao*.
doi:10.1371/journal.pone.0108906.g002

potential *miR167* loci, all predicted gene products formed typical microRNA stem–loop structures (Fig. S7). The three *EgrmiR160* genes encode a mature RNA identical to that in *Arabidopsis*. The three miR167 genes produce two different mature RNA forms (Table S6) whereas in *Arabidopsis* three different mature miR167 forms were detected. We also identified a potential *TAS3* locus in the *Eucalyptus* genome (Table S6).

We used these newly identified *Eucalyptus* small RNAs as probes to search *in silico* for their target sites in *EgrARF* genes. Ten of the 17 *EgrARF* genes were found to be potential targets of these three small RNAs (Table S7). We identified highly conserved target sites for *EgrmiR160* in *EgrARF10*, *16A*, *16B* and *17*, for *EgrmiR167* in *EgrARF6A* and *B*, and for *EgrTAS3* in *EgrARF2A*, *2B*, *3* and *4* (Table S7). The targeting of three different small RNA to their corresponding target genes was highly conserved between *Arabidopsis* and *Eucalyptus* suggesting common regulation of plant growth and development. For example, *miR160*, a highly conserved miRNA group across the plant kingdom, is known to target *ARF10*, *ARF16* and *ARF17* to regulate various aspects of plant development [30,52,53]. In *Arabidopsis*, miR167 regulates lateral root outgrowth [57], adventitious rooting [58], ovule and anther growth, flower maturation [20,29] and jasmonic acid homeostasis [59] by targeting both *AtARF6* and *AtARF8*. Very recently, it has also been shown that miR167 regulates flower development and female sterility in tomato [60]. Because *Eucalyptus* is a woody perennial plant, one could expect that some small RNAs (for instance miR160 and *Tasi* 3) could be involved in the regulation of wood formation through targeting of ARF genes preferentially expressed in cambial cells or developing xylem.

3.4 Expression of *EgrARFs* in different *Eucalyptus* organs and tissues and in response to environmental cues

To start investigating the functions of the *EgrARF* genes, we assessed their transcript expression levels in various *Eucalyptus* organs and tissues by qRT-PCR, with special attention to vascular

tissues (Fig. 3, Fig. S8 and Fig. S9). Transcript accumulation was detected for 16 *EgrARFs* in all 13 organs and tissues tested (Fig. 3), except for *EgrARF24*, which was detected only in shoot tips and young leaves (Fig. S8A). The very restricted expression profile of *EgrARF24* is surprising first because this gene belongs to a woody-preferential clade and second, because its poplar orthologs *PtrARF2.5* and *PtrARF2.6* could be detected in xylem based on microarray expression data [38], *PtrARF2.6* being highly expressed in developing wood (http://popgenie.org/). It should be noted however that this gene is truncated in *E. grandis*, it has lost domain III, whereas *PtrARF2.6* and their grapevine ortholog still have domain III and IV.

Heatmap representation (Fig. 3) indicated that *EgrARF* genes were expressed across various tissues and organs, but different members displayed preference to particular tissues and/or organs and could therefore be clustered into three main expression groups. Group A is the smallest with only two members *EgrARF10* (predicted repressor) and *EgrARF19A* (predicted activator) showing a relatively higher expression in vascular cambium as compared to other tissues and/or organs. *EgrARF10* was expressed at higher level in cambium (both mature and juvenile) than in differentiating xylem and/or phloem (Fig. 3 and Fig. S9). Its ortholog in *Populus*, *PtrARF10.1*, is highly expressed in developing xylem tissues [38], suggesting that *AtARF10* orthologs in trees might be involved in wood cell differentiation having a different/supplementary role as compared to that of the *Arabidopsis* sister pair *AtARF10 – AtARF16* whose mutants exhibit root cap defects and abnormal root gravitropism [30]. *EgrARF19A* was expressed at similar levels in the three vascular tissues (Fig. 3 and Fig. S9). Group B is the largest with eight genes (*EgrARF4, 6B, 6A, 3, 1, 9A, 9B, 17*) expressed in all tissues including vascular and non-vascular tissues (Fig. 3). The expression of *EgrARF3* and *EgrARF4* is highest in root, stem and phloem and differs from the specific expression of their *Arabidopsis* orthologs *AtARF3* and *AtARF4* associated with developing reproductive and vegetative tissues. This suggests that they might be involved in other processes than the control of the abaxial identity of the gynoecium, and lateral organs shown in *Arabidopsis* [26]. Group C includes six genes (*EgrARF2A*, *2B*, *5*, *16A*, *16B*, *19B*) preferentially expressed in leaves, floral buds and fruits and virtually absent from vascular tissues and particularly from cambium and xylem (Fig. 3 and Fig. S9). As its *Arabidopsis*

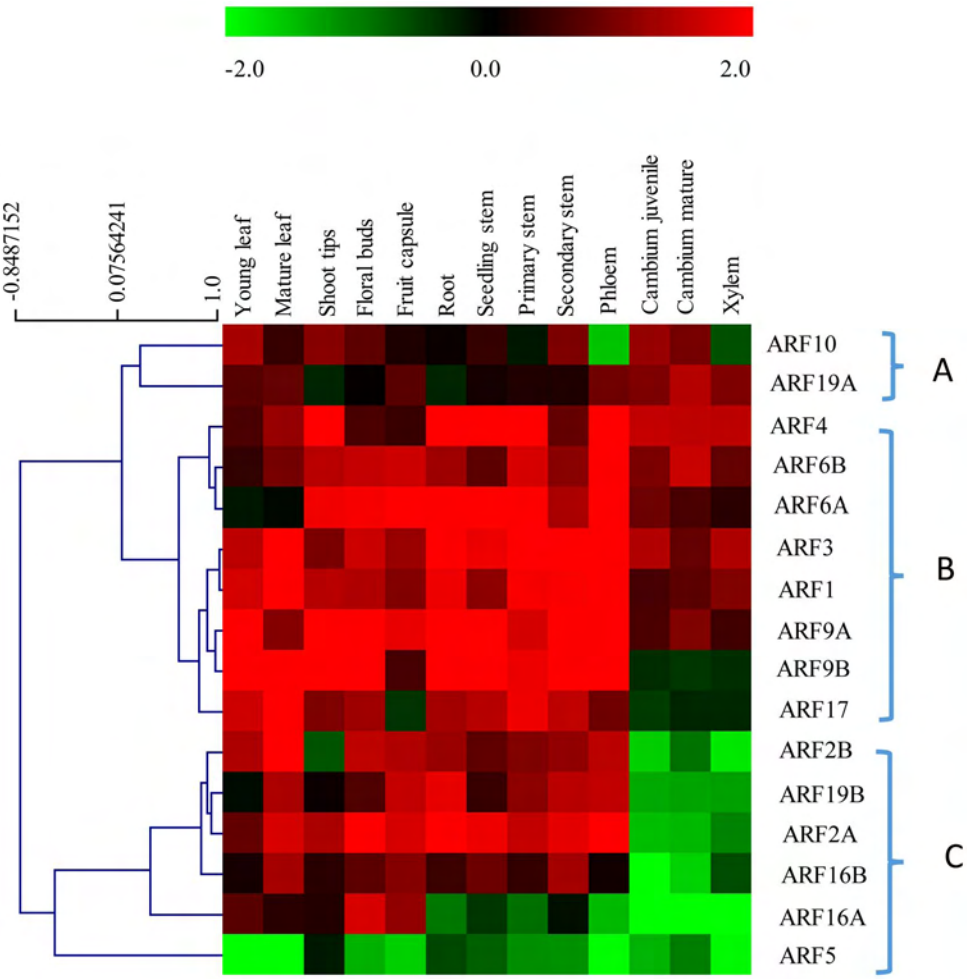


Figure 3. Expression profiles of 16 *EgrARF* genes in various organs and tissues. The heat map was constructed by using the relative expression values determined by qRT-PCR of 16 *EgrARF* genes (indicated on the right) in 13 tissues and organs (indicated at the top) normalized with a control sample (*in vitro* plantlets). In the heat map, red and green indicate relatively high and lower expression (log_2ratios) than in the control, respectively. Each measurement is the mean of three independent samples. The heat map and the hierarchical clustering were generated by MultiExperiment Viewer (MEV).
doi:10.1371/journal.pone.0108906.g003

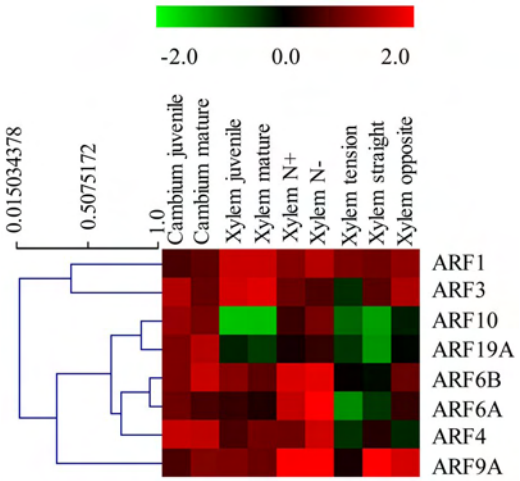


Figure 4. Effect of environmental cues and developmental stages on *EgrARF* expression. The heat map was constructed by using the relative expression values determined by qRT-PCR of *EgrARF* genes (indicated on the right) in various tissues and conditions (indicated at the top) normalized with a control sample (*in vitro* plantlets). In the heat map, red and green indicates relatively higher expression and lower expression (log_2ratios) than in the control, respectively. The heat map and the hierarchical clustering were generated by MultiExperiment Viewer (MEV).
doi:10.1371/journal.pone.0108906.g004

ortholog, *EgrARF19B* was highly expressed in root [24]. It should be noted that the activator *EgrARF5* is highly expressed in all samples analysed, with the highest expression in *in vitro* plantlets. Because *in vitro* plantlets were used to normalize the expression data in the heatmap, the expression of *EgrARF5* appeared in green in all other samples (Fig. 3). Its expression profile normalized using a different sample is given in Fig. S8.

Thirteen of the sixteen *EgrARF* genes examined (Fig. 3 and Fig. S9) exhibited higher expression in phloem than in xylem and/or cambium, suggesting that in *Eucalyptus* more *EgrARF* genes are involved in phloem than in xylem differentiation and/or function. *EgrARF5* was equally expressed in phloem and xylem. In *Arabidopsis*, *ARF5/MONOPTEROS (MP)* is known to play a critical role in the specification of vascular stem cells [27] but its role in secondary growth driven by vascular cambium activity has not been explored hitherto. *EgrARF10* and *EgrARF19A* were the only two genes more expressed in cambium and/or xylem than in other organs or tissues, supporting their possible involvement during the differentiation of meristematic cambium cells into xylem cells. No obvious difference in transcript levels were observed between juvenile and mature stages neither in cambium nor in differentiating xylem (Fig. 3 and Fig. 4).

We further examined the responsiveness to bending stress of the eight *EgrARF* genes which showed moderate to high expression in vascular tissues (Fig. 4). Half of *EgrARFs* were down-regulated in tension wood as compared to the control upright xylem, including three predicted repressors (*EgrARF3*, *4*, and *9A*) and one predicted activator (*EgrARF6A*). Conversely, in opposite xylem, four genes were up-regulated, including three predicted activators

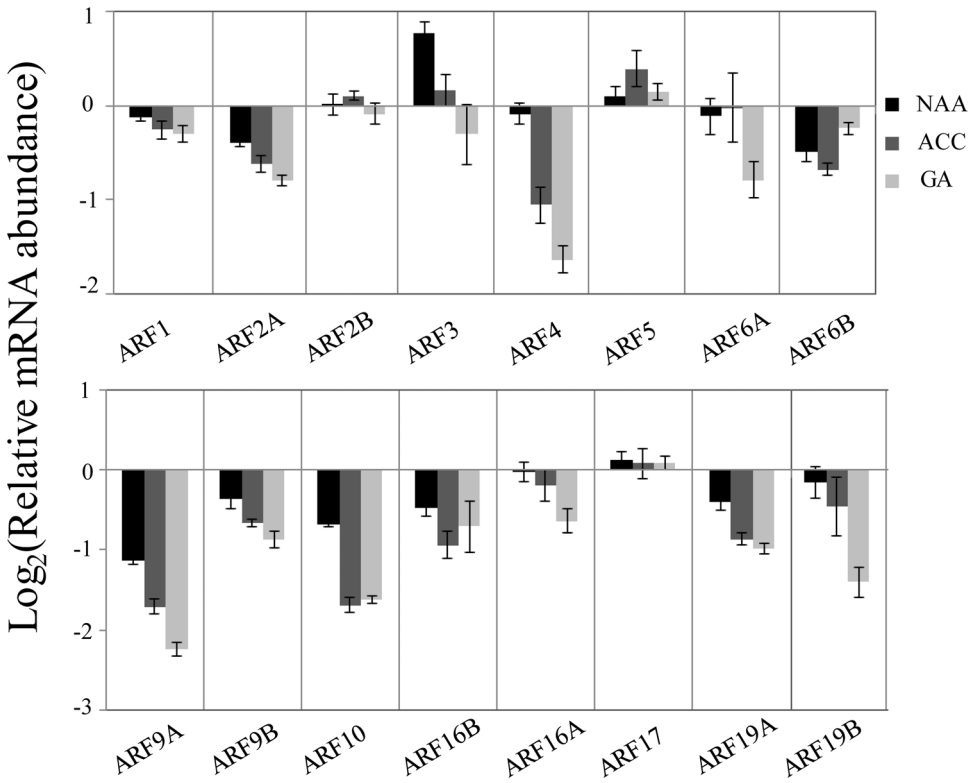


Figure 5. *EgrARF* genes expression levels in young tree stems after long term hormones treatments. Hormone treatments are detailed in the Methods section. NAA, 1-naphthaleneacetic acid, a synthetic auxin usually used in *in vitro* culture. ACC, a precursor of ethylene biosynthesis. GA, gibberellic acid. Relative mRNA abundance was compared to expression in mock-treated young tree stems. Error bars indicated the SE of mean expression values from three independent experiments.
doi:10.1371/journal.pone.0108906.g005

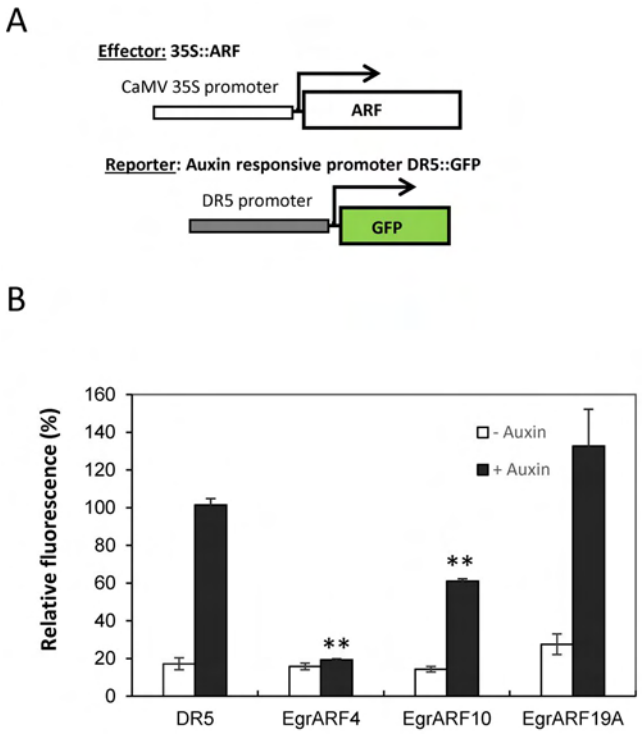


Figure 6. *Egr*ARF transcriptional activities in tobacco protoplasts. (**A**) Schemes of the effector and reporter constructs used to analyse the function of *Egr*ARFs in auxin-responsive gene expression. The effector constructs express the *Egr*ARF of interest driven by the 35S promoter. The reporter construct consists of a reporter gene expressing GFP driven by the auxin-responsive promoter DR5 (DR5::GFP). (**B**) Effector and reporter constructs were co-expressed in tobacco protoplasts in the presence or absence of a synthetic auxin (50 μM 2, 4-D). GFP fluorescence was quantified 16 h after transfection by flow cytometry. A mock effector construct (empty vector) was used as a control. In each experiment, protoplast transformations were performed in independent biological triplicates. Two independent experiments were performed and similar results were obtained; the figure indicates the data from one experiment. Error bars represent SE of mean fluorescence. Significant statistical differences (student T test, $P<0.001$) to control are marked with **.
doi:10.1371/journal.pone.0108906.g006

(*EgrARF6A*, *6B*, *19A*) and one repressor (*EgrARF10*). Only one gene (*EgrARF4*) was repressed. In general, *EgrARF* gene expression was repressed in tension xylem and induced in opposite xylem, except in the case of *EgrARF4*, which was down-regulated in both tension and opposite xylem (Fig. 4). These results are consistent with a study performed in *Populus* where seven *ARF* genes were detected in a poplar tension wood EST database, while the majority of genes were down-regulated in tension wood as compared to opposite wood [61].

Recent studies indicated that high nitrogen fertilization affects xylem development and alters fibre structure and composition in *Populus* [62,63] and induces some overlapping effects with tension wood on xylem cell walls. Interestingly, *EgrARF4* and *EgrARF6A* were down regulated in tension wood, but were down regulated when nitrogen was in excess (Fig. 4).

3.5 Effects of long-term hormone treatments on *EgrARFs* transcript levels

Several hormones are known to regulate cambium activity and xylem formation [[31], [64] and references therein]. For instance, application of exogenous ethylene (ACC) on young poplar trees during 12 days was shown to stimulate cambial activity, while xylem cell size was decreased [65]. We performed similar long-term hormonal treatments (15 days) by growing young *Eucalyptus* trees on medium supplemented with either auxin, gibberellin or ethylene in order to evaluate the consequences on the transcripts levels of the *EgrARF* genes in stems (organs enriched in xylem). The phenotypes of the *Eucalyptus* trees after hormonal treatments were typical of each hormone: gibberellin stimulated plant growth resulting in longer stems, ethylene reduced plant growth and led to epinastic leaves, whereas auxin induced shortened and bolded roots (Fig. S10). All *EgrARF* transcripts except *EgrARF24* were detected in young tree stems and the expression levels of 13 were altered and mainly down-regulated by long-term hormonal treatments (Fig. 5). Although long-term hormonal treatments likely have both direct and indirect effects on *ARFs* expression, it is interesting to note distinct and differential behaviours: Five *ARFs* exhibited a kind of "hormonal preference" response since their transcripts levels were altered in stems treated only by one of the three hormones. For instance, *EgrARF3* was up-regulated only in auxin treated samples; *EgrARF5*, only in ethylene treated samples, whereas *EgrARF6A*, *EgrARF16A* and *EgrARF19B* were altered only in gibberellin treated samples. Most of the other *ARFs* were modulated at different degrees by the direct and/or indirect actions of each of three hormones with the notable exception of *EgrARF4* that was down-regulated in stems treated by ethylene and gibberellic acid but not affected in those treated by auxin.

3.6 Transcriptional activities of EgrARF4, EgrARF 10 and EgrARF19A

We decided to characterize the transcriptional activity of three *ARF* members: *EgARF10* and *19A* which were preferentially expressed in cambium/xylem, and *EgARF4* whose expression was modulated in xylem in response to mechanical stress and to nitrogen fertilization. For this purpose, tobacco protoplasts were co-transfected with an effector construct expressing the full-length coding sequence of the *ARFs* under the *35SCaMV* promoter and a reporter construct carrying the auxin-responsive *DR5* promoter fused to *GFP* coding sequence (Fig. 6A). *DR5* is a synthetic auxin-responsive promoter made of nine inverted repeats of the conserved Auxin-Responsive Element, (TGTCTC box), fused to a *35SCaMV* minimal promoter. This reporter construct has been widely used to assess auxin responsive transcriptional activation or repression *in vivo* and *in planta* [15,47]. The *DR5*-driven GFP showed low basal activity which was induced up to 4-fold by exogenous auxin treatment (Fig. 6B). Co-transfection with the effector genes *EgrARF4* and *EgrARF10* resulted in a very significant ($p<0.001$) repression of auxin-induced reporter gene. Expression of 80% and 38%, respectively hereby confirming their predicted repressors roles. On the other hand, the values obtained for *EgrARF19A* suggested that it could be an activator as predicted by its sequence analysis, but this tendency was not strongly supported by the student-T test.

4 Conclusions

The *ARF* family in *E. grandis* contains 17 members (5 activators and 12 repressors) and is slightly contracted as compared to most angiosperm *ARF* families studied hitherto. In contrast to these species, it is characterized by the absence of whole genome, segmental and/or tandem duplication events. Indeed, whole genome duplication in *Eucalyptus* occurred 109.9 Mya ago, considerably earlier than those detected in other rosids

and 95% of the paralogs were lost [40]. The absence of tandem duplication is remarkable especially because *E. grandis* has the largest number of genes in tandem repeats (34% of the total number of genes) reported among sequenced plant genomes. Indeed, tandem duplication shaped functional diversity in many gene families in *Eucalyptus*. The *ARF* family thus evolved in a very different way. Our data suggests that genomic truncation and alternative splicing were preeminent mechanisms leading to the diversity of domain architecture, shaping and increasing the functional diversity of the *ARF* family in *Eucalyptus*, thereby compensating for the lack of extensive gene duplication found in other species. Comparative phylogenetic studies pointed out the presence of a new clade, maintained preferentially in woody and perennial plants. Finally, large scale expression profiling allowed identifying candidates potentially involved in the auxin-regulated transcriptional programs underlying wood formation.

Supporting Information

Figure S1 Procedure used for identifying *ARF* genes in *Eucalyptus grandis*. *Arabidopsis* ARF protein sequences were used to search their orthologs in the predicted *Eucalyptus* proteome by using in BLASTP. Sixty-four *Eucalyptus* proteins identified in this initial search were further examined by manual curation using protein motif scanning and the FgeneSH program to complete partial sequences. Redundant and invalid genes were eliminated based on gene structure, integrity of conserved motifs and EST support. Manual curation resulted in 17 complete Eucalyptus ARF protein sequences. These 17 protein sequences were used in two subsequent additional searches: first, a BLASTP search against the *Eucalyptus* proteome to identify exhaustively all divergent *Eucalyptus ARF* gene family members and, second, tBLASTn searches against the *Eucalyptus* genome for any possible unpredicted genes. To confirm our findings, we used poplar ARF proteins and repeated the complete search procedure described above and obtained identical results.
(TIFF)

Figure S2 Locations of the 17 *EgrARF* genes on the 11 *Eucalyptus grandis* chromosomes.
(TIFF)

Figure S3 Multiple sequence alignment of predicted amino acid sequences of *EgrARF* and *AtARF* proteins. The multiple sequence alignment was obtained with the MUSCLE software [66]. The highly conserved domains and nuclear localization signals (NLSs) proteins were noted on the bottom of the alignment with different colours.
(PDF)

Figure S4 Comparative analysis of predicted *ARF* alternative variants between *Eucalyptus grandis* and *Arabidopsis thaliana*. The alternative spliced protein sequences were extracted from Phytozome except for *AtARF4* (obtained from Finet *et al*. (2013), the motif structures were predicted by Pfam (http://pfam.xfam.org/).
(PDF)

Figure S5 Structure of the *ARF* alternative transcripts in *E. globulus*. The *E. globulus* alternative transcripts were obtained from a compendium of RNASeq data. The material and methods are described in Table S4. The illumina reads sequences are provided in File S1 in the FastQ format.
(PDF)

Figure S6 Comparative Phylogenetic relationships between ARF proteins from poplar, *Eucalyptus*, grapevine, *Arabidopsis*, tomato and rice. Full-length protein sequences were aligned using the Clustal_X program. The phylogenetic tree was constructed by using the MEGA5 program and the neighbour-joining method with predicted full-length ARF proteins. Bootstrap supports are indicated at each node.
(PDF)

Figure S7 Predicted stem-loop structures of three *EgrmiR160* and three *EgrmiR167*. The part of the stem-loop from which the mature microRNA derives is highlighted in yellow.
(TIFF)

Figure S8 Expression profiles of *EgrARF5* and *EgrARF24* in various organs and tissues. Relative mRNA abundance of *EgrARF5* and *EgrARF24* was compared to expression in the control sample of mature leaves and *in vitro* plantlets, respectively. Error bars indicate the SE of mean expression values from three independent experiments.
(TIFF)

Figure S9 Expression profiles of *EgrARF* genes in tissues involved in secondary growth. Relative mRNA abundance was compared to expression in the control sample (*in vitro* plantlets).
(TIFF)

Figure S10 Young *Eucalyptus grandis* trees phenotypes in response to various long-term hormonal treatments. 10 μM NAA, or 20 μM gibberellic acid or 100 μM ACC were added to the medium of 65-d-old young tree, and phenotypes were observed 14 days later.
(TIFF)

Table S1 Primers for *EgrARF* genes and reference genes used in qRT-PCR experiments.
(PDF)

Table S2 Protein identity matrix between *Egr*ARF and *At*ARF.
(PDF)

Table S3 Protein identity matrix among *Egr*ARF.
(PDF)

Table S4 Comparison of the number of alternative transcripts predicted in phytozome *for E. grandis* to those found in a large compendium of transcriptomic data from in *E. globulus*.
(PDF)

Table S5 Comparison of the number of *EgrARF24* putative orthologs in other species.
(PDF)

Table S6 Potential small RNAs targeting *EgrARF* genes.
(PDF)

Table S7 Small RNAs target site prediction in *EgrARF* genes.
(PDF)

File S1 Sequences of the Illumina reads from RNA Seq used to predict the *E. globulus* alternative transcripts. The origin of the material and the procedure are described in Table S4.
(ZIP)

Acknowledgments

The authors are grateful to J. P Combier for his help and advices for miRNA and target prediction, to E. Camargo (Unicamp, Brazil), J.M.

Gion and E. Villar (CIRAD, FR), F. Melun and L. Harvengt (FCBA, France), C. Araujo and L. Neves (AltriFlorestal, Portugal) and C. Marques (RAIZ, Portugal) for kindly providing and/or allowing collection of *Eucalyptus* organ and tissue samples, C. Graça, and V. Carocha for help with sample collection and RNA extraction, N Ladouce for performing Fluidigm experiments. Thanks to the Bioinfo Genotoul Platform (http://bioinfo.genotoul.fr) for access to bioinformatic resources and to the Genome and transcriptome Platform (http://get.genotoul.fr/) for advice and technical assistance with high-throughput Biomark Fluidigm qRT-PCR amplifications. Thanks to Y. Martinez (FR3450) for assistance with microscopy analysis to C. Pecher and A. Zakaroff-Girard for technical assistance and expertise in flow cytometry (Toulouse RIO imaging platform http://tri.ups-tlse.fr).

Author Contributions

Conceived and designed the experiments: HY HCW MS JGP. Performed the experiments: HY HSC CD IM BS. Analyzed the data: HY HCW MB MS JGP. Contributed reagents/materials/analysis tools: JAPP AAM. Wrote the paper: JGP YH HCW.

5 References

1. Woodward AW, Bartel B (2005) Auxin: regulation, action, and interaction. Ann Bot 95: 707-735.
2. Miyashima S, Sebastian J, Lee JY and Helariutta Y (2013) Stem cell function during plant vascular development. EMBO J 32: 178–193.
3. Tuominen H, Puech L, Fink S, Sundberg B (1997) A radial concentration gradient of indole-3-acetic acid is related to secondary xylem development in hybrid aspen. Plant Physiol 115: 577–585.
4. Uggla C, Moritz T, Sandberg G, Sundberg B (1996) Auxin as a positional signal in pattern formation in plants. Proc Natl Acad Sci U S A 93: 9282–9286.
5. Nilsson J, Karlberg A, Antti H, Lopez-Vernaza M, Mellerowicz E, et al. (2008) Dissecting the molecular basis of the regulation of wood formation by auxin in hybrid aspen. Plant Cell 20: 843–855.
6. Pilate G, Déjardin A, Laurans F, Leplé JC (2004) Tension wood as a model for functional genomics of wood formation. New Phytol 164: 63–72.
7. Paux E, Carocha V, Marques C, Mendes de Sousa A, Borralho N, et al. (2005) Transcript profiling of Eucalyptus xylem genes during tension wood formation. New Phytol 167: 89–100.
8. Timell TE (1969) The chemical composition of tension wood. Svensk Papperstidn 72: 173–181.
9. Andersson-Gunneras S, Mellerowicz EJ, Love J, Segerman B, Ohmiya Y, et al. (2006) Biosynthesis of cellulose-enriched tension wood in *Populus*: global analysis of transcripts and metabolites identifies biochemical and developmental regulators in secondary wall biosynthesis. Plant J 45: 144–165.
10. Baba K, Karlberg A, Schmidt J, Schrader J, Hvidsten TR, et al. (2011) Activity-dormancy transition in the cambial meristem involves stage-specific modulation of auxin response in hybrid aspen. Proc Natl Acad Sci U S A 108: 3418–3423.
11. Sundberg B, Uggla C, Tuominen H (2000) Cambial growth and auxin gradients. Oxford: BIOS Scientific Publishers. 169–188 p.
12. Guilfoyle TJ, Hagen G (2007) Auxin response factors. Curr Opin Plant Biol 10: 453–460.
13. Tiwari SB, Hagen G, Guilfoyle T (2003) The roles of auxin response factor domains in auxin-responsive transcription. Plant Cell 15: 533–543.
14. Hagen G, Guilfoyle T (2002) Auxin-responsive gene expression: genes, promoters and regulatory factors. Plant Mol Biol 49: 373–385.
15. Ulmasov T, Hagen G, Guilfoyle TJ (1999) Activation and repression of transcription by auxin-response factors. Proc Natl Acad Sci U S A 96: 5844–5849.
16. Gray WM, Kepinski S, Rouse D, Leyser O and Estelle M (2001) Auxin regulates SCF(TIR1)-dependent degradation of AUX/IAA proteins. Nature 414: 271–276.
17. Szemenyei H, Hannon M and Long JA (2008) TOPLESS mediates auxin-dependent transcriptional repression during *Arabidopsis* embryogenesis. Science 319: 1384–1386.
18. Vernoux T, Brunoud G, Farcot E, Morin V, Van den Daele H, et al. (2011) The auxin signalling network translates dynamic input into robust patterning at the shoot apex. Mol Syst Biol 7: 508.
19. Paponov IA, Paponov M, Teale W, Menges M, Chakrabortee S, et al. (2008) Comprehensive transcriptome analysis of auxin responses in *Arabidopsis*. Mol Plant 1: 321–337.
20. Nagpal P, Ellis CM, Weber H, Ploense SE, Barkawi LS, et al. (2005) Auxin response factors *ARF6* and *ARF8* promote jasmonic acid production and flower maturation. Development 132: 4107–4118.
21. Pufky J, Qiu Y, Rao MV, Hurban P, Jones AM (2003) The auxin-induced transcriptome for etiolated *Arabidopsis* seedlings using a structure/function approach. Funct Integr Genomics 3: 135–143.
22. Ellis CM, Nagpal P, Young JC, Hagen G, Guilfoyle TJ, et al. (2005) *AUXIN RESPONSE FACTOR 1* and *AUXIN RESPONSE FACTOR 2* regulate senescence and floral organ abscission in *Arabidopsis thaliana*. Development 132: 4563–4574.
23. Okushima Y, Mitina I, Quach HL, Theologis A (2005) *AUXIN RESPONSE FACTOR 2* (*ARF2*): a pleiotropic developmental regulator. Plant J 43: 29–46.
24. Okushima Y, Overvoorde PJ, Arima K, Alonso JM, Chan A, et al. (2005) Functional genomic analysis of the *AUXIN RESPONSE FACTOR* gene family members in *Arabidopsis thaliana*: unique and overlapping functions of *ARF7* and *ARF19*. Plant Cell 17: 444–463.
25. Sessions A, Nemhauser JL, McColl A, Roe JL, Feldmann KA, et al. (1997) *ETTIN* patterns the *Arabidopsis* floral meristem and reproductive organs. Development 124: 4481–4491.
26. Pekker I, Alvarez JP, Eshed Y (2005) Auxin response factors mediate *Arabidopsis* organ asymmetry via modulation of KANADI activity. Plant Cell 17: 2899–2910.
27. Hardtke CS, Berleth T (1998) The *Arabidopsis* gene *MONOPTEROS* encodes a transcription factor mediating embryo axis formation and vascular development. EMBO J 17: 1405–1411.
28. Harper RM, Stowe-Evans EL, Luesse DR, Muto H, Tatematsu K, et al. (2000) The *NPH4* locus encodes the auxin response factor *ARF7*, a conditional regulator of differential growth in aerial *Arabidopsis* tissue. Plant Cell 12: 757–770.
29. Wu MF, Tian Q, Reed JW (2006) Arabidopsis *microRNA167* controls patterns of *ARF6* and *ARF8* expression, and regulates both female and male reproduction. Development 133: 4211–4218.
30. Wang JW, Wang LJ, Mao YB, Cai WJ, Xue HW, et al. (2005) Control of root cap formation by MicroRNA-targeted auxin response factors in *Arabidopsis*. Plant Cell 17: 2204–2216.
31. Miyashima S, Sebastian J, Lee JY, Helariutta Y (2013) Stem cell function during plant vascular development. EMBO J 32: 178–193.
32. Xing H, Pudake RN, Guo G, Xing G, Hu Z, et al. (2011) Genome-wide identification and expression profiling of *auxin response factor* (*ARF*) gene family in maize. BMC Genomics 12: 178.
33. Wang D, Pei K, Fu Y, Sun Z, Li S, et al. (2007) Genome-wide analysis of the *auxin response factors* (*ARF*) gene family in rice (*Oryza sativa*). Gene 394: 13–24.
34. Kumar R, Tyagi AK, Sharma AK (2011) Genome-wide analysis of *auxin response factor* (*ARF*) gene family from tomato and analysis of their role in flower and fruit development. Mol Genet Genomics 285: 245–260.
35. Zouine M, Fu Y, Chateigner-Boutin AL, Mila I, Frasse P, et al. (2014) Characterization of the tomato *ARF* gene family uncovers a multi-levels post-transcriptional regulation including alternative splicing. PloS one 9: e84203.
36. Ha CV, Le DT, Nishiyama R, Watanabe Y, Sulieman S, et al. (2013) The auxin response factor transcription factor family in soybean: genome-wide identification and expression analyses during development and water stress. DNA Res 20: 511–524.
37. Mun JH, Yu HJ, Shin JY, Oh M, Hwang HJ, et al. (2012) Auxin response factor gene family in *Brassica rapa*: genomic organization, divergence, expression, and evolution. Mol Genet Genomics 287: 765–784.
38. Kalluri UC, Difazio SP, Brunner AM, Tuskan GA (2007) Genome-wide analysis of *Aux/IAA* and *ARF* gene families in *Populus trichocarpa*. BMC Plant Biol 7: 59.
39. Wan S, Li W, Zhu Y, Liu Z, Huang W, et al. (2014) Genome-wide identification, characterization and expression analysis of the auxin response factor gene family in *Vitis vinifera*. Plant Cell Rep 33: 1365–1375.
40. Myburg AA, Grattapaglia D, Tuskan GA, Hellsten U, Hayes RD, et al. (2014) The genome of *Eucalyptus grandis*. Nature 510: 356–362.
41. Voorrips RE (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. J Hered 93: 77–78.
42. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, et al. (2014) Pfam: the protein families database. Nucleic Acids Res 42: D222–230.
43. Cassan-Wang H, Soler M, Yu H, Camargo EL, Carocha V, et al. (2012) Reference genes for high-throughput quantitative RT-PCR analysis of gene expression in organs and tissues of *Eucalyptus* grown in various environmental conditions. Plant Cell Physiol 53: 2101–2116.
44. Udvardi MK, Czechowski T, Scheible WR (2008) Eleven golden rules of quantitative RT-PCR. Plant Cell 20: 1736–1737.
45. Arvidsson S, Kwasniewski M, Riano-Pachon DM, Mueller-Roeber B (2008) QuantPrime-a flexible tool for reliable high-throughput primer design for quantitative PCR. BMC Bioinformatics 9: 465.
46. Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res 29: e45.
47. Audran-Delalande C, Bassa C, Mila I, Regad F, Zouine M, et al. (2012) Genome-wide identification, functional analysis and expression profiling of the *Aux/IAA* gene family in tomato. Plant Cell Physiol 53: 659–672.
48. Shen C, Wang S, Bai Y, Wu Y, Zhang S, et al. (2010) Functional analysis of the structural domain of ARF proteins in rice (*Oryza sativa L.*). J Exp Bot 61: 3971–3981.
49. Finet C, Berne-Dedieu A, Scutt CP and Marletaz F (2013) Evolution of the *ARF* gene family in land plants: old domains, new tricks. Mol Biol Evol 30: 45–56.

50. Wu J, Wang F, Cheng L, Kong F, Peng Z, et al. (2011) Identification, isolation and expression analysis of *auxin response factor* (*ARF*) genes in *Solanum lycopersicum*. Plant Cell Rep 30: 2059–2073.
51. Remington DL, Vision TJ, Guilfoyle TJ, Reed JW (2004) Contrasting modes of diversification in the *Aux/IAA* and *ARF* gene families. Plant Physiol 135: 1738–1752.
52. Liu PP, Montgomery TA, Fahlgren N, Kasschau KD, Nonogaki H, et al. (2007) Repression of *AUXIN RESPONSE FACTOR10* by microRNA160 is critical for seed germination and post-germination stages. Plant J 52: 133–146.
53. Mallory AC, Bartel DP, Bartel B (2005) MicroRNA-directed regulation of *Arabidopsis AUXIN RESPONSE FACTOR17* is essential for proper development and modulates expression of early auxin response genes. Plant Cell 17: 1360–1375.
54. Williams L, Carles CC, Osmont KS, Fletcher JC (2005) A database analysis method identifies an endogenous trans-acting short-interfering RNA that targets the *Arabidopsis ARF2*, *ARF3*, and *ARF4* genes. Proc Natl Acad Sci U S A 102: 9703–9708.
55. Zhang B, Pan X, Cannon CH, Cobb GP, Anderson TA (2006) Conservation and divergence of plant microRNA genes. Plant J 46: 243–259.
56. Lu S, Sun YH, Chiang VL (2008) Stress-responsive microRNAs in *Populus*. Plant J 55: 131–151.
57. Gifford ML, Dean A, Gutierrez RA, Coruzzi GM, Birnbaum KD (2008) Cell-specific nitrogen responses mediate developmental plasticity. Proc Natl Acad Sci U S A 105: 803–808.
58. Gutierrez L, Bussell JD, Pacurar DI, Schwambach J, Pacurar M, et al. (2009) Phenotypic plasticity of adventitious rooting in *Arabidopsis* is controlled by complex regulation of *AUXIN RESPONSE FACTOR* transcripts and microRNA abundance. Plant Cell 21: 3119–3132.
59. Gutierrez L, Mongelard G, Flokova K, Pacurar DI, Novak O, et al. (2012) Auxin controls *Arabidopsis* adventitious root initiation by regulating jasmonic acid homeostasis. Plant Cell 24: 2515–2527.
60. Liu N, Wu S, Van Houten J, Wang Y, Ding B, et al. (2014) Down-regulation of *AUXIN RESPONSE FACTORS 6* and *8* by microRNA 167 leads to floral development defects and female sterility in tomato. J Exp Bot 65: 2507–2520.
61. Jin H, Do J, Moon D, Noh EW, Kim W, et al. (2011) EST analysis of functional genes associated with cell wall biosynthesis and modification in the secondary xylem of the yellow poplar (*Liriodendron tulipifera*) stem during early stage of tension wood formation. Planta 234: 959–977.
62. Pitre FE, Lafarguette F, Boyle B, Pavy N, Caron S, et al. (2010) High nitrogen fertilization and stem leaning have overlapping effects on wood formation in poplar but invoke largely distinct molecular pathways. Tree Physiol 30: 1273–1289.
63. Pitre FE, Pollet B, Lafarguette F, Cooke JE, MacKay JJ, et al. (2007) Effects of increased nitrogen supply on the lignification of poplar wood. J Agri Food Chem 55: 10306–10314.
64. Sorce C, Giovannelli A, Sebastiani L, Anfodillo T (2013) Hormonal signals involved in the regulation of cambial activity, xylogenesis and vessel patterning in trees. Plant Cell Rep 32: 885–898.
65. Love J, Björklund S, Vahala J, Hertzberg M, Kangasjärvi J, et al. (2009) Ethylene is an endogenous stimulator of cell division in the cambial meristem of Populus. Proc Natl Acad Sci U S A 106: 5984–5989.
66. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucl Acids Res 32: 1792–1797.

Chapter III

Comprehensive Genome-wide Analysis of Aux/IAA Gene Family in Eucalyptus: Evidence for EgrIAA4's Role in Wood Formation

Chapter III: Comprehensive Genome-wide Analysis of *Aux/IAA* Gene Family in *Eucalyptus*: Evidence for *EgrIAA4*'s Role in Wood Formation

Hong Yu[1], Marçal Soler[1], Hélène San Clemente[1], Isabelle Mila[2,3], Jorge A. P. Paiva[4,5], Alexander A. Myburg[6], Mondher Bouzayen[2,3], Jacqueline Grima-Pettenati[1(*)] and Hua Cassan-Wang[1]

[1] *LRSV Laboratoire de Recherche en Sciences Végétales, UMR5546, Université Toulouse III, CNRS, BP 42617, Auzeville, 31326 Castanet Tolosan, France*

[2]*Université de Toulouse, INP-ENSA Toulouse, Génomique et Biotechnologie des Fruits, Avenue de l'Agrobiopole BP 32607, Castanet-Tolosan F-31326, France*

[3] *INRA, UMR990 Génomique et Biotechnologie des Fruits, Chemin de Borde Rouge, Castanet-Tolosan, F-31326, France*

[4] *Instituto de Investigação Científica e Tropical (IICT/MNE), Palácio Burnay, Rua da Junqueira, 30, 1349-007 Lisboa, Portugal*

[5] *IBET - Instituto de Biologia Experimental e Tecnológica, Apartado 12, 2781-901 Oeiras, Portugal*

[6] *Department of Genetics, Forestry and Agricultural Biotechnology Institute (FABI), Genomics Research Institute (GRI), University of Pretoria, Private Bag X20, Pretoria, 0028, South Africa*

corresponding author:

Drs Jacqueline Grima-Pettenati
Laboratoire de Recherches en Sciences Végétales, LRSV
UMR 5546 CNRS/ Université Toulouse III (UPS)
24 Chemin de Borde Rouge, BP 42617, Auzeville,
31326 Castanet Tolosan, France
Emails: grima@lrsv.ups-tlse.fr
Tel: +33534323813

1 Abstract

Auxin plays a pivotal role in various plant growth and development processes, including vascular differentiation. The modulation of auxin responsiveness through the auxin perception and signaling machinery is believed to be a major regulatory mechanism to control cambium activity and wood-formation. In order to get more insights into the role of *Aux/IAAs* key regulators of auxin response in these processes, we performed a genome-wide identification and characterization of *Aux/IAA* family in *Eucalyptus grandis* a tree of high worldwide economic importance. The size of the family in *Eucalyptus* is slightly contracted as compared to *Populus* and *Arabidopsis*, but all the phylogenetic groups are represented. High-throughput expression profiling among different organs and tissues highlighted several *Aux/IAAs* expressed in vascular cambium and/or developing xylem, some showing differential expression in response to developmental (juvenile *versus* mature) and/or to environmental (tension stress) cues. Based on expression profiles, we selected the most promising candidate gene *EgrIAA4* for functional characterization. We showed that *Egr*IAA4 was localized in nucleus and functioned as an auxin-responsive repressor. Overexpressing a stabilized version of *Egr*IAA4 in *Arabidopsis* impeded dramatically plant growth and fertility and induced auxin-insensitive phenotypes such as inhibition of primary root elongation, lateral root emergence and agravitropism. More interestingly, the lignified secondary walls of the interfascicular fibers appeared very late whereas that of the xylary fibers was virtually undetectable, indicating that *EgrIAA4* plays crucial roles in fiber development and secondary cell wall deposition.

Key words: *Aux/IAA*, auxin, *Eucalyptus*, expression analysis, wood formation

Abbreviations:

Aux/IAA, auxin/indole-3-acetic acid; TIR1, transport inhibitor response1; AFB, auxin signaling F-box; ARF, auxin response factor; AuxRE, auxin-responsive *cis*-element; CaMV, Cauliflower mosaic virus; EAR, ethylene-responsive element-binding factor-associated amphiphilic repression; GFP, green fluorescent protein; MS medium,

Murashige and Skoog medium; NLS, nuclear localization signal; qRT–PCR, quantitative reverse transcription–PCR.

2 Introduction

The plant hormone auxin plays an important role in regulating plant growth and development processes like embryogenesis, apical dominance, lateral root formation, tropism, fruit development and vascular differentiation (Friml 2003). Auxin also plays a crucial role in the specification of vascular stem cells (Miyashima et al. 2013), and in the regulation of the activity of the vascular cambium (for a review see Bhalerao and Fischer (2014)] a lateral meristem that contributes to secondary radial growth of wood in trees. Measurements of auxin levels across wood-forming tissues evidenced a radial auxin concentration gradient where high auxin concentrations localize to the cambium, intermediate concentrations to the xylem elongation zone and low auxin concentrations to the maturation zone (Tuominen et al. 1997; Uggla et al. 1998; Uggla et al. 1996). It was proposed that this gradient regulates cambial activity and differentiation of cambial derivatives by providing positional information to cells within the tissue (Schrader et al. 2003; Sundberg et al. 2000; Uggla et al. 1998). However, the hypothesis that auxin acts as a kind of morphogen still lacks strong experimental support (Bhalerao and Fischer 2014; Nilsson et al. 2008). Moreover, the expression patterns of most of the auxin-responsive genes display limited correlation with the auxin concentration across the wood-forming gradient, questioning a strong and direct impact of auxin levels on radial patterning (Nilsson et al. 2008). In addition, there is only subtle variation of the absolute auxin levels between the active and the dormant cambium in trees, suggesting instead seasonal fluctuation of auxin sensitivity (Schrader et al. 2003; Schrader et al. 2004; Uggla et al. 2001; Uggla et al. 1996). Indeed, reduced auxin responsiveness of the dormant cambium correlates with reduced expression levels of components of the auxin perception machinery, implying that auxin signalling controls cambial activity by modulation of auxin responsiveness (Baba et al. 2011).

Wood is a highly variable material both developmentally and environmentally regulated

(Plomion et al. 2001). For instance, in reaction to mechanical stress, a local increase in cambial cell division induces the formation of tension wood in the upper side of the leaning angiosperm tree stems. Tension wood had specific anatomical specificities as the presence of a characteristic inner gelatinous cell-wall layer (G layer) (Pilate et al. 2004; Timell 1969). Auxin has been proposed to be implicated in the tension response, and application of either exogenous auxin or auxin transport inhibitors was shown to induce the formation of the G-fibers (Mellerowicz et al. 2001; Morey and Cronshaw 1968). For a long time, it has been assumed that auxin distribution was involved in the regulation of tension wood but Hellgren et al. (2004) have shown that auxin levels were homogeneously balanced under gravitational stress in *Populus* bended stems. Auxin may rather exert its influence *via* components of its signalling pathway, as suggested by changes in expression of a large set of auxin-related genes (Andersson-Gunneras et al. 2006) including some members of the aspen *Aux/IAA* gene family (Moyle et al. 2002; Paux et al. 2005).

The perception and signalling of auxin involve central regulators such as the Transport Inhibitor Response 1 (TIR1)/Auxin Signalling F-Box (ABFs) proteins, Auxin/Indole Acetic Acid (Aux/IAA) proteins, and Auxin Response Factor (ARF) proteins (Calderon Villalobos et al. 2012; Mockaitis and Estelle 2008). Aux/IAA proteins are direct targets of TIR1 and of its paralogs AFBs (Dharmasiri et al. 2005; Kepinski and Leyser 2005; Tan et al. 2007). At low intracellular auxin concentrations, Aux/IAA proteins act as transcriptional repressors of the auxin-mediated gene expression, by interacting and sequestering ARF proteins thus preventing them from regulating the transcription of their target genes (Guilfoyle and Hagen 2007). In contrast, high intracellular auxin levels foster interactions between Aux/IAA proteins and TIR1 E3 ubiquitin–ligase complexes, resulting in the degradation of Aux/IAA proteins by the 26S proteasome (Gray et al. 2001; Tan et al. 2007; Woodward and Bartel 2005). As a consequence, ARF proteins are released from their Aux/IAA interactants and can regulate the transcription of their auxin-responsive target genes.

The *Aux/IAA* genes were first identified in soybean and pea and described as early

auxin-responsive genes (Ainley et al. 1988; Theologis et al. 1985; Walker and Key 1982). This plant specific transcription factor family displays 29 members in *Arabidopsis* (Overvoorde et al. 2005), whose vast majority encode short-lived nuclear proteins. Canonical Aux/IAA contains four highly conserved domains (called I, II III and IV). Domain I contains a conserved Leu repeat motif (LxLxLx) similar to the EAR motif, which is responsible for the repressing activity (Tiwari et al. 2004) and can also interact with the co-repressor TOPLESS (Szemenyei et al. 2008). Domain II has a five highly conserved amino acid motif (VGWPP) that leads to the rapid degradation of Aux/IAA through interaction with a component of the ubiquitin–proteasome protein (TIR1/AFBs) degradation pathway (Dharmasiri et al. 2005; Gray et al. 2001; Kepinski and Leyser 2005). This interaction is abolished by mutations within the core sequence of domain II (VGWPP) resulting in accumulation of the mutated protein and leading to defects in auxin responses (Gray et al. 2001; Ouellet et al. 2001; Reed 2001; Tian et al. 2003). Domain III and IV can mediate homo-dimerization and hetero-dimerization with other Aux/IAA family members, as well as dimerization with ARFs which also contain these two similar domains (Kim et al. 1997; Ouellet et al. 2001; Ulmasov et al. 1997). High-throughput protein-protein interactions study between 29 Aux/IAAs and 23 ARFs in *Arabidopsis* showed that the majority of Aux/IAA proteins interact with themselves and with ARF activators (Vernoux et al. 2011).

Our current knowledge of the diverse roles of *Aux/IAA* in plant growth and development comes mainly from the characterization of gain-of-function mutants in *Arabidopsis* (Fukaki et al. 2002; Rouse et al. 1998; Tian and Reed 1999; Watahiki and Yamamoto 1997; Yang et al. 2004) and from down-regulation in *Solanaceae* species (Su et al. 2014; Bassa et al. 2012; Deng et al. 2012a; Deng et al. 2012b; Chaabouni et al. 2009; Wang et al. 2005).

Although, several studies have stressed out the importance of the auxin perception and signalling machinery in the regulation of cambial activity, cambial dormancy, secondary cell wall deposition and tension wood formation (Bhalerao and Fischer 2014), the involvement of *Aux/IAAs* in these processes remains largely under

investigated. The work from Nilsson et al. (2008) showed that changes in endogenous auxin levels in wood-forming tissues modulate the expression of a few key regulators such as *Aux/IAAs* that control the global gene expression patterns essential for normal secondary xylem development. Moreover, these authors have shown that overexpressing a stabilized version of *PttIAA3m* (mutation in the degron domain II) in *Populus* led to a reduction in cambium cell divisions and in a decrease of secondary xylem's width (Nilsson et al. 2008). To the best of our knowledge, this is the only example demonstrating the role of an *Aux/IAA* gene in xylem development in a woody species.

The *Eucalyptus grandis* genome (Myburg et al. 2014) was recently made available being the second forest tree genome sequenced and offering unique opportunities to analyze the characteristics of the *Aux/IAA* family in the most planted hardwood worldwide, which in contrast to *Populus* does not present cambium dormancy. In this study, we performed a comprehensive genome-wide identification and characterization of the *Aux/IAA* gene family in *E. grandis*. In addition to comparative phylogenetic analyses, genomic organization and prediction of protein structural motifs, we investigated by RT-qPCR the expression profiles of the 24 *Aux/IAA* members among different organ/tissues, at different developmental stages and in response to environmental cues, with special focus on wood-forming tissues. Based on these phylogenetic and expression results, we identified *EgrIAA4* as the best candidate potentially involved in the regulation of wood formation in *Eucalyptus*. Overexpression of a stabilized version of *EgrIAA4* (*EgrIAA4*m) in transgenic *Arabidopsis* led to auxin-related aberrant phenotypes and strongly affected interfascicular and xylary fiber formation, thereby confirming the hypothesized role of *EgrIAA4* in the regulation of cambium and wood forming tissues.

3 Results

3.1 Identification and sequence analysis of *Aux/IAA* gene family in *E. grandis*

The procedure to identify all members of *Aux/IAA* family in *E. grandis* is illustrated in

Fig. S1. After manual curation, a total of 26 *E. grandis Aux/IAA* genes were identified. We named them according to their putative *Arabidopsis* orthologs (henceforth referred to as *EgrIAA* genes) based on their phylogenetic relationships (Fig. S2). Two genes, *EgrIAA29A* and *29B*, encode protein sequences very similar to *EgrIAA29* but both lack domains III and IV that are crucial for Aux/IAA activity (Fig. S3).

The coding sequences of the vast majority (83%) of the *EgrIAA* genes are disrupted by three or four introns, and the introns' positions and phases are well conserved (Fig. 1). However, variations were observed for some members, implicating mainly loss of one or more introns (*EgrIAA3A*, *19* and *33A*) and gain of one additional intron (*EgrIAA9B*). The *EgrIAA* exons-introns patterns are similar to their orthologs in *Arabidopsis* (Fig. S2).

The predicted EgrIAA proteins range in size from 154 (*Egr*IAA20) to 370 amino acids (*Egr*IAA26B) with the corresponding molecular mass ranging from 17 to 41 kDa (Table 1). The theoretical isoelectric points also vary widely from 4.7 (*Egr*IAA33B) to 9.5 (*Egr*IAA31), indicating that they may function in different microenvironments. Pair-wise analysis of *Egr*IAA protein sequences showed that the identity levels greatly vary from 83.7% (*Egr*IAA3A and 3B) to 26.4% (*Egr*IAA9A and 28) (Table S1). A similar wide variation of identity values was reported in *Arabidopsis* (Overvoorde et al. 2005) and tomato (Audran-Delalande et al. 2012). Sequence alignment of the predicted proteins and MEME protein motif analyses showed that 17 of the 24 *Egr*IAA proteins harbour the typical four highly conserved domains of the canonical Aux/IAA proteins (Fig. 1; Fig. S4), comprising (i) a N-terminal repressor domain I with a conserved leucine repeat motif (LxLxLx) which can recruit the co-repressor TOPLESS, (ii) a degron domain II, which is responsible for the stability of Aux/IAA proteins, and (iii) a carboxy-terminal dimerization (CTD) domain, consisting of two highly conserved dimerization subdomains III and IV, similar to those found in ARFs (Reed 2001). However, some members displayed poor conservation or even missing domains I and/or II. For instance, the consensus sequence T/LE**L**r**L**G**L**PG in domain I, is not well in domain II is not well conserved in *Egr*IAA31 (QDWPP) and is missing in *Egr*IAA20

Table 1. *Aux/IAA* gene family in *E. grandis*

Gene name[a]	Number of predicted alternative transcript	Accession no[b.]	Chromosome[c]	Genome location[d]	ORF[e] (bp)	Deduced polypetide[f]			Exon No.
						length (aa)	MW (kDa)	PI	
EgrIAA1	2	*Eucgr.K03314.1*	11	42,167,198..42,168,788	570	189	21	5.39	4
EgrIAA3A	2	*Eucgr.H03171.1*	8	46,743,317..46,744,932	612	203	23	6.02	3
EgrIAA3B	1	*Eucgr.H00216.1*	8	2,425,940..2,428,241	615	204	23	5.58	4
EgrIAA4	1	*Eucgr.H04336.1*	8	62,350,663..62,352,118	576	191	22	5.63	4
EgrIAA9A	2	*Eucgr.H02407.1*	8	32,905,397..32,912,178	1107	368	39	6.05	5
EgrIAA9B	4	*Eucgr.F02172.1*	6	29,479,509..29,487,212	1080	359	39	6.18	6
*EgrIAA11**	1	*Eucgr.K01426.1*	11	17,332,089..17,336,565	930	309	33	5.88	4
EgrIAA13	1	*Eucgr.H02914.1*	8	42,398,802..42,401,976	801	266	29	8.8	5
EgrIAA14	1	*Eucgr.H03170.1*	8	46,720,521..46,722,685	732	243	26	6.45	5
EgrIAA15A	1	*Eucgr.J03016.1*	10	37,360,412..37,362,334	648	215	23	5.76	5
*EgrIAA15B**	1	*Eucgr.C01083.1*	3	17,083,069..17,086,237	645	214	23	7.54	5
EgrIAA16	1	*Eucgr.H04335.1*	8	62,338,061..62,341,249	813	270	29	8.28	5
*EgrIAA17**	1	*Eucgr.K03313.1*	11	42,152,995..42,155,812	807	268	29	6.06	5
EgrIAA19	1	*Eucgr.F02578.1*	6	35,331,299..35,341,010	585	194	22	8.33	3
EgrIAA20	1	*Eucgr.K00561.1*	11	6,262,537..6,264,036	465	154	17	4.99	4
EgrIAA26A	3	*Eucgr.F03080.1*	6	40,229,065..40,232,796	990	329	36	8.38	5
EgrIAA26B	2	*Eucgr.J02934.1*	10	36,614,769..36,618,150	1113	370	41	8.88	5
EgrIAA27	3	*Eucgr.F03050.1*	6	39,988,867..39,991,412	1050	349	37	6.27	4
*EgrIAA28**	2	*Eucgr.C02984.1*	3	56,397,825..56,400,087	1044	347	38	6.72	5
EgrIAA29	1	*Eucgr.C01734.1*	3	29,257,063..29,258,993	651	216	25	6.6	4

EgrIAA[T]29A	1	*Eucgr.C01731.1*	3	29,214,163 - 29,215,275	390	129	14	9.87	2
EgrIAA[T]29B	1	*Eucgr.C01732.1*	3	29,230,438 - 29,231,106	369	122	14	9.45	2
*EgrIAA31**	1	*Eucgr.H04141.1*	8	59,329,525..59,330,844	687	228	25	9.51	4
EgrIAA32	1	*Eucgr.B02853.1*	2	52,269,360..52,271,045	654	217	25	4.83	4
EgrIAA33A	1	*Eucgr.C02329.1*	3	43,423,640..43,424,409	531	176	19	8.71	2
EgrIAA33B	1	*Eucgr.C01373.1*	3	22,002,554..22,005,153	651	216	24	4.75	4

*Using FGENESH+ to complete the complete sequence

[a]Designation given to *E. grandis* in this work.

[b]Accession number in phytozome

[cd]Location of the *EgrIAA* genes to the Chromosome

[e]Length of open reading frame in base pairs

[f]The number of amino acids, molecular weight (kilodaltons), and isoelectric point (pI) of the deduced polypeptides

[T] Truncated gene

32, 33A and 33B (Fig. 1; Fig. S4). Rapid degradation of Aux/IAA proteins is essential for auxin signalling and an amino acid substitution in the degron sequence was shown to cause Aux/IAA protein accumulation leading to auxin response defects in *Arabidopsis* (Woodward and Bartel 2005). Therefore, the Aux/IAA proteins with either no or degenerated degron sequence are to be more stable, and their molecular properties may be distinct from those of canonical Aux/IAA proteins.

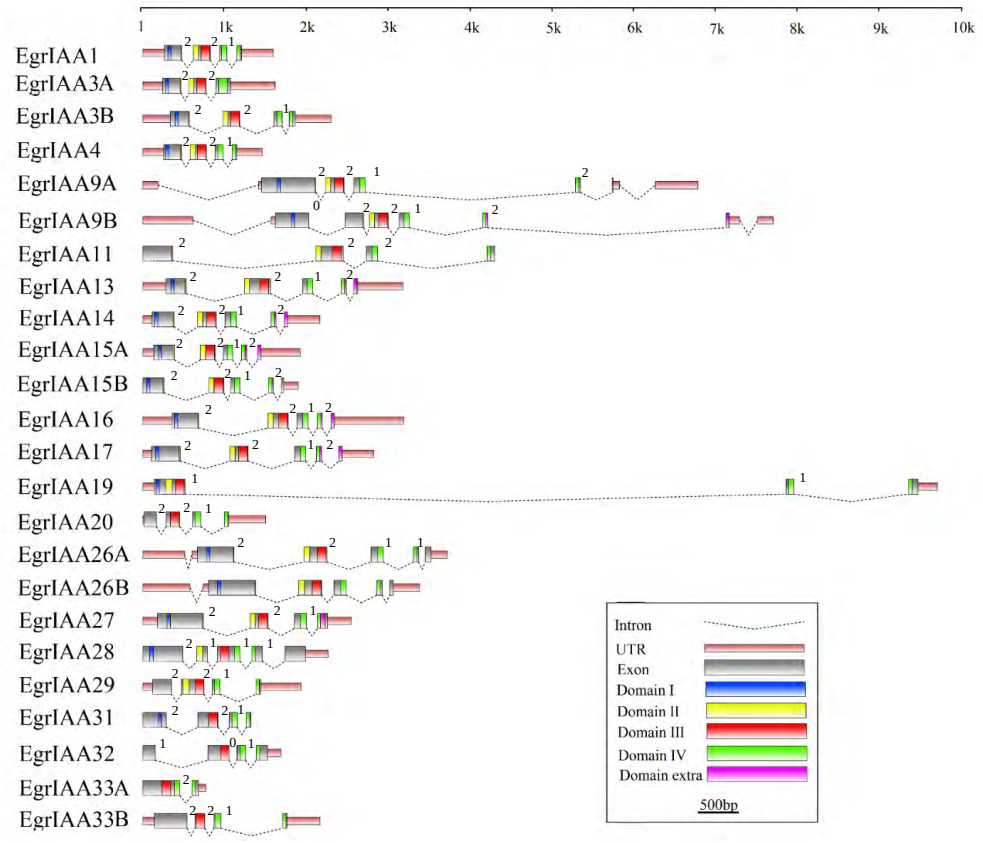


Fig. 1 Gene structures of the *EgrIAA* family members. The sizes of exons and introns are indicated by the scale at the top. The protein domains are indicated by different colours. The numbers 0, 1 and 2 represent the phases of the introns.

Two types of putative nuclear localization signals (NLS) were detected in most of the *E. grandis* Aux/IAA proteins. The first one is a bipartite structure comprising a conserved basic doublet KR between domains I and II and being associated with the presence of basic amino acids in domain II (Fig. S4). The second one is a basic residue-rich region located in domain IV that resembles to SV40-type NLS (Fig. S4). Most of

the *Egr*IAA proteins contain the two types of NLS and are most likely targeted to the nucleus consistent with their transcriptional activity. However,-family members such as *Egr*IAA29 and 33B contain a degenerated SV40-type NLS whereas *Egr*IAA20, *Egr*IAA32, *Egr*IAA33A and *Egr*IAA33B lack the bipartite NLS. Using transient expression assays with tomato Aux/IAA, Audran-Delalande et al. (2012) have shown that despite of having a degenerated NLS, *Sl*-IAA29 protein was specifically targeted to the nucleus. In contrast, in the absence of the bipartite structure like for instance in *Sl*-IAA32, the accumulation of the protein was not restricted to the nuclear compartment (Audran-Delalande et al. 2012).

3.2 Comparative phylogenetic analysis and chromosomal distribution of *EgrIAA* genes

The *Aux/IAA* family in *E. grandis* is slightly contracted with only 24 members *versus* 29 and 35 in *Arabidopsis* and *Populus*, respectively. However, its size is quite similar to that of tomato, which contains only 25 members (Table 2). To investigate the phylogenetic relationships of *Aux/IAA* family members in different species, we constructed a phylogenetic tree using predicted full-length amino acid sequences of

Table 2. Number of Aux/*IAA* family gene members in angiosperm species

	Aux/IAA	Reference
Eucalyptus grandis	24	This study
Solanum lycopersicon	25	Audran-Delalande et al. (2012)
Arabidopsis thaliana	29	Overvoorde et al. (2005)
Cucumis sativus	29	Gan et al. (2013)
Populus trichocarpa	35	Kalluri et al. (2007)
Zea mays	34	Ludwig et al. (2013)
Oryza sativa	31	Jain et al. (2006)

Aux/IAA from *Eucalyptus*, *Arabidopsis* and *Populus.* All the Aux/IAA members could be grouped into 11 groups named A-K (Fig. 2), and *Egr*IAAs are distributed among all groups and have representatives even within the four groups of non-canonical Aux/IAAs (groups H, I, J and K) lacking domain I and/or II (Overvoorde et al. 2005). As compared to *Arabidopsi*s and *Populus*, *Eucalyptus* has the fewest members of

Aux/IAA proteins in each group except in group E, which notably contains two *Eucalyptus* members (*Egr*IAA15A and 15B).

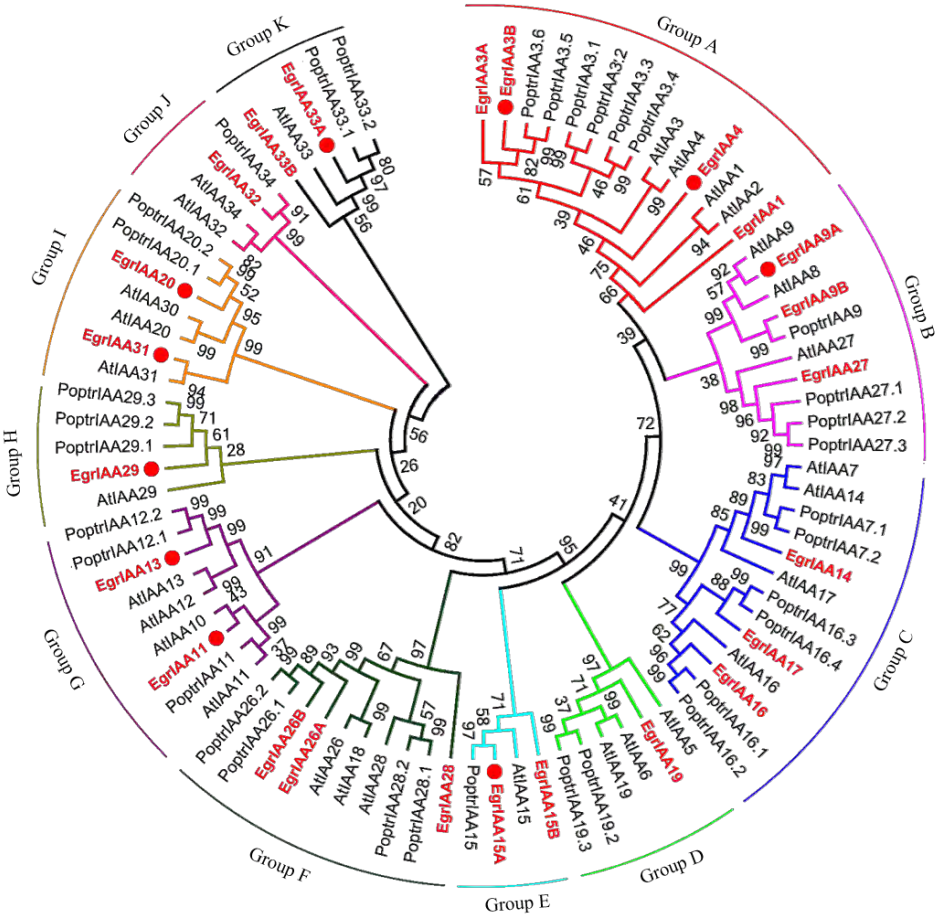


Fig. 2 Phylogenetic analysis between *Eucalyptus*, *Arabidopsis* and *Populus* Aux/IAA proteins. Full-length protein sequences were aligned using the Clustal_X program. The phylogenetic tree was constructed using the MEGA5 program and the neighbor-joining method with predicted Aux/IAA proteins. Bootstrap support is indicated at each node. Each Aux/IAA group (A to K) is indicated by a specific color. EgrIAAs are noted in red and bold. *EgrIAAs* exhibiting vascular preferential expression (cluster III, Fig 4) are marked with a red solid circle.

In silico mapping of the gene loci showed that the 24 *EgrIAA* and the two truncated genes were unevenly distributed among 6 out of the 11 *E. grandis* chromosomes with one to eight genes per chromosome-and with chromosomes 1, 4, 5, 7 and 9 being devoid of *Aux/IAA* genes (Table 1; Fig. S5). Similarly, in *Populus*, the *PtrIAAs* were present in only ten of the nineteen chromosomes with one to five genes per chromosome (Kalluri

et al. 2007), whereas *Arabidopsis Aux/IAAs* are scattered on all of the five chromosomes (Overvoorde et al. 2005). Notably, one cluster of recent tandem was detected on chromosome 3. *EgrIAA29*, *29A* and *29B* are located within a 45 kb fragment on chromosome 3 and result from a recent tandem duplication event, which led to two truncated versions and probably inactive versions of *EgrIAA29* (Fig. S3). Three pairs of *EgrIAAs* were identified as the results of segmental duplications (Myburg et al. 2014) although they are located very close to each other (within a distance < 25 kb) on chromosomes 8 (*EgrIAA4/16* and *EgrIAA3A/14)* and 11 (*EgrIAA1/17)* respectively (Fig. S5). These six genes are phylogenetically related, the more distant genes sharing 61% protein identity. Members of protein pairs shares identity levels higher than 63% (Table S1). However, each member of the protein pairs belongs to a distinct phylogenetic group. *EgrIAA1*, *3A* and *4* belong to group A, whereas *EgrIAA17*, *14* and *16* are members of group C (Fig. S5 and Fig. S2). Altogether, these data suggest that one ancestor gene went through ancient tandem duplication to give a first pair of *EgrIAAs* that subsequently went through segmental duplication events to generate the other two pairs. It seems that *Aux/IAA* family tandem duplication events in *Eucalyptus* occurred prior to the chromosomal block duplication in *Aux/IAA* family in a similar way to that proposed for *Arabidopsis* (Remington et al. 2004).

As duplication and alternative splicing are the two main mechanisms involved in diversification of function within gene families, we performed an *in silico* survey of the alternative transcripts predicted in the *E. grandis* genome JGI assembly v1.0, annotation v1.1 (http://www.phytozome.net/eucalyptus). Among the 24 *EgrIAAs*, eight of them showed more than one splice variant (Table 1). This proportion of alternative transcripts is similar to that found in *Arabidopsis*. In *Eucalyptus*, most of the alternative transcripts arise from members of groups B and F, while in *Arabidopsis*, the majority of them belong to groups B and G (Table S2, Fig. 2). Some alternative transcripts are lacking domains III and/or IV (Fig. 1, Fig. S6), likely contributing to shape the functional diversity of the family.

3.3 Expression profiling of *EgrIAA* genes in different tissues/organs and environmental cues

To gain insights into the developmental patterns of expression of *EgrIAA* genes, we assessed their transcript levels by qRT-PCR in thirteen different organs and tissues. The 24 *EgrIAA* genes could be detected in all the tissues tested, and most of them showed preferential expression in some distinct tissues (Fig. 3). The relative transcript accumulations of *EgrIAA* genes are presented as a heat map, and hierarchical clustering allowed grouping all the expression patterns into three distinct clusters (Fig. 3). Many members of the three clusters exhibited differential expression between young and

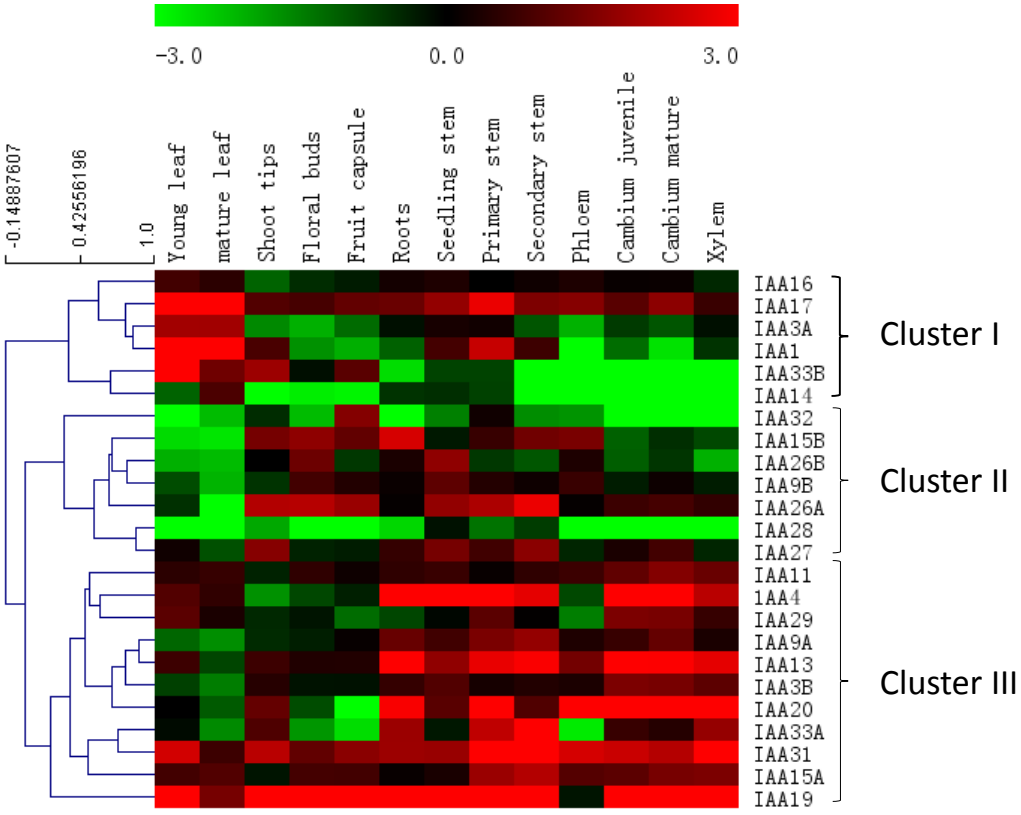


Fig. 3 Expression profiles of the 24 *EgrIAA* genes in various organs and tissues. The hierarchical clustered heat map was constructed using the $\log_2$ ratios of relative expression values normalized with a control sample (*in vitro* plantlets) of 24 *EgrIAA* genes (indicated on the right) in thirteen tissues and organs (indicated at the top). The genes belonging to cluster III are preferentially expressed in vascular tissues (phloem, cambium and/or xylem) and in organs such as stems and roots.

mature leaves. Most of them showed higher expression in young leaves whereas only *EgrIAA14* displayed higher expression level in mature leaves. Members of both cluster I and II displayed highest expression in non-vascular tissues, but they diverged regarding their differential expression in leaves. The members of cluster I showed high expression in leaves, while those of cluster II showed very low expression in these organs. Cluster III was the largest cluster, containing eleven members that displayed preferential expression in vascular tissues (phloem, cambium and xylem).

Within cluster III, *EgrIAA13*, 20 and *31* displayed relatively high expression in all three vascular tissues whereas *EgrIAA4* and *33A* showed higher expression levels in cambium and xylem than in phloem. *EgrIAA29* was preferentially expressed in cambium as compared to phloem and xylem. Most of the members of cluster III showed higher expression in juvenile xylem *versus* mature xylem (Fig. 4A) whereas no obvious differences in their expression patterns were detected between juvenile and mature cambium (Fig. 3). Of particular note, *EgrIAA4* was the only gene showing higher expression in mature xylem as compared to juvenile xylem (Fig. 4A). Most of *EgrIAA* genes from cluster III are responsive to bending mechanical stress (Fig. 4B). Four *EgrIAA* genes were down-regulated (*EgrIAA4, 11*, *13* and *20*) and two (*EgrIAA31* and *33A*) were up-regulated in tension wood as compared to straight wood (Fig. 4B). Interestingly, *EgrIAA11* and *20* were down-regulated in both tension and opposite woods, whereas *EgrIAA33A* was up-regulated in these tissues (Fig. 4B).

3.4 Candidates *EgrIAA* genes potentially involved in wood formation

In order to identify the best candidate(s) potentially involved in wood formation for functional characterization *in planta,* we defined several criteria based on (i) transcript abundance in vascular tissues, (ii) high expression in cambium and/or xylem *versus* low expression in phloem, (iii) responsiveness to bending, given that response to bending involves dramatic changes both in xylem development and secondary cell wall structure and composition, (iv) differential expression between mature *versus* juvenile xylem known to exhibit different biochemical and physical properties. The combination of all these criteria presented as a Venn diagram (Fig. S7) highlighted *EgrIAA4* and

EgrIAA33A as the most suited candidates matching the three first and main criteria. Ultimately, *EgrIAA4* was selected for further functional characterization because it was the only gene more expressed in mature than in juvenile xylem (Fig. S7).

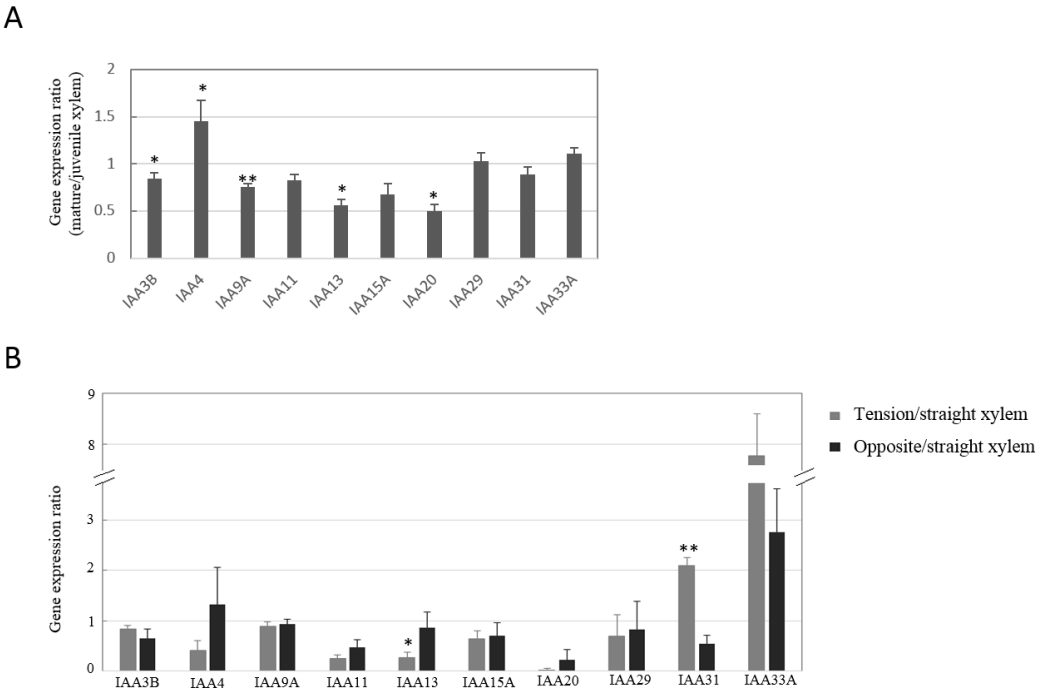


Fig. 4 Effects of developmental and environmental cues on the transcript levels of ten *EgrIAAs* preferentially expressed in vascular tissues. Ratios of the relative mRNA abundance between (A) mature and juvenile xylem, (B) tension *versus* straight xylem, and opposite *versus* straight xylem. Each relative mRNA abundance was normalized with a control sample (*in vitro* plantlets). Error bars indicate the standard error (SE) of mean expression values from three independent experiments. Asterisks indicate values found to be significantly different (student's t-test, p value) * $p<0.05$, ** $p<0.01$.

3.5 Nuclear localization and transcriptional activity of *EgrIAA4*

When transiently expressed as GFP fusion proteins in BY-2 tobacco protoplasts, *Egr*IAA4 protein was exclusively located in the nucleus (Fig. 5A) as predicted by the presence of two well conserved NLS (Fig. S4). The impact of *Egr*IAA4 on the transcriptional activity of target genes was assessed in tobacco protoplasts co-transfected with an effector construct expressing the full-length coding sequence of

EgrIAA4 under the 35SCaMV promoter and a reporter construct carrying the DR5::GFP, an auxin-responsive promoter fused to the GFP coding sequence). The DR5::GFP reporter construct is commonly used to assess auxin-dependent transcriptional activity

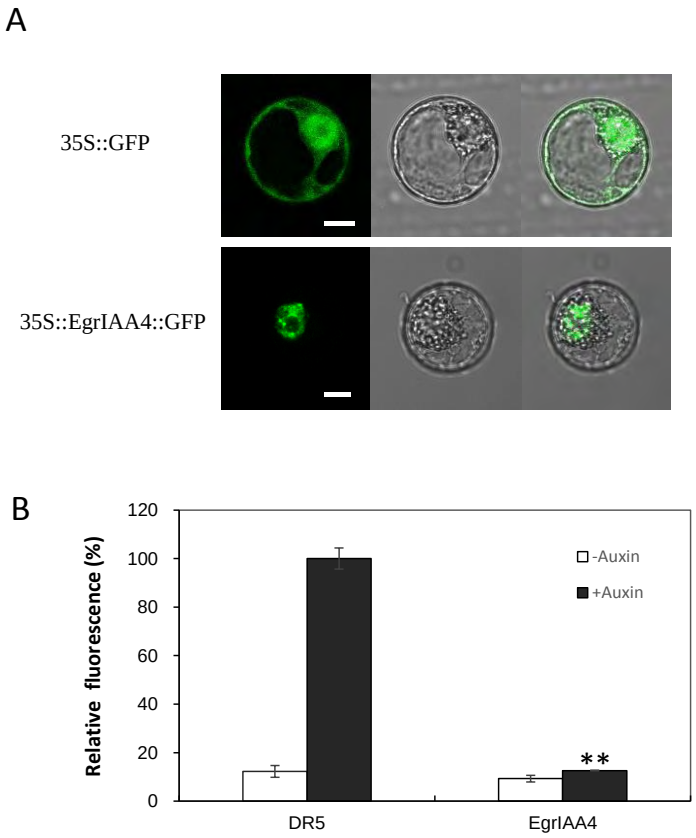


Fig. 5 Subcellular localization and repressor activity of EgrIAA4 protein on a synthetic DR5 promoter. (A) Subcellular localization of EgrIAA4-GFP protein in BY-2 tobacco protoplasts. The merged pictures of the green fluorescence channel (left panels) and the corresponding bright field (middle panels) are shown in the right panels. Scale bar, 10 μm. (B) Repressor activity of EgrIAA4 protein on a synthetic DR5 promoter. Effector and reporter constructs were co-expressed in tobacco protoplasts in the presence or absence of a synthetic auxin (50 μM 2, 4-D). A mock effector construct (empty vector) was used as control. In each experiment, protoplast transformations were performed in independent biological triplicates. Three independent experiments were performed and similar results were obtained; the figure indicates the data from one experiment. Error bars represent SE of mean fluorescence. Significant statistical differences (student T test, $P<0.001$) to control are marked with **.

in planta (Audran-Delalande et al. 2012; Ulmasov et al. 1999). The DR5-driven GFP showed a low basal activity which was induced up to 9-fold by exogenous auxin treatment (Fig. 5B). Co-transfection with the effector plasmid *EgrIAA4* resulted in a strong repression (87%) of the auxin-induced expression of the reporter gene (Fig. 5B) indicating that *Egr*IAA4 is able to mediate auxin response *in vivo* and that it functions as a strong transcriptional repressor.

3.6 Overexpression of *EgrIAA4m* affects plant growth and development

To gain insights into the role of *EgrIAA4*, we overexpressed in *Arabidopsis* a mutated version of the gene, referred to as *EgrIAA4m* (see Material and Methods and Fig. S8A), encoding a stabilized form of the protein. The mutation introduced in the degron domain was hypothesized to prevent auxin-mediated *Egr*IAA4 protein degradation. Then *EgrIAA4m* was cloned under the control of the 35SCaMV promoter into the Gateway *pFAST-G02* expression vector and transformed into *A. thaliana* Col-0. This *pFAST-G02* vector contains an *OLE1-GFP* selection marker that allows direct screening of transformed seeds displaying fluorescence under UV light (Shimada et al. 2010). Among the ten seeds sown, six germinated giving independent *EgrIAA4m* transformants, all exhibiting similar and dramatically reduced sizes as compared to controls. We selected two phenotypically representative transformants (*EgrIAA4m_1.3*, *ErIAA4m_2.3*) overexpressing *EgrIAA4m* (Fig. S8B) for further characterization. T2 seeds from both lines germinated and half of the seedlings exhibited a drastic phenotype remaining very tiny (around one cm rosette only) and displaying very small and curled-down leaves turning yellow or brown just before the plants die without flowering. (Fig. S8C). The surviving transgenic plants (T2) exhibited several distinctive phenotypes in their aerial parts (Fig. 6A, B). During the seedling stage, the cotyledons showed curled-up shape, while the leaves had a curled-down shape and very short petioles (Fig. 6A). During the vegetative stage, the *EgrIAA4m* transgenic plants from both lines showed significantly reduced plant height and rosette diameter (Fig. 6B), and the inflorescence stems were obviously thinner than those of wild type plants. In addition, *EgrIAA4m* plants showed smaller siliques and much lower fertility than wild-type (no seeds or less

than 20 seeds per plant).

3.7 *EgrIAA4m* affects root development and gravitropism

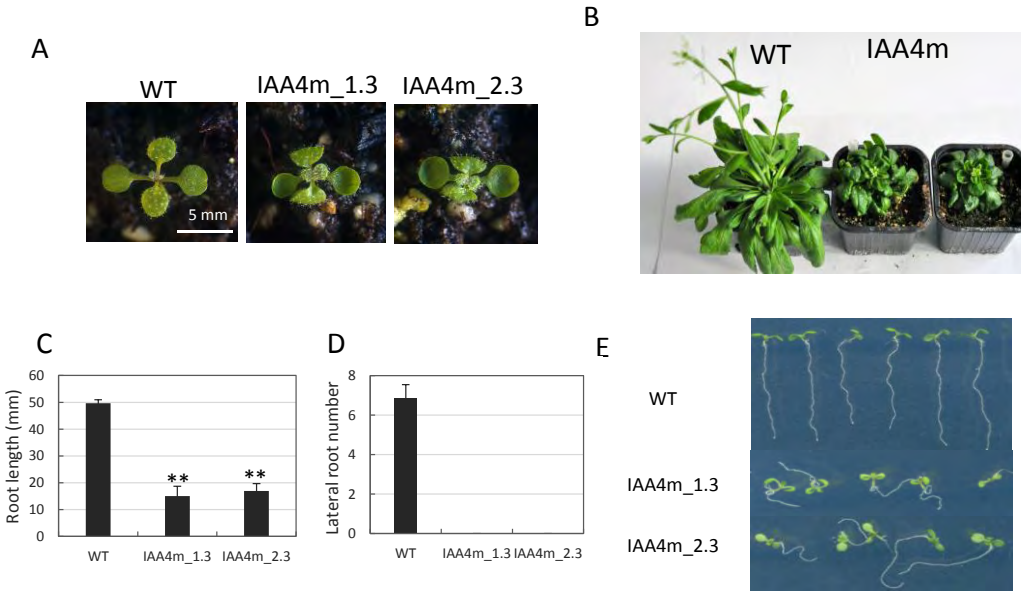


Fig. 6 Phenotypic characterization of *EgrIAA4m* transgenic plants. (A) Nine-day-old plantlets of wild-type and two representative *EgrIAA4m* lines plantlets. (B) 45-day-old wild-type and *EgrIAA4m* transgenic plants. (C, D) Primary root length and lateral root numbers in 10-day-old wild type and *EgrIAA4m* transgenic seedlings grown on 1/2 MS medium. Error bars represent SE of mean of primary root length (n>10). Significant statistical differences (student T test, P<0.001) to control are marked with **. (E) Three-day-old seedlings grown on vertical plates with 1/2 MS medium.

To examine the root phenotype of the transgenic plants, seeds collected from T1 plants were grown *in vitro* on 1/2 MS medium and the primary root length and the number of lateral roots were compared ten days after germination between wild-type and *EgrIAA4m* plants. The primary roots of *EgrIAA4m* lines were significantly shorter (around 15 mm) than those of wild type seedlings, which were approximately 49.7 mm long (Fig. 6C). In addition, no lateral roots were observed in *EgrIAA4m* transgenic seedlings, while wild-type seedlings had well-developed lateral roots of 6.9 mm (Fig. 6D). Seventeen days after germination, there was still no lateral root emergence in *EgrIAA4m* transgenic plants. These results indicate that *EgrIAA4m* overexpression

inhibits both primary root elongation and lateral root emergence.

Gravitropic response is a typical auxin-related phenotype, and several *Aux/IAA Arabidopsis* mutants showed agravitropism in roots and/or hypocotyls, such as *axr5/IAA1*, *axr2/IAA7*, *slr/IAA14*, *axr3/IAA17*and *msg2/IAA19* (Fukaki et al. 2002; Nagpal et al. 2000; Rouse et al. 1998; Watahiki and Yamamoto 1997; Yang et al. 2004). To assess the gravitropic response of the *EgrIAA4m* transgenic plant, we grew them on vertically oriented 1/2 MS plates. Ten days after germination, the hypocotyls and roots of the *EgrIAA4m* transgenic plants grew in random directions (agravitropically) as compared to the wild-type seedlings that all grew vertically (Fig. 6E). Then we reoriented the plates by 90 degrees and checked the roots' response. After 48 hours, the wild-type *Arabidopsis* roots changed their growth direction of 90 degrees, while the *EgrIAA4m* transgenic plants did not respond to the gravity change (data not shown), indicating that transgenic roots lost completely their ability to perceive gravity.

3.8 *EgrIAA4m* negatively regulates xylem differentiation in *Arabidopsis* stem

We further investigated the impact of *EgrIAA4m* overexpression on xylem and interfascicular fibers' development in *Arabidopsis* inflorescence stem (Fig. 7). The lignification patterns were obtained by staining the stem cross-sections of plants grown for 37 (Fig. 7A, B, C, D) and 47 days (E, F) with phloroglucinol-HCl. As compared to 37 day-old wild-type plants (Fig. 7A, C), the proportion of lignified tissues stained in red by phloroglucinol were dramatically reduced in transgenic lines (Fig. 7B and D) suggesting that, at the same age, the activity of both fascicular and interfascicular cambium was greatly reduced and/or delayed in the transgenic plants (Fig. 7A-F). Closer examination revealed the virtual absence of lignified interfascicular fibers in 37 day-old *EgrIAA4m* transgenic *Arabidopsis* (Fig. 7D) as compared to control (Fig. 7C). However, a very weak and discontinuous light reddish staining appeared in the interfascicular zone of 47 day-old transgenic plants (Fig. 7F), suggesting that a very limited lignification had occurred. In addition, the xylary fibers in the vascular bundle regions were completely absent and the xylem vessels were much smaller as compared

to control plants (Fig. 7C-F). These results indicate that overexpression of *EgrIAA4m* strongly and negatively affected xylem bundles and interfascicular fibers development.

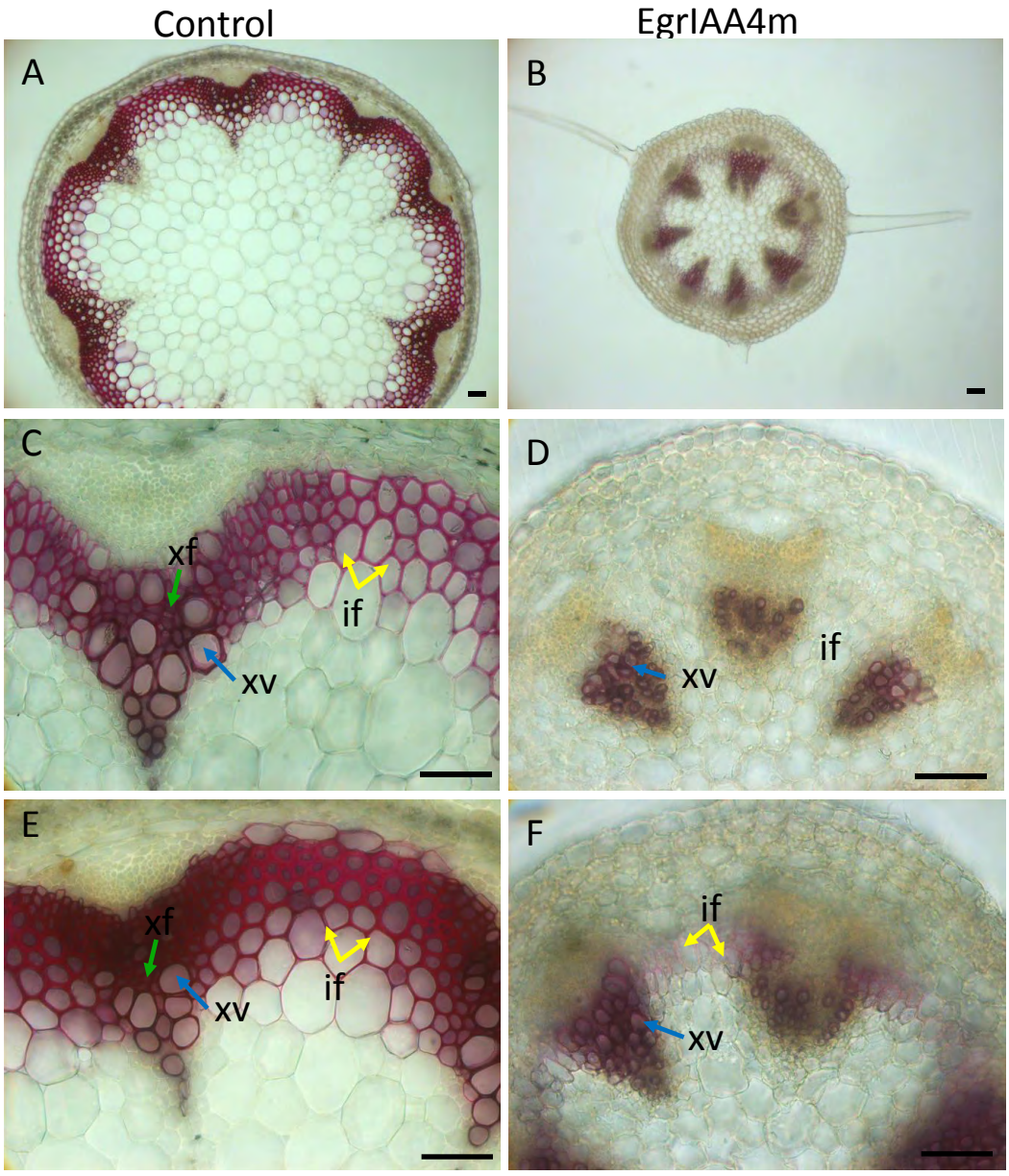


Fig. 7 Histochemical analysis of basal stem cross sections of *EgrIAA4m* transgenic *Arabidopsis*. Sections of wild-type and *EgrIAA4m* transgenic plants were stained with phloroglucinol-HCl. if, interfascicular fiber; xf, xylary fiber; xv, xylem vessel. Scale bar: 50 μm.

4 Discussion

The *Aux/IAA* family in *E. grandis* contains 24 members and therefore is slightly

contracted as compared to most angiosperms *Aux/IAA* families studied hitherto (Table 2). With 25 members, the size gene family in tomato is the closest to *Eucalyptus* but it lacks non-canonical members in group I (Audran-Delalande et al. 2012), which were shown here to be present in *Eucalyptus*. In fact, all the eleven phylogenetic groups defined in *Arabidopsis* were represented in *Eucalyptus* but with less members in each group except in group E that contains two members whereas *Arabidopsis* and *Populus* have only one member each. In *Arabidopsis*, transcripts of *AtIAA15*, the unique member of group E, were never been detected suggesting that it may be a pseudogene (Remington et al. 2004), while in *E. grandis* the transcripts of both *EgrIAA15A* and *15B* were detected by qRT-PCR, indicating they are likely functional genes. Whereas in *Arabidopsis* and *Populus* members of the *Aux/IAA* family arose predominantly through large-scale genomic duplication events (Kalluri et al. 2007; Remington et al. 2004), *Aux/IAA* family members in the *E. grandis* genome are concerned with only very few segmental duplications events. One cluster of recent tandem duplication was detected on chromosome 3, but the duplication of *EgrIAA29* led to two truncated genes lacking both domains III and IV. To the best of our knowledge, such a structure had never been reported in other species and these truncated genes are most likely not functional.

The virtual absence of tandem duplication in the *Eucalyptus Aux/IAA* family is especially striking because *E. grandis* has the largest number of genes in tandem repeats reported among sequenced plant genomes (34% of the total number of genes; Myburg et al. 2014) and tandem duplication shaped functional diversity in many gene families in *Eucalyptus* such as in the *MYB* (Soler et al. 2014) and *NAC* (Hussey et al. 2014) transcription factor families. On the other hand, this situation is quite similar to *ARF* family in *E. grandis* whose size is also slightly contracted as compared to other angiosperms genome studied (Yu et al. 2014). Because Aux/IAA proteins regulate auxin-mediated gene expression through interaction with ARF proteins, a proper dosage relationship is probably needed (Remington et al. 2004).

The expression profiles of *EgrIAAs* in various tissues and organs showed that some genes have preferential expression patterns, contrasting to ARF family members that

exhibited more constitutive expression in the same panel of samples (Yu et al. 2014). According to the current model, Aux/IAA proteins regulate auxin-mediated gene expression by protein-protein interactions with ARFs, so the preferential expression pattern of Aux/IAA genes may have a primary role in their physiological functions (Muto et al. 2007). All the *EgrIAA* genes in groups G, H and I showed vascular tissue preferential expression, and most of their orthologs in *Arabidopsis* and/or *Populus* also showed high/preferential expression in xylem cells (Table S3) (Hruz et al. 2008; Kalluri et al. 2007; Moyle et al. 2002; Nilsson et al. 2008), and/or induce vascular defects in transgenic plants (Nilsson et al. 2008; Sato and Yamamoto 2008) suggesting that these phylogenetic groups are regulating cambium activity and/or xylem development. Overexpression of *AtIAA20*, *30* and *31* in *Arabidopsis* resulted in reduced and discontinued vascular strands in cotyledons (Sato and Yamamoto 2008). Alteration of auxin responsiveness through overexpression of a stabilized version of *PttIAA3* (*EgrIAA20* 's putative ortholog*)* in transgenic aspen reduced the cambial cell division activity, caused spatial deregulation of cell division of the cambial initials and led to reductions in not only radial but also axial dimensions of fibers and vessels (Nilsson et al. 2008).

Most of the *EgrIAAs* in cluster III showed responsiveness to bending stress, but their behaviour seems quite different from that of their potential orthologs in *Populus* (Moyle et al. 2002). For example, in our conditions (three weeks of bending stress), *EgrIAA2*0 was down-regulated in tension wood, whereas its *Populus* orthologs *PttIAA3* and *4* were not affected (Moyle et al. 2002). An opposite pattern was evidenced for *EgrIAA9A* that was not affected by long-term bending, while its *Populus* ortholog *PttIAA2* was down-regulated in tension wood (Moyle et al. 2002). In fact, in a previous study, we had shown that the transcript level of *EgrIAA9A* increased sharply between six and 24 hours and then decreased between 24 h and one week, to reach a level similar to that of the control (Paux et al. 2005). *Aux/IAAs* are thus responding to bending in a time-dependent way explaining the apparent discrepancies between the responses of *Populus* and *Eucalyptus Aux/IAA* orthologs. Indeed, Moyle et al. (2002) analysed an early response

to bending (from 30 minutes to eleven hours) whereas we analysed a late response after three weeks of bending. Our results are indeed more consistent with those of Anderson-Gunneras et al. (2006) who analyzed *Populus* gene expression after three weeks bending. We found that *EgrIAA4, 11* and *20* were all down-regulated in tension wood as their *Populus* orthologs (Anderson-Gunneras et al. 2006).

EgrIAA4, which belongs to subgroup A, was considered as the best candidate to regulate cambium activity and wood formation based on its expression profile, and was chosen for functional characterization *in planta*. *EgrIAA4* is the ortholog of *Arabidopsis* sister pair *AtIAA3* and *AtIAA4*. Only the function of *AtIAA3* has been studied through gain-of-function experiments. *AtIAA3* mutation *shy2* led to plants with shorter hypocotyls, fewer lateral roots, and slower gravitropic response (Reed et al. 1998; Tian and Reed 1999). *AtIAA3* was shown recently to be part of the auxin-signalling module (SLR/IAA14–ARF7–ARF19 and SHY2/IAA3–ARFs) regulating lateral root formation (Goh et al. 2012). Likewise the down-regulation of *Sl-IAA3*, the ortholog of *AtIAA3* in the tomato, leads to reduced growth of primary and inhibition of lateral roots (Chaabouni et al. 2009). Overexpression of an auxin-insensitive version of *EgrIAA4* in *Arabidopsis* led to very similar but more severe auxin aberrant phenotypes. The gravitropic response was completely lost and the growth and development were severely affected with strongly reduced height and diameter of the inflorescence stem, reduced leaf size and fertility and absence of lateral roots. Notably, *EgrIAA4m* overexpression also led to phenotypic alterations of the vascular system not reported yet for its *Arabidopsis'*s ortholog. The development of both xylem and interfascicular fibers were dramatically delayed and reduced suggesting a reduced cambium activity in response to the altered auxin responsiveness induced by *EgrIAA4m* expression. In the transgenic plants, the lignified secondary walls of the interfascicular fibers appeared very late whereas that of the xylary fibers was virtually undetectable, indicating that *EgrIAA4* plays crucial roles in fiber development and secondary cell walls deposition consistent with its expression profile.

5 Materials and Methods

5.1 Identification of *Aux/IAA* gene family in *Eucalyptus*

Firstly, we used 29 *Arabidopsis* Aux/IAA proteins as queries to do BLASTP against *Eucalyptus grandis* genome (JGI assembly v1.0, annotation v1.1, http://www.phytozome.net/eucalyptus), and identified 55 potential *E. grandis* Aux/IAA proteins. Then, the Pfam database (http://pfam.sanger.ac.uk/search) and NCBI conserved domain database (http://www.ncbi.nlm.nih.gov/cdd) web server were used to examine the conserved domains (Finn et al. 2010; Marchler-Bauer et al. 2011). The incomplete gene models were completed by FGENESH+ (http://linux1.softberry.com), and redundant and invalid gene models were removed. The corrected *E.grandis* Aux/IAA proteins were used as query sequences in two subsequent additional search: 1) BLASTP against *E. grandis* proteome for exhaustive identification of divergent *E. grandis* gene family member, 2) TBLASTN searches against *E. grandis* genome for seeking any possible no predicted genes. Finally 26 *EgrIAA*s were identified in *E. grandis* genome (*E.grandis* genome V1.1, May 2012). For validation, we also used *Populus* Aux/IAA proteins as queries to do the complete search procedure described above, and we obtained exactly the same final 26 *EgrAux/IAA* genes. Gene information on chromosomal location was retrieved from the *E. grandis* genome browser (http://www.phytozome.net/eucalyptus) with manual curation and we mapped their loci using MapChart 2.2 (Voorrips 2002). Basic physical and chemical parameters of Aux/IAA proteins were calculated by online ProtParam tool (http://web.expasy.org/protparam/).

5.2 Sequence, gene structure and phylogenetic analysis

Conserved protein motifs of *Egr*IAA were determined by MEME-MAST programs (http://meme.sdsc.edu/meme) (Bailey et al. 2009). The exon-intron structures were extracted from Phytozome with manual curation and visualized by Fancy Gene v1.4 (http://bio.ieo.eu/fancygene/). Multiple protein sequences alignment was performed using Clustal_X2 program (Version 2.0.11). All predicted protein sequences were used

to do phylogenetic analysis and the phylogenetic trees were constructed with MEGA5 program by neighbour-joining method with 1000 bootstrap.

5.3 Plant materials and growth conditions

All *Eucalyptus* organ/tissues provenance and preparation are described in Cassan-Wang et al. (2012). *A. thaliana* ecotype Col-0 plants were grown in growth chamber: 16 h day/8 h night for long days, 22/20 ℃ day/night temperature, 70% relative humidity, 200 µmol photons $m^{-1}s^{-1}$ light intensity (intense luminosity). The plants were watered every two days and fertilized weekly. Seeds were surface-sterilized for 1 min in 70% ethanol, 10 min in 25% bleach, rinsed five times in sterile water and plated on Murashige and Skoog (MS) medium containing 1.0% sucrose solidified with 1% agar.

5.4 RNA isolation, cDNA synthesis and qRT-PCR

Total RNAs were extracted from 100–200 mg of frozen material as described by (Southerton et al. 1998) and treated by Turbo DNA-*free*™ kit (Ambion, Carlsbad, USA) to remove the genomic DNA contamination. RNA concentration and purity were determined by using a NanoDrop spectrophotometer ND-1000 (Thermo Scientific, Waltham, USA) and the integrity was assessed by using the Agilent 2100 Bioanalyzer (Agilent, Santa Clara, USA). Only samples with RNA Integrity Number (RIN) over 7 were used for reverse transcription. cDNA was reversed transcribed from 2 µg total RNA using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster, USA).

Primers were designed using the software QuantPrime (qPCR primer design tool: http://www.quantprime.de/ (Arvidsson et al. 2008) and the sequences are shown in Table S4. Oligonucleotides were synthesized by Sigma Life Science. qRT-PCR was performed by the Genotoul service in Toulouse (http://genomique.genotoul.fr/) using the BioMark® 96:96 Dynamic Array integrated fluidic circuits (Fluidigm Corporation, San Francisco, USA) as described in (Cassan-Wang et al. 2012). Only primers that produced a linear amplification and qPCR products with a single-peak melting curves were used for further analysis. The efficiency of each pair of primers was determined

from the data of amplification Ct value plot with a serial dilution of mixture cDNA and the equation $E = 10^{(-1/slope)} - 1$. $E^{-\Delta\Delta Ct}$ method was used to calculate relative mRNA fold change compared to control sample using formula $(E_{target})^{\Delta Ct\_target\ (control-sample)}/(E_{reference})^{\Delta Ct\_reference\ (control-sample)}$ (Pfaffl 2001) and five reference genes (IDH, PP2A1, PP2A3, EF-1a and SAND, Table S4) were used for data normalization. *In vitro* plantlets were used as control sample.

5.5 *EgrIAA4* amplification and gain-of-function transgenic *Arabidopsis* construction

The *EgrIAA4* gene was amplified by polymerase chain reaction (PCR) using Phusion Taq (Thermo, Waltham, USA) by a gene specific primer pair: 5'-CACCATGGCAGCTCAAGGAGAGGAT-3' and 5'-AACCTCTGATGACCCTTTCATGATT-3' which was based on the prediction from *E. grandis* genome v.1.1 (Eucgr.H04336.1). To investigate the function of this gene, we create a mutation in the auxin degradation domain changing the 80th amino acid from proline to serine (P-to-S) in domain II by using overlap PCR. The overlap primers are 5'-**ATC**GGACCGGACTCCACCCCACGACTTGTGCCTTA-3' and 5'-CGTGGGGTGG**AGT**CCGGTCCGATCCTACCGAAA-3'. The underlined sequences are the overlap region and the bold bases are the mutated nucleotides. The *EgrIAA4m* fragment was cloned into the *pENTR/D-TOPO* vector (Invitrogen, Carlsbad, USA) to sequence. After sequencing, we recombined *EgrIAA4m* fragment into the destination vector *pFAST-G02* using LR Clonase (Invitrogen, Carlsbad, USA). pFAST-*EgIAA4m* vector was transformed into *A. tumefaciens* strain GV3101, and then transformed into *A. thaliana* ecotype Col-0 using the floral dip method (Clough and Bent 1998).

5.6 Transient expression using a single cell system

Protoplast for transfection were obtained from suspension-cultured tobacco (*Nicotiana tabacum*) BY-2 cells according to the method of Leclercq et al. (2005). Protoplast were transfected by a modified polyethylene glycol method as described by Abel and Theologis (1994). For nuclear localization of the selected Aux/IAA, the full length

cDNAs were fused in frame at the C-terminus with GFP in the *pK7FWG2.0* vector (Karimi et al. 2002) under the control of 35SCaMV promoter. Transfected protoplasts were incubated for 16 h at 25 ℃ and examined for GFP florescence signals using a Leica TCS SP2 laser scanning confocal microscope. Images were obtained with a x40 water immersion objective. For co-transfection assays, the full-length cDNAs of the selected *Aux/IAA* were cloned into pGreen vector under 35SCaMV promoter to create the effector constructs. The reporter constructs used a synthetic auxin-responsive promoter DR5 fused to the GFP reporter gene. Tobacco BY-2 protoplasts were co-transfected with the reporter and effector constructs as described in Audran-Delalande et al. (2012). After 16h incubation, GFP expression was quantified by flow cytometry (LSR Fortessa, BD Bioscience), and the data were analysed using Cell Quest software BD FacsDiva software. Transfection assays were performed in three independent replicates and 3000-4000 protoplasts were gated for each sample. GFP fluorescence corresponds to the average fluorescence intensity of the protoplasts population after subtraction of auto-fluorescence determined with non-transformed protoplasts. 50 µM 2, 4-D was used for auxin treatment.

5.7 Microscopy analysis

The *Arabidopsis* inflorescence stems at the basal end (~1 cm) and hypocotyls were harvested at 37 and 47 days old, and then stored in 70% ethanol. The cross sections were prepared using vibratome Leica VT1000 S (Leica, Paris, France). Lignin polymers are the characteristic components of secondary cell wall (SCW) and are normally absent from primary cell wall, therefore we used lignin deposition detection techniques to screen for SCW phenotype. Cross sections of inflorescence stem and hypocotyl (~80 µm) were stained with phloroglucinol-HCl, which stains specifically lignin polymer precursors coniferaldehyde and p-coumaraldehyde in the SCW giving violet-red color. Phloroglucinol-HCl was directly applied on the slide and images were recorded with a CCD camera (Photonic Science, Robertsbridge, UK, http://www.photonic-science.co.uk).

Acknowledgements

This work was supported by the Centre National pour la Recherche Scientifique (CNRS), the University Paul Sabatier Toulouse III (UPS), the French Laboratory of Excellence project " TULIP" (ANR-10-LABX-41; ANR-11-IDEX-0002-02), the Plant KBBE TreeForJoules project [ANR-2010-KBBE-007-01 (FR) and P-KBBE/AGR_GPL/0001/2010 (FCT, PT)], and the project microEGo (PTDC/AGR-GPL/098179/2008; FCT, PT). HY were supported by PhD grants from the China Scholarship council. MS was supported by a Postdoc fellowship “Beatriu de Pinós” thanks to the Departament d'Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya. JAPP acknowledges FCT for the research contract *Ciência 2008* program and the postdoctoral fellowship SFRH/BPD/92207/2013. The authors are grateful to C. Dunand and Q. Li for the help of bioinformatics analysis, to E. Camargo, J.M. Gion and E. Villar (CIRAD, FR), F. Melun and L. Harvengt (FCBA, France), C. Araujo and L. Neves (AltriFlorestal, Portugal) and C. Marques (RAIZ, Portugal) for kindly providing and/or allowing collection of *Eucalyptus* organ and tissue samples, C. Graça, and N. Saidi for help with sample collection and RNA extraction, N Ladouce and V. Carocha for performing Fluidigm experiments. Thanks to the Bioinfo Genotoul Platform (http://bioinfo.genotoul.fr) for access to resources and to the Genome and transcriptome Platform (http://get.genotoul.fr/) for advice and technical assistance with high-throughput Biomark Fluidigm qRT-PCR amplifications. Thanks to Y. Martinez (FR3450) for assistance with microscopy analysis (Plateforme Imagerie TRI http://tri.ups-tlse.fr/). Thanks to C. Pecher and A. Zakaroff-Girard for their technical assistance and expertise in flow cytometry (Cytometry and cell sorting platform, INSERM UPS UMR 1048, Toulouse RIO imaging platform).

6 References

Abel, S. and Theologis, A. (1994) Transient transformation of *Arabidopsis* leaf protoplasts: a versatile experimental system to study gene expression. *Plant J.* 5: 421-427.

Ainley, W.M., Walker, J.C., Nagao, R.T. and Key, J.L. (1988) Sequence and characterization of two auxin-regulated genes from soybean. *J Biol Chem.* 263: 10658-10666.

Andersson-Gunneras, S., Mellerowicz, E.J., Love, J., Segerman, B., Ohmiya, Y., Coutinho, P.M., et al. (2006) Biosynthesis of cellulose-enriched tension wood in *Populus*: global analysis of transcripts and metabolites identifies biochemical and developmental regulators in secondary wall biosynthesis. *Plant J.* 45: 144-165.

Arvidsson, S., Kwasniewski, M., Riano-Pachon, D.M. and Mueller-Roeber, B. (2008) QuantPrime-a flexible tool for reliable high-throughput primer design for quantitative PCR. *BMC Bioinformatics* 9: 465.

Audran-Delalande, C., Bassa, C., Mila, I., Regad, F., Zouine, M. and Bouzayen, M. (2012) Genome-wide identification, functional analysis and expression profiling of the *Aux/IAA* gene family in tomato. *Plant Cell Physiol.* 53: 659-672.

Baba, K., Karlberg, A., Schmidt, J., Schrader, J., Hvidsten, T.R., Bako, L., et al. (2011) Activity-dormancy transition in the cambial meristem involves stage-specific modulation of auxin response in hybrid aspen. *Proc Natl Acad Sci U S A.* 108: 3418-3423.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., et al. (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 37: W202-208.

Bassa, C., Mila, I., Bouzayen, M. and Audran-Delalande, C. (2012) Phenotypes associated with down-regulation of *Sl-IAA27* support functional diversity among *Aux/IAA* family members in the tomato. *Plant Cell Physiol.* 53: 1583-1595.

Bhalerao, R.P. and Fischer, U. (2014) Auxin gradients across wood-instructive or incidental? *Physiol Plant.* 151: 43-51.

Calderon Villalobos, L.I., Lee, S., De Oliveira, C., Ivetac, A., Brandt, W., Armitage, L., et al. (2012) A combinatorial TIR1/AFB-Aux/IAA co-receptor system for differential sensing of auxin. *Nat Chem Biol.* 8: 477-485.

Cassan-Wang, H., Soler, M., Yu, H., Camargo, E.L., Carocha, V., Ladouce, N., et al. (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: 2101-2116.

Chaabouni, S., Jones, B., Delalande, C., Wang, H., Li, Z., Mila, I., et al. (2009) Sl-IAA3, a tomato Aux/IAA at the crossroads of auxin and ethylene signalling involved in differential growth. *J Exp Bot.* 60: 1349-1362.

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16: 735-743.

Deng, W., Yan, F., Liu, M., Wang, X. and Li, Z. (2012a) Down-regulation of *SlIAA15* in tomato altered stem xylem development and production of volatile compounds in leaf exudates. *Plant Signal Behav.* 7: 911-913.

Deng, W., Yang, Y., Ren, Z., Audran-Delalande, C., Mila, I., Wang, X., et al. (2012b) The tomato *SlIAA15* is involved in trichome formation and axillary shoot development. *New Phytol.* 194: 379-390.

Dharmasiri, N., Dharmasiri, S. and Estelle, M. (2005) The F-box protein TIR1 is an auxin receptor. *Nature.* 435: 441-445.

Finn, R.D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J.E., et al. (2010) The Pfam protein families database. *Nucleic Acids Res.* 38: D211-222.

Friml, J. (2003) Auxin transport - shaping the plant. *Curr Opin Plant Biol.* 6: 7-12.

Fukaki, H., Tameda, S., Masuda, H. and Tasaka, M. (2002) Lateral root formation is blocked by a gain-of-function mutation in the *SOLITARY-ROOT/IAA14* gene of *Arabidopsis*. *Plant J.* 29: 153-168.

Gan, D., Zhuang, D., Ding, F., Yu, Z., and Zhao, Y. (2013) Identification and expression analysis of primary auxin-responsive *Aux/IAA* gene family in cucumber (*Cucumis sativus*). *J Genet*. 92: 513-21.

Goh, T., Kasahara, H., Mimura, T., Kamiya, Y. and Fukaki, H. (2012) Multiple AUX/IAA-ARF modules regulate lateral root formation: the role of *Arabidopsis* SHY2/IAA3-mediated auxin signalling. *Philos Trans R Soc Lond B Biol Sci.* 367: 1461-1468.

Gray, W.M., Kepinski, S., Rouse, D., Leyser, O. and Estelle, M. (2001) Auxin regulates SCF(TIR1)-dependent degradation of AUX/IAA proteins. *Nature* 414: 271-276.

Guilfoyle, T.J. and Hagen, G. (2007) Auxin response factors. *Curr Opin Plant Biol.* 10: 453-460.

Hellgren, J.M., Olofsson, K. and Sundberg, B. (2004) Patterns of auxin distribution during gravitational induction of reaction wood in poplar and pine. *Plant Physiol.* 135: 212-220.

Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., et al. (2008) Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Adv Bioinformatics.* 2008: 420747.

Hussey, S.G., Saïdi, M.N., Hefer, C.A., Myburg, A.A. and Grima-Pettenati, J. (2014)

Structural, evolutionary and functional analysis of the NAC domain protein family in *Eucalyptus*. *New Phytol. in press*.

Jain, M., Kaur, N., Garg, R., Thakur, J.K., Tyagi, A.K., and Khurana, J.P. (2006) Structure and expression analysis of early auxin-responsive *Aux/IAA* gene family in rice (Oryza sativa). *Funct Integr Genomics*. 6: 47-59.

Kalluri, U.C., Difazio, S.P., Brunner, A.M. and Tuskan, G.A. (2007) Genome-wide analysis of *Aux/IAA* and *ARF* gene families in *Populus trichocarpa*. *BMC Plant Biol.* 7: 59.

Karimi, M., Inze, D. and Depicker, A. (2002) GATEWAY vectors for *Agrobacterium*-mediated plant transformation. *Trends Plant Sci.* 7: 193-195.

Kepinski, S. and Leyser, O. (2005) The *Arabidopsis* F-box protein TIR1 is an auxin receptor. *Nature.* 435: 446-451.

Kim, J., Harter, K. and Theologis, A. (1997) Protein-protein interactions among the Aux/IAA proteins. *Proc Natl Acad Sci U S A.* 94: 11786-11791.

Leclercq, J., Ranty, B., Sanchez-Ballesta, M.T., Li, Z., Jones, B., Jauneau, A., et al. (2005) Molecular and biochemical characterization of LeCRK1, a ripening-associated tomato CDPK-related kinase. *J Exp Bot.* 56: 25-35.

Ludwig, Y., Zhang, Y., and Hochholdinger F. (2013) The maize (*Zea mays L.) AUXIN/INDOLE-3-ACETIC ACID* gene family: phylogeny, synteny, and unique root-type and tissue-specific expression patterns during development. *PLoS One*. 8: e78859.

Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., et al. (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* 39: D225-229.

Mellerowicz, E.J., Baucher, M., Sundberg, B. and Boerjan, W. (2001) Unravelling cell wall formation in the woody dicot stem. *Plant Mol Biol.* 47: 239-274.

Miyashima, S., Sebastian, J., Lee, J.Y. and Helariutta, Y. (2013) Stem cell function during plant vascular development. *EMBO J.* 32: 178-193.

Mockaitis, K. and Estelle, M. (2008) Auxin receptors and plant development: a new signaling paradigm. *Annu Rev Cell Dev Biol.* 24: 55-80.

Morey, P.R. and Cronshaw, J. (1968) Induction of tension wood by 2,4-dinitrophenol and auxins. *Protoplasma.* 65: 393-405.

Moyle, R., Schrader, J., Stenberg, A., Olsson, O., Saxena, S., Sandberg, G., et al. (2002) Environmental and auxin regulation of wood formation involves members of the *Aux/IAA* gene family in hybrid aspen. *Plant J.* 31: 675-685.

Muto, H., Watahiki, M.K., Nakamoto, D., Kinjo, M. and Yamamoto, K.T. (2007) Specificity and similarity of functions of the *Aux/IAA* genes in auxin signaling of

Arabidopsis revealed by promoter-exchange experiments among *MSG2/IAA19*, *AXR2/IAA7,* and *SLR/IAA14*. *Plant Physiol.* 144: 187-196.

Myburg, A.A., Grattapaglia, D., Tuskan, G.A., Hellsten, U., Hayes, R.D., Grimwood, J., et al. (2014) The genome of *Eucalyptus grandis*. *Nature.* 509: 356-362.

Nagpal, P., Walker, L.M., Young, J.C., Sonawala, A., Timpte, C., Estelle, M., et al. (2000) AXR2 encodes a member of the Aux/IAA protein family. *Plant Physiol.* 123: 563-574.

Nilsson, J., Karlberg, A., Antti, H., Lopez-Vernaza, M., Mellerowicz, E., Perrot-Rechenmann, C., et al. (2008) Dissecting the molecular basis of the regulation of wood formation by auxin in hybrid aspen. *Plant cell.* 20: 843-855.

Ouellet, F., Overvoorde, P.J. and Theologis, A. (2001) *IAA17/AXR3:* biochemical insight into an auxin mutant phenotype. *Plant cell.* 13: 829-841.

Overvoorde, P.J., Okushima, Y., Alonso, J.M., Chan, A., Chang, C., Ecker, J.R., et al. (2005) Functional genomic analysis of the *AUXIN/INDOLE-3-ACETIC ACID* gene family members in *Arabidopsis thaliana*. *Plant cell.* 17: 3282-3300.

Paux, E., Carocha, V., Marques, C., Mendes de Sousa, A., Borralho, N., Sivadon, P., et al. (2005) Transcript profiling of *Eucalyptus* xylem genes during tension wood formation. *New Phytol.* 167: 89-100.

Pfaffl, M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29: e45.

Pilate, G., Déjardin, A., Laurans, F. and Leplé, J.C. (2004) Tension wood as a model for functional genomics of wood formation. *New Phytol.* 164: 63-72.

Plomion, C., Leprovost, G. and Stokes, A. (2001) Wood formation in trees. *Plant Physiol.* 127: 1513-1523.

Reed, J.W. (2001) Roles and activities of Aux/IAA proteins in *Arabidopsis*. *Trends Plant Sci.* 6: 420-425.

Reed, J.W., Elumalai, R.P. and Chory, J. (1998) Suppressors of an *Arabidopsis thaliana phyB* mutation identify genes that control light signaling and hypocotyl elongation. *Genetics* 148: 1295-1310.

Remington, D.L., Vision, T.J., Guilfoyle, T.J. and Reed, J.W. (2004) Contrasting modes of diversification in the *Aux/IAA* and *ARF* gene families. *Plant Physiol.* 135: 1738-1752.

Rouse, D., Mackay, P., Stirnberg, P., Estelle, M. and Leyser, O. (1998) Changes in auxin response from mutations in an *AUX/IAA* gene. *Science.* 279: 1371-1373.

Sato, A. and Yamamoto, K.T. (2008) Overexpression of the non-canonical *Aux/IAA* genes causes auxin-related aberrant phenotypes in *Arabidopsis*. *Physiol Plant.* 133: 397-405.

Schrader, J., Baba, K., May, S.T., Palme, K., Bennett, M., Bhalerao, R.P., et al. (2003) Polar auxin transport in the wood-forming tissues of hybrid aspen is under simultaneous control of developmental and environmental signals. *Proc Natl Acad Sci U S A.* 100: 10096-10101.

Schrader, J., Moyle, R., Bhalerao, R., Hertzberg, M., Lundeberg, J., Nilsson, P., et al. (2004) Cambial meristem dormancy in trees involves extensive remodelling of the transcriptome. *Plant J.* 40: 173-187.

Shimada, T.L., Shimada, 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.

Soler, M., Camargo, E.L.O., Carocha, V., Cassan-Wang, H., San Clemente, H., et al. (2014) The *Eucalyptus grandis R2R3-MYB* transcription factor family: evidence for woody growth related evolution and function. *New Phytol*. *in press.*

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.

Su, L., Bassa C., Audran C., Cheniclet C., Chevalier C., Bouzayen M., et al. (2014) The auxin *Sl-IAA17* transcriptional repressor controls fruit size via the regulation of endoreduplication-related cell expansion. *Plant Cell Physiol. in press.*

Sundberg, B., Uggla, C. and Tuominen, H. (2000) Cambial growth and auxin gradients. *Oxford, UK: BIOS Scientific Publishers*: 169-188.

Szemenyei, H., Hannon, M. and Long, J.A. (2008) TOPLESS mediates auxin-dependent transcriptional repression during *Arabidopsis* embryogenesis. *Science*. 319: 1384-1386.

Tan, X., Calderon-Villalobos, L.I., Sharon, M., Zheng, C., Robinson, C.V., Estelle, M., et al. (2007) Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature.* 446: 640-645.

Theologis, A., Huynh, T.V. and Davis, R.W. (1985) Rapid induction of specific mRNAs by auxin in pea epicotyl tissue. *J Mol Biol.* 183: 53-68.

Tian, Q., Nagpal, P. and Reed, J.W. (2003) Regulation of *Arabidopsis* SHY2/IAA3 protein turnover. *Plant J.* 36: 643-651.

Tian, Q. and Reed, J.W. (1999) Control of auxin-regulated root development by the *Arabidopsis thaliana SHY2/IAA3* gene. *Development* 126: 711-721.

Timell, T. (1969) The chemical composition of tension wood. *Svensk Papperstidning* 72: 173–181.

Tiwari, S.B., Hagen, G. and Guilfoyle, T.J. (2004) Aux/IAA proteins contain a potent transcriptional repression domain. *Plant cell.* 16: 533-543.

Tuominen, H., Puech, L., Fink, S. and Sundberg, B. (1997) A radial concentration

gradient of indole-3-acetic acid is related to secondary xylem development in Hybrid Aspen. *Plant Physiol.* 115: 577-585.

Uggla, C., Magel, E., Moritz, T. and Sundberg, B. (2001) Function and dynamics of auxin and carbohydrates during earlywood/latewood transition in scots pine. *Plant Physiol.* 125: 2029-2039.

Uggla, C., Mellerowicz, E.J. and Sundberg, B. (1998) Indole-3-acetic acid controls cambial growth in scots pine by positional signaling. *Plant Physiol.* 117: 113-121.

Uggla, C., Moritz, T., Sandberg, G. and Sundberg, B. (1996) Auxin as a positional signal in pattern formation in plants. *Proc Natl Acad Sci U S A.* 93: 9282-9286.

Ulmasov, T., Hagen, G. and Guilfoyle, T.J. (1999) Activation and repression of transcription by auxin-response factors. *Proc Natl Acad Sci U S A.* 96: 5844-5849.

Ulmasov, T., Murfett, J., Hagen, G. and Guilfoyle, T.J. (1997) Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. *Plant cell.* 9: 1963-1971.

Vernoux, T., Brunoud, G., Farcot, E., Morin, V., Van den Daele, H., Legrand, J., et al. (2011) The auxin signalling network translates dynamic input into robust patterning at the shoot apex. *Mol Syst Biol.* 7: 508.

Voorrips, R.E. (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. *J Hered.* 93: 77-78.

Walker, J.C. and Key, J.L. (1982) Isolation of cloned cDNAs to auxin-responsive poly(A)RNAs of elongating soybean hypocotyl. *Proc Natl Acad Sci U S A.* 79: 7185-7189.

Wang, H., Jones, B., Li, Z., Frasse, P., Delalande, C., Regad, F., et al. (2005) The tomato *Aux/IAA* transcription factor *IAA9* is involved in fruit development and leaf morphogenesis. *Plant cell.* 17: 2676-2692.

Watahiki, M.K. and Yamamoto, K.T. (1997) The *massugu1* mutation of *Arabidopsis* identified with failure of auxin-induced growth curvature of hypocotyl confers auxin insensitivity to hypocotyl and leaf. *Plant Physiol.* 115: 419-426.

Woodward, A.W. and Bartel, B. (2005) Auxin: regulation, action, and interaction. *Annals of Botany* 95: 707-735.

Yang, X., Lee, S., So, J.H., Dharmasiri, S., Dharmasiri, N., Ge, L., et al. (2004) The IAA1 protein is encoded by *AXR5* and is a substrate of SCF(TIR1). *Plant J.* 40: 772-782.

Yu, H., Soler, M., Mila, I., San Clemente, H., Savelli, B., Dunand, C., et al. (2014) Genome-Wide Characterization and Expression Profiling of the *AUXIN RESPONSE FACTOR (ARF)* Gene Family in *Eucalyptus grandis*. *PLoS One*. 9: e108906.

Chapter IV

Wood formation associated IAAs/ARFs CGs selection and functional characterization

Chapter IV: Wood formation associated *IAAs/ARFs* candidate genes selection and functional characterization

1 Introduction

The goal of the work presented in this chapter was refine the selection of the most promising candidate genes involved in the differentiation of secondary xylem in order to start their functional characterization *in planta.* Capitalizing on the genome-wide analyses of the *Aux/IAA*s and *ARFs* families in the *Eucalyptus grandis* genome and on their expression profiling in a large panel of organs, tissues and experimental conditions, we were able to select 13 genes (3 *ARFs* and 10 *IAAs* presented in chapters II and III and summarized in Table IV-1) highly and/or preferentially expressed in secondary xylem and/or cambium and for some of them differentially expressed between contrasting wood tissues. For instance, *EgrIAA31* displayed a high expression in tension wood while low expression in opposite wood, suggesting that this gene could be involved in either the regulation of active cell division occurring in tension wood and/or in the regulation of secondary cell wall formation since these contrasting woods have very different SCW composition and structure. Although *EgrARF10* showed a relatively low expression in xylem, it was selected because it was highly expressed in cambium, suggesting that it could be involved in the regulation of cambium identity and/or xylem cell specification.

In addition, we seek for *Eucalyptus Aux/IAA* and *ARF* genes in EUCATOUL database (*Eucalyptus* wood EST database: a dataset of 3928 wood-related unigenes (2479 contigs and 1449 singletons) (Paux et al. 2004; Foucart et al. 2006; Rengel et al. 2009). Among the 17 *EgrARF* and 24 *EgrIAA*, only four genes were detected including three IAA (*EgIAA3B*, *9B*, *13*) and one ARF (*EgARF6A*), and two of them (*EgIAA3B*, and 13) were included in wood associated CGs (Table IV-1).

Table IV-1 Wood formation associated CGs in *Eucalyptus*

Gene name	Ortholog in *Arabidopsis*	Expression in *Eucalyptus*		*In silico*			Bibliography
				EUCAWOOD EST	*Arabidopsis* Genevestigator*		
		xylem/ cambium	opposite/ tension		expression in xylem	TE	
EgrIAA3B	AtIAA3/4	++B		3			
EgrIAA4		+++A	3.3		++B	-2.9	
EgrIAA9A	AtIAA8/9	++B			+++B	0.9	AtIAA8/SlIAA9
EgrIAA11	AtIAA11	++B			+++A	0.2	
EgrIAA13	AtIAA12/13	+++A	3.2	2	+++A	1	AtIAA12/13
EgrIAA20	AtIAA20/30	+++A	7				AtIAA20/30/PttIAA3
EgrIAA29	AtIAA29	++A			+++A	2.5	
EgrIAA31	AtIAA31	+++B	0.3				AtIAA31
EgrARF4	AtARF4	++B			+++A	-0.5	AtARF4
EgrARF10	AtARF10	++B			+++A	-2.9	AtARF10
EgrARF19A	AtARF19	++B			++A	0.9	AtARF19

A Preferential expression

B Non-preferential expression

++ Medium expression

+++ High expression

TE Tracheary element induction

* *AtARF2, 3* and *AtIAA2, 16, 18, 19, 26, 27, 28* also showed high and/or preferential expression in xylem, but not their orthologs in *Eucalyptus*.

In order to gather information on the expression patterns of their potential orthologs in *Arabidopsis*, we performed an *in silico* gene expression profiling analysis by surveying Genevestigator database for xylem-associated member of *Aux/IAA* and *ARF* families. We seek for orthologs showing high and/or specific expression in xylem (hypocotyls of adult plants) and/or differentially expressed during *in vitro* tracheary element (TE) induction experiment. The later expression data have been obtained from an *in vitro* xylem vessel element induction system from *Arabidopsis* (Kubo et al 2005). In this system, about 50% of suspension cultured cells of *Arabidopsis* ecotype Col-0 differentiated into xylem vessel elements showing thickened SCW within 7 d in the presence of 1 μM brassinolide and 10 mM boric acid. The data are reported on Table IV-1.

Some *ARF* and *Aux/IAA* genes have been reported involved or potentially involved in vascular tissue growth and development (Table IV-2). Most of the studies were done on the model plants *Arabidopsis*. For instance *AtARF5* loss-of-function mutants *mp/arf5* displayed discontinuous and reduced vascular patterns in cotyledons (Przemeck et al. 1996). *AtIAA12* gain-of-function mutant *bdl/iaa12* mutants displayed strong phenotype similar to *mp*, showing reduced and incomplete vascular venation in the cotyledons (Hamann et al. 1999). Overexpression of the non-canonical *Aux/IAA* genes *IAA20*, *IAA30* or *IAA31*, (lacking the degron domain II, and therefore encoding more stable proteins), also caused incomplete vasculatures in seedling cotyledons (Sato and Yamamoto 2008). During the development of *Arabidopsis* inflorescence stems, the expression of four *Aux/IAAs* (*IAA13*, *IAA26*, *IAA27*, and *IAA28*) was up-regulated in the mature stem where the secondary walled cells are abundantly developed. Several *ARF* genes were differentially regulated during secondary xylem formation, among them *ARF2*, *ARF4*, and *ARF12* showed the most dramatic expression changes, suggesting putative roles in xylem formation (Ko et al. 2004).

Several studies in poplar also pointed out the involvement of these two gene families in wood formation. For instance overexpression of the mutated *PttIAA3m* lead to a reduction in cambium cell divisions and to a decrease in width of secondary xylem (Nilsson et al. 2008). In additional, some *Aux/IAA* genes in tomato, maize and cotton were also proposed to be involved in xylem development. In tomato, down-regulation of *Sl-IAA9* led to pronounced vascular venation due to more procambial cell differentiated into xylem cells (Wang et al. 2005a). In suppressed S*l-IAA15* lines, the number of xylem cells in stem also increased (Deng et al. 2012). According to the comparative phylogenic analyses described in Chapter II and III, we identified their corresponding orthologs in *Eucalyptus* (Table IV-2).

Table IV-2 Bibliography summary and comparative analysis of *ARF* and *Aux/IAA* genes potentially involved in wood formation

Gene	Expression in Eucalyptus	*Arabidopsis*		Poplar		Tomato	Maize
		Expression in Genevestigator	Bibliography	Microarray Data (Kalluri et al. 2007)	Bibliography		
EgrARF5			*ARF5* mutants (*mp*) display discontinuous and reduced leaf vascular patterns (Przemeck et al. 1996)				
EgrARF10	High expression in cambium	*AtARF10* shows high and preferential expression in hypocotyl xylem cells	*AtARF10* expressed in vascular tissue (Liu et al. 2007)				
EgrARF19A	High expression in cambium and xylem	*AtARF10* shows high and preferential expression in root xylem cells	*AtARF19* expressed in vascular tissue (Wilmoth et al. 2005)				
EgrIAA9A	Xylem and cambium preferential expression	*AtIAA8* and *9* shows high expression in hypocotyl xylem cells	*AtIAA8* was up-regulated in xylem (Oh et al. 2003). *IAA8* transcript is localized to developing vasculature and its expression precedes the appearance of secondary cell walls in TEs (Groover et al 2003) and involved in LR formation (Arase et al. 2012)	*PoptrIAA9* also showed high expression in xylem	*PttIAA2* showed high expression in cambium and xylem (Moyle et al. 2002)	Down-regulation of *SlIAA9* altered vascular venation patterning (Wang et al 2005)	*RUM1* mutant cause lignin deposition and thicker cell wall in maize root pith cells probably by decreased polar auxin transport (Zhang et al. 2014b)
EgrIAA27	High expression in stem	*AtIAA27* shows high expression in hypocotyl xylem cells	*AtIAA27* was up-regulated in mature stem (Ko, et al. 2004)	*PoptrIAA27.2* also showed high expression in xylem			

EgrIAA11	Xylem and cambium preferential expression	*AtIAA11* showed preferential expression in root xylem cells		*PoptrIAA11* showed high expression in xylem	*PttIAA5* preferentially strong expressed in secondary wall forming xylem cells (Moyle et al. 2002)	
EgrIAA13	Vascular tissue preferential expression	*AtIAA12* and A*tIAA13* display high/specific expression in hypocotyl xylem cells	*iaa12/bdl* seeding showed reduced cotyledons vasculature (Hamann et al. 1999), *AtIAA12* and *13* was up-regulated in xylem (Oh et al. 2003), and up-regulated in mature stem (Ko, et al. 2004)			
EgrIAA15A	High expression in vascular tissue					Suppressed *SlIAA15* lines increased xylem cells in tomato stem (Deng et al. 2012)
EgrIAA29	Preferential expression in xylem and cambium	*AtIAA29* showed preferential expression in root xylem cells		*PoptrIAA29* showed high expression in xylem		
EgrIAA20	High and preferential expression in vascular tissue		overexpressing these three genes (A*tIAA20, 30, 31*) result in reduced and discontinued vascular in cotyledons (Sato and Yamamoto 2008)	*PoptrIAA20.2* showed preferential expression in xylem	*PttIAA3* and *4* showed cambium and differentiating xylem preferential expression (Moyle et al. 2002). *PttIAA3m* transgenic poplar showed reduced xylem development (Nilsson et al. 2008).	
EgrIAA31	Vascular preferential expression					

In the course of this PhD, we carried on the subcellular localization of eight of the 13 candidate genes, and assessed their transcriptional activities using co-transfection assays tobacco protoplasts.

Heterologous expression in *Arabidopsis* was adopted to explore their physiological roles in plants. We made overexpression constructs of either native or mutated versions as well as dominant repression constructs. We analyzed in depth the effects of *EgrIAA4m* (chapter III), *EgrIAA9Am* and *EgrIAA20* (this chapter) in transgenic *Arabidopsis* lines. Physiological and molecular experiments demonstrated the impacts of over-expressing these genes on plant growth and development. Anatomical analyses revealed altered vascular patterning and xylem development and biochemical analysis further confirmed their impacts on plant cell wall compositions in transgenic plants.

2 Results

2.1 Subcellular localization

We examined the subcellular localisation of eight proteins including three *ARF* and five *Aux/IAA* genes. When transiently expressed as GFP fusion proteins in tobacco protoplasts, two ARFs (*Egr*ARF4 and 10) and two Aux/IAA (9A and 13) were located exclusively in the nucleus (Fig. IV-1), consistent with their predicted function as transcription factors. However, *Egr*IAA3B and 20 were localized both in the nucleus and cytoplasm, as reported for other Aux/IAA proteins (Arase et al. 2012; Audran-Delalande et al 2013). Surprisingly, both the N- and C- terminal GFP fusion proteins of *Egr*ARF19A were not localized in the nucleus. Co-transformation with a nuclear marker (simian virus 40 (SV40) large-T antigen) (David S. Goldfarb 1986) fused with mRFP confirmed that *Egr*ARF19A which harbours a weak bipartite NLS as its ortholog in *Arabidopsis*, was localized outside the nucleus (Fig. IV-2). It is worth noting that *Egr*ARF19A localization was different to that of the control expressing GFP protein alone which was localized everywhere including in the nucleus and the cytoplasm.

*Egr*ARF19A seemed to be localized in some particular subcellular organelles in cytoplasm (Fig. IV-2) and further experiments would be necessary to determine which

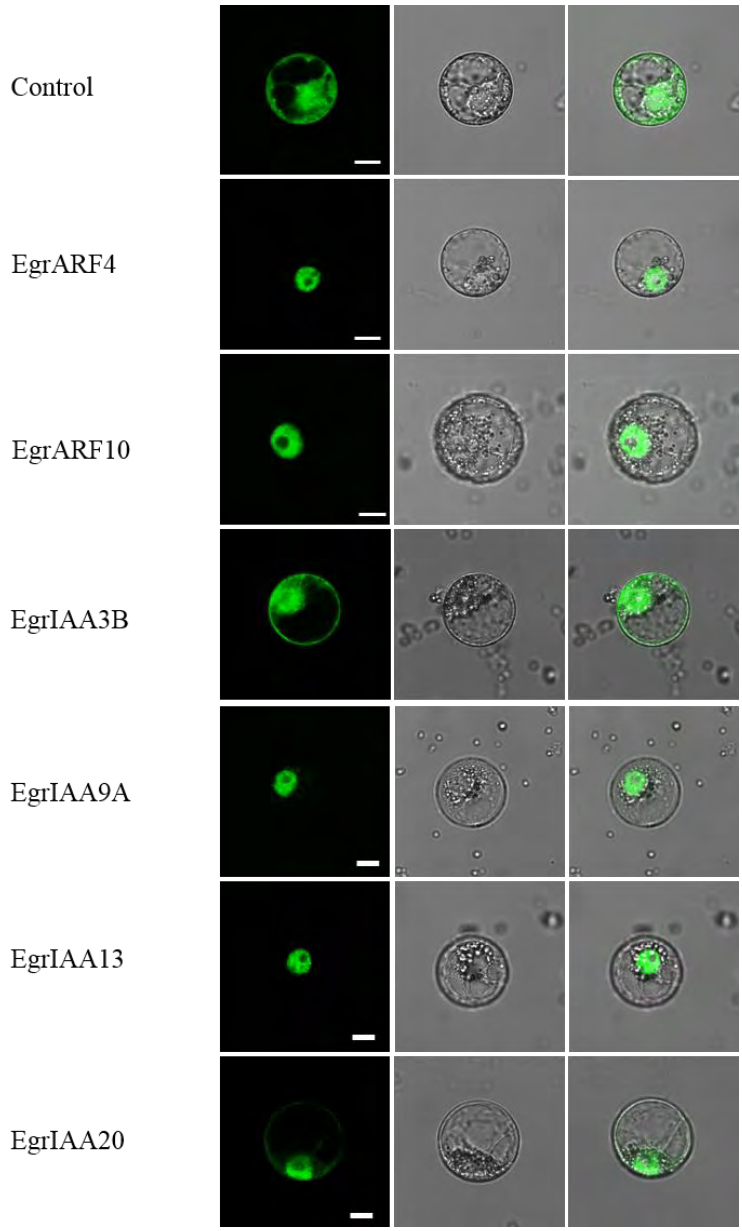


Fig. IV-1 Subcellular localization of selected *Egr*ARF and *Egr*IAA proteins. Selected *Egr*ARF and *Egr*IAA proteins were fused with green fluorescent protein (GFP) at their C-terminal and transiently expressed in BY-2 tobacco protoplasts under the control of the cauliflower mosaic virus 35S promoter. The protoplasts were then analysed by confocal laser-scanning microscopy (left panels) and by bright-field microscopy (middle panels). Merged images are shown in the right panels. Control cells expressed GFP alone. Scale bar, 10 μm.

organelle. We tested the hypothesis that *Egr*ARF19A proteins could migrate into the nucleus in response to auxin signals and monitored the fluorescence emitted by the *Egr*ARF19A-GFP, 8h, 18h and 25h after auxin treatment (50 µM 2,4-D), but no nuclear localization was observed in these experimental conditions.

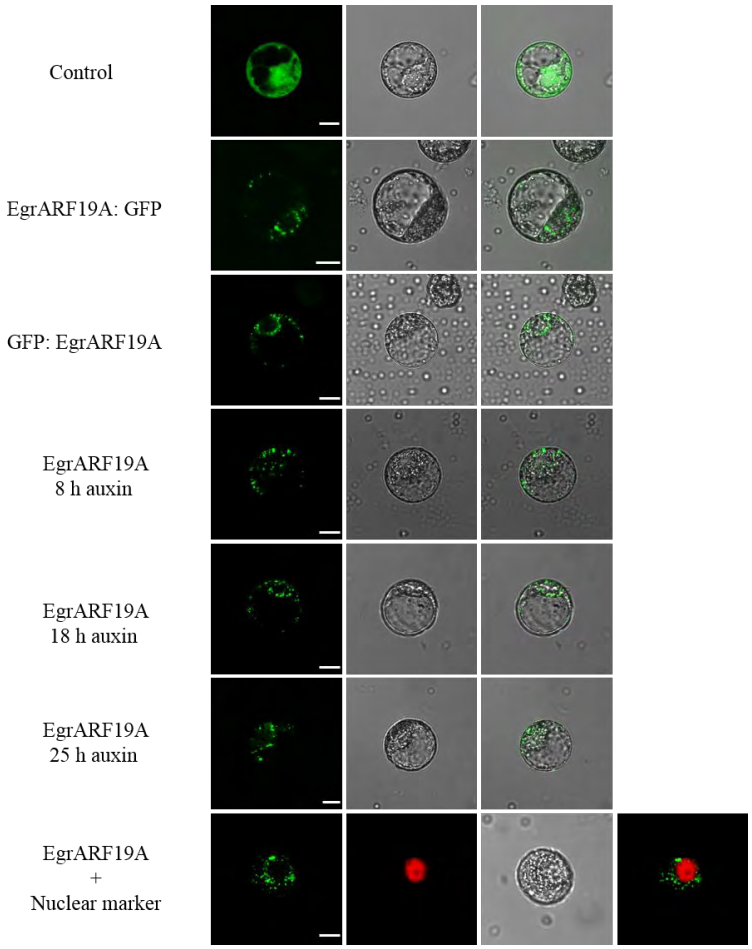


Fig. IV-2 Subcellular localisation of *Egr*ARF19A. *Egr*ARF19A protein was fused with green fluorescent protein (GFP) and transiently expressed in BY-2 tobacco protoplasts under the control of the cauliflower mosaic virus 35S promoter. Control cells expressed GFP alone. Scale bar, 10 µm. 50 µm 2, 4-D was used for auxin treatment.

2.2 Transcriptional activity

We also tested the ability of three *EgrARF* genes (*EgrARF4*, *10* and *19A*) and three *EgrIAA* genes (*EgrIAA4*, *9A* and *20*) to modulate transcription using the auxin

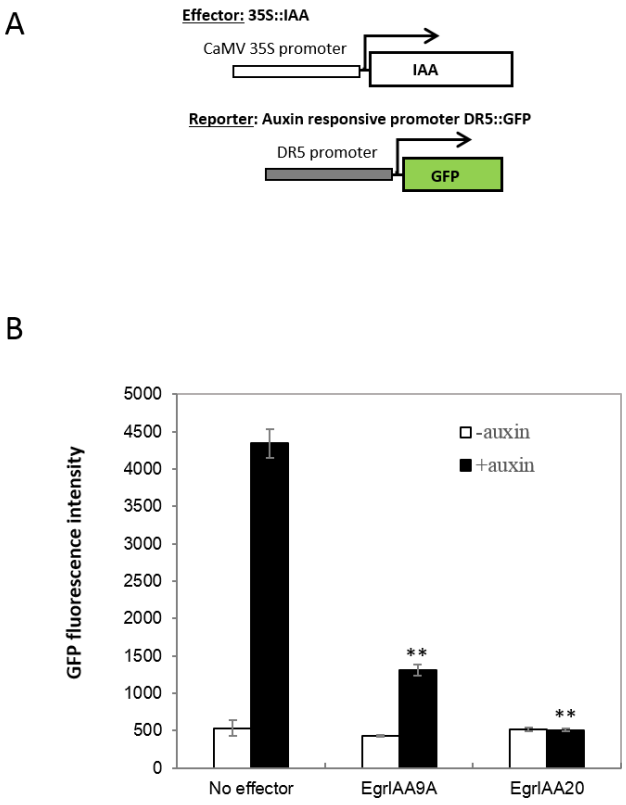


Fig. IV-3 Transcriptional activities of *Egr*IAA in tobacco protoplasts. (A) Schemes of the effector and reporter constructs used to analyse the function of *Egr*IAAs in auxin-responsive gene expression. (B) Effector and reporter constructs were co-expressed in tobacco protoplasts in the presence or absence of a synthetic auxin. GFP fluorescence was quantified 16 h after transfection by flow cytometry. A mock effector construct (empty vector) was used as a control. In each experiment transformations were performed in independent triplicates. Three independent experiments were performed and similar results were obtained; the figure indicates the data from one experiment. Error bars represent SE of mean fluorescence. Asterisks indicate values found to be significantly (student's t-test) different from the wild type. * $p<0.05$, ** $p<0.01$.

responsive reporter construct, DR5::GFP in tobacco protoplasts and assaying GFP expression by flow cytometry. Transcriptional activities of the three ARF proteins (*Egr*ARF4, 10 and 19A) has been described in chapter II (Fig. II-6), *Egr*ARF4 and 10 function as strong repressors of auxin response, while *Egr*ARF19A seems to be a weak

activator. All three tested *Egr*IAAs mediated auxin response *in vivo* and functioned as strong transcriptional repressors. *Egr*IAA4 (reported in Chapter III), *Egr*IAA9A and *Egr*IAA20 repressed auxin-induced reporter gene expression by 87%, 70% and 88% respectively as compared to the control co-transformed with an empty vector (Fig. IV-3).

2.3 Strategies for functional characterization by heterologous expression of *Eucalyptus ARF* or *IAA* transcription factors in *Arabidopsis*

We used reverse genetic approach to explore the role of some candidate genes using heterologous expression in *Arabidopsis*. Despite its herbaceous nature, *Arabidopsis* is a good model to study wood formation (Zhang et al. 2011; Chaffey et al. 2002; Nieminen et al. 2004). We describe below the different construction we made explaining the rationale behind them. These construction were made in the framework of European project but only some of them were transformed into plants as detailed in 2.4.

2.3.1 'gain-of-function' strategy for *Aux/IAA* candidate genes

In *Arabidopsis*, all of the *Aux/IAA* T-DNA insertion mutants characterized (12 out of 29) failed to show any obvious visible phenotype (Overvoorde et al. 2005). All other loss-of-function mutants using RNAi or antisense exhibited only subtle phenotypes (reviewed in Reed 2001), probably because of gene function redundancy or feedback regulatory loops that enabled the mutant plants to compensate for the absence of a particular Aux/IAA protein. Recent progress toward discovering the functions of *Aux/IAA* genes has come from the fortuitous discovery of gain-of-function mutations in several of these genes *(IAA3/SHY2*, *IAA6/SHY1*, *IAA7/AXR2*, *IAA12/BDl*, *IAA14/SLR1*, *IAA17/AXR3*, *IAA18*, *IAA19/MSG2*, and *IAA28*). All these mutations occurred in the highly conserved domain II and conferred increased stability to their corresponding proteins. Aux/IAA proteins are very short living proteins, and their degradation is essential for auxin signaling. Mutations in domain II increase protein stability. Some

non-canonical Aux/IAA proteins that lack of domain II have longer half-lives (Dreher et al. 2006).

So, we decided to create gain-of-function mutants by changing one amino acid in domain II (VGWPP to VGWPS) for *EgrIAA3B*, *4*, *9A*, *15A*, *13*, *29*, *31* and *33A*. For the non-canonical member *EgrIAA20* lacking of domain II and for *EgrIAA31* harboring a partially conserved domain II (QDWPP), we over-expressed the native forms.

We used the vector *pFAST-G02* (Shimada et al 2010) to overexpress the *EgrIAAs* under the 35SCamV promoter. Because time was limited, we generated transgenic plants corresponding to four *Aux/IAA* genes *EgrIAA4m*, *9Am*, *13m and 20*. For each gene at least two independent lines were characterized (Fig. IV-4).

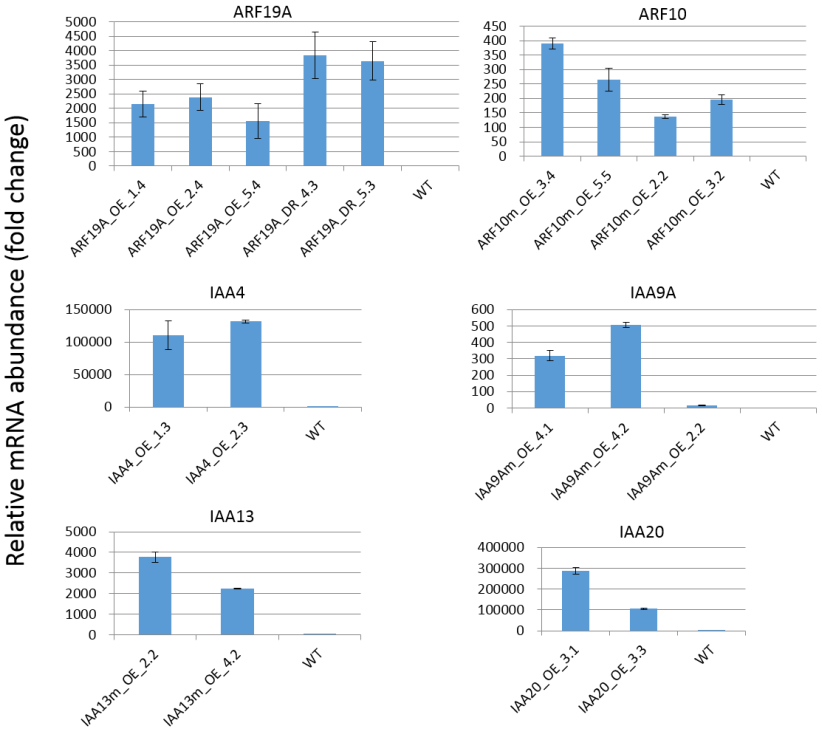


Fig. IV-4 Transcript levels of the candidate gene in transgenic *Arabidopsis* under the control of the 35SCaMV promoter and the wild type by RT-PCR. A total of three technic replicates were used for each RT-PCR experiment. Bars shows averages ratio of the fragment intensity of interest relative to that of AtUBQ10. Error bars represent SE.

2.3.2 Appropriate reverse genetic strategies for three selected *EgrARFs*

Many loss-of-function mutants of ARF (T-DNA insertion mutant) did not show any phenotypes, probably due to functional redundancy. Similarly, some transgenic plants over expressing ARF failed to give any phenotype due to the endogenous post-transcriptional regulation by miRNAs. Mutations in their miRNA target sites lead to distinct phenotypes revealing their functions in plant growth and development (Liu et al. 2007; Hendelman et al. 2012). Two of the selected *Eucalyptus* ARF genes, *EgrARF4* and *EgrARF10* were predicted to be targeted by miRNA; *EgrARF4* has two TAS3 target sites and *EgrARF10* has one miR160 target site. Therefore, we choose to introduce miRNA resistant-mutation in *EgrARF4* and *EgrARF10* to generate 'gain-of-function' transgenic plants. For the third selected gene *EgrARF19A*, no miRNA target site was detected, we chose to overexpress the native form and also to make a dominant repression construct to transform the TF in a strong repressor.

The transgenic *Arabidopsis* (overexpression, dominant repression, and miRNA resistant mutations) were generated as described in Materials and Methods. We obtained ten independent *EgrIAA19A* overexpression lines (*EgrARF19A*_OE) and twelve *EgrARF19A* dominant repression lines (*EgrARF19A*_DR), four independent miRNA resistant *EgrARF10* lines (*EgrARF10m*) (Fig. IV-4). For technical reasons and limited time, we failed the *EgrARF4* overlap PCR because it contains two miRNA target sites.

2.4 Preliminary analyses of the *IAA* and *ARF* transgenic lines

We generated transgenic *Arabidopsis* lines for four *IAA* (*EgrIAA4, 9A, 13* and *20*) and two *ARF* (*EgrARF10 and 19A*). Among all these transgenic *Arabidopsis,* only three mutants (*EgrIAA4m, 9Am* and *20*) exhibited distinct xylem related phenotypes. The functional characterization of ***EgrIAA4m* was reported in Chapter III, and that of *EgrIAA9Am* and *EgrIAA20* will be addressed in the following sections**. The two *EgrIAA13m* transgenic lines grew smaller plants as compared to the wild type without

no other obvious phenotype on plant growth and development.

Both of the *EgrARF19A* overexpression and dominant repression transgenic plants showed no obvious phenotypes. This is may be not unexpected because this TF was not localized in the nucleus and exhibited very weak transcriptional activity. Some *EgrARF10m* transgenic plants showed severely impaired leaf development and formed needle-like leaves, but seeds collect was impossible since they were sterile. These transgenic plants were not further characterized in the framework of this PhD due to time constraints.

2.5 Functional characterization of *EgrIAA9A* in transgenic *Arabidopsis*

2.5.1 Over-expression of *EgrIAA9Am* in *Arabidopsis* led to smaller, shorter plant with enhanced apical dominance

To evaluate the function of *EgrIAA9A* on plant growth and development, a mutated version in which the 229th amino acid was mutated from proline to serine (P-to-S) in the highly conserved domain II was overexpressed in transgenic *Arabidopsis* plants. Two independent transgenic lines: *EgrIAA9Am_OE_4.1* and *EgrIAA9Am_OE_4.2* were studied.

Both independent lines were smaller than wild-type. The rosette diameter measured at bolting time was reduced of 30% in the two transgenic lines as compared to controls (Fig. IV-5A and C). The number of rosette's leaves and the size of the leaves were also significantly reduced as compared to wild type (Fig. IV-5D). The differences in rosette size, in leaves number and in leaves size were maintained throughout all developmental stages.

Neither bolting nor flowering times were altered in the two transgenic lines, however the inflorescence stems of the transgenic plants grew slower and adult plants were substantially shorter than wild-type plants (Fig. IV-5B). The inflorescence stem diameters were also reduced of one third in the transgenic plants. In addition, the

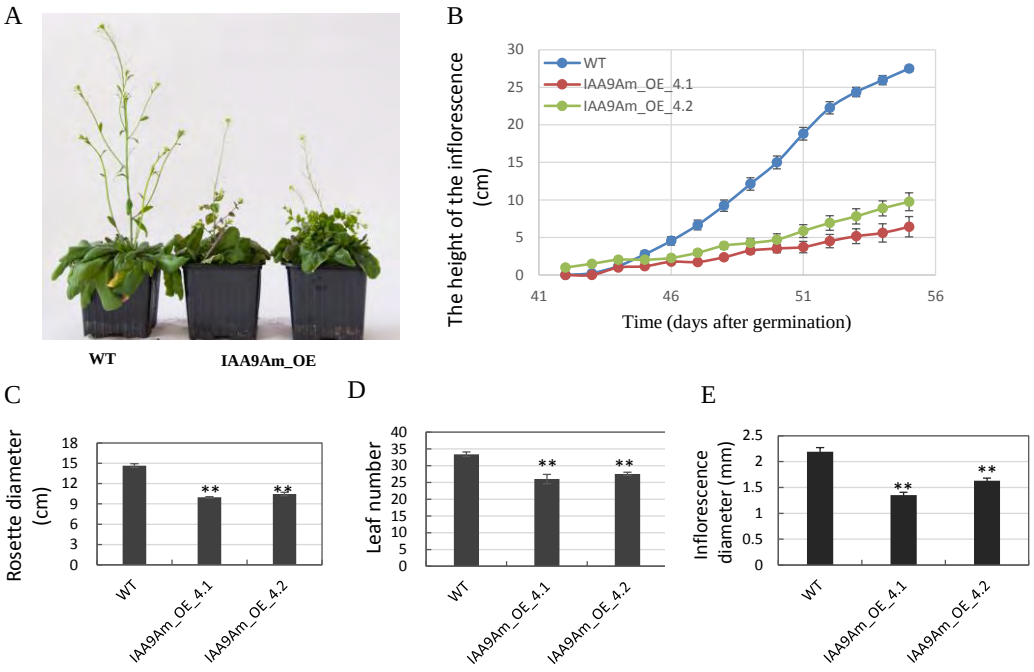


Fig. IV-5 Phenotypic characteristics of *EgrIAA9Am_OE* transgenic plants in short days. (A) 50-day-old wild type and *IAA9m_OE* transgenic *Arabidopsis*. (B) Growth curve of the *Arabidopsis* inflorescence stem. (C, D) Rosette diameter and leaf number at bolting time. (E) The diameter of the inflorescence stem base when the first silique fully development. Asterisks indicate values found to be significantly (student's t-test) different from the wild type. * $p<0.05$, ** $p<0.01$.

transgenic lines produce dramatically fewer lateral stems both at the rosette level and on the primary inflorescences (Fig. IV-5A), indicating enhanced apical dominance. In *Arabidopsis*, *IAA9* T-DNA insertion mutant exhibited no visible phenotype (Overvoorde et al, 2005) whereas in tomato the down-regulation of *IAA9* resulted in bigger plant size due to enhanced stem elongation and reduced apical dominance (Wang et al 2005a).

2.5.2 *EgrIAA9Am* transgenic plants showed altered root development and reduced auxin sensitivity

Many *Aux/IAA* members were shown to regulate root growth and development such as

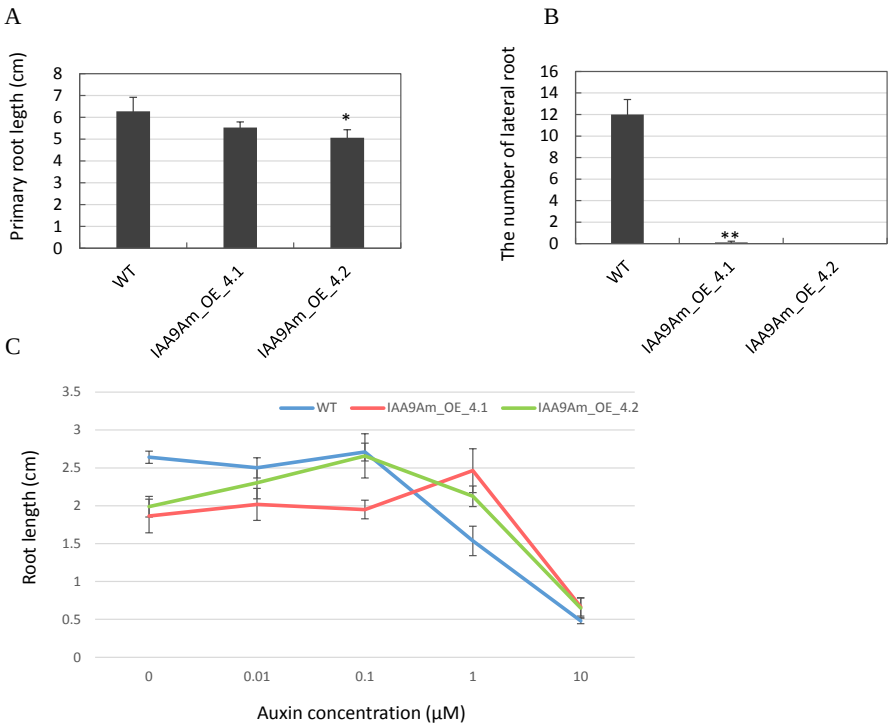


Fig. IV-6 Root phenotype of the *EgrIAA9Am_OE* transgenic *Arabidopsis*. (A, B) The length of primary root and the number of lateral root of the *Arabidopsis* seedlings at 10 days after germination (n =10). (C) The effect of 0-1.0 μM NAA on the transgenic and wild type *Arabidopsis* roots (n =10). Error bars indicate standard error of the mean. Asterisks indicate values found to be significantly (student's t-test) different from the wild type. * $p<0.05$, ** $p<0.01$.

axr5/IAA1, *axr2/IAA7*, *slr/IAA14*, *axr3/IAA17*and *msg2/IAA19* (Fukaki et al. 2002; Nagpal et al. 2000; Rouse et al. 1998; Watahiki and Yamamoto 1997; Yang et al. 2004). We therefore examined the root development of the transgenic plants growing on 1/2 MS medium. We measured the primary root length and the lateral root number of 10-day-old seedlings. Both independent *EgrIAA9Am_OE* lines exhibited shorter primary root (Fig. IV-6A). The average of primary root length was 5.5 mm in line *EgrIAA9Am_OE_4.1* and 5 mm in line *EgrIAA9Am_OE_4.2* as compared to the 6.3 mm root length of wild-type. At this stage, wild-type had well developed lateral roots (on average 12 per plants) but the two *EgrIAA9Am_OE* line grew barely no lateral roots (Fig. IV-6B). Together the data indicate that *EgrIAA9Am* overexpression prevents

lateral root formation and slightly reduce primary root elongation.

To assess auxin-responsiveness of *EgrIAA9Am_OE* plants, we examined auxin dose response on root elongation. Wild type and *EgrIAA9Am* transgenic seeds were sowed on 1/2MS medium containing 0, 0.01, 0.1, 1 and 10 μM NAA and root length was measured 5 days after germination. The wild type root elongation was significantly inhibited at 1 μM, while the *EgrIAA9Am* transgenic plants roots were inhibited by a 10-fold higher concentration (10 μM) than the wild type plants (Fig. IV-6C). This result indicates that the *EgrIAA9Am_OE* transgenic plants are less sensitive to the inhibitory effect of the high auxin concentrations on primary root elongation.

2.5.3 *EgrIAA9Am_OE* plants showed increased secondary cell wall deposition and lignification proportion in *Arabidopsis* inflorescence

We further investigated the impact of *EgrIAA9m* overexpression on xylem formation in transgenic plants. Phenotyping was performed on inflorescence stems of plants grown under short-day growth conditions when the first siliques were fully developed. Under these conditions, the basal part of the stem abundantly develops wood-like cells undergoing SCW thickening (xylem vessel cells, xylary fiber cells, and interfascicular fiber cells).

The overall organization of the vascular bundles and interfascicular fibers was not altered, but a more intense and thicker phloroglucinol staining was observed in *EgrIAA9Am_OE* transgenic lines in both regions as compared to the control (Fig. IV-7A and C), suggesting a higher lignin content and an increased SCW deposition in xylem vessels, xylary fibers and interfascicular fiber cells (Fig. IV-7B and D). A strong phloroglucinol staining was also detected in the phloem cap cells (Fig. IV-7D, green arrows) in transgenic lines suggesting a transition of phloem cap cells to phloem sclereids (highly lignified). The proportion lignified area quantified using image J software, was increased of 37% in stems of transgenic lines as compared to controls (Fig. IV-7E). It should be noted that stem diameter was reduced (around 33%) in

transgenic lines, most likely a consequence of reduced cell size for different cell types (including vessel, fibers and parenchyma cells) in transgenic lines (Fig. IV-7B and D).

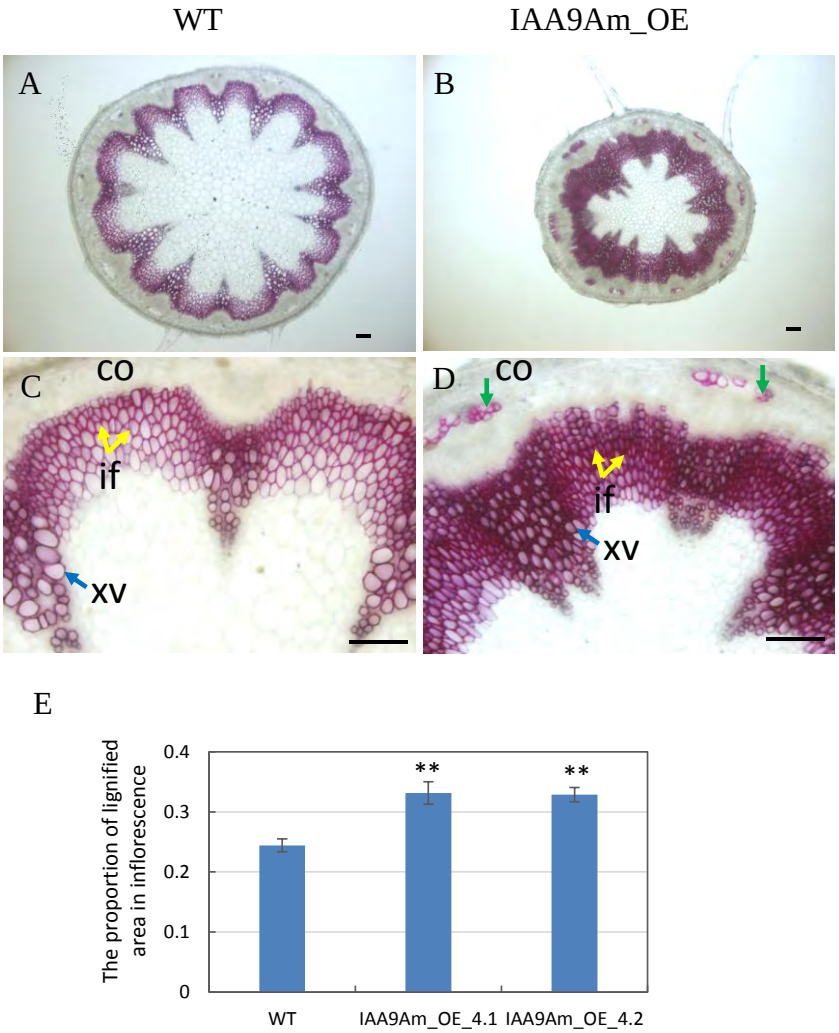


Fig. IV-7 Microscopic analysis of cross sections of *EgrIAA9Am_OE* transgenic and wild type *Arabidopsis* inflorescence base stem. Sections of wild-type plant (A, C) and *EgrIAA9Am_OE* transgenic plants (B, D) were stained with phloroglucinol-HCl. Ectopic lignification in phloem cap cells are indicated by green arrows. The proportion of lignified area of stem cross sections were displayed (E). Error bars indicate standard error of the mean. co, cortex; if, interfascicular fiber; xv, xylem vessel. Scale bar, 100 μm.

2.5.4 *EgrIAA9Am* greatly altered secondary xylem formation in *Arabidopsis* hypocotyl

The most extensive secondary xylem (wood formation) was found to develop in the *Arabidopsis* hypocotyls grown under short-day conditions (Chaffey et al. 2002). In the first phase of secondary xylem development, only vessel elements differentiate, and the rest of the xylem cells remain as parenchyma cells. This phase I is visualized as a central core in which only the vessels are stained by phloroglucinol due to their lignified SCW (Fig. IV-8A and C). During phase II, all xylem cells develop into either vessels or fibers, both have lignified thick SCW stained by phloroglucinol into red-violet. The fiber cells are smaller and show brighter phloroglucinol staining than the vessels, due to the

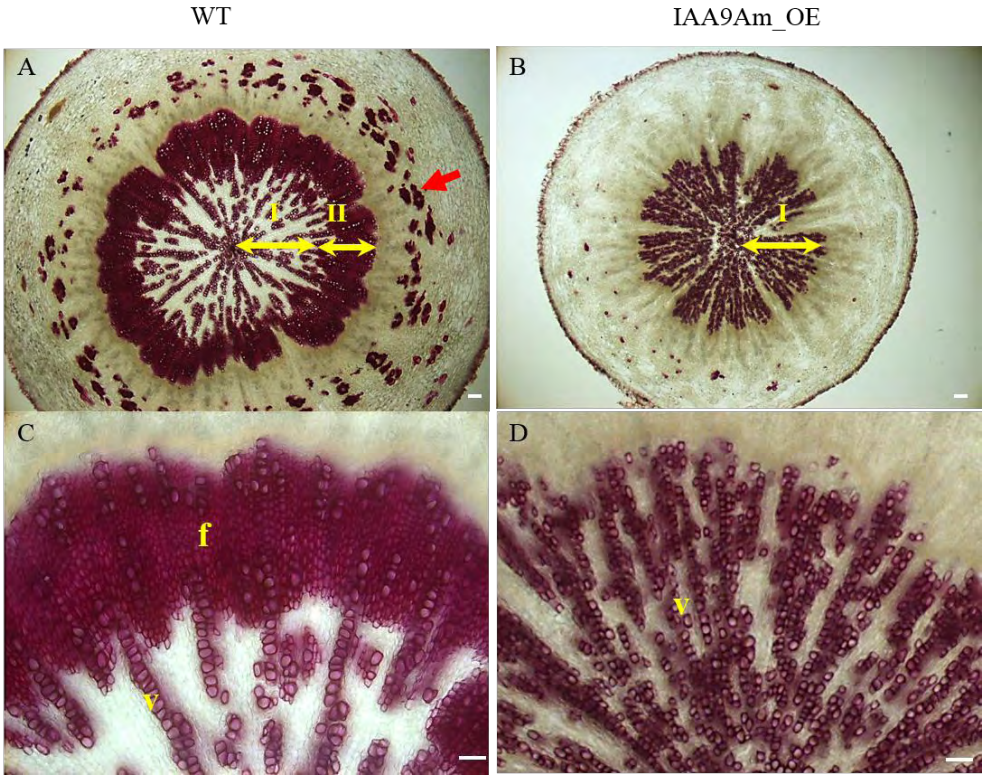


Fig. IV-8 Microscopic analysis of cross sections of *EgrIAA9Am_OE* transgenic *Arabidopsis* and wild type hypocotyl. Sections of wild-type plant (A, C) and *EgrIAA9Am_OE* transgenic plants (B, D) were stained with phloroglucinol-HCl. I, xylem I; II, xylem II; f, fiber; v, vessel. Scale bar, 100 μm.

presence of Syringyl lignin (S) whereas the lignin in vessels are made mainly of G units (Fig. IV-8A and C). It is during this later stage of secondary xylem development that the anatomy of the *Arabidopsis* hypocotyl most closely resembles the xylem of woody species (Chaffey et al. 2002).

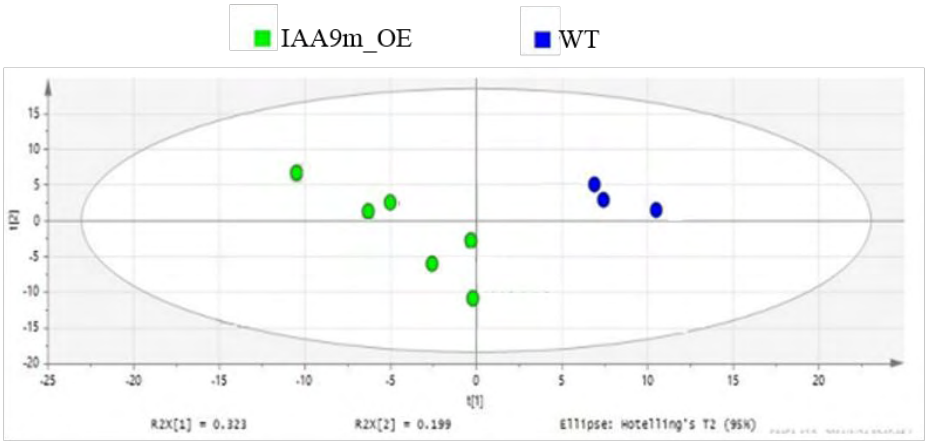
EgrIAA9m_OE lines presented strikingly altered secondary xylem formation in hypocotyls showing only phase I, in which much more vessel elements were differentiated with smaller size as compared to controls (Fig. IV-8B and D). The phase II of xylem development was completely absent in the transgenic lines at this stage (approximately two weeks after bolting). It is also worth noting that much less lignified phloem fibers were found in *EgrIAA9m_OE* lines as compared to the wild-type (Fig. IV-8A, indicated by arrow).

Unexpectedly, these observations are in sharp contrast with those made at the basal part of inflorescence stem which showed more lignified cells in *EgrIAA9Am_OE* lines. These discrepancies in xylem patterning between the two organs suggest that different regulating mechanisms may exist between stem and hypocotyl. These results underlie the different origins of xylem in the two organs and deserve more investigations.

2.5.5 *EgrIAA9Am* overexpression affects lignin composition of secondary cell walls in *Arabidopsis* hypocotyls

GC/MS pyrolysis (Py-GC/MS) analysis provides a chemical finger-print of the cell walls and now is widely used for rapid determination of lignin chemistry (Gerber et al. 2012; Meier et al. 2005). We performed Py-GC/MS analysis to evaluate and identify SCW modifications of the transgenic plants using hypocotyls as target material. The Py-GC/MS chromatograms were evaluated using Simca software by multivariate OPLS-DA method (orthogonal projections to latent structures discriminant analysis) (Worley et al. 2013), which yielded in a clear separation of the two transgenic lines from wild type (Fig. IV-9A).

A



B

Genotype	Carbohydrate	Guaiacyl	Syringyl	p-hydroxyphenol	Phenolic	Lignin	S
WT	77.5±2.0	8.5±1.2	2.9±0.7	1.7±0.2	1.1±0.1	14.2±2.0	0.34
IAA9Am_OE	78.0±1.2	**6.7±1.0***	**1.9±0.4***	1.5±0.4	**0.8±0.2****	**11.0±1.9***	**0.29**

Fig. IV-9 Chemical composition of *EgrIAA9Am_OE* transgenic *Arabidopsis* and wild type hypocotyl by Py-GC/MS analysis. (A) Orthogonal projections to latent structures discriminant analysis (OPLS-DA) scores plot of spectra acquired from *Arabidopsis* hypocotyl. (B) The chemical compounds proportion of *Egr*IAA9Am_OE transgenic *Arabidopsis* and WT (n>12). Asterisks indicate values found to be significantly (student's t-test) different from the wild type. * $p<0.05$, ** $p<0.01$.

The transgenic plants showed a significantly reduced total lignin content as compared to wild type plants. The S lignin content decreased dramatically (65%) whereas the G units content also decreased but to a lesser extent (21%) resulting in a significantly decreased S/G ratio (Fig. IV-9B). These results are in agreement with the histological observations from hypocotyl. Indeed, the virtual absence of fibers (rich is S units) as a consequence of the lack of phase II of secondary xylem development, would contribute to a decreased S/G ratio since the vessels lignin are rich in G units and depleted in S units.

2.6 Functional characterization of *EgrIAA20* in transgenic *Arabidopsis*

2.6.1 Overexpression of *EgrIAA20* in *Arabidopsis* results in smaller, shorter and bushier plants

The *EgrIAA20_OE* transgenic *Arabidopsis* displayed helical twisting and backward rolling leaves since early stages of development, visible ~20 day after germination (Fig. IV-10A). The number of rosette leaves was significantly reduced in transgenic plants, as well as the rosette diameter as a consequence of obviously smaller leaves (Fig. IV-10B and C). The differences in rosette sizes were maintained throughout all development stages. The bolting time was not altered in transgenic lines but the inflorescence stem of transgenic lines grew much slower, especially that of the strong line *EgrIAA20_OE_3.1* (Fig. IV-11A and B). The main inflorescence stem was obviously thinner and shorter as compared to wild type control (Fig. IV-11C). In the

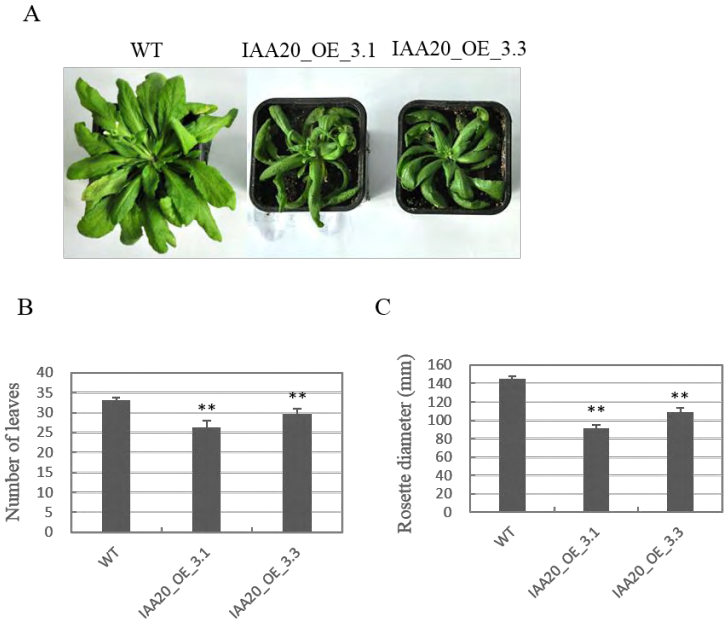


Fig. IV-10 Phenotype of 42 days old *EgrIAA20_OE* transgenic plants and wild type. The twisting leaves were displayed (A), the leaf number (B) and rosette diameter (C) were measured at bolting time (n>10). Error bars indicate standard error of the mean. Asterisks indicate values found to be significantly (student's t-test) different from the wild type. * $p<0.05$, ** $p<0.01$.

severe transgenic line *EgrIAA20_OE_3.1*, the main stem grew very slowly and a lot of lateral stems emerged from the base of the stem (rosette level) which grew rapidly becoming in many cases even longer than the main stem, indicating that the apical dominance was severely reduced in *EgrIAA20_OE* lines. While the bolting time was unaffected, the flowering time in the *EgrIAA20_OE* lines was retarded, and flower development was impaired. The flowers remained as closed buds for a long time and very often failed to reach the anthesis stage. Closer observations showed that the stamens were dramatically shorter than the wild- type (Fig. IV-11D). The fecundity of *EgrIAA20_OE* lines was seriously decreased resulting in a small number of seeds collected per plant (0-20 seeds per plant). Artificial pollination of the transgenic lines

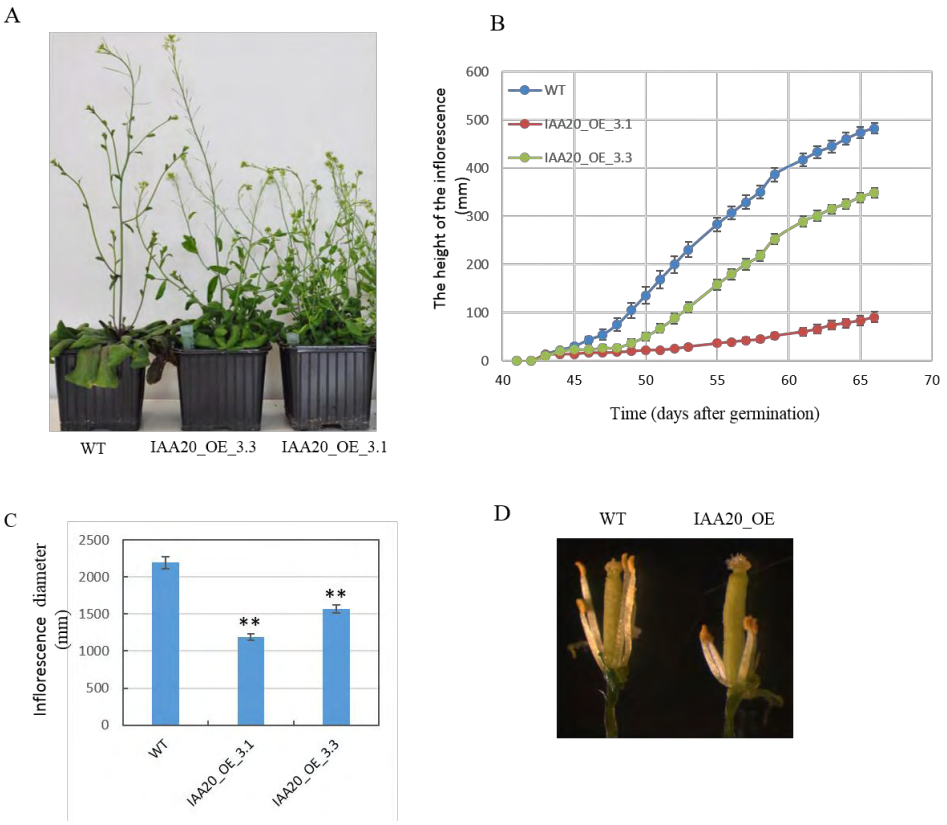


Fig. IV-11 Phenotypic characteristics of *EgrIAA20_OE* transgenic plants in short days. (A) 53 DAG WT and transgenic *Arabidopsis*. (B) Growth curve of the *Arabidopsis* inflorescence stem (n>10). (C) Stem diameter was measured when the first silique fully developed (n>6). (D) Flower bud of *EgrIAA20_OE* and WT transgenic *Arabidopsis*. Asterisks indicate values found to be significantly (student's t-test) different from the wild type. * $p<0.05$, ** $p<0.01$.

using the wild-type pollen allowed seed formation, indicating that the ovules were still functional in the transgenic lines. Thus the reduced fertility could be a consequence of an impaired stamen development and/or of the shorter stamens formed. Indeed, an uneven position of stamen relative to the pistil is known to prevent the normal pollination and fertilization.

2.6.2 Overexpression of *EgrIAA20* affected root development and root gravitropic response

We also examined the root growth phenotype in the transgenic lines. We compared the primary root length and number of lateral roots between wild type and *EgrIAA20_OE* lines 10 days after germination on 1/2 MS medium. The *EgrIAA20_OE lines* showed significantly reduced primary roots elongation as compared to the wild-type (Fig. IV-12A). Few lateral roots were formed in *EgrIAA20_OE* lines at this stage, in sharp contrast to the wild type plants which had already 12 lateral roots in average (Fig. IV-12B). This result indicate that the over expression *EgrIAA20* gene inhibit primary root elongation and lateral root formation.

Gravitropic response is a typical auxin-related phenotype, and several Aux/IAA mutants show agravitropism in *Arabidopsis* roots and/or hypocotyl, such as *axr5/IAA1*, *axr2/IAA7*, *slr/IAA14*, *axr3/IAA17*and *msg2/IAA19* (Fukaki et al. 2002; Nagpal et al. 2000; Rouse et al. 1998; Watahiki and Yamamoto 1997; Yang et al. 2004). To assess the gravitropic response of the *EgrIAA20* transgenic plant, we grew them on vertically oriented 1/2 MS medium plates. Eight days after germination we reoriented these plates by 90 d*Egr*ees to test their roots response to gravity. Fourty eight hours later, all the wild type *Arabidopsis* roots changes 90 d*Egr*ees according to the gravity changes, while all the *EgrIAA20* transgenic plants shows no changes to the gravity change (Fig. IV-

12C, D), indicating that transgenic plants roots lost the gravitropic response.

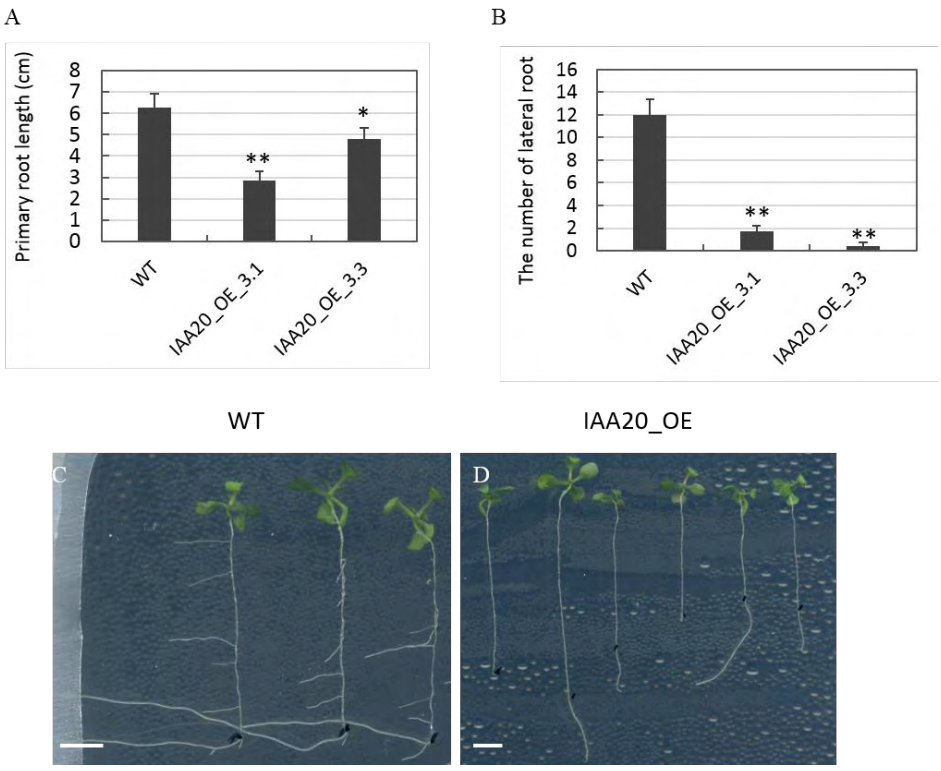
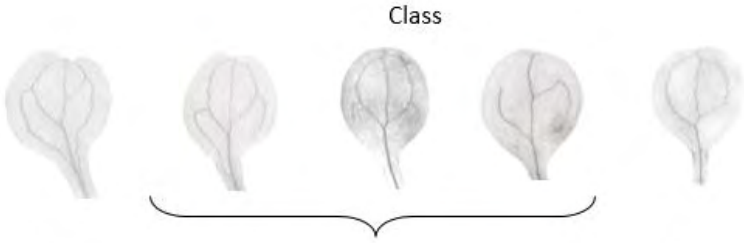


Fig. IV-12 Root phenotype of the *EgrIAA20_OE* transgenic *Arabidopsis*. (A, B) The length of primary root and the number of lateral root of the *Arabidopsis* seedlings at 10 days after germination (n =10). (C, D) Eight day-old *Arabidopsis* root response to gravity after 48 h reoriented the plate to 90 d*Egr*ee. Error bars indicate standard error of the mean. Scale bar, 0.5 cm. Asterisks indicate values found to be significantly (student's t-test) different from the wild type. * $p<0.05$, ** $p<0.01$.

2.6.3 *EgrIAA20* impaired vascular patterning in cotyledons

During embryogenesis and primary growth, the conductive tissues (xylem and phloem) are formed from the procambial tissues associated with the apical meristem. The vascular patterning in cotyledons was successfully used to screen regulators governing vascular tissues development. Overexpression of *EgrIAA20* led to severe defects in the cotyledons vascular patterning (Fig. IV-13). The vascular patterning was assessed by the number of secondary vein loops originating from the mid-vein. In wild type



Genotype	I	II	III
WT	13 (45)	16 (55)	0
IAA20_OE_3.1	0	18 (64)	10 (36)
IAA20_OE_3.3	0	4 (50)	4 (50)

Fig. IV-13 Cotyledon venation pattern in *EgrIAA20_OE* transgenic plants. Values in brackets indicate the percentage contribution of each class.

seedlings, all the cotyledon venation patterns belonged to classes I and II, and 45% of them exhibited a more complex pattern (class I with four complete loops). In contrast, cotyledon venation patterns of all the transgenic plants belonged to classes II and III, none of them showed the complete vascular patterning with four loops (class I), indicating that overexpressing *EgrIAA20* prevented the complete vascular patterning during embryogenesis.

This incomplete vascular patterning in cotyledon phenotype mimicks that of *Arabidopsis ARF5* loss-of-function mutant *monopteros/arf*5, and AtIAA20 shows interaction with AtARF5 in *Arabidopsis* (Vernoux et al 2010), suggesting that *Egr*IAA20 may interact with ARF5 to control this development process. To assess this

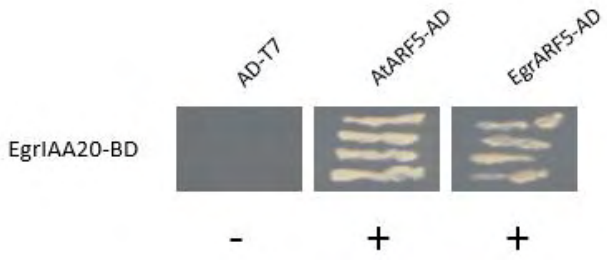


Fig. IV-14 Protein-protein interaction between *Egr*IAA20 and AtARF5, *Egr*ARF5. Yeast of co-transformed *Egr*IAA20-BD and AtARF5 or *Egr*ARF5 grew on selected medium lacking Trp, Leu, His, and Ade (THLA) and then scratched again on a TLHA plate. AD-T7 were used as negative controls.

hypothesis, we verified the protein-protein interaction between *Egr*IAA20 and AtARF5 using yeast two hybrid system and the results confirmed that there was a strong interaction between these two proteins (Fig. IV-14).

2.6.4 Overexpressing of *EgrIAA20* altered xylem formation during secondary growth in inflorescence stem and hypocotyl

In *Arabidopsis* secondary growth can occur in the inflorescence stem, hypocotyl and root. We assessed the overexpression of *EgrIAA20* in *Arabidopsis* secondary growth in the basal part of inflorescence stems and in the hypocotyls of plants grown in short day conditions.

The overall organization of vascular bundles and interfascicular fibers was not altered but a more intense phloroglucinol staining was observed in sections made at the basal part of the inflorescence stems of *EgrIAA20_OE* plants (Fig. IV-15B and D), suggesting a higher lignification level in both xylem vessels and interfascicular fibres. A higher proportion of lignified area was quantified in both transgenic lines as compared to controls (Fig. IV-15E). Moreover, the increase in the proportion of lignified area correlated well to the level of *EgrIAA20* transcript, the strong line *EgrIAA20_OE_3.1* had higher lignified area than the weak line *EgrIAA20_OE_3.3* (Fig. IV-4, Fig, IV-15E). It should be noted that the *EgrIAA20_OE* lines have thinner stem probably due to smaller cell size (Fig. IV-15B and C).

We next evaluated the secondary growth in hypocotyls. The phase I of xylem development in which only vessels differentiate and mature appeared not affected in transgenic lines (Fig. IV-16A-D); on the other hand, the phase II was altered in both transgenic lines. The onset of phase II development appeared earlier than in wild-type, leading to an increased proportion of phase II -xylem relative to phase I-xylem. Moreover, only a thin layer of fibres differentiated and matured in transgenic lines, giving rise to a narrow discontinuous ring where phase II xylem starts (Fig. IV-16B and D). This is in sharp contrast to the controls, in which both fiber and vessels were

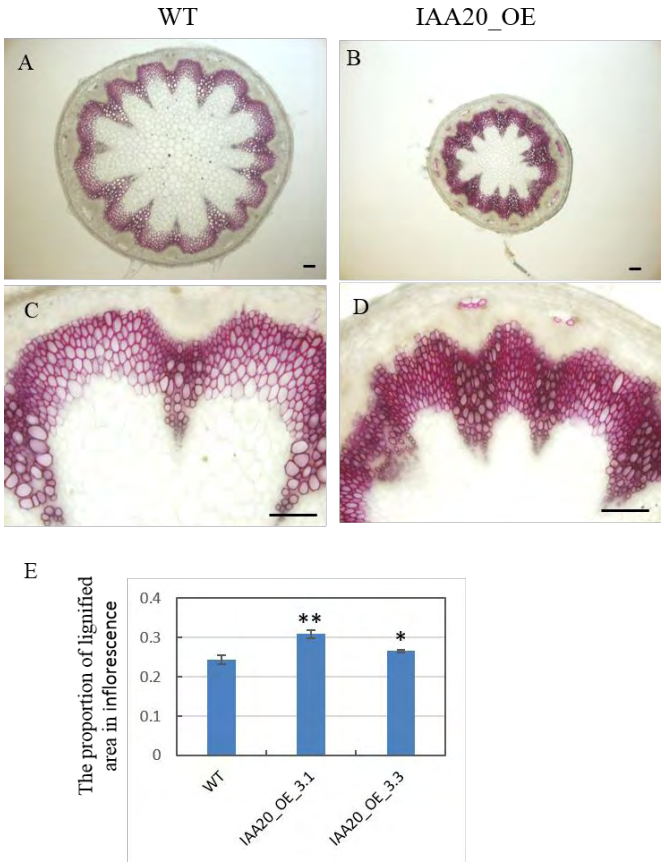


Fig. IV-15 Microscopic analysis of cross sections of *EgrIAA20_OE* transgenic and wild type *Arabidopsis* inflorescence base stem. Sections of wild-type plant (A, C) and *EgrIAA20_OE* transgenic plants (B, D) were stained with phloroglucinol-HCl. The proportion of lignified area of stem cross sections were displayed (E). Error bars indicate standard error of the mean (n>6). Asterisks indicate values found to be significantly (student's t-test) different from the wild type. * $p<0.05$, ** $p<0.01$. Scale bar, 100 μm.

simultaneously differentiated and matured forming a thick continuous ring (Fig. IV-16A and C). In addition, the transgenic lines differentiated much more vessel elements in phase II xylem as compared to the control (Fig. IV-16C and D). It should be noted that transgenic lines had a smaller hypocotyl diameter as compared to the control, indicating a globally reduced secondary growth in *EgrIAA20_OE* lines.

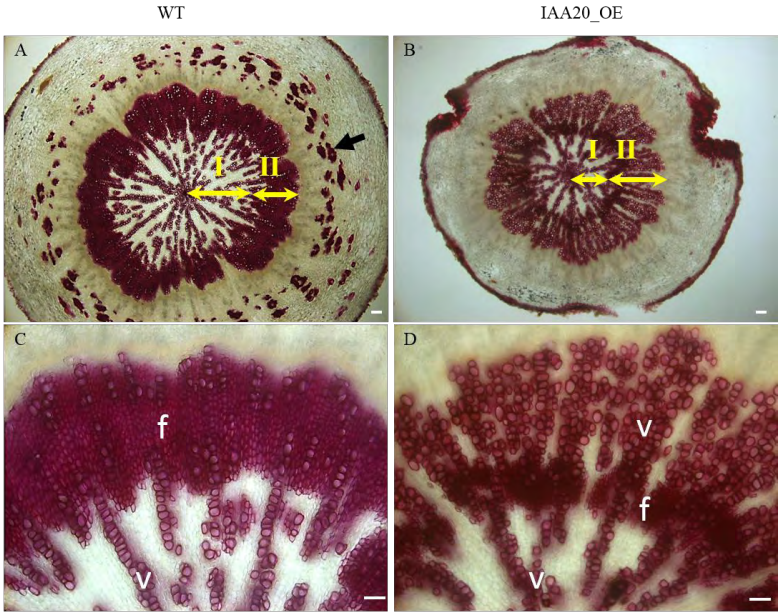


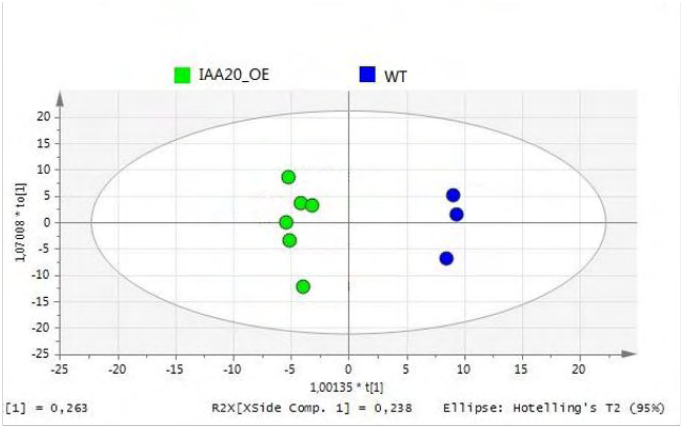
Fig. IV-16 Microscopic analysis of cross sections of *EgrIAA20_OE* transgenic *Arabidopsis* and wild type hypocotyl. Sections of wild-type plant (A, C) and *EgrIAA20_OE* transgenic plants (B, D) were stained with phloroglucinol-HCl. I, xylem I; II, xylem II; f, fiber; v, vessel. Scale bar, 100 μm.

2.6.5 *EgrIAA20* overexpression affects lignin composition of secondary cell walls in *Arabidopsis* hypocotyls

We further analyzed the chemical composition of cell walls in these *Arabidopsis* hypocotyl by PY-GC/MS. Py-GC/MS chromatograms were evaluated using Simca software by multivariate OPLS-DA method (orthogonal projections to latent structures discriminant analysis), which clearly separated the two transgenic plants from the controls (Fig. IV-17A). In particular, the transgenic plants showed less S lignin and a reduced phenolic content while H lignin was increased as compared to wild type plants. As a consequence of the reduced S lignin, the S/G was largely reduced (Fig. IV-17B). These results are in agreement with the histological results, which showed dramatically reduced fiber cells (higher S/G ratio) in transgenic *Arabidopsis* hypocotyls (Fig. IV-16B and D).

Taking together, these data showed that overexpressing *EgrIAA20* led to distinct auxin related phenotype including diminished primary growth, reduced apical dominance, reduced the secondary growth in both inflorescence stem and hypocotyl that led to narrow stems and thinner hypocotyls, and in particular an abnormal xylem development in hypocotyls.

A



B

Genotype	Carbohydrate	Guaiacyl	Syringyl	p-hydroxyphenol	Phenolic	Lignin	S/G	C/L
WT	77.5±2.0	8.5±1.2	2.9±0.7	1.7±0.2	1.1±0.1	14.2±2.0	0.34±0.04	5.55±0.82
EgrIAA20_OE	76.9±3.0	8.3±0.8	**2.0±0.4***	**1.9±0.2***	**1.0±0.1***	13.2±2.4	**0.25±0.03****	6.03±1.4

Fig. IV-17 Chemical composition of *EgrIAA20_OE* transgenic *Arabidopsis* and wild type hypocotyl by Py-GC/MS analysis. (A) Orthogonal projections to latent structures discriminant analysis (OPLS-DA) scores plot of spectra acquired from *Arabidopsis* hypocotyl. (B) The chemical compounds proportion of *EgrIAA20_OE* transgenic *Arabidopsis* and WT (n>12). Asterisks indicate values found to be significantly (student's t-test) different from the wild type. * p<0.05, ** p<0.01.

3 Discussion

The functional analyses of gain of function lines for *EgrIAA9A* and *20* revealed that these two selected candidate genes are indeed regulating wood formation.

We have shown that *Egr*IAA9A belongs to a distinct phylogenetic clade whose members are particularly long proteins (50% longer, ~300 amino acids) as compared to the majority of the other AtIAAs (~200 amino acids). It is strictly localized in the

nucleus and behaves as a strong repressor of auxin-dependent transcription. It was chosen for functional characterization because it belongs to IAA cluster III (more expressed in xylem/cambium than in phloem, Chapter III) that also contains genes involved in the biosynthesis of the main components of the SCWs (data not shown) and also because it was more expressed in juvenile than in mature xylem, two types of wood with distinct SCW properties. Its potential orthologs in *Arabidopsis* are the sister pair *AtIAA8/AtIAA9*. No gain of function mutant of *AtIAA8* nor *AtIAA9* was reported, whereas loss of function mutants (T-DNA insertion mutants *iaa8-1, iaa9-1*) and double mutant *iaa8/iaa9* did not exhibit any obvious phenotype (Overvoorde et al 2005) except the presence of more lateral roots in *iaa8-1* (Arase et al 2012). Despite the fact that no phenotype was reported at the level of the vascular system, a large body of evidence supports the hypothesis that the orthologs of *EgrIAA9 in Arabidopsis* and other species are involved in the regulation of xylem formation: *(i)* both *AtIAA8* and *AtIAA9* show high expression in hypocotyl xylem and inflorescence stem (Genevestigator database survey), *(ii) AtIAA8* and *AtIAA9* transcript levels, measured at the base of the inflorescence stem under artificial weight treatment shown to induce secondary growth, were the highest among all other *IAAs* (Ko et al, 2004) but they underwent differential expression strands suggesting different roles. *AtIAA8* transcript level reaches it maximum at a developmental stage corresponding to the beginning of the lignification of the interfascicular fibres and decreases at a more mature stage, whereas *AtIAA9* transcript level is highest at this latter stage (Ko et al 2004), *(iii)* the promoter of *AtIAA8* is active in the vascular strands (Groover et al 2003), (iv) Zinnia *zIAA8* is highly induced during *in vitro* tracheary element induction (Groover et al 2003), *(iv)* Aspen *PttIAA2* shows high expression in cambium and xylem (Moyle et al 2002), *(v)* poplar *PtrIAA9* also displayed high expression in xylem (Kalluri et al 2007) and finally *(vi)* down-regulation of the tomato *SlIAA9* results in enhanced venation in leaves and thick peduncles (Wang et al 2005). During the course of our study, it was reported that in the monocot plant maize, the gain of function mutation in *RUM1* (ortholog of *EgrIAA9A*) led to defects in root xylem organisation and to ectopic lignification and wall thickening in the pith cells (Zhang et al 2014).

Overexpressing a stabilized version of *EgrIAA9Am* in *Arabidopsis* led to auxin-related phenotypes, such as dwarfism, inhibition of lateral root emergence, reduced auxin sensitivity in root growth, enhanced apical dominance. In line with *EgrIAA9Am* expression profile, secondary xylem formation was altered in both the basal part of the inflorescence stem and hypocotyl. Surprisingly, the alterations of secondary xylem formation exhibit opposite patterns between the inflorescence stems and the hypocotyls: secondary xylem formation was globally enhanced in the inflorescence stems of transgenic lines showing an increased lignification staining both in the xylem bundles and the interfascicular fibres, whereas in the hypocotyls, secondary xylem (fibres and vessels) normally produced in the phase II was completely absent in transgenic lines although in phase I xylem, more lignified vessels were produced in comparison to controls. To the best of our knowledge such opposite vascular patterns between the inflorescence stems and the hypocotyls has not been reported before, but histochemical analysis of both organs is not systematically done and such distinct patterns may have escaped to previous observations. These puzzling results suggest that different auxin-regulated differentiation mechanisms may occur between inflorescence stems and hypocotyls, the development of the latter resembling more to root ontogeny than to stem ontogeny. Moreover, it also should be kept in mind that *EgrIAA9Am* was expressed in *Arabidopsis* which contains two potential orthologs *AtIAA8* and *9* having different behaviours during secondary xylem formation in inflorescence stems artificially treated to induce secondary growth (Ko et al 2004) and thus likely playing different roles during this process. *EgrIAA9m* might interfere with both endogenous genes leading to the complexes and antagonist vascular phenotypes, the most striking and dramatic being the inhibition of secondary growth and xylem formation (Phase II) in hypocotyls. These results are very promising although further investigations are needed to better understand the role of *EgrIAA9A* and of its *Arabidopsis* orthologs. To this end, the respective expression levels of *Arabidopsis* endogenous *AtIAA8* and *9* will need to be assessed as well as other auxin-related genes in *EgrIAA9Am* gain of function lines both in hypocotyls and basal inflorescence stems and in parallel, auxin levels should be evaluated at different development stages of xylem differentiation. In line with this, it

is worth noting that co-expression analysis showed that *AtIAA9* is tightly co-expressed with *WAT1* gene encoding a tonoplast localized auxin efflux facilitator. Mutations in the fibre deficient mutant *wat1* cause a reduction in polar auxin transport and misregulation of genes associated with auxin response of which *AtIAA9* was found to be down-regulated (Ranocha et al. 2010, 2013).

EgrIAA20 encode a non-canonic IAA protein, as its orthologs in *Arabidopsis*, the sister pair AtIAA20/AtIAA30, it lacks the domain II (d*Egr*on), which interacts with the SCF^{TIR1} complex and lead to the rapid degradation of Aux/IAA proteins by 26S proteasome. In *Arabidopsis* mutation in domain II cause higher steady-state levels of Aux/IAA proteins and the members who lack domain II (such AtIAA20 and AtIAA30) are more stable proteins. Thus, *Egr*IAA20 was supposed to be more stable as compared to the canonical *Egr*IAAs. We also showed that *Egr*IAA20 was not exclusively located in the nucleus suggesting the possibility of a regulatory mechanism allowing its migration to the nucleus to strongly repress the transcription of auxin-responsive genes. In sharp contrast with *AtIAA20* that shows very low basic expression in all tissues examined (Genevestigator database survey), *EgrIAA20* expression was found preferentially expressed in vascular tissues, and particular highly expressed in vascular cambium in *Eucalyptus*. No report described the loss of function mutant for *AtIAA20* or *AtIAA30* (no T-DNA insertion mutation for AtIAA20 was available and the *AtIAA30* T-DNA insertion mutant *iaa30-1* was not yet characterized (Overvoorde et al 2005). Gain of function mutations by overexpressing *AtIAA20* and *AtIAA30* in *Arabidopsis* resulted in semi-dwarf plants with agravitropic and impaired root growth. The cotyledon venation was also impressively reduced (Sato and Kotaro 2008), mimicking the phenotype of loss of function mutant for *ARF5* (*mp/arf5*) and we also showed that the *Egr*IAA20 protein was able to interact with ARF5. Overexpressing *EgrIAA20* resulted in auxin related growth phenotypes such as reduced apical dominance, inhibition of lateral root formation and agravitropic root response. The inflorescence stems and hypocotyls were thinner and vascular patterning in cotyledons was impaired. Histological analysis of hypocotyls sections showed altered secondary xylem formation

formed during phase II, harboring virtually no fibers but more vessels than controls. Biochemical analysis by Py-GC/MS further confirmed the significantly lower S lignin content and the reduced S/G ratio due to fewer fibers rich in S units. These data strongly suggest that *EgrIAA20* regulates fibre development during secondary xylem differentiation. It is worth noting that although the poplar *PtrIAA20.1*, *PtrIAA20.2* and aspen *PttIAA3* are phylogenetically close to *EgrIAA20*, they contain the domain II in contrast to *Egr*IAA20 and its *Arabidopsis* orthologs, suggesting that they are not stable, and may have different roles *in planta*. Overexpressing a stabilized version of *PttIAA3 (OE PttIAA3m)* in poplar resulted in reduced secondary growth eg. decreased xylem radial growth and reduced xylem/phloem ratio.

Although preliminary, the results obtained are new and promising suggesting that IAA9A is regulating secondary xylem formation in hypocotyl whereas IAA20 is more specifically involved in the regulation of fibres. Further investigations will be necessary to get further insights into the roles and mechanisms of action of these two genes during secondary xylem formation, and will be presented in the general discussion.

4 Materials and Methods

4.1 Plant materials and culture condition

The *Arabidopsis* seeds used for *in vitro* culture were surface-sterilized for 1 min in 70% ethanol, 10 min in 25% bleach, rinsed five times in sterile water and plated on 1/2 Murashige and Skoog (MS) medium containing 1.0% sucrose solidified with 1% agar. The culture conditions were: 16 h day/8 h night cycle, 22/20℃ day/night temperature, 40% relative humidity. For auxin dose-response experiments, sterilized seeds were germinated on 1/2 MS medium containing 0, 0.01, 0.1, 1, 10 μM NAA respectively. The primary root length was measured five days after germination. For gravitropic response, sterilized seeds were planted on vertically oriented 1/2 MS plates. Eight days after germination, we reoriented the plate by 90 d*Egr*ees for another 48 h to check their roots response to gravity.

For pyrolysis and histochemical analysis, the *Arabidopsis* were planted in short day conditions: 9 h day/15 h night cycle, 22/20℃ day/night temperature, 70% relative humidity, 200 μmol photons $m^{-1}s^{-1}$ light intensity (intense luminosity). The plants were watered every two days and fertilized weekly.

4.2 Gene expression analysis

All the *Eucalyptus ARFs* and *Aux/IAAs* gene expression in different tissues/organs were performed in chapter II and III. In addition, *Eucalyptus* wood EST database EUCATOUL (http://www.polebio.lrsv.ups-tlse.fr/eucatoul/db.php) was used for checking their EST support. All the *Arabidopsis ARF* and *Aux/IAA* gene expression profiles and their expression changes during the tracheary element (TE) induction experiment were analyzed by surveying Genevestigator database (https://www.genevestigator.com/gv/plant.jsp).

4.3 Transient expression of CGs in protoplast system for subcellular localization and transcription activation analysis

For the subcellular localization, the CDS sequences of the *EgrARF4*, *10*, *19A* and *EgrIAA3B*, *4*, *9A*, *13*, *20* were first cloned into the Gateway entry vector pENTR-TOPO (Fig. SIV-1), then recombined them into pK7FWG2 vector fused with GFP at C-terminal and expressed under the control of the 35S CaMV promoter (Fig. SIV-2A). *EgrARF19A* was also recombined into pK7WGF2 vector fused with GFP at N-terminal (Fig. SIV-2B). The empty vector pK7WGF2.0 was used as control. The constructs were introduced into tobacco BY-2 protoplasts by polyethylene glycol-mediated transfection as described in Materials and Methods in chapter III. After 16 h incubation, the transfected protoplasts were examined for green florescence signals using a Leica TCS SP2 laser scanning confocal microscope. Images were obtained with a 40x 1.25 numerical aperture water-immersion objective.

For testing the ability of *ARF* transcription factors to up or down regulate the expression

of auxin responsive promoter DR5, the full-length cDNAs of the *EgrARF4*, *10*, *19A* and *EgrIAA9A*, *13*, *20* transcription factors were cloned in pGreen vector under CaMV35S promoter to create the effector constructs(Fig. SIV-3). The reporter constructs use a synthetic auxin-responsive promoter DR5 followed by GFP reporter gene. Tobacco BY-2 protoplasts were co-transfected with the reporter and effector constructs as described in (Audran-Delalande et al. 2012). After 16 h incubation, GFP expression was quantified by flow cytometry (LSR-*Fortessa*, BD Bioscience). Data were analysed using Facsdiva software. Transfection assays were performed in three independent replicates and 400-1000 protoplasts were gated for each sample. GFP fluorescence corresponds to the average fluorescence intensity of the protoplasts population after subtraction of auto-fluorescence determined with non-transformed protoplasts. 50 μM 2, 4-D was used for auxin treatment.

4.4 Genetic transformation in *Arabidopsis* (vectors, primers, OE, OE mutated protein, dominant repression)

Total RNA was isolated from *Eucalyptus* xylem using CTAB method as described by Southerton *et al*. (1998). The cDNA first strand was synthesized from 2 μg total RNA using SuperScript III reverse transcriptase (Invitrogen, USA), according to the manufacturer's instructions. We designed primer pairs of the corresponding gene coding sequences using Primer3 v.04.0 (http://frodo.wi.mit.edu/), and the primers sequences are shown in Table SIV-1. To get gain-of-function transgenic plants, we did site target mutations in the miRNA target sites of *EgrARF4* and *EgrARF10*, as well as the domain II for *EgrIAA3B*, *4*, *9A*, *13* and *29*. The mutation sites were introduced by using overlap PCR. The CDS sequence was cloned into the Gateway entry vector pENTR-TOPO, and then recombined into the overexpression vector pFAST-G02 (Fig. SIV-4) (Shimada et al. 2010) and/or dominant repression vector pH35GEAR (Fig. SIV-5) (Kubo et al. 2005), respectively. All binary expression constructs were transferred to *Agrobacterium tumefaciens* strain GV3101 and the flower dip method was used to transform *Arabidopsis thaliana* Col-0 (Clough and Bent 1998). For overexpression

construction, the vector pFAST-G02 contains a GFP marker fused with a seed-specific promoter OLE1, and this allowed us to select the transgenic seeds using the binocular microscope under UV light (Shimada et al. 2010). For the dominant repression construction, we used vector PH35GEAR which contains a repression domain (**L**D**L**D**L**E**L**R**L**G) after the target protein and hygromycin marker for transgenic plants selection. PCR analysis of corresponding targeted genes were applied to confirm transgenic plants for each line.

Underline sequences (CACC) represent the extra sequence needed for the pENTR/D-TOPO cloning; sequences in red represent the overlap region for overlap PCR; sequences highlight in yellow represent the mutation base.

4.5 Histochemical analysis

The histological comparative analysis of SCW between wild type and transgenic plants was done at the stage of the first green siliques newly fully developed. At this stage, the basal part of the inflorescence stem and hypocotyl abundantly develops cells undergoing secondary wall thickening (xylem vessel cells, fascicular, and interfascicular fiber cells). The *Arabidopsis* inflorescence stems at the basal end (~1 cm) and hypocotyls were harvested and stored in 70% ethanol. The cross sections were prepared using vibratome Leica VT1000 S. Lignin polymers are the characteristic components of SCW and are normally absent from primary cell wall, therefore we used lignin deposition detection techniques to screen for SCW phenotype. Cross sections of inflorescence stem and hypocotyl were (~80 µm) were stained with phloroglucinol-HCl which stains specifically lignin polymer precursors coniferaldehyde and p-coumaraldehyde in the SCW giving violet-red color. Phloroglucinol-HCl was directly applied on the slide and images were recorded with a CCD camera (Photonic Science, http://www.photonic-science.co.uk).

4.6 Pyrolysis analysis

EgrIAA9A and *20* overexpression transgenic (two different lines for each) and wild type

Arabidopsis were planted 3 different batches in short days (5 plants for each line in one batch), and 10 days different between each batch. The hypocotyls were harvested in liquid nitrogen when the first silique was fully development. Then the hypocotyls were freeze-dried and ball-milled (MM400; Retsch) at 30 Hz in stainless steel jars (1.5 mL) for 2 min with one ball (diameter of 7 mm). A total of 50 – 70 μg (XP6, Mettler-Toledo, Switzerland) powder was transferred to auto sampler containers (Eco-cup SF, Frontier Laboratories, Japan) for the Py-GC/MS. The sample was carried to oven pyrolyzer by an auto sampler (PY-2020iD and AS-1020E, FrontierLabs, Japan) and analyzed by a GC/MS system (Agilent, 7890A/5975C, Agilent Technologies AB, Sweden). The pyrolysis oven was set to 450 ℃, the interface to 340 ℃ and the injector to 320 ℃. The pyrolysate was separated on a capillary column with a length of 30 m, diameter of 250 mm, and film thickness of 25 mm (JandW DB-5; Agilent Technologies Sweden). The gas chromatography oven temperature program started at 40 ℃, followed by an temperature ramp of 32 ℃/min to 100 ℃, 6 ℃/min to 118.75 ℃, 15 ℃/min to 250 ℃, and 32 ℃/min to 320 ℃. Total run time was 19 min and full-scan spectra were recorded in the range of 35 to 250 mass-to-charge ratio. Data processing, including peak detection, int*Egr*ation, normalization, and identification, was done as described by Gerber et al. (2012).

The relative amounts of S-, G-, and H-lignin and the carbohydrates were further expressed as the percentage of the total compounds amounts. Orthogonal projections to latent structures discriminant analysis (OPLS-DA) analysis of each individual replicate was performed using SIMCA-P+ (12.0).

4.7 Protein-protein interaction analysis by yeast two hybrid system

For Y2H experiments, the cDNA sequence of *EgrIAA20* was cloned into pBD vector (Fig. SIV-5A), the cDNA sequence of *EgrARF5* and *AtARF5* were cloned into pAD (Fig. SIV-5B) vectors, respectively. pBD Y2H vector (BD-bait) was used for the bait-protein construction. It contains a functional copy of the *Trp1* gene, thus restoring in AH109 tryptophan autotrophy. pAD Y2H vector (AD-prey) was used for the prey-

protein construction. It contains a functional copy of the *Leu2* gene, thus restoring in AH109 leucine autotrophy. Then, the BD-bait and AD-prey were co-transformed into yeast AH109, co-transformants being selected on culture lacking both tryptophane and leucine. If there is no interaction between the bait and prey protein, yeast will grow only on the selection medium lacking Trp and Leu. If there was an interaction, yeast can grow on the selection medium lacking Trp, Leu, His and Ade. Manipulation and analysis of the Y2H assay followed the manufacturer's instructions (Clontech Yeast Protocols Handbook).

General discussion and perspectives

General discussion and perspectives

In woody plants, auxin has been proposed to play a key role in the development of secondary xylem cells, a differentiation process involving cell division, expansion, secondary cell wall formation and cell death (Miyashima et al. 2013; Sundberg et al. 2000). In comparison to the knowledge accumulated on the role of auxin in the control of early vascular patterning in which some important actors and mechanisms involved either in the maintenance of pro(cambial) identity or xylem specification have been identified (reviewed in Berleth et al 2000; Elo et al 2009; Reinhardt 2003), our knowledge of the auxin actors underlying secondary xylem differentiation is very limited. Due to the growing importance of wood as a renewable and environmentally cost effective alternative to fossil fuels, and the economic and ecological importance of trees, a better understanding of the auxin regulation of wood formation is timely.

Auxin is believed to exert its function in wood formation through its perception/signaling pathway, of which Auxin Response Factors (ARFs) and Aux/IAAs are two well-known mediators of auxin responsive gene expression (Guilfoyle and Hagen 2007; Tiwari et al. 2003). These two families have been studied in several annual plants and more deeply in *Arabidopsis*, but remain largely under investigated in tree species.

Taking advantage of the recent release of the *E. grandis* genome (Myburg et al, 2014), we have performed a genome-wide comprehensive survey of two major transcription families involved in the auxin perception and signaling pathway. One striking feature of the *E. grandis* is the very high proportion of genes found in tandem repeats (34% of the total number of genes). Indeed it is the highest rate of tandem duplication reported among sequenced plant genomes (Myburg et al, 2014). In *Eucalyptus*, tandem duplication shaped the functional diversity of many gene families like for instance the MYB transcription factor family (Soler et al, 2014*)*. However tandem duplication is distributed unevenly between gene families and indeed, the *ARF* and *IAA* families evolved in a very different way. Both families were slightly contracted as compared to

most angiosperms and monocots families studied hitherto, including the model plant *Arabidopsis*. The functional diversity of the *ARF* family in *Eucalyptus* was mainly shaped by alternative splicing and gene truncation leading to a diversity of domain architecture, thereby compensating for the lack of extensive gene duplication mechanisms found in other species. In contrast to these species, it is characterized by the absence of whole genome, segmental and/or tandem duplication events. Indeed, whole genome duplication in *Eucalyptus* occurred 109.9 Mya ago, considerably earlier than those detected in other rosids and 95% of the paralogs were lost. The *IAA* family has a different evolutionary history from that of the *ARF*, and the number of alternative transcripts was similar to that of *Arabidopsis*.

Noteworthy, comparative phylogenetic studies pointed out the presence of a new clade absent in *Arabidopsis* and maintained preferentially in woody and perennial plants. It is likely that the genes belonging to this clade have been lost in herbaceous annual species, and more work will be needed to decipher the role of these genes. Although in our qRT-PCR experiments, *EgrARF24* was not found to be expressed in xylem, unlike its ortholog in poplar, it was found later to be expressed in this tissue in RNA-Seq data (unpublished). This apparent discrepancies may be due to the presence of an alternative transcript expressed in xylem and not detected by the primers chosen for the qRT-PCR, this will need further verification.

We used large–scale expression profiling to identify *ARF* and *Aux/IAA* genes highly and/or preferentially expressed in vascular cambium and/or xylem to identify those potentially involved in the auxin-regulated transcriptional programs underlying wood formation. Among the 13 genes exhibiting such expression patterns, we generated gain of function transgenic *Arabidopsis* lines for six of them (four IAA (*EgrIAA4, 9A, 13 and 20*) and two ARF (*EgrARF10 and 19A*)). We functionally characterized lines overexpressing *EgrIAA4m, 9Am* and *20.* All three mutants exhibited distinct alterations of the lignified vascular tissues:

The most dramatic macroscopic phenotype was obtained by overexpressing a stabilized

version of *EgrIAA4*, the plants were dwarf with strongly reduced fertility and exhibited strong auxin insensitive phenotypes such as inhibition of primary root elongation and lateral root emergence and roots agravitropism. *EgrIAA4* is the potential ortholog of the *Arabidopsis* sister pair *AtIAA3/AtIAA4.* Loss of function mutants (T-DNA insertion *mutant iaa4-1)* exhibited no phenotype *(*Overvoorde *et al 2005)* whereas *AtIAA3* gain of function mutant (EMS mutation in the d*Egr*on) exhibited a similar macroscopic phenotype to the one induced by *EgrIAA4m* especially a reduced stature and short hypocotyls (Reed et al., 1998; Tian and Reed, 1999). However, no phenotype was reported at the level of the vascular system, probably because the authors did not look specifically for it. Unfortunately, the hypocotyls were so tiny that we did not succeed to obtain hypocotyl sections and the material was insufficient for chemical analyses. So, we made sections to observe lignified vascular tissues patterning at two developmental stages at the base of the inflorescence which is known to undergo secondary growth (Altamura et al, 2001). Overexpression of *EgrIAA4m* strongly and negatively affected the interfascicular fibres and the xylary fibres development, as shown by the virtual absence of lignified secondary cell wall formation.

Similar phenotypes have been reported but for loss of function mutants suggesting antagonist roles: (i) for the NAC master regulators of the SCW known to switch on the expression of secondary wall genes in xylem fibres, (ii) for REV a class III homeodomain leucine zipper transcription factor involved in organ polarity (ii) for WAT1, which encodes for a tonoplast localized auxin efflux facilitator. In a *nst1, nst3/snd1* double mutant, secondary cell wall deposition in interfascicular fibres is completely abolished (Mitsuda et al 2007; Mitsuda and Ohme-Takagi 2008). In Rev loss-of-function mutants the development of interfascicular fibres is impaired (Zhong and Ye 1999; Prigge et al 2005). REV is likely involved in cell fate determination rather than directly promoting secondary cell wall deposition. As observed for *rev* mutants, mutations in *wat1* cause a similar reduction in polar auxin transport and misregulation of genes associated with auxin response (Ranocha et al 2010, 2013). Secondary cell wall deposition in xylary and interfascicular fibers is strongly reduced in *wat1* mutants.

However, as there is residual lignification of fibres indicating fibre cell identity seems not to be affected by wat1 mutation and that WAT1 acts more specifically on late fibre differentiation.

Whether *EgrIAA4* is involved more in interfascicular and xylary fibres specification or in later differentiation and SCW deposition stages is yet to be solved. However, this gene is differentially expressed between contrasting wood samples (tension *versus* opposite wood, mature *versus* juvenile wood) whose secondary cell walls have very different structure and composition. It is for instance, the only *Eucalyptus IAA* to be more expressed in mature wood than in juvenile wood. These results together with the residual lignification (although very light) of the SCW of the interfascicular fibres at later stages of development in the *EgrIAA4m* lines, indirectly support a role in fibre differentiation although more investigations are needed to decipher the mechanisms of action of this gene.

Similar questions on the exact roles and mechanisms of action are still to be solved for *EgrIAA9A* and *EgrIAA20* which seem to regulate the whole secondary xylem formation and secondary xylem fiber differentiation in hypocotyls, respectively.

Given the phenotypes of the three mutants, it would be important to assess endogenous auxin distribution, this can be done by crossing the mutants lines to transgenic *Arabidopsis* lines expressing the auxin induced marker DR5-Venus (Heisler et al 2005) and visualization of its fluorescence at different stages of development of the vascular system.

It would be also useful to assess the transcripts levels of genes involved in auxin transport/signalling, those known to affect vascular system patterning as well of transcription factors regulating fibre and vessels differentiation such the NACs, and SCW biosynthetic genes. A systematic search for the transcripts levels of the three *IAA* genes in the transcriptome of mutants such as Wat1, Rev, NAC… when available as well as a without a priori search for their partners using an *Eucalyptus* two yeast hybrid

library available in the lab would help getting more insights into the regulation of the three *IAAs*.

The use of an inducible promoter or even preferably of the promoter of the gene fused to its coding sequence and tagged by a fluorescent protein would an interesting alternative to the use of the 35SCamV promoter to allow less strong and pleiotropic effects and in the latter case would allow to localize more precisely the sites of gene expression. In the case of *EgrIAA4m* lines where the hypocotyls were particularly tiny, it should allow comparative examination of xylem patterning between inflorescence stems and hypocotyls. These strategies would be suitable at minima for the three genes studied here and moreover a systematic comparison between inflorescence stems and hypocotyls should be done for all transgenic lines as suggested by the highly contrasting patterns of secondary xylem deposition in the two organs as observed in *EgrIAA9m* overexpressing lines. These observations should be done at different stages of development of the vascular system.

To overcome problems linked with heterologous transformation system, as well as the frequent presence of sister pairs orthologs for one *Eucalyptus IAA*, a good alternative would be to express the three *IAA* genes in an homologous system. Since the *Eucalyptus* transformation is a very time consuming process with low efficiency of transformation, our team has set up a hairy root transformation system and shown that secondary xylem could be easily obtain. This promising and simplified tool will be valuable to further functionally characterize the three *IAA* genes especially because the anatomy of the hypocotyl vascular system is similar to that of the root (Chaffey et al 2002).

References

References

Abel, S. and Theologis, A. (1996) Early genes and auxin action. *Plant physiology* 111: 9-17.

Agusti, J., Herold, S., Schwarz, M., Sanchez, P., Ljung, K., Dun, E.A., et al. (2011a) Strigolactone signaling is required for auxin-dependent stimulation of secondary growth in plants. *Proceedings of the National Academy of Sciences of the United States of America* 108: 20242-20247.

Agusti, J., Lichtenberger, R., Schwarz, M., Nehlin, L. and Greb, T. (2011b) Characterization of transcriptome remodelinmeleg during cambium formation identifies MOL1 and RUL1 as opposing regulators of secondary growth. *PLoS genetics* 7: e1001312.

Aharoni, A., Keizer, L.C., Van Den Broeck, H.C., Blanco-Portales, R., Munoz-Blanco, J., Bois, G., et al. (2002) Novel insight into vascular, stress, and auxin-dependent and -independent gene expression programs in strawberry, a non-climacteric fruit. *Plant physiology* 129: 1019-1031.

Ainley, W.M., Walker, J.C., Nagao, R.T. and Key, J.L. (1988) Sequence and characterization of two auxin-regulated genes from soybean. *The Journal of biological chemistry* 263: 10658-10666.

Alejandro, S., Lee, Y., Tohge, T., Sudre, D., Osorio, S., Park, J., et al. (2012) AtABCG29 is a monolignol transporter involved in lignin biosynthesis. *Current biology : CB* 22: 1207-1212.

Aloni, R. (2001) Foliar and axial aspects of vascular differentiation: Hypotheses and evidence. *J Plant Growth Regul* 20: 22-34.

Aloni, R. and Zimmermann, M.H. (1983) The Control of Vessel Size and Density Along the Plant Axis - a New Hypothesis. *Differentiation* 24: 203-208.

Altamura MM, Possenti M, Matteucci A, Baima S, Ruberti I, Morelli G (2001) Development of the vascular system in the inflorescence stem of Arabidopsis. *New Phytol* 151:381–389

Amor, Y., Haigler, C.H., Johnson, S., Wainscott, M. and Delmer, D.P. (1995) A membrane-associated form of sucrose synthase and its potential role in synthesis of cellulose and callose in plants. *Proceedings of the National Academy of Sciences of the United States of America* 92: 9353-9357.

Andersson-Gunneras S, Mellerowicz EJ, Love J, Segerman B, Ohmiya Y, et al. (2006) Biosynthesis of cellulose-enriched tension wood in Populus: global analysis of transcripts and metabolites identifies biochemical and developmental regulators in secondary wall biosynthesis. *Plant journal* 45: 144-165.

Arase F, Nishitani H, Egusa M, Nishimoto N, Sakurai S, et al. (2012) IAA8 Involved in

Lateral Root Formation Interacts with the TIR1 Auxin Receptor and ARF Transcription Factors in Arabidopsis. *PLoS One* 7: e43414.

Arioli, T., Peng, L., Betzner, A.S., Burn, J., Wittke, W., Herth, W., et al. (1998) Molecular analysis of cellulose biosynthesis in *Arabidopsis*. *Science* 279: 717-720.

Audran-Delalande, C., Bassa, C., Mila, I., Regad, F., Zouine, M. and Bouzayen, M. (2012) Genome-wide identification, functional analysis and expression profiling of the Aux/IAA gene family in tomato. *Plant & cell physiology* 53: 659-672.

Axtell, M.J. and Bartel, D.P. (2005) Antiquity of microRNAs and their targets in land plants. *The Plant cell* 17: 1658-1673.

Baba, K., Karlberg, A., Schmidt, J., Schrader, J., Hvidsten, T.R., Bako, L., et al. (2011) Activity-dormancy transition in the cambial meristem involves stage-specific modulation of auxin response in hybrid aspen. *Proceedings of the National Academy of Sciences of the United States of America* 108: 3418-3423.

Baima, S., Possenti, M., Matteucci, A., Wisman, E. and Altamura, M.M. et al. (2001) The Arabidopsis ATHB-8 HD-Zip protein acts as a differentiation-promoting transcription factor of the vascular meristems. *Plant Physiol* 126, 643–655.

Barbez, E., Kubes, M., Rolcik, J., Beziat, C., Pencik, A., Wang, B., et al. (2012) A novel putative auxin carrier family regulates intracellular auxin homeostasis in plants. *Nature* 485: 119-122.

Barbier-Brygoo, H., Ephritikhine, G., Klämbt, D., Maurel, C., Palme, K., Schell, J., et al. (1991) Perception of the auxin signal at the plasma membrane of tobacco mesophyll protoplasts. *The Plant Journal* 1: 83-93.

Barlow P. W., Volkmann D., Baluška F. (2004) Polarity in roots. In Polarity in plants (ed. Lindsey K.), pp. 192–241. Oxford, UK: Blackwell Publishing Ltd.

Bassa,C., Mila,I., Bouzayen,M. and Audran-Delalande,C., (2012) Phenotypes associated with down-regulation of Sl-IAA27 support functional diversity among Aux/IAA family members in the tomato. *Plant Cell Physio*l 53: 1583-1595.

Baucher, M., Halpin, C., Petit-Conil, M., Boerjan, W. (2003) Lignin: genetic engineering and impact on pulping. *Crit Rev Biochem Mol Biol* 38:305-50.

Bellini, C., Pacurar, D.I., Perrone, I. (2014) Adventitious roots and lateral roots: similarities and differences. *Annu Rev Plant Biol* 65: 639-666.

Bennett, M.J., Marchant, A., Green, H.G., May, S.T., Ward, S.P., Millner, P.A., et al. (1996) *Arabidopsis* AUX1 gene: a permease-like regulator of root gravitropism. *Science* 273: 948-950.

Berleth T, J.G. (1993) The role of the MONOPTEROSgene in organizing the basal body region of the *Arabidopsis* embryo. *Development* 122: 575–587.

Berleth, T. and Mattsson, J. (2000) Vascular development: tracing signals along veins.

Current opinion in plant biology 3: 406-411.

Bhalerao, R.P. and Fischer, U. (2014) Auxin gradients across wood-instructive or incidental? *Physiologia plantarum* 151: 43-51.

Blakeslee, J.J., Peer, W.A. and Murphy, A.S. (2005) Auxin transport. *Current opinion in plant biology* 8: 494-500.

Boerjan, W., Ralph, J. and Baucher, M. (2003) Lignin biosynthesis. *Annual review of plant biology* 54: 519-546.

Brett, C.T. (2000) Cellulose microfibrils in plants: biosynthesis, deposition, and int*Egr*ation into the cell wall. *International review of cytology* 199: 161-199.

Brooker, M.I.H. (2000) A new Classification of the Genus *Eucalyptus* L'Her. (Myrtaceae). *Australian Systematic Botany* 13: 79-148.

Brown, D.M., Zhang, Z., Stephens, E., Dupree, P. and Turner, S.R. (2009) Characterization of IRX10 and IRX10-like reveals an essential role in glucuronoxylan biosynthesis in *Arabidopsis*. *Plant J* 57: 732-746.

Burton, R.A. and Fincher, G.B. (2014) Plant cell wall engineering: applications in biofuel production and improved human health. *Curr Opin Biotechnol* 26:79-84.

Capek, P., Alfoldi, J. and Liskova, D. (2002) An acetylated galactoglucomannan from Picea abies L. Karst. *Carbohydrate research* 337: 1033-1037.

Cardozo, T. and Pagano, M. (2004) The SCF ubiquitin ligase: insights into a molecular machine. *Nature reviews. Molecular cell biology* 5: 739-751.

Carlsbecker, A., Le, J.Y., Roberts, C.J., Dettmer, J. and Lehesranta, S. et al. (2010) Moreno-Risueno MA, Vate ń A, Thitamadee S et al.: Cell signalling by microRNA165/6 directs gene dose- dependent root cell fate. *Nature* 465: 316-321.

Carpita, N.C. (2012) Progress in the biological synthesis of the plant cell wall: new ideas for improving biomass for bioenergy. *Curr Opin Biotechnol* 23: 330-337.

Cecchetti, V., Altamura, M. M., Falasca, G., Costantino, P., Cardarelli, M. (2008) Auxin regulates Arabidopsis anther dehiscence, pollen maturation, and filament elongation. Plant cell 20: 1760-1774.

Chaffey, N., Cholewa, E., Regan, S., Sundberg, B. (2002) Secondary xylem development in Arabidopsis: a model for wood formation. *Physiol Plant* 114:594-600.

Chapman, E.J. and Estelle, M. (2009) Mechanism of auxin-regulated gene expression in plants. *Annual review of genetics* 43: 265-285.

Chen, J.G., Shimomura, S., Sitbon, F., Sandberg, G. and Jones, A.M. (2001a) The role of auxin-binding protein 1 in the expansion of tobacco leaf cells. *The Plant journal : for cell and molecular biology* 28: 607-617.

Chen, J.G., Ullah, H., Young, J.C., Sussman, M.R. and Jones, A.M. (2001b) ABP1 is

required for organized cell elongation and division in *Arabidopsis* embryogenesis. *Genes & development* 15: 902-911.

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of *Arabidopsis* thaliana. *The Plant journal : for cell and molecular biology* 16: 735-743.

Cohen, J.D. (1996) In vitro Tomato Fruit Cultures Demonstrate a Role for Indole-3-acetic Acid in Regulating Fruit Ripening. *Journal of the American Society for Horticultural Science* 121: 520-524.

Colón-Carmona A., Chen D.L., Yeh K.C., Abel S. (2000) Aux/IAA proteins are phosphorylated by phytochrome in vitro. *Plant Physiol.* 124: 1728-38.

David, K.M., Couch, D., Braun, N., Brown, S., Grosclaude, J. and Perrot-Rechenmann, C. (2007) The auxin-binding protein 1 is essential for the control of cell cycle. *The Plant journal : for cell and molecular biology* 50: 197-206.

David S. Goldfarb, J.G., Gary Schoolnik & Roger D. Kornberg (1986) Synthetic peptides as nuclear localization signals. *Nature* 322: 641 - 644.

De Rybel, B., Moller, B., Yoshida, S., Grabowicz, I., Barbier de Reuille, P., Boeren, S., et al. (2013) A bHLH complex controls embryonic vascular tissue establishment and indeterminate growth in *Arabidopsis*. *Developmental cell* 24: 426-437.

De Rybel, B., Vassileva, V., Parizot, B., Demeulenaere, M., Grunewald, W. et al. (2010) A novel Aux/IAA28 signaling cascade activates GATA23-dependent specification of lateral root founder cell identity. *Curr. Biol.* 20: 1697–1706.

Déjardin A., Laurans F., Arnaud D., Breton C., Pilate G. and Leplé J.C. (2010) Wood formation in Angiosperms. Comptes Rendus Biologies *333: 325-334.*

Demura, T. and Fukuda, H. (2007) Transcriptional regulation in wood formation. *Trends in plant science* 12: 64-70.

Deng, W., Yan, F., Liu, M., Wang, X., Li, Z. (2012) Down-regulation of SlIAA15 in tomato altered stem xylem development and production of volatile compounds in leaf exudates. Plant Signal Behav 7: 911-913.

De Jong, M., Mariani, C. and Vriezen, W.H. (2009) The role of auxin and gibberellin in tomato fruit set. *Journal of Experimental Botany*. 60: 1523 –1532.

Dharmasiri, N., Dharmasiri, S. and Estelle, M. (2005a) The F-box protein TIR1 is an auxin receptor. *Nature* 435: 441-445.

Dharmasiri, N., Dharmasiri, S., Weijers, D., Lechner, E., Yamada, M., Hobbie, L., et al. (2005b) Plant development is regulated by a family of auxin receptor F box proteins. *Developmental cell* 9: 109-119.

Doblin, M.S., Kurek, I., Jacob-Wilk, D. and Delmer, D.P. (2002) Cellulose biosynthesis in plants: from genes to rosettes. *Plant & cell physiology* 43: 1407-1420.

Donner, T.J., Sherr, I. and Scarpella, E. (2009) Regulation of preprocambial cell state acquisition by auxin signaling in *Arabidopsis* leaves. *Development* 136: 3235-3246.

Dreher, K.A., Brown, J., Saw, R.E. and Callis, J. (2006) The *Arabidopsis* Aux/IAA protein family has diversified in degradation and auxin responsiveness. *The Plant cell* 18: 699-714.

Du, J. and Groover, A. (2010) Transcriptional regulation of secondary growth and wood formation. *Journal of intEgrative plant biology* 52: 17-27.

Du, J., Mansfield, S.D. and Groover, A.T. (2009) The Populus homeobox gene ARBORKNOX2 regulates cell differentiation during secondary growth. *The Plant journal : for cell and molecular biology* 60: 1000-1014.

Du, J., Miura, E., Robischon, M., Martinez, C. and Groover, A. (2011) The Populus Class III HD ZIP transcription factor POPCORONA affects cell differentiation during secondary growth of woody stems. *PloS one* 6: e17458.

Dubrovsky, J. G., Rost, T. L., Colon-Carmona, A., Doerner, P. (2001) Early primordium morphogenesis during lateral root initiation in Arabidopsis thaliana. *Planta* 214: 30-36.

Eames A.J., MacDaniels L.H. (1947) An Introduction to Plant Anatomy 2nd edn. New York: McGraw-Hill.

Ellis, C.M., Nagpal, P., Young, J.C., Hagen, G., Guilfoyle, T.J. and Reed, J.W. (2005) AUXIN RESPONSE FACTOR1 and AUXIN RESPONSE FACTOR2 regulate senescence and floral organ abscission in *Arabidopsis* thaliana. *Development* 132: 4563-4574.

Elo, A., Immanen, J., Nieminen, K., Helariutta, Y. (2009) Stem cell function during plant vascular development. *Semin Cell Dev Biol* 20:1097-1106.

Esau K. (1965) Vascular Differentiation in Plants. New York: Holt, Rinehart, and Winston.

Etchells, J.P., Turner, S.R. (2010) The PXY-CLE41 receptor ligand pair defines a multifunctional pathway that controls the rate and orientation of vascular cell division. *Development* 137:767-774.

Etchells, J. P., Provost, C. M., Mishra, L. (2013) Turner, S. R. WOX4 and WOX14 act downstream of the PXY receptor kinase to regulate plant vascular proliferation independently of any role in vascular organisation. *Development* 140: 2224-2234.

Evans, M., Ishikawa, H. and Estelle, M. (1994) Responses of *Arabidopsis* roots to auxin studied with high temporal resolution: Comparison of wild type and auxin-response mutants. *Planta* 194: 215-222.

Fahlgren, N., Montgomery, T.A., Howell, M.D., Allen, E., Dvorak, S.K., Alexander, A.L., et al. (2006) Regulation of AUXIN RESPONSE FACTOR3 by TAS3 ta-siRNA affects developmental timing and patterning in *Arabidopsis*. *Curr Biol* 16: 939-944.

Foucart, C., Paux, E., Ladouce, N., San-Clemente, H., Grima-Pettenati, J. and Sivadon, P. (2006) Transcript profiling of a xylem vs phloem cDNA subtractive library identifies new genes expressed during xylogenesis in *Eucalyptus*. *New Phytol* 170: 739-752.

Friml, J., Wisniewska, J., Benkova, E., Mendgen, K. and Palme, K. (2002) Lateral relocation of auxin efflux regulator PIN3 mediates tropism in *Arabidopsis*. *Nature* 415: 806-809.

Fukaki, H., Tameda, S., Masuda, H. and Tasaka, M. (2002) Lateral root formation is blocked by a gain-of-function mutation in the SOLITARY-ROOT/IAA14 gene of *Arabidopsis*. *Plant Journal* 29: 153-168.

Fukaki, H., Nakao, Y., Okushima, Y., Theologis, A., Tasaka, M. (2005) Tissue-specific expression of stabilized SOLITARY-ROOT/IAA14 alters lateral root development in Arabidopsis. Plant Journal 44: 382-395.

Fukaki, H., Taniguchi, N. and Tasaka, M. (2006) PICKLE is required for SOLITARY-ROOT/IAA14-mediated repression of ARF7 and ARF19 activity during *Arabidopsis* lateral root initiation. *The Plant journal : for cell and molecular biology* 48: 380-389.

Fukaki, H., Okushima, Y., Tasaka, M. (2007) Auxin-mediated lateral root formation in higher plants. *Int. Rev. Cytol.* 256: 111–137.

Fukaki, H. and Tasaka, M. (2009) Hormone interactions during lateral root formation. *Plant Mol. Biol.* 69: 437–449.

Fukuda, H. (1997) Tracheary Element Differentiation. *The Plant cell* 9: 1147-1156.

Furuya N., Takahashi S., Mirayashi M., (1970) The chemical composition of gelatinous layer from the tension wood of Populus eura-americana. *Mokuzai Gakkaishi* 16: 26–30.

Galweiler, L., Guan, C.H., Muller, A., Wisman, E., Mendgen, K., Yephremov, A., et al. (1998) Regulation of polar auxin transport by AtPIN1 in *Arabidopsis* vascular tissue. *Science* 282: 2226-2230.

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. *Journal of Analytical and Applied Pyrolysis* 95: 95-100.

Gindl, W., Grabner, M., and Wimmer, R. (2014). The influence of temperature on latewood lignin content in treeline Norway spruce compared with maximum density and ring width. *Trees* 14: 409-414.

Goetz, M., Hooper, L.C., Johnson, S.D., Rodrigues, J.C., Vivian-Smith, A. and Koltunow, A.M. (2007) Expression of aberrant forms of AUXIN RESPONSE FACTOR8 stimulates parthenocarpy in *Arabidopsis* and tomato. *Plant Physiol* 145: 351-366.

Goh, T., Kasahara, H., Mimura, T., Kamiya, Y. and Fukaki, H. (2012) Multiple AUX/IAA-ARF modules regulate lateral root formation: the role of *Arabidopsis* SHY2/IAA3-

mediated auxin signalling. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 367: 1461-1468.

Goicoechea, M., Lacombe, E., Legay, S., Mihaljevic, S., Rech, P., Jauneau, A., et al. (2005) EgMYB2, a new transcriptional activator from *Eucalyptus* xylem, regulates secondary cell wall formation and lignin biosynthesis. *The Plant journal : for cell and molecular biology* 43: 553-567.

Gray, W.M., Kepinski, S., Rouse, D., Leyser, O. and Estelle, M. (2001) Auxin regulates SCF(TIR1)-dependent degradation of AUX/IAA proteins. *Nature* 414: 271-276.

Greenham, K., Santner, A., Castillejo, C., Mooney, S., Sairanen, I., Ljung, K., et al. (2011) The AFB4 auxin receptor is a negative regulator of auxin signaling in seedlings. *Current biology : CB* 21: 520-525.

Guilfoyle, T., Hagen, G., Ulmasov, T. and Murfett, J. (1998) How does auxin turn on genes? *Plant Physiol* 118: 341-347.

Guilfoyle, T.J. and Hagen, G. (2007) Auxin response factors. *Current opinion in plant biology* 10: 453-460.

Habrant, A., Laurans, F., Dégardin, A., Leplé, J.-C. and Pilate, G., et al. (2003) Physical and chemical changes of poplar cell walls during tension wood formation, in: 12th Int. Symp. on Wood and Pulping Chemistry (ISWPC), Madison, Wisconsin (USA), 8–12 June.

Haigler, C.H., Ivanova-Datcheva, M., Hogan, P.S., Salnikov, V.V., Hwang, S., Martin, K., et al. (2001) Carbon partitioning to cellulose synthesis. *Plant molecular biology* 47: 29-51.

Hamann, T., Benkova, E., Baurle, I., Kientz, M. and Jurgens, G. (2002) The *Arabidopsis* BODENLOS gene encodes an auxin response protein inhibiting MONOPTEROS-mediated embryo patterning. *Genes Dev* 16: 1610-1615.

Hamann, T., Mayer, U. and Jurgens, G. (1999) The auxin-insensitive bodenlos mutation affects primary root formation and apical-basal patterning in the *Arabidopsis* embryo. *Development* 126: 1387-1395.

Hardtke, C.S. and Berleth, T. (1998) The *Arabidopsis* gene MONOPTEROS encodes a transcription factor mediating embryo axis formation and vascular development. *EMBO J* 17: 1405-1411.

Harper, R.M., Stowe-Evans, E.L., Luesse, D.R., Muto, H., Tatematsu, K., Watahiki, M.K., et al. (2000) The NPH4 locus encodes the auxin response factor ARF7, a conditional regulator of differential growth in aerial *Arabidopsis* tissue. *Plant Cell* 12: 757-770.

Heisler MG, Ohno C, Das P, Sieber P, Reddy GV, Long JA, Meyerowitz EM (2005) Patterns of auxin transport and gene expression during primordium development revealed by live imaging of the Arabidopsis inflorescence meristem. *Curr Biol* 15: 1899–1911.

Hellgren, J.M., Olofsson, K. and Sundberg, B. (2004) Patterns of auxin distribution during gravitational induction of reaction wood in poplar and pine. *Plant physiology* 135: 212-220.

Hendelman, A., Buxdorf, K., Stav, R., Kravchik, M., Arazi, T. (2012) Inhibition of lamina outgrowth following Solanum lycopersicum AUXIN RESPONSE FACTOR 10 (SlARF10) derepression. *Plant Mol Biol* 78: 561-576.

Hertel, R., Thomson, K.S. and Russo, V.E. (1972) In-vitro auxin binding to particulate cell fractions from corn coleoptiles. *Planta* 107: 325-340.

Himanen, K., Boucheron, E., Vanneste, S., de Almeida Engler, J., Inze, D. and Beeckman, T. (2002) Auxin-mediated cell cycle activation during early lateral root initiation. *The Plant cell* 14: 2339-2351.

Hirakawa, Y., Kondo, Y. and Fukuda, H. (2010) TDIF peptide signaling regulates vascular stem cell proliferation via the WOX4 homeobox gene in *Arabidopsis*. *The Plant cell* 22: 2618-2629.

Hobbie L, M.M., Hurwitz LR, Pierro A, Liu NY, Bandyopadhyay A, Estelle M. (2000) The axr6mutants of *Arabidopsis* define a gene involved in auxin response and early development. *Development* 127: 23–32.

Hu, Z., Keceli, M.A., Piisila, M., Li, J., Survila, M., Heino, P., et al. (2012) F-box protein AFB4 plays a crucial role in plant growth, development and innate immunity. *Cell research* 22: 777-781.

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 biology* 11: 173.

Ilegems, M., Douet, V., Meylan-Bettex, M., Uyttewaal, M., Brand, L., Bowman, J.L., et al. (2010) Interplay of auxin, KANADI and Class III HD-ZIP transcription factors in vascular tissue formation. *Development* 137: 975-984.

Jain, M., Kaur, N., Garg, R., Thakur, J.K., Tyagi, A.K. and Khurana, J.P. (2006) Structure and expression analysis of early auxin-responsive Aux/IAA gene family in rice (Oryza sativa). *Functional & intEgrative genomics* 6: 47-59.

Jones, B., Frasse, P., Olmos, E., Zegzouti, H., Li, Z.G., Latche, A., et al. (2002) Down-regulation of DR12, an auxin-response-factor homolog, in the tomato results in a pleiotropic phenotype including dark green and blotchy ripening fruit. *Plant J* 32: 603-613.

Jourez B., Riboux A., Leclercq A. (2001) Anatomical characteristics of tension wood and opposite wood in young inclined stems of poplar (Populus euramericana cv‘Ghoy’). *IAWA Journal* 22: 133–157.

Kalluri, U.C., Difazio, S.P., Brunner, A.M. and Tuskan, G.A. (2007) Genome-wide

analysis of Aux/IAA and ARF gene families in Populus trichocarpa. *BMC plant biology* 7: 59.

Kaneda, M., Schuetz, M., Lin, B.S., Chanis, C., Hamberger, B., Western, T.L., et al. (2011) ABC transporters coordinately expressed during lignification of *Arabidopsis* stems include a set of ABCBs associated with auxin transport. *Journal of experimental botany* 62: 2063-2077.

Kepinski, S. and Leyser, O. (2004) Auxin-induced SCFTIR1-Aux/IAA interaction involves stable modification of the SCFTIR1 complex. *Proceedings of the National Academy of Sciences of the United States of America* 101: 12381-12386.

Kim, J., Harter, K. and Theologis, A. (1997) Protein-protein interactions among the Aux/IAA proteins. *Proceedings of the National Academy of Sciences of the United States of America* 94: 11786-11791.

Ko, J.H., Han, K.H., Park, S. and Yang, J. (2004) Plant body weight-induced secondary growth in *Arabidopsis* and its transcription phenotype revealed by whole-transcriptome profiling. *Plant physiology* 135: 1069-1083.

Kong, Y., Zhou, G., Avci, U., Gu, X., Jones, C., Yin, Y., et al. (2009) Two poplar glycosyltransferase genes, PdGATL1.1 and PdGATL1.2, are functional orthologs to PARVUS/AtGATL1 in *Arabidopsis*. *Molecular plant* 2: 1040-1050.

Koroleva, O.A., Tomlinson, M., Parinyapong, P., Sakvarelidze, L., Leader, D., Shaw, P., et al. (2004) CycD1, a putative G1 cyclin from Antirrhinum majus, accelerates the cell cycle in cultured tobacco BY-2 cells by enhancing both G1/S entry and progression through S and G2 phases. *The Plant cell* 16: 2364-2379.

Kubo, M., Udagawa, M., Nishikubo, N., Horiguchi, G., Yamaguchi, M., Ito, J., et al. (2005) Transcription switches for protoxylem and metaxylem vessel formation. *Genes Dev* 19: 1855-1860.

Kumar, R., Tyagi, A.K. and Sharma, A.K. (2011) Genome-wide analysis of auxin response factor (ARF) gene family from tomato and analysis of their role in flower and fruit development. *Mol Genet Genomics* 285: 245-260.

Lavenus, J., Goh, .T, Roberts, I., Guyomarc'h, S., Laplaze, L. et al (2013) Lateral root development in Arabidopsis: fifty shades of auxin. *Trends Plant Sci* 18:450-458.

Lee, C., O'Neill, M.A., Tsumuraya, Y., Darvill, A.G. and Ye, Z.H. (2007a) The irregular xylem9 mutant is deficient in xylan xylosyltransferase activity. *Plant & cell physiology* 48: 1624-1634.

Lee, C., Teng, Q., Huang, W., Zhong, R. and Ye, Z.H. (2009a) The F8H glycosyltransferase is a functional paralog of FRA8 involved in glucuronoxylan biosynthesis in *Arabidopsis*. *Plant & cell physiology* 50: 812-827.

Lee, C., Teng, Q., Huang, W., Zhong, R. and Ye, Z.H. (2009b) The poplar GT8E and GT8F glycosyltransferases are functional orthologs of *Arabidopsis* PARVUS

involved in glucuronoxylan biosynthesis. *Plant & cell physiology* 50: 1982-1987.

Lee, C., Zhong, R., Richardson, E.A., Himmelsbach, D.S., McPhail, B.T. and Ye, Z.H. (2007b) The PARVUS gene is expressed in cells undergoing secondary wall thickening and is essential for glucuronoxylan biosynthesis. *Plant & cell physiology* 48: 1659-1672.

Leyser, H.M., Pickett, F.B., Dharmasiri, S. and Estelle, M. (1996) Mutations in the AXR3 gene of *Arabidopsis* result in altered auxin response including ectopic expression from the SAUR-AC1 promoter. *The Plant journal : for cell and molecular biology* 10: 403-413.

Leyser, O. (2003) Regulation of shoot branching by auxin. *Trends Plant Sci* 8: 541–545.

Li, E., Bhargava, A., Qiang, W., Friedmann, M.C., Forneris, N., Savidge, R.A., et al. (2012) The Class II KNOX gene KNAT7 negatively regulates secondary wall formation in *Arabidopsis* and is functionally conserved in Populus. *The New phytologist* 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. *The Plant journal : for cell and molecular biology* 67: 328-341.

Li, S., Bashline, L., Lei, L. and Gu, Y. (2014) Cellulose synthesis and its regulation. *The Arabidopsis book / American Society of Plant Biologists* 12: e0169.

Liebminger, E., Grass, J., Altmann, F., Mach, L. and Strasser, R. (2013) Characterizing the link between glycosylation state and enzymatic activity of the endo-beta1,4-glucanase KORRIGAN1 from *Arabidopsis* thaliana. *The Journal of biological chemistry* 288: 22270-22280.

Liscum, E. and Reed, J. W. (2002) Genetics of Aux/IAA and ARF action in plant growth and development. *Plant Mol Biol* 49:387-400.

Little, C.H.A. and Savidge, R.A. (1987) The Role of Plant-Growth Regulators in Forest Tree Cambial Growth. *Plant Growth Regul* 6: 137-169.

Liu, P.P., Montgomery, T.A., Fahlgren, N., Kasschau, K.D., Nonogaki, H. and Carrington, J.C. (2007) Repression of AUXIN RESPONSE FACTOR10 by microRNA160 is critical for seed germination and post-germination stages. The Plant journal 52: 133-146.

Ljung, K., Bhalerao, R.P. and Sandberg, G. (2001) Sites and homeostatic control of auxin biosynthesis in *Arabidopsis* during vegetative growth. *The Plant journal : for cell and molecular biology* 28: 465-474.

Love, J., Bjorklund, S., Vahala, J., Hertzberg, M., Kangasjarvi, J. and Sundberg, B. (2009) Ethylene is an endogenous stimulator of cell division in the cambial meristem of Populus. *Proc Natl Acad Sci U S A* 106: 5984-5989.

Lucas, M., Godin, C., Jay-Allemand, C., Laplaze, L. (2008) Auxin fluxes in the root apex co-regulate gravitropism and lateral root initiation. *J Exp Bot* 59: 55-56.

Ludwig-Muller, J. (2011) Auxin conjugates: their role for plant development and in the evolution of land plants. *Journal of experimental botany* 62: 1757-1773.

Mano, Y. and Nemoto, K. (2012) The pathway of auxin biosynthesis in plants. *Journal of experimental botany* 63: 2853-2872.

Mauriat, M., Le, P.G., Rozenberg, P., Delzon S., Bréda, N. et al (2014) Tree biotechnology: Wood formation in tress. CRC Press.

Mauriat, M. and Moritz, T. (2009) Analyses of GA20ox- and GID1-over-expressing aspen suggest that gibberellins play two distinct roles in wood formation. *The Plant journal : for cell and molecular biology* 58: 989-1003.

McCarthy, R.L., Zhong, R., Fowler, S., Lyskowski, D., Piyasena, H., Carleton, K., et al. (2010) The poplar MYB transcription factors, PtrMYB3 and PtrMYB20, are involved in the regulation of secondary wall biosynthesis. *Plant & cell physiology* 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.

McCann, M.C., Stacey, N.J., Dahiya, P., Milioni, D. and Sado, P.E. et al. (2001) Zinnia. Everybody needs good neighbors. *Plant Physiology* 127: 1380–1382.

Meier, D., Fortmann, I., Odermatt, J. and Faix, O. (2005) Discrimination of genetically modified poplar clones by analytical pyrolysis–gas chromatography and principal component analysis. *Journal of Analytical and Applied Pyrolysis* 74: 129-137.

Mele, G., Ori, N., Sato, Y., Hake, S. (2003) The knotted1-like homeobox gene BREVIPEDICELLUS regulates cell differentiation by modulating metabolic pathways. *Genes Dev* 17:2088-2093.

Mellerowicz, E.J., Baucher, M., Sundberg, B. and Boerjan, W. (2001) Unravelling cell wall formation in the woody dicot stem. *Plant molecular biology* 47: 239-274.

Mitsuda, N., Iwase, A., Yamamoto, H., Yoshida, M., Seki, M., Shinozaki, K., et al. (2007) NAC transcription factors, NST1 and NST3, are key regulators of the formation of secondary walls in woody tissues of *Arabidopsis*. *Plant Cell* 19: 270-280.

Mitsuda N., Ohme-Takagi M. (2008) NAC transcription factors NST1 and NST3 regulate pod shattering in a partially redundant manner by promoting secondary wall formation after the establishment of tissue identity. *Plant Journal* 56:768-778.

Miyashima, S., Sebastian, J., Lee, J.Y. and Helariutta, Y. (2013) Stem cell function during plant vascular development. *The EMBO journal* 32: 178-193.

Moura, J.C.M.S., Bonine, C.A.V., de Oliveira Fernandes Viana, J., Dornelas, M.C., and

Mazzafera, P. (2010). Abiotic and biotic stresses and changes in the lignin content and composition in plants. *J IntEgr Plant Biol* 52: 360-366.

Myburg, A.A., Grattapaglia, D., Tuskan, G.A., Hellsten, U., Hayes, R.D., Grimwood, J., et al. (2014) The genome of *Eucalyptus* grandis. *Nature* 509: 356-362.

Nagpal, P., Walker, L.M., Young, J.C., Sonawala, A., Timpte, C., Estelle, M., et al. (2000) AXR2 encodes a member of the Aux/IAA protein family. *Plant physiology* 123: 563-574.

Nagpal, P., Ellis, C. M., Weber, H., Ploense, S. E., Barkawi, L. S. et al. (2005) Auxin response factors ARF6 and ARF8 promote jasmonic acid production and flower maturation. *Development* 132: 4107-4118.

Nieminen, K., Immanen, J., Laxell, M., Kauppinen, L., Tarkowski, P., Dolezal, K., et al. (2008) Cytokinin signaling regulates cambial development in poplar. *Proceedings of the National Academy of Sciences of the United States of America* 105: 20032-20037.

Nieminen, K.M., Kauppinen, L., 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., et al. (2008) Dissecting the molecular basis of the regulation of wood formation by auxin in hybrid aspen. *The Plant cell* 20: 843-855.

Normanly, J. (2010) Approaching cellular and molecular resolution of auxin biosynthesis and metabolism. *Cold Spring Harbor perspectives in biology* 2: a001594.

Novaes, E., Osorio, L., Drost, D. R., Miles, B. L. and Boaventura-Novaes, C. R. (2009) Quantitative genetic analysis of biomass and wood chemistry of Populus under different nitrogen levels. *New Phytol* 182: 878-890.

Oh, S., Park, S., Han, K.H. (2003)Transcriptional regulation of secondary growth in Arabidopsis thaliana. Journal of Experimental Botany 54: 2609-2722.

Ohashi-Ito, K. and Fukuda, H. (2003) HD-zip III homeobox genes that include a novel member, ZeHB-13 (Zinnia)/ATHB-15 (*Arabidopsis*), are involved in procambium and xylem cell differentiation. *Plant & cell physiology* 44: 1350-1358.

Ohashi-Ito, K. and Fukuda, H. (2010) Transcriptional regulation of vascular cell fates. *Curr Opin Plant Biol* 13: 670-676.

Ohtani, M., Nishikubo, N., Xu, B., Yamaguchi, M., Mitsuda, N., Goue, N., et al. (2011) A NAC domain protein family contributing to the regulation of wood formation in poplar. *The Plant journal : for cell and molecular biology* 67: 499-512.

Okushima, Y., Mitina, I., Quach, H.L. and Theologis, A. (2005) AUXIN RESPONSE FACTOR 2 (ARF2): a pleiotropic developmental regulator. *Plant J* 43: 29-46.

Okushima, Y., Overvoorde, P.J., Arima, K., Alonso, J.M., Chan, A., Chang, C., et al. (2005b) Functional genomic analysis of the AUXIN RESPONSE FACTOR gene

family members in *Arabidopsis* thaliana: unique and overlapping functions of ARF7 and ARF19. *The Plant cell* 17: 444-463.

Ouellet, F., Overvoorde, P.J. and Theologis, A. (2001) IAA17/AXR3: biochemical insight into an auxin mutant phenotype. *The Plant cell* 13: 829-841.

Overvoorde, P.J., Okushima, Y., Alonso, J.M., Chan, A., Chang, C., Ecker, J.R., et al. (2005) Functional genomic analysis of the AUXIN/INDOLE-3-ACETIC ACID gene family members in *Arabidopsis* thaliana. *The Plant cell* 17: 3282-3300.

Overvoorde, P., Fukaki, H., Beeckman, T. (2010) Auxin control of root development. Cold Spring Harb Perspect Biol 6: a001537.

Paredez, A.R., Persson, S., Ehrhardt, D.W. and Somerville, C.R. (2008) Genetic evidence that cellulose synthase activity influences microtubule cortical array organization. *Plant physiology* 147: 1723-1734.

Parry, G., Calderon-Villalobos, L.I., Prigge, M., Peret, B., Dharmasiri, S., Itoh, H., et al. (2009) Complex regulation of the TIR1/AFB family of auxin receptors. *Proceedings of the National Academy of Sciences of the United States of America* 106: 22540-22545.

Pauly M., Keegstra K. (2008) Cell-wall carbohydrates and their modification as a resource for biofuels. *The plant journal* 54: 559-568.

Pauly, M., Gille, S., Liu, L., Mansoori, N., de Souza, A., Schultink, A., et al. (2013) Hemicellulose biosynthesis. *Planta* 238: 627-642.

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: 263-280.

Pekker, I., Alvarez, J.P. and Eshed, Y. (2005) Auxin response factors mediate *Arabidopsis* organ asymmetry via modulation of KANADI activity. *Plant Cell* 17: 2899-2910.

Peret, B., De Rybel, B., Casimiro, I., Benkova, E., Swarup, R. (2009) Arabidopsis lateral root development: an emerging story. *Trends Plant Sci* 14: 399-408.

Perrot-Rechenmann, C. (2010) Cellular responses to auxin: division versus expansion. *Cold Spring Harbor perspectives in biology* 2: a001446.

Persson, S., Caffall, K.H., Freshour, G., Hilley, M.T., Bauer, S., Poindexter, P., et al. (2007) The *Arabidopsis* irregular xylem8 mutant is deficient in glucuronoxylan and homogalacturonan, which are essential for secondary cell wall int*Egr*ity. *The Plant cell* 19: 237-255.

Pesquet, E., Zhang, B., Gorzsás, A., Puhakainen, T., Serk, H. (2013) Non-cell-autonomous postmortem lignification of tracheary elements in Zinnia elegans. *Plant cell* 25: 1314-1328.

Pilate, G., Chabbert, B., Cathala, B., Yoshinaga, A. and Leplé, J.C. et al. (2004)

Lignification and tension wood. *C R Biol* 327:889-901.

Pitre, F.E., Pollet, B., Lafarguette, F., Cooke, J.E. and MacKay, J.J. et al. (2007) Effects of increased nitrogen supply on the lignification of poplar wood. *J Agric Food Chem* 5:10306-1030614.

Pitre, F. E., Lafarguette, F., Boyle, B., Pavy, N. and Caron, S. et al. (2010) High nitrogen fertilization and stem leaning have overlapping effects on wood formation in poplar but invoke largely distinct molecular pathways. *Tree Physiol* 30:1273-1289.

Plomion, C., Leprovost, G. and Stokes, A. (2001) Wood formation in trees. *Plant Physiol* 127: 1513-1523.

Przemeck, G.K., Mattsson, J., Hardtke, C.S., Sung, Z.R. and Berleth, T. (1996) Studies on the role of the *Arabidopsis* gene MONOPTEROS in vascular development and plant cell axialization. *Planta* 200: 229-237.

Rademacher, E.H., Moller, B., Lokerse, A.S., Llavata-Peris, C.I., van den Berg, W. and Weijers, D. (2011) A cellular expression map of the *Arabidopsis* AUXIN RESPONSE FACTOR gene family. *The Plant journal : for cell and molecular biology* 68: 597-606.

Ranocha P, Denanc é N, Vanholme R, Freydier A, Martinez Y et al. (2010) Walls are thin 1 (WAT1), an Arabidopsis homolog of Medicago truncatula NODULIN21, is a tonoplast-localized protein required for secondary wall formation in fibers. *Plant Journal* 63: 469-483.

Ranocha, P., Dima, O., Nagy, R., Felten, J. and Corratge-Faillie, C. et al. (2013) Arabidopsis WAT1 is a vacuolar auxin transport facilitator required for auxin homoeostasis. *Nature Communications* 4: 2625

Reed, J.W. (2001) Roles and activities of Aux/IAA proteins in *Arabidopsis*. *Trends in plant science* 6: 420-425.

Reinhardt, D., Mandel, T., Kuhlemeier, C. (2000) Auxin regulates the initiation and radial position of plant lateral organs. *Plant cell* 12: 507-18.

Reinhardt D. (2003) Vascular patterning: more than just auxin? *Curr Biol* 13: R485–R487

Rengel, D., San Clemente, H., Servant, F., Ladouce, N., Paux, E., Wincker, P., et al. (2009) A new genomic resource dedicated to wood formation in *Eucalyptus*. *BMC Plant Biol* 9: 36.

Prigge, M.J., Otsuga, D., Alonso, J.M., Ecker, J.R., Drews, G.N. and Clark, S.E. (2005) Class III homeodomain-leucine zipper gene family members have overlapping, antagonistic, and distinct roles in *Arabidopsis* development. *The Plant cell* 17: 61-76.

Robischon, M., Du, J., Miura, E. and Groover, A. (2011) The Populus class III HD ZIP, popREVOLUTA, influences cambium initiation and patterning of woody stems. *Plant physiology* 155: 1214-1225.

Rogg, L.E., Lasswell, J. and Bartel, B. (2001) A gain-of-function mutation in IAA28 suppresses lateral root development. *The Plant cell* 13: 465-480.

Rouse, D., Mackay, P., Stirnberg, P., Estelle, M. and Leyser, O. (1998) Changes in auxin response from mutations in an AUX/IAA gene. *Science* 279: 1371-1373.

Ruan, Y. L., Patrick, J. W., Bouzayen, M., Osorio, S., Fernie, A. R. (2012) Molecular regulation of seed and fruit set. Trends Plant Sci 17: 656-665.

Sachs, T. (1986) Cellular interactions in tissue and organ development. *Symposia of the Society for Experimental Biology* 40: 181-210.

Sachs, T. (1981) The control of patterned differentiation of vascular tissues. *Adv. Bot. Res.* 9: 151–262.

Sagar, M., Chervin, C., Mila, I., Hao, Y., Roustan, J. P., et al. (2013) SlARF4, an auxin response factor involved in the control of sugar metabolism during tomato fruit development. *Plant Physiol* 161: 1362-1374.

Santner A., Calderon-Villalobos L. and Estelle M. (2009) Plant hormones are versatile chemical regulators of plant growth. Nat Chem Biol. 5: 301-7.

Sato, A. and Yamamoto, K.T. (2008) Overexpression of the non-canonical Aux/IAA genes causes auxin-related aberrant phenotypes in *Arabidopsis*. *Physiologia plantarum* 133: 397-405.

Sauer, M., Balla, J., Luschnig, C., Wisniewska, J., Reinohl, V., Friml, J., et al. (2006) Canalization of auxin flow by Aux/IAA-ARF-dependent feedback regulation of PIN polarity. *Gene Dev* 20: 2902-2911.

Sauer, M. and Kleine-Vehn, J. (2011) AUXIN BINDING PROTEIN1: the outsider. *The Plant cell* 23: 2033-2043.

Savidge, R.A. (1983) The role of plant hormones in higher plant cellular differentiation. II. Experiments with the vascular cambium, and sclereid and tracheid differentiation in the pine, Pinus contorta. *The Histochemical journal* 15: 447-466.

Saxena, I.M. and Brown, R.M., Jr. (2005) Cellulose biosynthesis: current views and evolving concepts. *Annals of botany* 96: 9-21.

Scarpella, E., Barkoulas, M. and Tsiantis, M. (2010) Control of Leaf and Vein Development by Auxin. *Cold Spring Harbor perspectives in biology* 2.

Scarpella, E., Marcos, D., Friml, J., Berleth, T. (2006) Control of leaf vascular patterning by polar auxin transport. *Genes Dev* 8: 1015-1027.

Schrader, J., Baba, K., May, S.T., Palme, K., Bennett, M., Bhalerao, R.P., et al. (2003) Polar auxin transport in the wood-forming tissues of hybrid aspen is under simultaneous control of developmental and environmental signals. *Proceedings of the National Academy of Sciences of the United States of America* 100: 10096-10101.

Searson, M.J., Thomas, D.S., Montagu, K.D., and Conroy, J.P. (2004). Wood density and

anatomy of water-limited eucalypts. *Tree Physiol* 24: 1295-1302.

Séguin A. (2011) How could forest trees play an important role as feedstock for bioenergy production? Current Opinion in Environmental Sustainability 3: 90-94.

Sessions, A., Nemhauser, J.L., McColl, A., Roe, J.L., Feldmann, K.A. and Zambryski, P.C. (1997) ETTIN patterns the *Arabidopsis* floral meristem and reproductive organs. *Development* 124: 4481-4491.

Sethaphong, L., Haigler, C.H., Kubicki, J.D., Zimmer, J., Bonetta, D., DeBolt, S., et al. (2013) Tertiary model of a plant cellulose synthase. *Proceedings of the National Academy of Sciences of the United States of America* 110: 7512-7517.

Shimada, T.L., Shimada, T. and Hara-Nishimura, I. (2010) A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of *Arabidopsis* thaliana. *The Plant journal : for cell and molecular biology* 61: 519-528.

Shimizu-Sato, S. and Mori, H. (2001) Control of outgrowth and dormancy in axillary buds. *Plant Physiol* 127: 1405–1413.

Simon, S. and Petrášek, J. (2011) Why plants need more than one type of auxin. *Plant Science* 180: 454-460.

Slabaugh, E., Davis, J.K., Haigler, C.H., Yingling, Y.G. and Zimmer, J. (2014) Cellulose synthases: new insights from crystallography and modeling. *Trends in plant science* 19: 99-106.

Soler M, Camargo ELO, Carocha V, Cassan-Wang H, San Clemente H, Savelli B, Hefer CA, Paiva JP; Myburg AA, Grima-Pettenati J. (2014). The *Eucalyptus* grandis R2R3-MYB transcription factor family: evidence for woody growth related evolution and function. *New Phytol* (*in press*).

Song, D., Shen, J. and Li, L. (2010) Characterization of cellulose synthase complexes in *Populus* xylem differentiation. *The New phytologist* 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 reports* 32: 885-898.

Suer, S., Agusti, J., Sanchez, P., Schwarz, M. and Greb, T. (2011) WOX4 imparts auxin responsiveness to cambium cells in *Arabidopsis*. *The Plant cell* 23: 3247-3259.

Sundberg, B., Uggla, C. and Tuominen, H. (2000) Cambial growth and auxin gradients. *Exptl Biol Rev*: 169.

Sundberg, E. Ostergaard, L. (2009) Distinct and dynamic auxin activities during reproductive development. *Cold Spring Harb Perspect Biol* 1: a001628.

Szemenyei, H., Hannon, M. and Long, J.A. (2008) TOPLESS mediates auxin-dependent transcriptional repression during *Arabidopsis* embryogenesis. *Science* 319: 1384-1386.

Szyjanowicz, P.M., McKinnon, I., Taylor, N.G., Gardiner, J., Jarvis, M.C. and Turner, S.R. (2004) The irregular xylem 2 mutant is an allele of korrigan that affects the secondary cell wall of *Arabidopsis* thaliana. *The Plant journal : for cell and molecular biology* 37: 730-740.

Tanaka, M., Takei, K., Kojima, M., Sakakibara, H., Mori, H. (2006) Auxin controls local cytokinin biosynthesis in the nodal stem in apical dominance. *Plant Journal* 45: 1028-1036.

Tatematsu, K., Kumagai, S., Muto, H., Sato, A., Watahiki, M.K., Harper, R.M., et al. (2004) MASSUGU2 encodes Aux/IAA19, an auxin-regulated protein that functions together with the transcriptional activator NPH4/ARF7 to regulate differential growth responses of hypocotyl and formation of lateral roots in *Arabidopsis* thaliana. *The Plant cell* 16: 379-393.

Taylor, N.G., Howells, R.M., Huttly, A.K., Vickers, K. and Turner, S.R. (2003) Interactions among three distinct CesA proteins essential for cellulose synthesis. *Proceedings of the National Academy of Sciences of the United States of America* 100: 1450-1455.

Taylor, N.G., Laurie, S. and Turner, S.R. (2000) Multiple cellulose synthase catalytic subunits are required for cellulose synthesis in *Arabidopsis*. *The Plant cell* 12: 2529-2540.

Theologis, A., Huynh, T.V. and Davis, R.W. (1985) Rapid induction of specific mRNAs by auxin in pea epicotyl tissue. *Journal of molecular biology* 183: 53-68.

Tian, Q., Nagpal, P. and Reed, J.W. (2003) Regulation of *Arabidopsis* SHY2/IAA3 protein turnover. *The Plant journal : for cell and molecular biology* 36: 643-651.

Tian, Q. and Reed, J.W. (1999) Control of auxin-regulated root development by the *Arabidopsis* thaliana SHY2/IAA3 gene. *Development* 126: 711-721.

Timpte, C. (2001) Auxin binding protein: curiouser and curiouser. *Trends in plant science* 6: 586-590.

Tiwari, S.B., Hagen, G. and Guilfoyle, T. (2003) The roles of auxin response factor domains in auxin-responsive transcription. *The Plant cell* 15: 533-543.

Tiwari, S.B., Hagen, G. and Guilfoyle, T.J. (2004) Aux/IAA proteins contain a potent transcriptional repression domain. *The Plant cell* 16: 533-543.

Tiwari, S.B., Wang, X.J., Hagen, G. and Guilfoyle, T.J. (2001) AUX/IAA proteins are active repressors, and their stability and activity are modulated by auxin. *The Plant cell* 13: 2809-2822.

Tromas A., Paque S., Stierlé V., Quettier A.L., Muller P., Lechner E., et al. (2013) Auxin-binding protein 1 is a negative regulator of the SCF(TIR1/AFB) pathway. *Nat Commun* 4: 2496.

Uggla, C., Magel, E., Moritz, T. and Sundberg, B. (2001) Function and dynamics of auxin

and carbohydrates during earlywood/latewood transition in scots pine. *Plant physiology* 125: 2029-2039.

Uggla, C., Mellerowicz, E.J. and Sundberg, B. (1998) Indole-3-acetic acid controls cambial growth in scots pine by positional signaling. *Plant physiology* 117: 113-121.

Uggla, C., Moritz, T., Sandberg, G. and Sundberg, B. (1996) Auxin as a positional signal in pattern formation in plants. *Proc Natl Acad Sci U S A* 93: 9282-9286.

Ulmasov, T., Murfett, J., Hagen, G. and Guilfoyle, T.J. (1997) Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. *Plant Cell* 9: 1963-1971.

Van Acker, R., Leple, J.C., Aerts, D., Storme, V., Goeminne, G., Ivens, B., et al. (2014) Improved saccharification and ethanol yield from field-grown transgenic poplar deficient in cinnamoyl-CoA reductase. *Proceedings of the National Academy of Sciences of the United States of America* 111: 845-850.

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.

Vanholme, R., Demedts, B., Morreel, K., Ralph, J. and Boerjan, W. (2010) Lignin biosynthesis and structure. *Plant physiology* 153: 895-905.

Vanneste, S., De Rybel, B., Beemster, G. T., Ljung, K., De Smet, I. et al. (2005) Cell cycle progression in the pericycle is not sufficient for SOLITARY ROOT/IAA14-mediated lateral root initiation in Arabidopsis thaliana. Plant Cell 17: 3035-3050.

Vernoux, T., Brunoud, G., Farcot, E., Morin, V., Van den Daele, H., L*Egr*and, J., et al. (2011) The auxin signalling network translates dynamic input into robust patterning at the shoot apex. *Molecular systems biology* 7: 508.

Vogt, T. (2010) Phenylpropanoid biosynthesis. *Molecular plant* 3: 2-20.

Walker, J.C. and Key, J.L. (1982) Isolation of cloned cDNAs to auxin-responsive poly(A)RNAs of elongating soybean hypocotyl. *Proceedings of the National Academy of Sciences of the United States of America* 79: 7185-7189.

Wang, D., Pei, K., Fu, Y., Sun, Z., Li, S., Liu, H., et al. (2007) Genome-wide analysis of the auxin response factors (ARF) gene family in rice (Oryza sativa). *Gene* 394: 13-24.

Wang, H., Jones, B., Li, Z., Frasse, P., Delalande, C., Regad, F., et al. (2005a) The tomato Aux/IAA transcription factor IAA9 is involved in fruit development and leaf morphogenesis. *The Plant cell* 17: 2676-2692.

Wang, J.W., Wang, L.J., Mao, Y.B., Cai, W.J., Xue, H.W. and Chen, X.Y. (2005b) Control of root cap formation by MicroRNA-targeted auxin response factors in *Arabidopsis*. *Plant Cell* 17: 2204-2216.

Wang, Y., Chantreau, M., Sibout, R. and Hawkins, S. (2013) Plant cell wall lignification and monolignol metabolism. *Frontiers in plant science* 4: 220.

Wang, Y., Deng, D., Bian, Y., Lv, Y. and Xie, Q. (2010) Genome-wide analysis of primary auxin-responsive Aux/IAA gene family in maize (Zea mays. L.). *Molecular biology reports* 37: 3991-4001.

Watahiki, M.K. and Yamamoto, K.T. (1997) The massugu1 mutation of *Arabidopsis* identified with failure of auxin-induced growth curvature of hypocotyl confers auxin insensitivity to hypocotyl and leaf. *Plant physiology* 115: 419-426.

Weijers, D., Benkova, E., Jager, K.E., Schlereth, A., Hamann, T., Kientz, M., et al. (2005) Developmental specificity of auxin response by pairs of ARF and Aux/IAA transcriptional regulators. *The EMBO journal* 24: 1874-1885.

Weng, J.K. and Chapple, C. (2010) The origin and evolution of lignin biosynthesis. *The New phytologist* 187: 273-285.

Wenzel, C. L., Schuetz, M., Yu, Q., Mattsson, J. (2007) Dynamics of MONOPTEROS and PIN-FORMED1 expression during leaf vein pattern formation in Arabidopsis thaliana. *Plant Journal* 49: 387-398.

Wilmoth JC, Wang S, Tiwari SB, Joshi AD, Hagen G. (2005) NPH4/ARF7 and ARF19 promote leaf expansion and auxin-induced lateral root formation. *Plant Journal* 43: 118-30.

Woo, E.J., Marshall, J., Bauly, J., Chen, J.G., Venis, M., Napier, R.M., et al. (2002) Crystal structure of auxin-binding protein 1 in complex with auxin. *The EMBO journal* 21: 2877-2885.

Worley, B., Halouska, S. and Powers, R. (2013) Utilities for quantifying separation in PCA/PLS-DA scores plots. *Analytical biochemistry* 433: 102-104.

Worley, C.K., Zenser, N., Ramos, J., Rouse, D., Leyser, O., Theologis, A., et al. (2000) degradation of Aux/IAA proteins is essential for normal auxin signalling. *The Plant journal : for cell and molecular biology* 21: 553-562.

Wu, A.M., Rihouey, C., Seveno, M., Hornblad, E., Singh, S.K., Matsunaga, T., et al. (2009) The *Arabidopsis* IRX10 and IRX10-LIKE glycosyltransferases are critical for glucuronoxylan biosynthesis during secondary cell wall formation. *The Plant journal : for cell and molecular biology* 57: 718-731.

Wu, M.F., Tian, Q. and Reed, J.W. (2006) *Arabidopsis* microRNA167 controls patterns of ARF6 and ARF8 expression, and regulates both female and male reproduction. *Development* 133: 4211-4218.

Xing, H., Pudake, R.N., Guo, G., Xing, G., Hu, Z., Zhang, Y., et al. (2011) Genome-wide identification and expression profiling of auxin response factor (ARF) gene family in maize. *Bmc Genomics* 12: 178.

Xu, T., Dai, N., Chen, J., Nagawa, S., Cao, M., Li, H., et al. (2014) Cell surface ABP1-

TMK auxin-sensing complex activates ROP GTPase signaling. *Science* 343: 1025-1028.

Yang, H. and Murphy, A.S. (2009) Functional expression and characterization of *Arabidopsis* ABCB, AUX 1 and PIN auxin transporters in Schizosaccharomyces pombe. *The Plant journal : for cell and molecular biology* 59: 179-191.

Yang, X.Q., Lee, S., So, J.H., Dharmasiri, S., Dharmasiri, N., Ge, L., et al. (2004) The IAA1 protein is encoded by AXR5 and is a substrate of SCFTIR1. *Plant Journal* 40: 772-782.

Zazimalova, E., Murphy, A.S., Yang, H., Hoyerova, K. and Hosek, P. (2010) Auxin transporters--why so many? *Cold Spring Harbor perspectives in biology* 2: a001552.

Zhang, J., Elo, A. and Helariutta, Y. (2011) *Arabidopsis* as a model for wood formation. *Curr Opin Biotechnol* 22: 293-299.

Zhang, J., Nieminen, K., Serra, J. A., Helariutta, Y. (2014a) The formation of wood and its control. *Curr Opin Plant Biol* 17: 56-63.

Zhang, Y., Paschold, A., Marcon, C., Liu, S., Tai, H. et al. (2014b) he Aux/IAA gene rum1 involved in seminal and lateral root formation controls vascular patterning in maize (Zea mays L.) primary roots. *J Exp Bot* 13: eru249.

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: 2763-2782.

Zhong, R., McCarthy, R.L., Haghighat, 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: e69219.

Zhong, R., Richardson, E.A. and Ye, Z.H. (2007) Two NAC domain transcription factors, SND1 and NST1, function redundantly in regulation of secondary wall synthesis in fibers of *Arabidopsis*. *Planta* 225: 1603-1611.

Zhong R, Ye Z-H. (1999) IFL1, a gene regulating interfascicular fiber differentiation in Arabidopsis, encodes a homeodomain-leucine zipper protein. *Plant Cell* 11:2139-2152.

Zhong, R. and Ye, Z.H. (2010) The poplar PtrWNDs are transcriptional activators of secondary cell wall biosynthesis. *Plant Signal Behav* 5: 469-472.

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.Q. and Ye, Z.H. (2001) Alteration of auxin polar transport in the *Arabidopsis* ifl1 mutants. *Plant physiology* 126: 549-563.

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*. *The Plant cell* 21: 248-266.

Zobel, B.J. and J.R. Sprague. 1998. Juvenile wood in forest trees. Springer-Verlag, Berlin, Germany.

Zouine M, Fu Y, Chateigner-Boutin AL, Mila I, Frasse P, et al. (2014) Characterization of the tomato ARF gene family uncovers a multi-levels post-transcriptional regulation including alternative splicing. PloS one 9: e84203.

Annexes

Annexes

Publications

Supplementary data for chapter II

Table SII-7. Small RNAs target site prediction in *EgrARF* genes.

Supplementary data for chapter III

Supplementary Fig. SIII-1 Procedure used for identifying *Aux/IAA* genes in *Eucalyptus grandis*.

Supplementary Fig. SIII-2 Phylogenetic analysis of *E. grandis* and *Arabidopsis* IAA proteins and the exon–intron organization of corresponding genes.

Supplementary Fig. SIII-3 Sequencing alignment of *Egr*IAA29 and two putatively pseudogenes, *Egr*IAA29A and 29B.

Supplementary Fig. SIII-4 Multiple sequence alignment of predicted amino acid sequences of *Egr*IAA proteins.

Supplementary Fig. SIII-5 Chromosomal positions of the *EgrIAA* genes.

Supplementary Fig. SIII-6 Predicted proteins of all the transcripts of *EgrIAA* and their conserved domains.

Supplementary Fig. SIII-7 Venn diagram of overall strategy to identify the best potential candidate(s) involved in wood formation.

Supplementary Fig. SIII-8 Protein sequence of *EgrIAA4* and the phenotype of gain-of-function transgenic *Arabidopsis*.

Supplementary Table SIII-1. Protein identity matrix among *Egr*IAA.

Table SIII-2. Alternative transcript comparison between *Eucalyptus* and *Arabidopsis.*

Table SIII-3. Bibliography of the genes which are potentially involved in wood formation in cluster III from the phylogenetic tree.

Supplementary Table SIII-4. The Primers for *EgrARF* genes and reference genes used in qRT-PCR experiments.

Supplementary data for chapter IV

Fig. SIV-1 Maps and features of pENTR/D-TOPO which can cloning blunt-end PCR products into the empty used for the gateway system.

Publications

Articles:

1. **Yu H**, Soler M, Mila I, San Clemente H, Dunand C, Paiva J, Myburg 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.
2. Li Q*, **Yu H***, Cao PB*, Fawal N*, Mathé C, Azar S, Wang H, Grima-Pettenati J, Marque C, Teulières C, Dunand C (2014) Explosive tandem and segmental duplications of multigenic families in *Eucalyptus grandis*. (Co-first author, submitted in Genome Biology and Evolution)
3. **Yu H**, Soler M, San Clemente H, Mila I, Myburg AA, Bouzayen M, Grima-Pettenati J and Cassan-Wang H (2014) Comprehensive Genome-wide Analysis of *Aux/IAA* Gene Family in *Eucalyptus*: Evidence for *EgrIAA4*'s Role in Wood Formation. (In preparation)
4. Cassan-Wang, H, Soler M, **Yu H**, Camargo EL, Carocha V, Ladouce N, Savelli B, Paiva JA, Leple JC and Grima-Pettenati J. 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. 2012, 53: 2101-2116.
5. Wang H, Soler M, **Yu H**, Camargo E, San Clemente H, Savelli B, Ladouce N, Paiva J, Grima-Pettenati J (2011) Master regulators of wood formation in *Eucalyptus*. BMC Proceedings. 5: P110.

Poster presentation:

1. Wang H, Soler M, **Yu H**, Plasencia A, San Clemente H, Ladouce N, Hefer C, Myburg A, Paiva JAP, Grima-Pettenati J.(2013) New transcription factors regulating lignified secondary cell wall formation in Eucalyptus. XIII Cell Wall Meeting, Nantes, July 7-12,

2013.

2. Soler M, **Yu H**, Cassan-Wang H, Grima-Pettenati J (2013) New transcription factors regulating secondary xylem differentiation In Eucalyptus. Third International Conference on Plant Vascular Biology Helsinki (FI), July 26-30, 2013.
3. Wang H, Soler M, Saïdi N, Hussey S, **Yu H**, Carocha V, Camargo E, Ladouce N, San Clemente H, Hefer C, Savelli B, Myburg Z, Paiva J, Grima-Pettenati J (2012) Master Regulators of Wood Formation in Eucalyptus. International Plant and Animal genome XX, San Diego (US-CA) January, 2012.
4. Wang H, **Yu H**, Soler M, San Clemente H, Camargo E, Saïdi N, Dunand C, Ladouce N, Paiva J, Grima-Pettenati J (2011) Auxin Signaling Transcription Factors in Eucalyptus Wood Formation. 9th SFBV meeting, Clermont-Ferrand, December 12-14, 2011.
5. Soler M, Camargo E, San Clemente H, Savelli B, Ladouce N, Bensussan M, **Yu H**, Hefer CA, Myburg AA, Paiva J, Wang H, Grima-Pettenati J (2011) R2R3-MYB transcription factors regulating wood formation in Eucalyptus. 9th SFBV meeting, Clermont-Ferrand, December 12-14, 2011.
6. Soler M, Camargo E, **Yu H**, San Clemente H, Savelli B, Ladouce N, Paiva J, Wang H, Grima-Pettenati J (2011) R2R3 MYB transcription factors controlling of wood formation in Eucalyptus. IUFRO 2011 Eucalyptus: Joining silvicultural and genetic strategies to minimize Eucalyptus environmental stresses, from research to practice, Porto Seguro (BR), November 14-18, 2011.
7. Carocha VJ, Soler M, Saidi M, **Yu H**, Camargo E, Graça C, Pinto J, San Clemente H, Ladouce N, Fevereiro P, Neves L, Araújo C, Wang H, Paiva J, Grima-Pettenati J (2011) Identification, phylogeny and tissue-specific expression of some members of multigene families involved in Eucalyptus lignin biosynthesis. IUFRO 2011 Eucalyptus: Joining silvicultural and genetic strategies to minimize Eucalyptus environmental stresses, from research to practice. Porto Seguro (BR), November 14-18, 2011.
8. Soler M, Camargo E, **Yu H**, San Clemente H, Savelli B, Ladouce N, Paiva J, Wang H, Grima-Pettenati J (2011) R2R3 MYB transcription factors controlling of wood formation

in Eucalyptus, IUFRO Tree Biotechnology, Arrial d'Ajuda (BR), June 26-July 2, 2011.

9. **Yu H**, San Clemente H, Dunand C, Soler M, Camargo E, Ladouce N, Paiva J, Grima-Pettenati J and Wang H. *Eucalyptus* Auxin Responses Mediators in Wood Formation. Poster, Nancy, France, 2011-05.
10. Soler M, Camargo E L O, San Clemente H, Savelli B, Ladouce N, Bensussan M, **Yu H**, Paiva J, Wang H, Grima-Pettenati J. Eucalyptus MYB transcription factors regulating wood formation. Poster, Nancy, France, 2011-05.

Supplementary data for chapter II

Figure SII-1. Procedure used for identifying *ARF* genes in *Eucalyptus grandis.* *Arabidopsis* ARF protein sequences were used to search their orthologs in the predicted *Eucalyptus* proteome by using in BLASTP. Sixty-four *Eucalyptus* proteins identified in this initial search were further examined by manual curation using protein motif scanning and the FgeneSH program to complete partial sequences. Redundant and invalid genes were eliminated based on gene structure, integrity of conserved motifs and EST support. Manual curation resulted in 17 complete Eucalyptus ARF protein sequences. These 17 protein sequences were used in two subsequent additional searches: first, a BLASTP search against the *Eucalyptus* proteome to identify exhaustively all divergent *Eucalyptus* ARF gene family members and, second, tBLASTn searches against the *Eucalyptus* genome for any possible unpredicted genes. To confirm our findings, we used poplar ARF proteins and repeated the complete search procedure described above and obtained identical results.

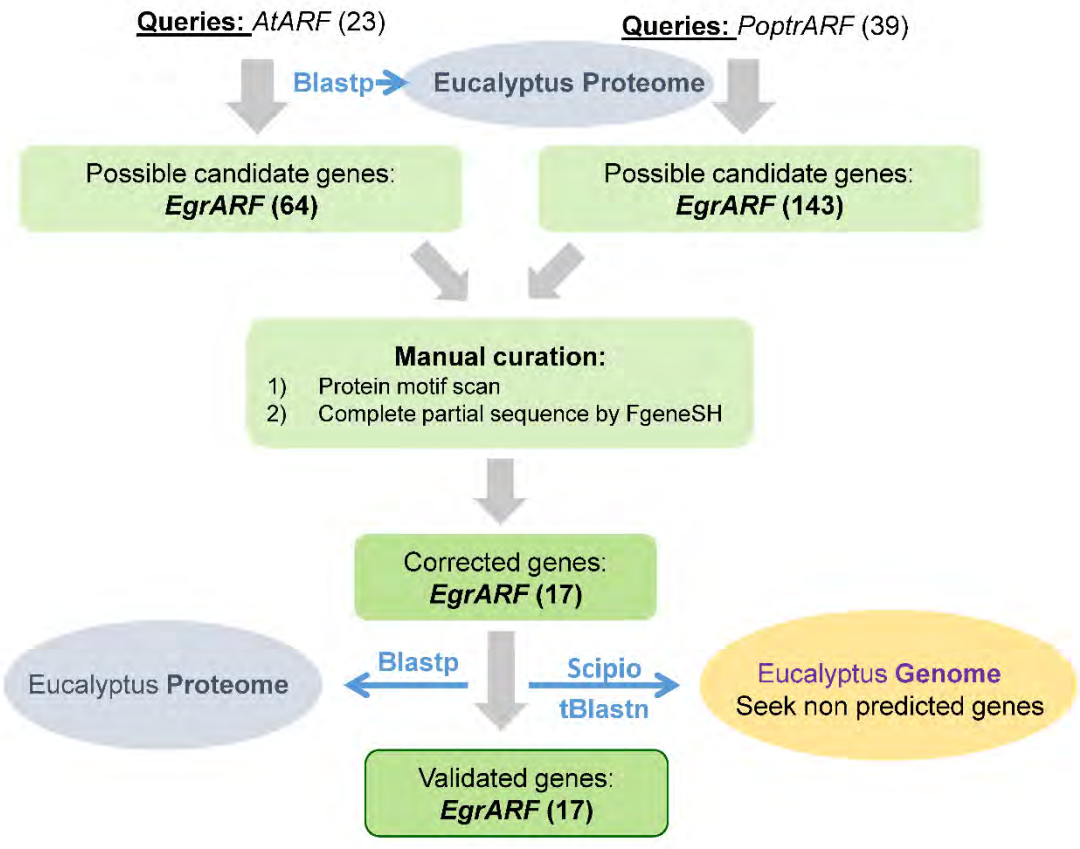


Figure SII-2. Locations of the 17 *EgrARF* genes on the 11 *Eucalyptus grandis* chromosomes.

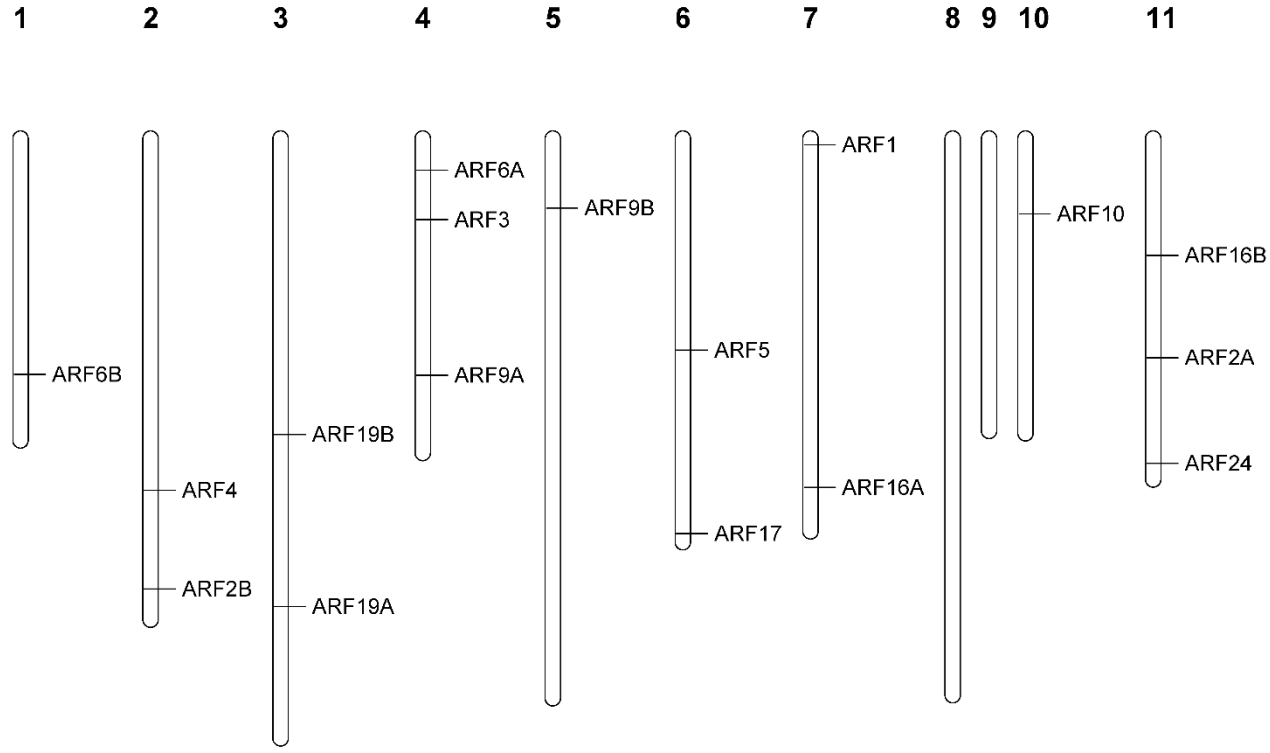
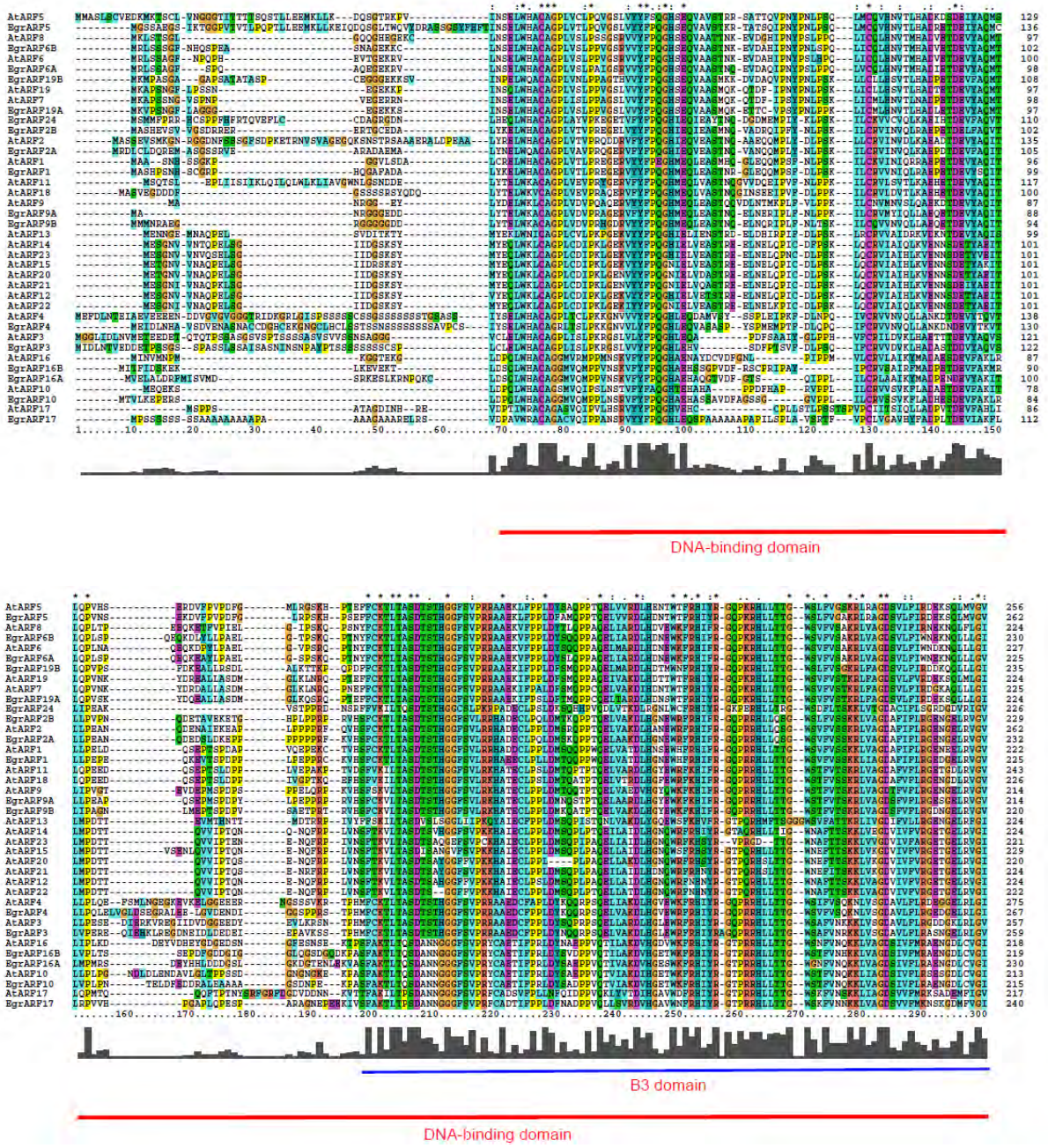
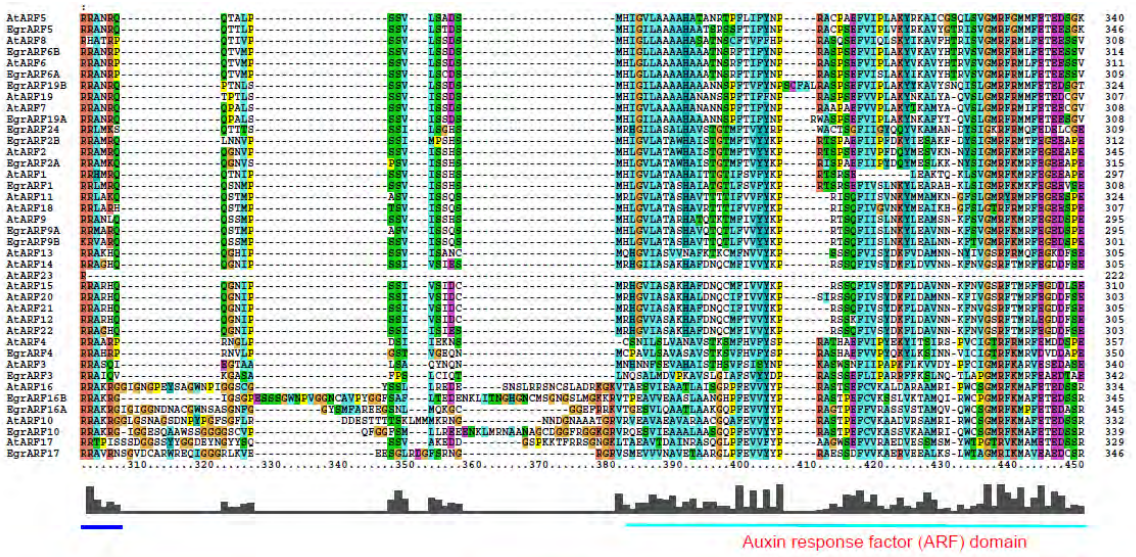
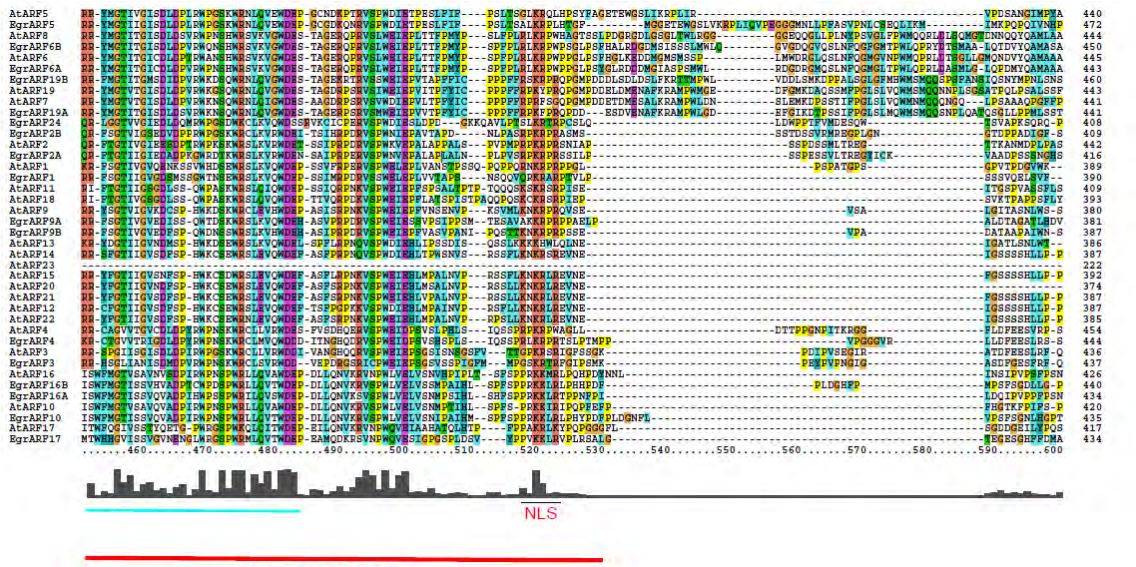
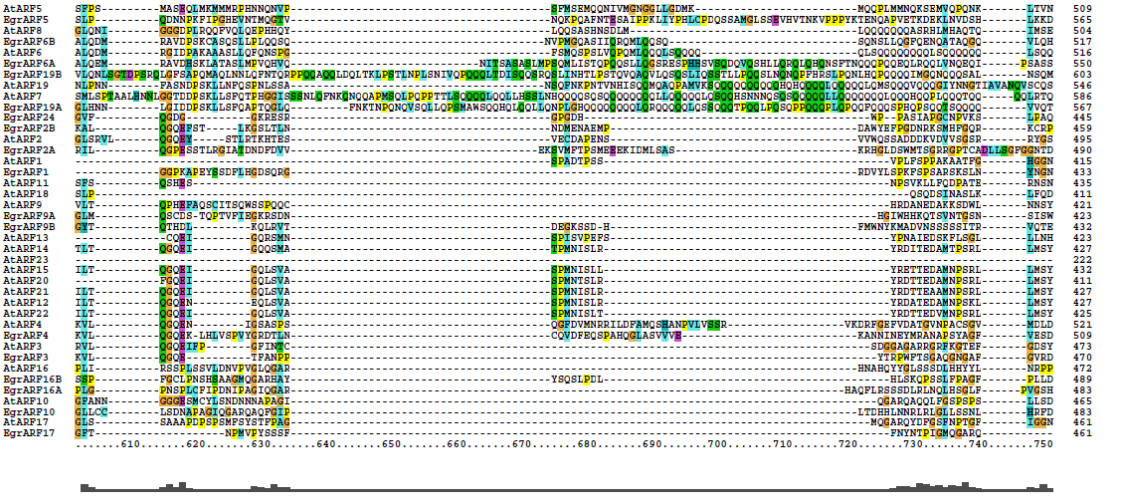


Figure SII-3. Multiple sequence alignment of predicted amino acid sequences of *EgrARF* and *AtARF* proteins. The multiple sequence alignment was obtained with the MUSCLE software. The highly conserved domains and nuclear localization signals (NLSs) proteins were noted on the bottom of the alignment with different colours.

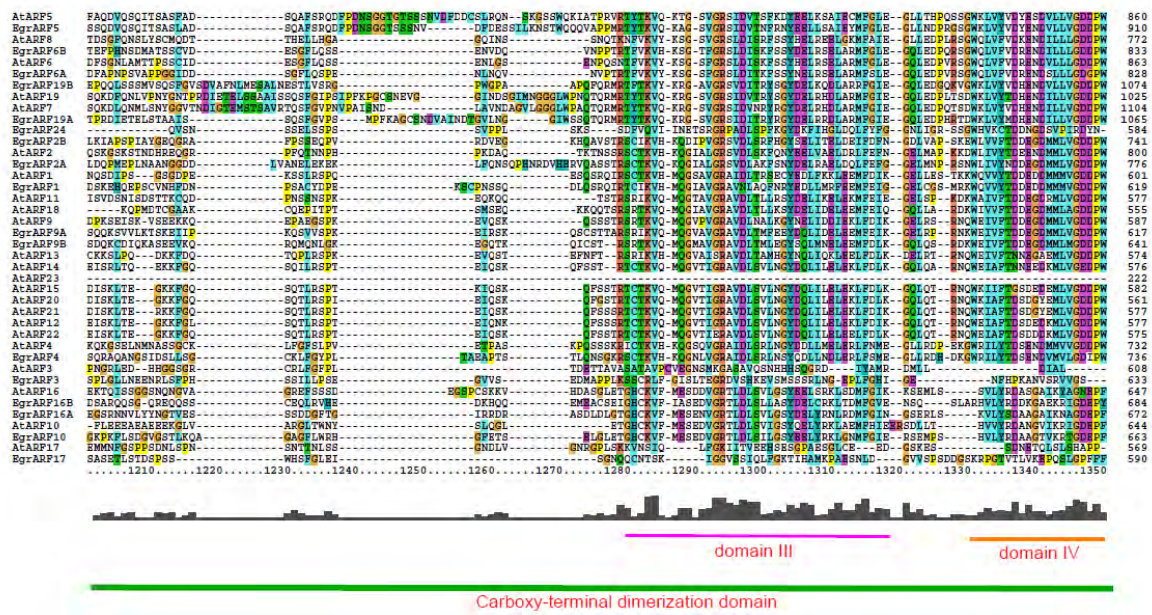



Auxin response factor (ARF) domain
DNA-binding domain


NLS
DNA-binding domain




Carboxy-terminal dimerization domain


domain III
domain IV
Carboxy-terminal dimerization domain

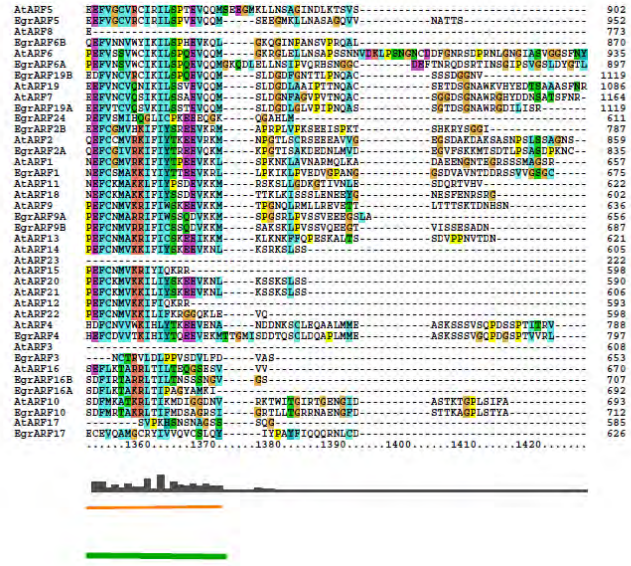

AtARF5
EgrARF5
AtARF8
EgrARF6B
AtARF6
EgrARF6A
EgrARF19B
AtARF19
AtARF7
EgrARF19A
EgrARP24
EgrARP2B
AtARF2
EgrARP2A
AtARF1
EgrARF1
AtARF11
AtARF18
AtARF9
EgrARF9A
EgrARF9B
AtARF13
AtARF14
AtARF23
AtARF15
AtARF20
AtARF21
AtARF12
AtARF22
AtARF4
EgrARF4
AtARF3
EgrARF3
AtARF16
EgrARF16B
EgrARF16A
AtARF10
EgrARF10
AtARF17
EgrARF17
1360
1370
1380
1390
1400
1410
1420

Figure SII-4. Comparative analysis of predicted *ARF* alternative variants between *Eucalyptus grandis* and *Arabidopsis thaliana*. The alternative spliced protein sequences were extracted from Phytozome except for *AtARF4 (*obtained from Finet *et al*. (2013), the motif structures were predicted by Pfam (http://pfam.xfam.org/).

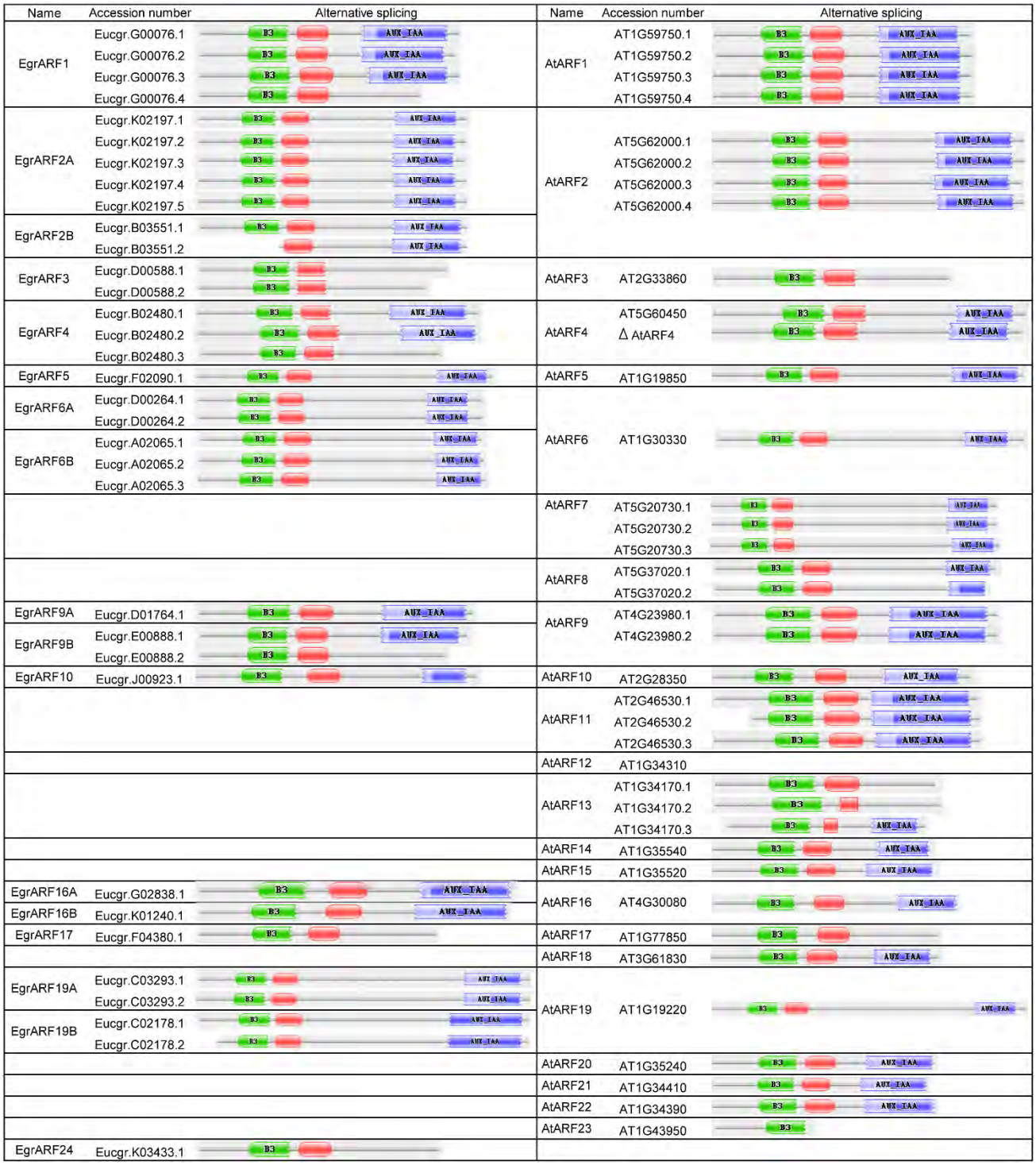


Figure SII-5. Structure of the *ARF* alternative transcripts in *E. globulus.* The *E. globulus* alternative transcripts were obtained from a compendium of RNASeq data. The material and methods are described in Table S1. The illumina reads sequences are provided in supplementary File S1 in the FastQ format.

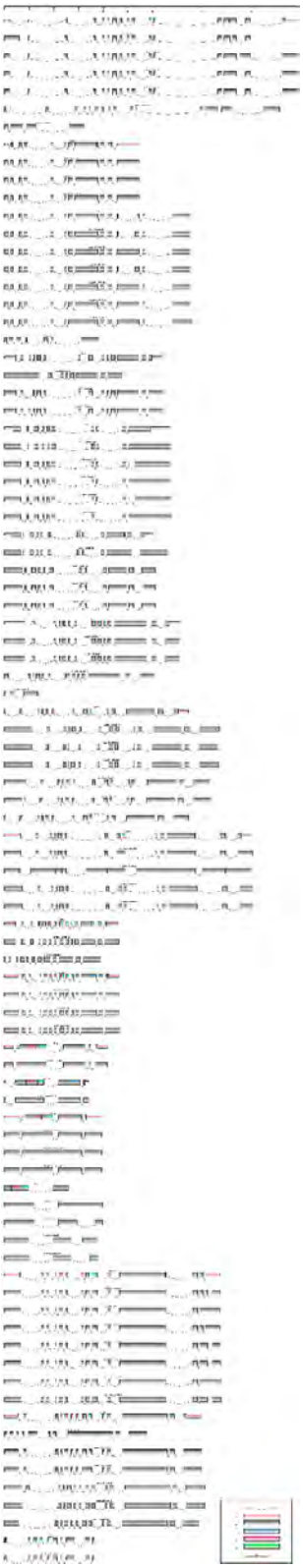


Figure SII-6. Comparative Phylogenetic relationships between ARF proteins from poplar, Eucalyptus, grapevine, Arabidopsis, tomato and rice. Full-length protein sequences were aligned using the Clustal_X program. The phylogenetic tree was constructed by using the MEGA5 program and the neighbour-joining method with predicted full-length ARF proteins. Bootstrap supports are indicated at each node.

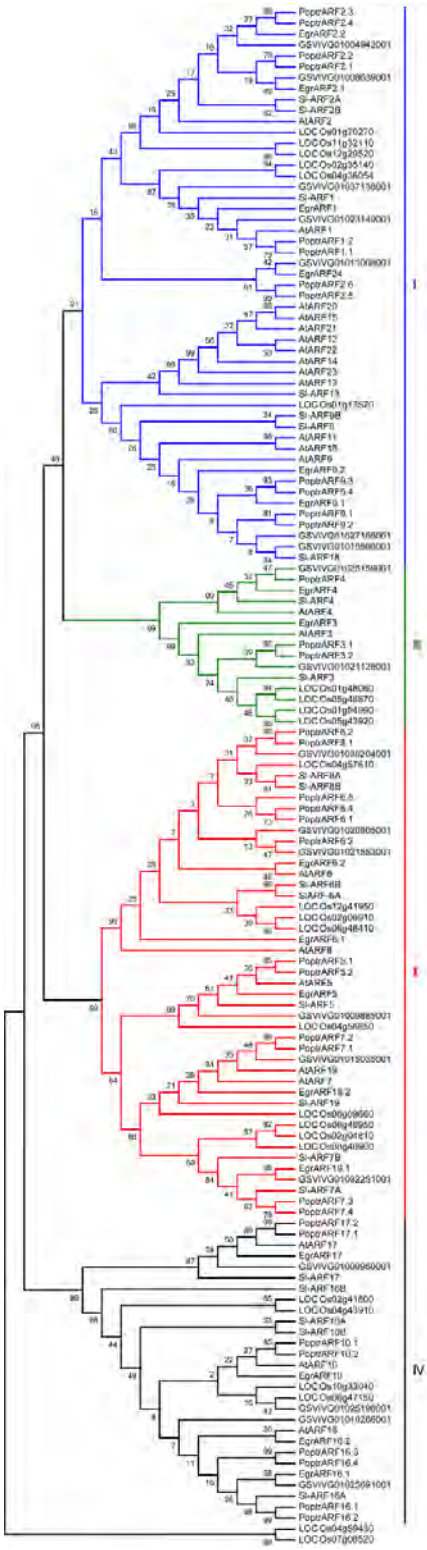


Figure SII-7. Figure SII-7. **Predicted stem-loop structures of three *EgrmiR160* and three *EgrmiR167.*** The part of the stem-loop from which the mature microRNA derives is highlighted in yellow.

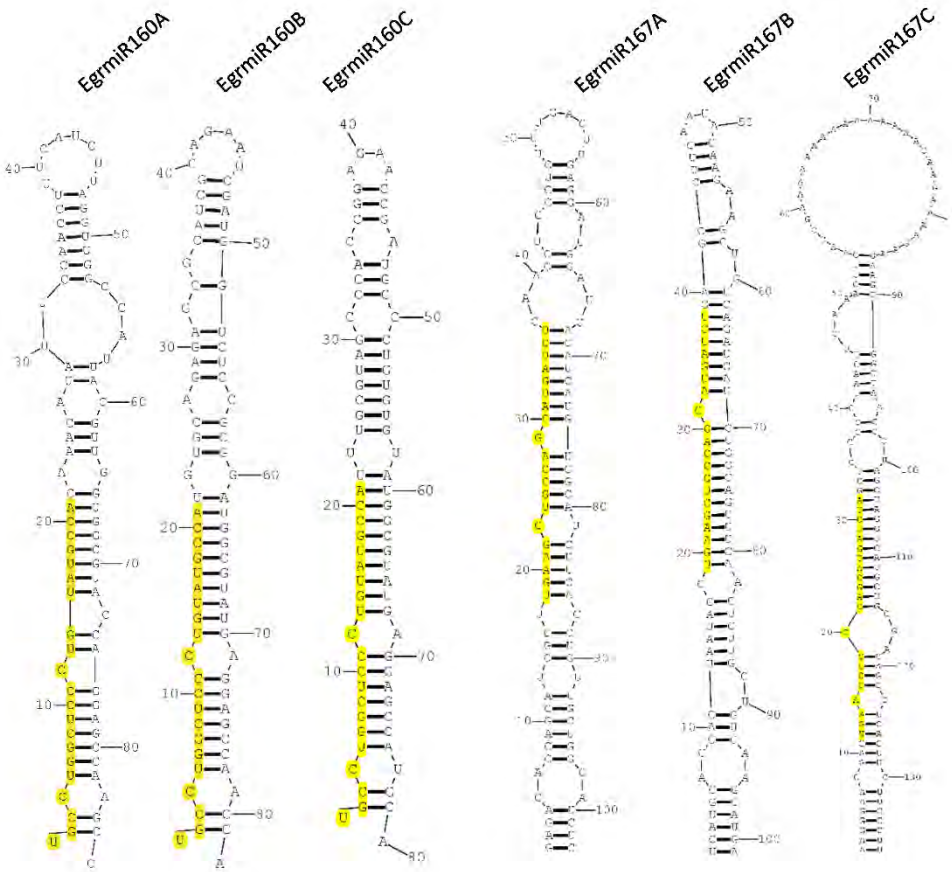


Figure SII-8. Expression profiles of *EgrARF5* and *EgrARF24* in various organs and tissues. Relative mRNA abundance of *EgrARF5* and *EgrARF24* was compared to expression in the control sample of mature leaves and *in vitro* plantlets, respectively. Error bars indicate the SE of mean expression values from three independent experiments.

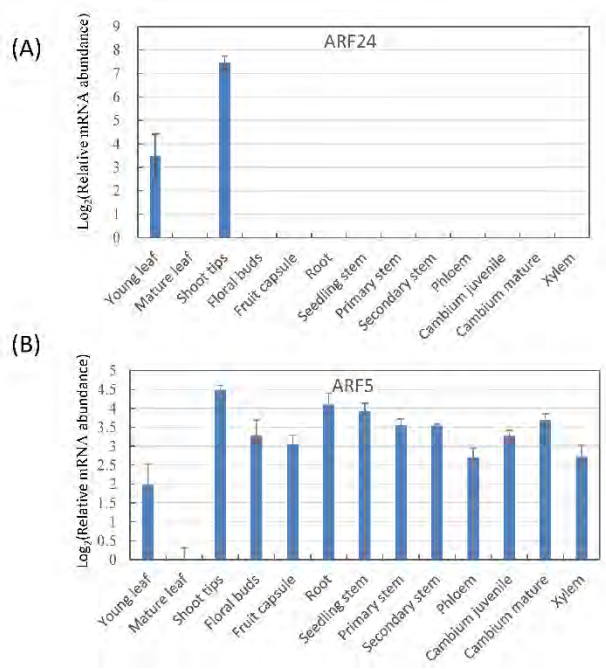


Figure SII-9. Expression profiles of *EgrARF* genes in tissues involved in secondary growth. Relative mRNA abundance was compared to expression in the control sample (*in vitro* plantlets).

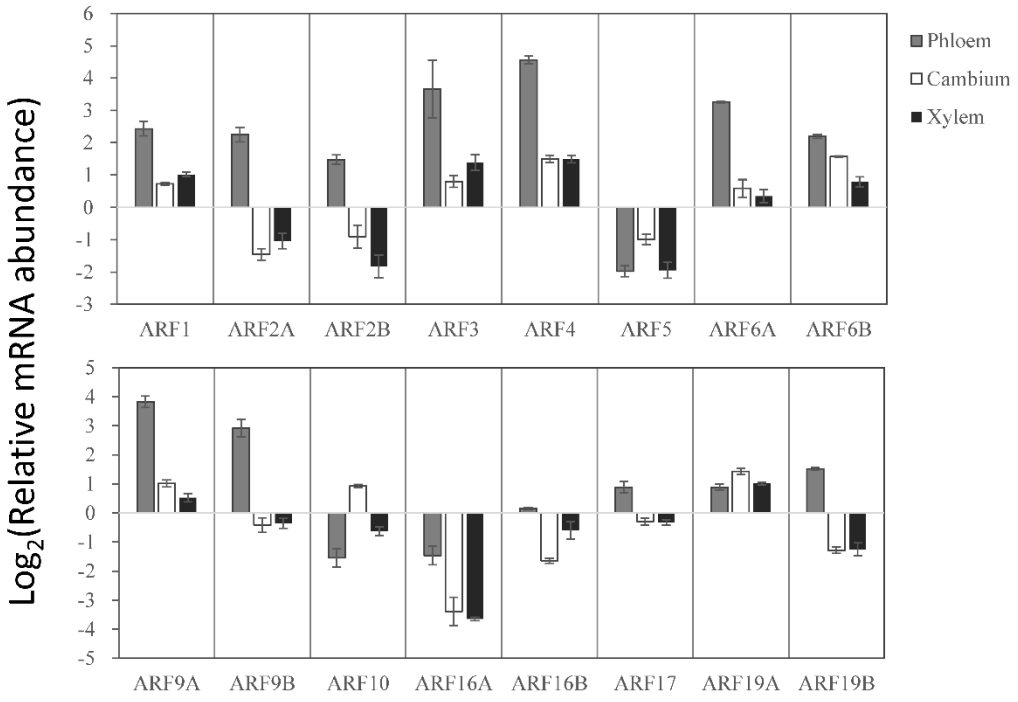


Figure SII-10. Young *Eucalyptus grandis* trees phenotypes in response to various long-term hormonal treatments. 10 µM NAA, or 20 µM gibberellic acid or 100 µM ACC were added to the medium of 65-d-old young tree, and phenotypes were observed 14 days later.

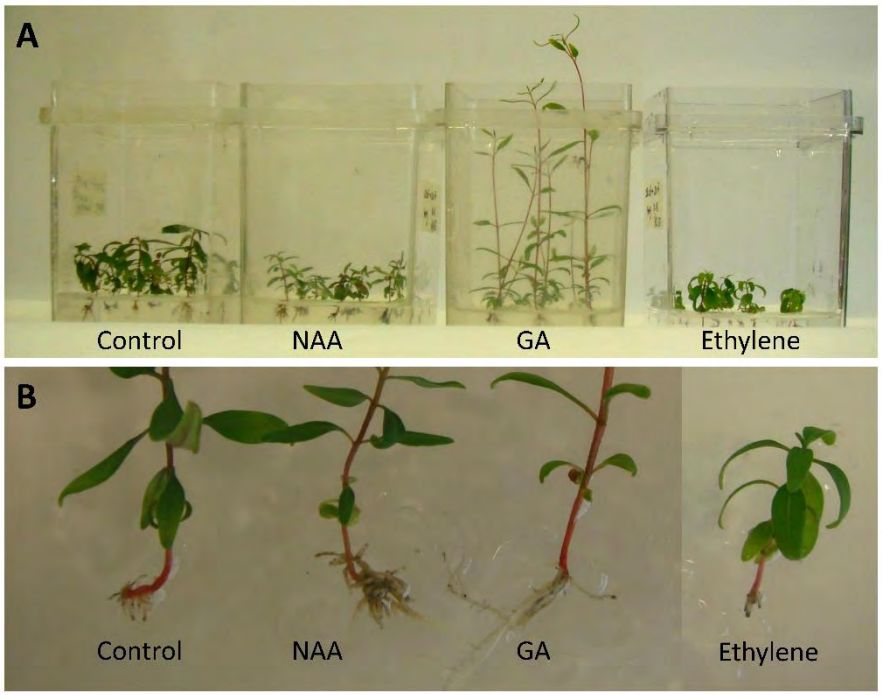


Table SII-1. Primers for *EgrARF* genes and reference genes used in qRT-PCR experiments

Name	Accession No.	Forward 5'-3'	Reverse 5'-3'	Efficiency
ARF1	Eucgr.G00076.1	TTTAAGAGCGCTGTGGTTCTCTG	AAGCATCAGCAAACGCACCTTG	1.90
ARF2A	Eucgr.K02197.1	GGCATGCCATTTCTACAGGAACC	TCAGCAGGGCTTATCCTAGGTTTG	1.97
ARF2B	Eucgr.B03551.1	ATTGAGGGCTGAACCAGAGACC	TGGTTAGGCACCGGAAGCAAAG	1.95
ARF3	Eucgr.D00588.1	ACATATATACCGGGCAGGACAGC	TCTTCCGGTTTACAAATGCACTCC	1.94
ARF4	Eucgr.B02480.1	AGCTGTCGCTAGTGCTGTATCC	AACTCTGCATGGCTTGCCCTTG	1.96
ARF5	Eucgr.F02090.1	GCACATGGCAACAGCAAGTAGC	ATCGACCTTCCGACTGATCCTG	1.97
ARF6A	Eucgr.D00264.1	TGAGCTTGCTCGCATGTTTAGCC	TGTTCACAAACTCCGGCCAAGG	1.89
ARF6B	Eucgr.A02065.1	GCTGGCAGCTTGTATTTGTAGACC	CCACACATTGTTGACGAACTCCTG	1.88
ARF9A	Eucgr.D01764.1	TCCAAAGCAGTCTGTGGTTTCACC	ATGCGACTCCTTGCAGTGGTAG	1.87
ARF9B	Eucgr.E00888.1	GTGACTCGTTCGTGTTCTTAAGGG	TGACGAGCAACTCGTTTCACTCC	2.00
ARF10	Eucgr.J00923.1	TAGCAGGGCGTGTCGTGTTATC	AATCCGGACGCTTGCGTTTCTC	1.90
ARF16A	Eucgr.G02838.1	AGGGACATGTTCGGCATAAACGG	AATAGCACCGGCAGCATCTGAG	1.84
ARF16B	Eucgr.K01240.1	CTCAGAAATCGGCCACTGCAAAG	TGAGGCAAGCAACGAGAGATCG	1.97
ARF17	Eucgr.F04380.1	ATGGAGGTGGTGGTGAATGCAG	AAACCACCTCGAACGGCAATCC	1.85
ARF19A	Eucgr.C03293.1	TCAGTTTCAAGCGGATGAGAAGCC	AAGGACAGCCACTCTGGGTTAG	1.78
ARF19B	Eucgr.C02178.1	TCAAACAGGACCTGGCTCGTAG	CGAACACAGTTCACGAAGTCCTC	1.93
ARF24	Eucgr.K03433.1	TCCAGCGACTTTGTGCAGGTTATC	AGCCTTTGAATGGGCTGAGGTC	1.92
EF-1α	Eucgr.B02473.1	ATGCGTCAGACTGTGGCTGTTG	TTGGTCACCTTGGCTCCACTTG	1.84
SAND	Eucgr.B02502.1	TTGATCCACTTGCGGACAAGGC	TCACCCATTGACATACACGATTGC	2.11
PP2A1	Eucgr.B03386.1	TCGAGCTTTGGACCGCATACAAG	ACCACAAGAGGTCACACATTGGC	1.98
PP2A3 5' end	Eucgr.B03031.1	CGGAAGAACTGGGTGTGTTT	CACAGAGGGTCTCCAATGGT	2.03
PP2A3 3' end	Eucgr.B03031.1	CAGCGGCAAACAACTTGAAGCG	ATTATGTGCTGCATTGCCCAGTC	2.02
IDH 5' end	Eucgr.F02901.1	AATCGACCTGCTTCGACCCTTC	TCGACCTTGATCTTCTCGAAACCC	1.92
IDH 3' end	Eucgr.F02901.1	TGCTGTGGCAGCTGAACTCAAG	ATGTTGTCCGCCAGTCACCTAC	1.86

Table SII-2. Protein identity matrix between EgrARF and AtARF

	EgrARF1	EgrARF2A	EgrARF2B	EgrARF3	EgrARF4	EgrARF5	EgrARF6A	EgrARF6B	EgrARF9A	EgrARF9B	EgrARF10	EgrARF16A	EgrARF16B	EgrARF17	EgrARF19A	EgrARF19B	EgrARF24
AtARF1	0.622	0.388	0.399	0.239	0.31	0.244	0.251	0.257	0.471	0.446	0.226	0.222	0.216	0.194	0.203	0.208	0.309
AtARF2	0.39	0.593	0.495	0.23	0.297	0.226	0.249	0.246	0.361	0.345	0.193	0.185	0.181	0.164	0.206	0.203	0.268
AtARF3	0.244	0.22	0.242	0.462	0.326	0.192	0.196	0.205	0.247	0.23	0.195	0.197	0.187	0.206	0.153	0.157	0.23
AtARF4	0.298	0.28	0.295	0.31	0.578	0.226	0.231	0.247	0.282	0.286	0.198	0.193	0.195	0.168	0.201	0.202	0.24
AtARF5	0.242	0.247	0.247	0.195	0.224	0.593	0.342	0.349	0.235	0.232	0.181	0.176	0.181	0.156	0.311	0.302	0.2
AtARF6	0.243	0.247	0.244	0.18	0.22	0.337	0.691	0.656	0.238	0.235	0.181	0.168	0.18	0.143	0.369	0.349	0.197
AtARF7	0.196	0.216	0.208	0.147	0.186	0.293	0.32	0.332	0.187	0.183	0.148	0.145	0.147	0.115	0.59	0.415	0.159
AtARF8	0.263	0.235	0.254	0.197	0.236	0.334	0.518	0.541	0.266	0.251	0.191	0.189	0.194	0.166	0.329	0.317	0.214
AtARF9	0.435	0.347	0.367	0.224	0.273	0.232	0.247	0.25	0.586	0.584	0.236	0.218	0.213	0.195	0.208	0.193	0.3
AtARF10	0.221	0.196	0.206	0.184	0.192	0.165	0.183	0.188	0.235	0.229	0.601	0.477	0.493	0.292	0.154	0.158	0.205
AtARF11	0.449	0.356	0.357	0.224	0.271	0.235	0.253	0.25	0.522	0.488	0.238	0.229	0.227	0.204	0.198	0.202	0.341
AtARF12	0.347	0.285	0.298	0.214	0.243	0.193	0.208	0.212	0.397	0.369	0.206	0.214	0.209	0.191	0.169	0.172	0.312
AtARF13	0.344	0.271	0.286	0.2	0.236	0.189	0.204	0.207	0.391	0.374	0.188	0.202	0.192	0.172	0.159	0.164	0.289
AtARF14	0.357	0.285	0.308	0.213	0.247	0.2	0.21	0.218	0.412	0.385	0.2	0.208	0.205	0.187	0.169	0.173	0.317
AtARF15	0.343	0.278	0.29	0.209	0.24	0.193	0.208	0.213	0.393	0.368	0.205	0.217	0.208	0.196	0.167	0.171	0.3
AtARF16	0.214	0.178	0.2	0.197	0.196	0.166	0.185	0.179	0.227	0.217	0.525	0.58	0.525	0.318	0.151	0.152	0.197
AtARF17	0.179	0.149	0.155	0.171	0.141	0.139	0.138	0.139	0.177	0.162	0.284	0.276	0.291	0.421	0.113	0.123	0.167
AtARF18	0.449	0.348	0.354	0.227	0.283	0.225	0.252	0.256	0.519	0.486	0.233	0.229	0.225	0.204	0.201	0.197	0.335
AtARF19	0.207	0.215	0.214	0.161	0.199	0.318	0.346	0.354	0.198	0.195	0.164	0.156	0.16	0.129	0.604	0.448	0.169
AtARF20	0.351	0.283	0.297	0.21	0.249	0.196	0.212	0.218	0.399	0.381	0.204	0.216	0.211	0.196	0.17	0.172	0.309
AtARF21	0.357	0.283	0.295	0.21	0.249	0.198	0.212	0.221	0.398	0.375	0.205	0.214	0.212	0.191	0.175	0.176	0.308
AtARF22	0.356	0.284	0.299	0.211	0.24	0.195	0.208	0.212	0.407	0.382	0.201	0.209	0.202	0.188	0.17	0.173	0.307
AtARF23	0.163	0.13	0.138	0.14	0.119	0.088	0.094	0.098	0.176	0.161	0.106	0.116	0.108	0.127	0.073	0.074	0.155

Table SII-3. Protein identity matrix among EgrARF

	EgrARF1	EgrARF2A	EgrARF2B	EgrARF3	EgrARF4	EgrARF5	EgrARF6A	EgrARF6B	EgrARF9A	EgrARF9B	EgrARF10	EgrARF16A	EgrARF16B	EgrARF17	EgrARF19A	EgrARF19B	EgrARF24
EgrARF1	ID	0.389	0.411	0.246	0.29	0.247	0.261	0.267	0.503	0.471	0.237	0.245	0.233	0.2	0.208	0.206	0.329
EgrARF2A	0.389	ID	0.515	0.222	0.276	0.247	0.261	0.262	0.372	0.371	0.202	0.204	0.206	0.167	0.215	0.218	0.296
EgrARF2B	0.411	0.515	ID	0.233	0.285	0.252	0.265	0.266	0.387	0.388	0.211	0.205	0.204	0.166	0.218	0.218	0.317
EgrARF3	0.246	0.222	0.233	ID	0.363	0.197	0.196	0.2	0.237	0.241	0.21	0.207	0.202	0.18	0.165	0.164	0.226
EgrARF4	0.29	0.276	0.285	0.363	ID	0.225	0.236	0.235	0.281	0.285	0.214	0.21	0.21	0.166	0.194	0.188	0.242
EgrARF5	0.247	0.247	0.252	0.197	0.225	ID	0.371	0.379	0.236	0.242	0.185	0.18	0.192	0.15	0.32	0.328	0.203
EgrARF6A	0.261	0.261	0.265	0.196	0.236	0.371	ID	0.678	0.263	0.265	0.191	0.187	0.193	0.155	0.35	0.337	0.217
EgrARF6B	0.267	0.262	0.266	0.2	0.235	0.379	0.678	ID	0.268	0.266	0.194	0.197	0.203	0.162	0.361	0.358	0.224
EgrARF9A	0.503	0.372	0.387	0.237	0.281	0.236	0.263	0.268	ID	0.668	0.238	0.236	0.239	0.198	0.21	0.208	0.321
EgrARF9B	0.471	0.371	0.388	0.241	0.285	0.242	0.265	0.266	0.668	ID	0.224	0.223	0.228	0.189	0.207	0.203	0.301
EgrARF10	0.237	0.202	0.211	0.21	0.214	0.185	0.191	0.194	0.238	0.224	ID	0.564	0.577	0.302	0.166	0.166	0.202
EgrARF16A	0.245	0.204	0.205	0.207	0.21	0.18	0.187	0.197	0.236	0.223	0.564	ID	0.556	0.301	0.162	0.156	0.204
EgrARF16B	0.233	0.206	0.204	0.202	0.21	0.192	0.193	0.203	0.239	0.228	0.577	0.556	ID	0.311	0.164	0.165	0.214
EgrARF17	0.2	0.167	0.166	0.18	0.166	0.15	0.155	0.162	0.198	0.189	0.302	0.301	0.311	ID	0.125	0.135	0.199
EgrARF19A	0.208	0.215	0.218	0.165	0.194	0.32	0.35	0.361	0.21	0.207	0.166	0.162	0.164	0.125	ID	0.476	0.159
EgrARF19B	0.206	0.218	0.218	0.164	0.188	0.328	0.337	0.358	0.208	0.203	0.166	0.156	0.165	0.135	0.476	ID	0.163
EgrARF24	0.329	0.296	0.317	0.226	0.242	0.203	0.217	0.224	0.321	0.301	0.202	0.204	0.214	0.199	0.159	0.163	ID

Table SII-4 Comparison of the number of alternative transcripts predicted in phytozome *for E. grandis* to those found in a large compendium of transcriptomic data from in *E. globulus*.

Gene name	*E. grandis*	*E. globulus*
EgrARF1	4	6
EgrARF2A	5	11
EgrARF2B	2	3
EgrARF3	2	5
EgrARF4	3	4
EgrARF5	1	4
EgrARF6A	2	6
EgrARF6B	3	4
EgrARF9A	1	2
EgrARF9B	2	3
EgrARF10	1	1
EgrARF16A	1	1
EgrARF16B	1	3
EgrARF17	1	4
EgrARF19A	2	7
EgrARF19B	2	6
EgrARF24	1	1

These numbers include the primary transcripts

Material and Methods for Alternative slipicing validation

TopHat and Cufflink Suite (Trapnell et al. 2012), currently used for differential gene and transcript expression analysis of RNA-seq experiments, were used to validate the *Eucalyptus ARF* genes (*E. grandis* genome v162 annotation). Ten developing xylem (DX) RNA-Seq libraries were produced and sequenced FASTERIS SA (Genève, CH; www.fasteris.com). RNA-Seq libraries were prepared from ten equimolar pools of high quality individual total RNA: a) four seasonal libraries produced from total RNA extracted from DX samples collected in 2008 on three ramets of non-related tree genotypes at Herdade do Zambujal, Pegões (Portugal) [February 26 th (Feb), May 23rd (May), September 5th (Sep) and December 3rd (Dec)] (Carocha et al., unpublished); b) four pulp yield pools (five trees by pool) obtained by pooling the total RNA from samples. collected at Carregal Fundeiro, Abrantes (Portugal), from 10 trees with very constrasting pulp yields in the coded AxB mapping population, and 10 trees with very constrasting pulp yields in a natural variation panel; c) two libraries from samples collected from an adult tree and a juvenile one, (same genotype) at Herdade do Zambujal, Pegões (Portugal). Sampling procedures and RNA extraction were described, respectively in Paux et al. (2004) and Cassan- Wang et al. (2012). All samples were kindly provided by RAIZ Institute (Portugal). The TruSeq™ SBSv5 sequencing kit (Illumina) was used for library sequencing, using the Illumina Hi-Seq 2000 instrument, on a multiplex runs with 1x100nt+7(index) cycles. For each RNA-Seq library, adaptors removal and a successive quality and contaminants (ribosomal sequences) filters were applied. The resulting filtered high quality reads were then mapped to the *E. grandis* genome sequence v162 using TopHat v1.3.1. Cufflinks v1.1.0 was used to determine the potential coding regions and their intron-exon structures. Finally, Cuffcompare v1.1.0 were then used to compare the assembled multiple libraries Cufflinks transcripts to the *E. grandis* v162 gene annotations. The sequences of the Illumina reads from RNA Seq used to predict the *E. globulus* alternative transcripts are provided in FastaQ format in supplementary File S1.

REF: Cole Trapnell, Adam Roberts, Loyal Goff, Geo Pertea, Daehwan Kim, David R Kelley, Harold Pimentel, Steven L Salzberg, John L Rinn & Lior Pachter (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature Protocols 7: 562–578.

Table SII-5 Comparison of the number of *EgrARF24* putative orthologs in other species

Species	Number of orthologs*
Brassica rapa	0
Gossypium raimondii	2
Malus domestica	2
Medicago truncatula	0
Citrus sinensis	1
Citrus clementina	1
Prunus persica	1
Solanum lycopersicum	0
Manihot esculenta	0
Ricinus communis	0
Linum usitatissimum	0
Phaseolus vulgaris	1
Glycine max	2
Cucumis sativus	0
Fragaria vesca	1
Arabidopsis lyrata	0
Capsella rubella	0
Thellungiella halophila	0
Carica papaya	1
Theobroma cacao	1
Solanum tuberosum	0
Mimulus guttatus v1.1	0
Aquilegia coerulea	1
Sorghum bicolor	0
Setaria italica	0
Panicum virgatum	0
Brachypodium distachyon	0
Selaginella moellendorffii	0
Populus trichocarpa	2
Arabidopsis thaliana	0
Zea mays	0
Oryza sativa	0
Vitis vinifera	1
Eucalyptus grandis	1

*The number of *EgrARF24* orthologs in other species are counted in phytozome by using "Filter homologs" with a cut off E value 1.0E-50

Table SII-6. Potential small RNAs targeting *EgrARF* genes

Small RNA gene	Small RNA sequence (5'-3')	Length (nt)	Genomic location	Predicted small RNA gene sequences (5'-3')
MIR160				
EgrMIR160a	UGCCUGGCUCCCUGUAUGCCA	21	scaffold_4:29,715,926..29,716,011	TGCCTGGCTCCCTGTATGCCACAAACATATTCCAACCTTTCATCTTAGGTTGGCCATTACGTTG GTGGCGTACCAGGAGCCAAGCC
EgrMIR160b	UGCCUGGCUCCCUGUAUGCCA	21	scaffold_7:46392610..46392691	TGCTTGGCTCCTCATACGCCATCCGCGGAGACCATCGATTCTCCGATGCGCTCTCTGCACATGG CATACAGGGAGCCAGGCA
EgrMIR160c	UGCCUGGCUCCCUGUAUGCCA	21	scaffold_10:12063249..12063328	TGCATGGCTCCTCATACGCCATACACACAGGCCATCGGTTCTCCGGTGGGCTACGCAAATGGCA TACAGGGAGCCAGGCA
MIR167				
EgrMIR167a	UGAAGCUGCCAGCAUGAUCU	20	scaffold_8:49,907,076..49,907,178	GAGATGCCAGCAACAGGTGAAGATGCCACATGATCTGATCCATCCTCAAGTCAAAGAGGGAGGT TGAGATCATGCTGGCAGCTTCAACCGATGCTGGTCTCTC
EgrMIR167b	UGAAGCUGCCAGCAUGAUCU	20	scaffold_8:62,733,419..62,733,519	TCATGCTTCACAGCAACAGTTGAAACTGCCAGATGATCTGACAGCTTCTTCTATTGAAGGCTCA GATCATGCTGGCAGCTTCAGCTATTAGTGGTGCATGA
EgrMIR167c	UGAAGCUGCCAGCAUGAUCUA	21	scaffold_11:42,258,705..42,258,841	AAGGGAACGAGTGAAGCTGCCAGCATGATCTAGCTCCGGCTAACTTGAAACTTCAACCGAAAAA AAAAAAAAAACAAGAAAAGAAGAAGTGGTTAACCCTAGCTAGGTCATGCTGCGACAGCCTCACT
EgrTAS3	UUCUUGACCUUGUAAGACCUU	21	scaffold_8:58712106..58712287	TTCTTGACCTTGTAAGACCTTTTCTTGACCTTGTAAGACCCCCCTCCGTTCCCGCCCCGTCACT TCTCCTTCCGGCCCCGGGCTCTGAATCGCGAGTGAGCGATGTAGAGTACGGTCGCCTGCCTTCC CCTTTCCCACCCAACTCATCTTCTCCTTCCTTGTCTATCCCTCCTGAGCTACTC

Table SII-7. Small RNAs target site prediction in *EgrARF* genes

miRNA Acc.	Target Acc.	Expectation[a] (E)	Target Accessibility[b] (UPE)	Alignment	Inhibition[c]	Multiplicity[d]
UGCCUGGCUCCCUGUAUGCCA EgrmiR160	Eucgr.J00923.1 EgrARF10	0	20.2	miRNA 20 CCGUAUGUCCCUCGGUCCGU 1 :::::::::::::::::::: Target 1804 GGCAUACAGGGAGCCAGGCA 1823	Cleavage	1
UGCCUGGCUCCCUGUAUGCCA EgrmiR160	Eucgr.G02838.1 EgrARF16A	0	20.2	miRNA 20 CCGUAUGUCCCUCGGUCCGU 1 :::::::::::::::::::: Target 1435 GGCAUACAGGGAGCCAGGCA 1454	Cleavage	1
UGCCUGGCUCCCUGUAUGCCA EgrmiR160	Eucgr.K01240.1 EgrARF16B	1	22.4	miRNA 21 ACCGUAUGUCCCUCGGUCCGU 1 :::.:::::::::::::::: Target 2095 UGGUAUGCAGGGAGCCAGGCA 2115	Cleavage	1
UGCCUGGCUCCCUGUAUGCCA EgrmiR160	Eucgr.F04380.1 EgrARF17	0.5	18.0	miRNA 21 ACCGUAUGUCCCUCGGUCCGU 1 ::::::.:::::::::::::::: Target 1218 UGGCAUGCAGGGAGCCAGGCA 1238	Cleavage	1
UGAAGCUGCCAGCAUGAUCUA EgrmiR167	Eucgr.D00264.1 EgrARF6A	3.5	22.2	miRNA 20 UCUAGUACGACCGUCGAAGU 1 :::::: :::::::::::: Target 2416 AGAUCAGGCUGGCAGCUUGU 2435	Cleavage	1
UGAAGCUGCCAGCAUGAUCUA EgrmiR167	Eucgr.A02065.1 EgrARF6B	3.5	23.4	miRNA 20 UCUAGUACGACCGUCGAAGU 1 :::::: :::::::::::: Target 3043 AGAUCAGGCUGGCAGCUUGU 3062	Cleavage	1
UUCUUGACCUUGUAAGACCUU EgrTAS3	Eucgr.K02197.1 EgrARF2A	3.5	17.3	miRNA 21 UUCCAGAAUGUUCCAGUUCUU 1 :.: ::::.:::::::: :::: Target 1249 AGGAUCUUGCAAGGUCCAGAA 1269	Cleavage	1
UUCUUGACCUUGUAAGACCUU EgrTAS3	Eucgr.B03551.1 EgrARF2B	1.5	18.5	miRNA 21 UUCCAGAAUGUUCCAGUUCUU 1 :::: :::.:::::::::::::: Target 1641 AAGGCCUUGCAAGGUCAAGAA 1661	Cleavage	1
UUCUUGACCUUGUAAGACCUU EgrTAS3	Eucgr.D00588.1 EgrARF3	0.5	20.5	miRNA 21 UUCCAGAAUGUUCCAGUUCUU 1 ::::::::.:::::::::::::: Target 1641 AAGGUCUUGCAAGGUCAAGAA 1661	Cleavage	1
UUCUUGACCUUGUAAGACCUU EgrTAS3	Eucgr.B02480.1 EgrARF4	0.5	24.1	miRNA 21 UUCCAGAAUGUUCCAGUUCUU 1 ::::::::.:::::::::::::: Target 1631 AAGGUCUUGCAAGGUCAAGAA 1651	Cleavage	1

[a]The complementarity score between small RNA and and their target transcript, if its score is greater than 4 it will be discarded

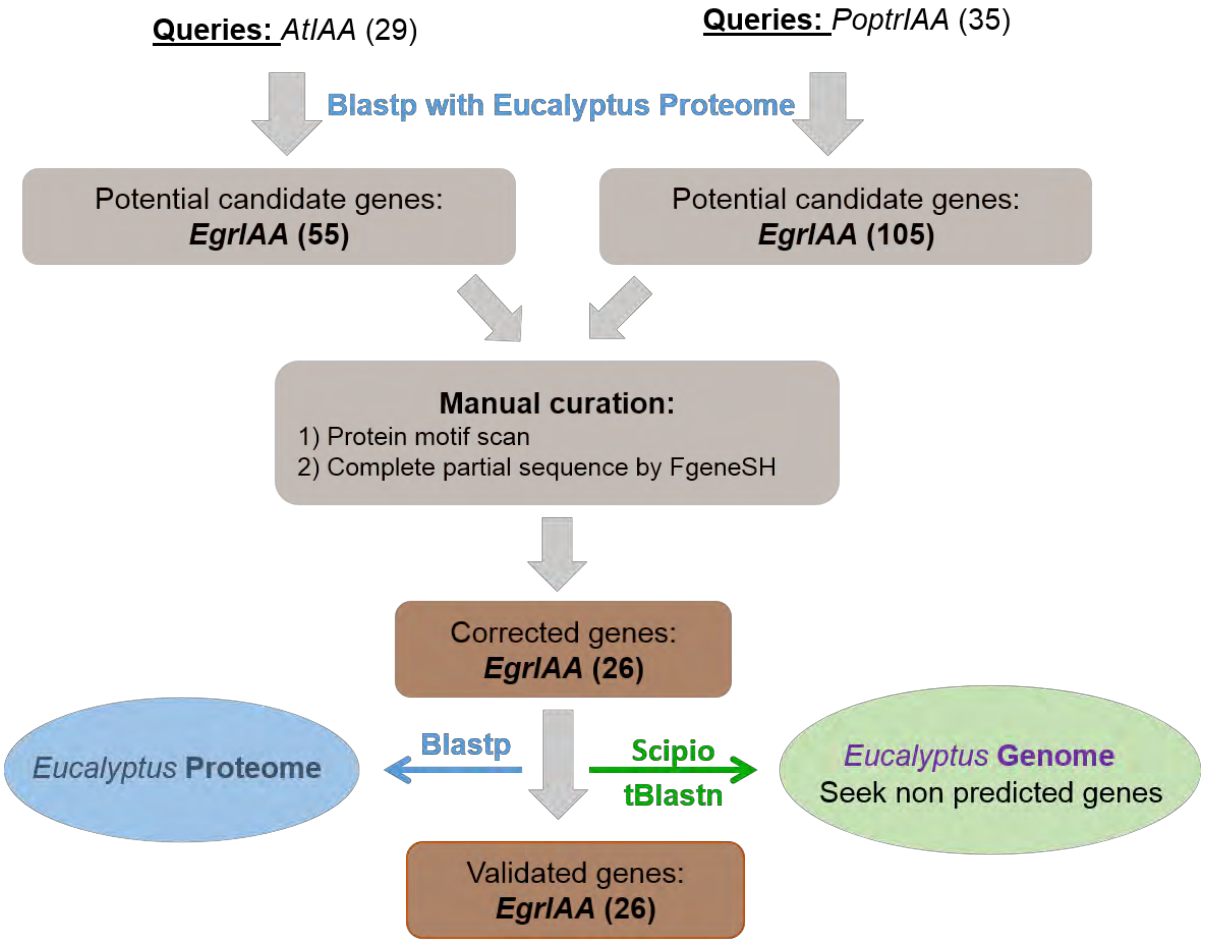
[b]The energy required to open (unpair) secondary structure around target site on target mRNA, The less energy means the more possibility that small

[c]The principle of the miRNA function, cleave mRNA or inhibits the translation of target genes

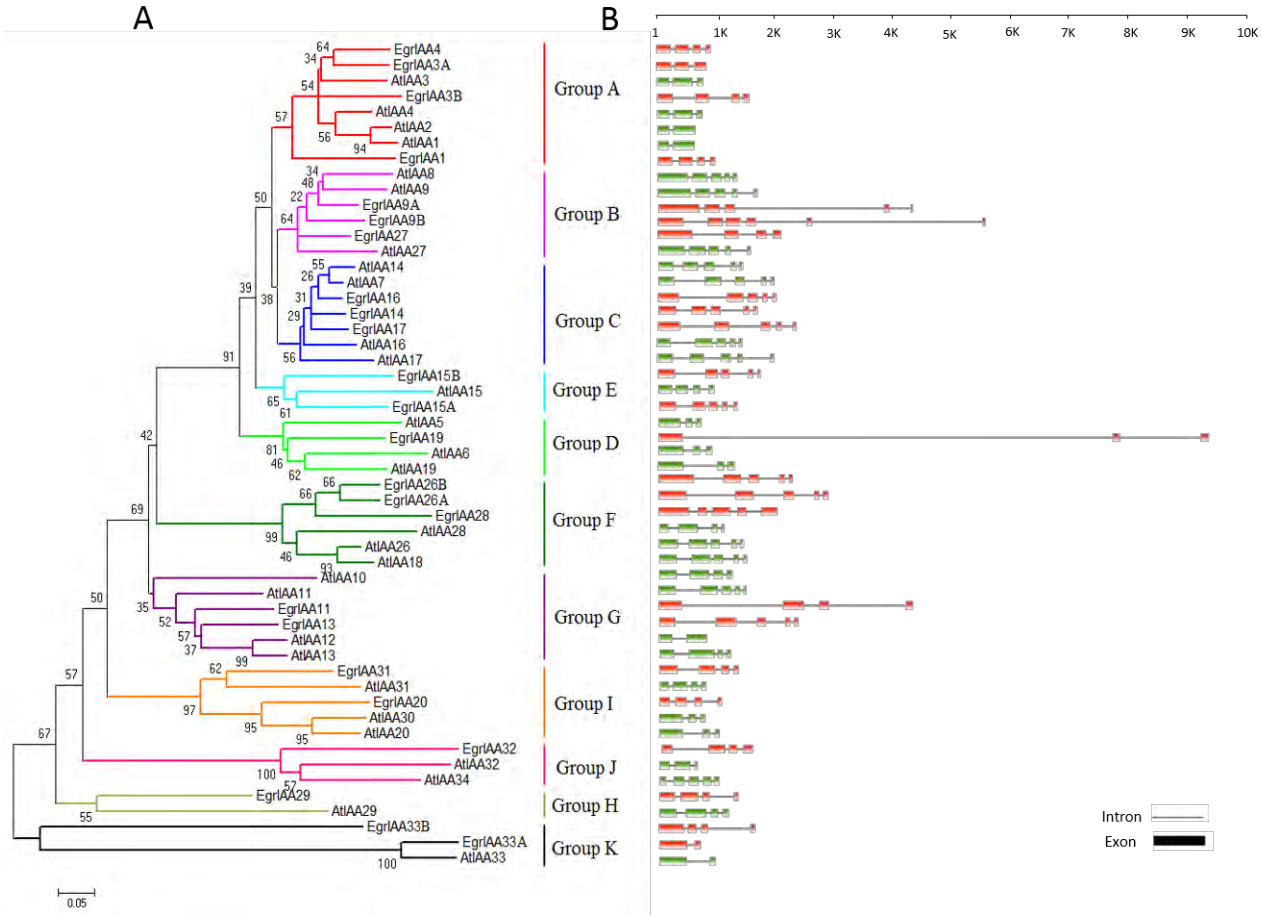
[d]The number of small RNA target sites on each mRNA

Supplementary data for chapter III

Supplementary Fig. SIII-1 Procedure used for identifying *Aux/IAA* genes in *Eucalyptus grandis*. *Arabidopsis* Aux/IAA protein sequences were used to search for related proteins in the predicted *E. grandis* proteome by using BLASTP (REF). Fifty-five *E. grandis* proteins identified in this initial search were further examined by manual curation using protein motif scanning and the FgeneSH program to complete partial sequences. Redundant and invalid genes were eliminated based on gene structure, the integrity of conserved motifs and EST support. Manual curation resulted in 26 EgrIAA protein sequences. We then used these 26 protein sequences in two subsequent additional searches: first, a BLASTP search against the *E. grandis* proteome to identify exhaustively all divergent *E. grandis* Aux/IAA gene family members and, second, tBLASTn searches against the *E. grandis* genome for any possible unpredicted genes. To confirm our findings, we also used poplar Aux/IAA proteins to repeat the complete search procedure described above with identical results obtained.



Supplementary Fig. SIII-2 Phylogenetic analysis of *E. grandis* and *Arabidopsis* IAA proteins and the exon–intron organization of corresponding genes. (A) Neighbor-joining phylogenetic tree of *E. grandis* and *Arabidopsis* IAA proteins. Bootstrap values are indicated at each node. (B) Exon–intron organization of the *EgrIAA* (in red) and *AtIAA* (in green) genes corresponding to the proteins in the phylogenetic tree. The sizes of exons and introns are indicated by the scale at the top.

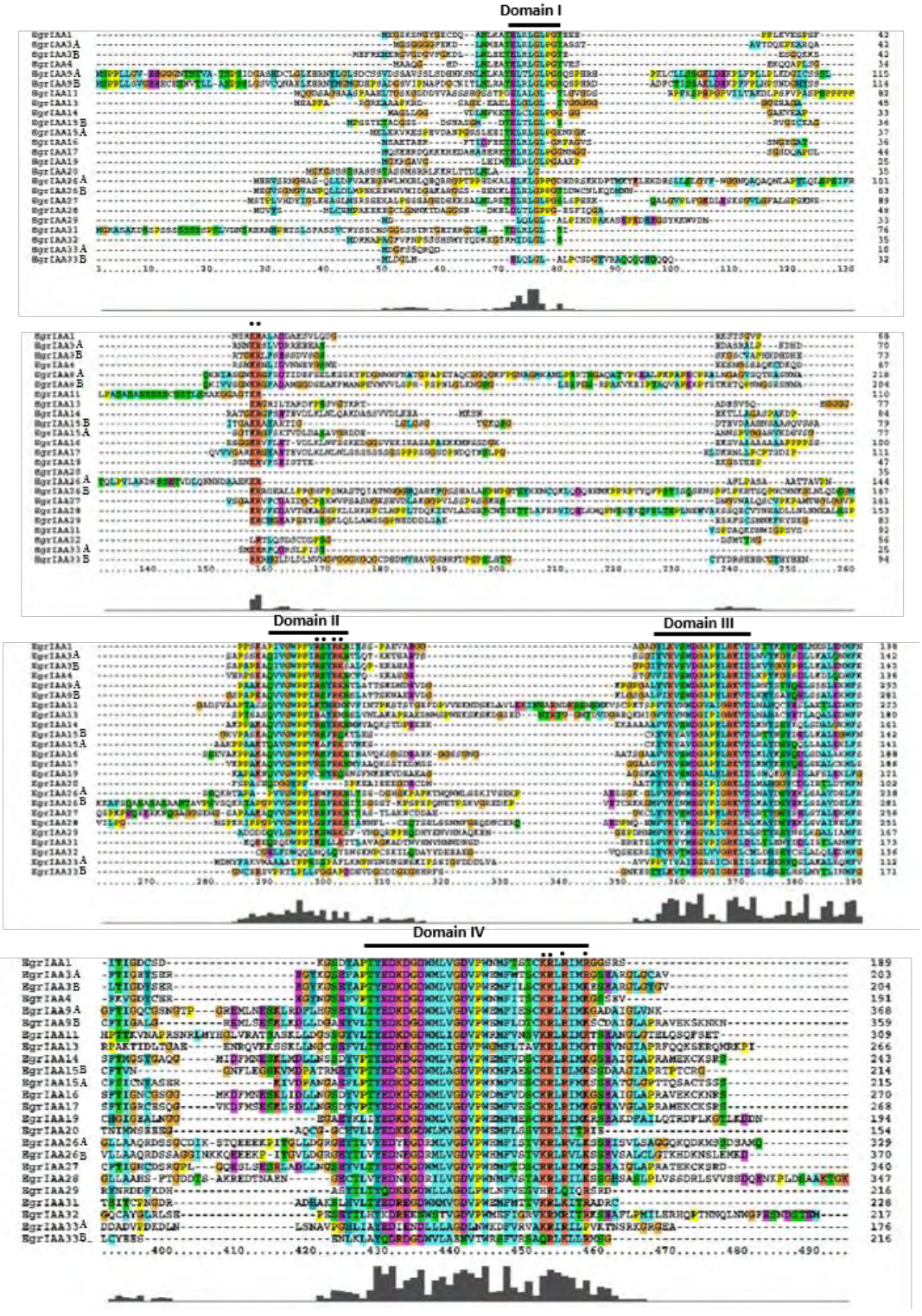


Supplementary Fig. SIII-3 Sequencing alignment of EgrIAA29 and two putatively pseudogenes, EgrIAA29A and 29B.

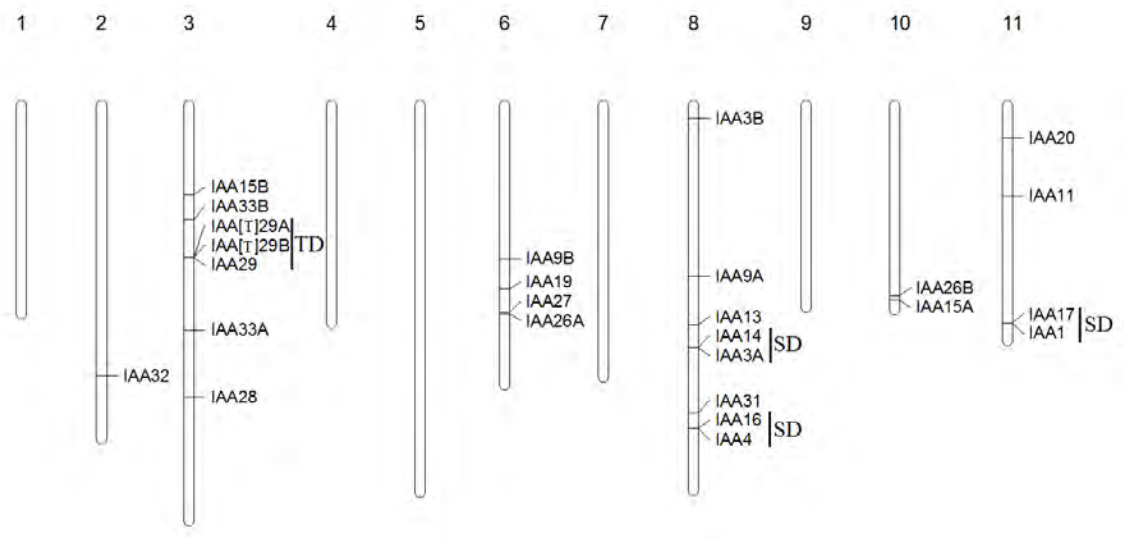
```
                1       10        20        30        40        50        60        70        80        90       100       110       120       130
                |--------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------|
EgrIAA29        MDLQLGLALPIHD-PAKASEPKDHEGSYKKWVDNKRCHGEAFGEYSSGKLQLLAWSSGPNEDDDLSAKRSKFSCSMNKFRYSEGADDDDQVLGWPPIKSWRKEFVHGQRPPHQDHYENVHHAQKENGEPD
EgrIAA29A       MDIQLGLALPIHDHPAKGSEPKDHEGSYKKWFNKKRCHGEAFGEHSSGKSQLLAWSSGPNEEDDPNARRGKFSCSINQFRYSEGADDNDRVVGWPPIKSWRKKFMHGQRPPRRDLHENVRQA
EgrIAA29B       MDLQLGLALPISDHSAKGSEPKDHEGLYKKWANNKRCHGEAFGQYSSGKSQSLAWSGRPNEEDDPSARRSKFYCSLNKFRYSEGADDNDQVVGWPPIQTWRKKFMHGQRPPRRDLYKIVRQA
Consensus       MDlQLGLALPIhDhpAKgSEPKDHEGsYKKW.#nKRCHGEAFG#ySSGKsQlLAWSsgPNE#DDpsArRsKFsCS.NkFRYSEGADD#DqVvGWPPIksWRKkFmHGQRPPrrDlyenVrqA.......

                131     140       150       160       170       180       190       200       210     217
                |--------+---------+---------+---------+---------+---------+---------+---------+-------|
EgrIAA29        HNMFVKVKMEGVAIVRKINLRTYRSYNSLKGALIAMFSRYNRDDFKDHASYTLTYQDKEGDWLLAGDLPWLNFVESVHRLQIQRSRD
EgrIAA29A
EgrIAA29B
Consensus       .....................................................................................
```

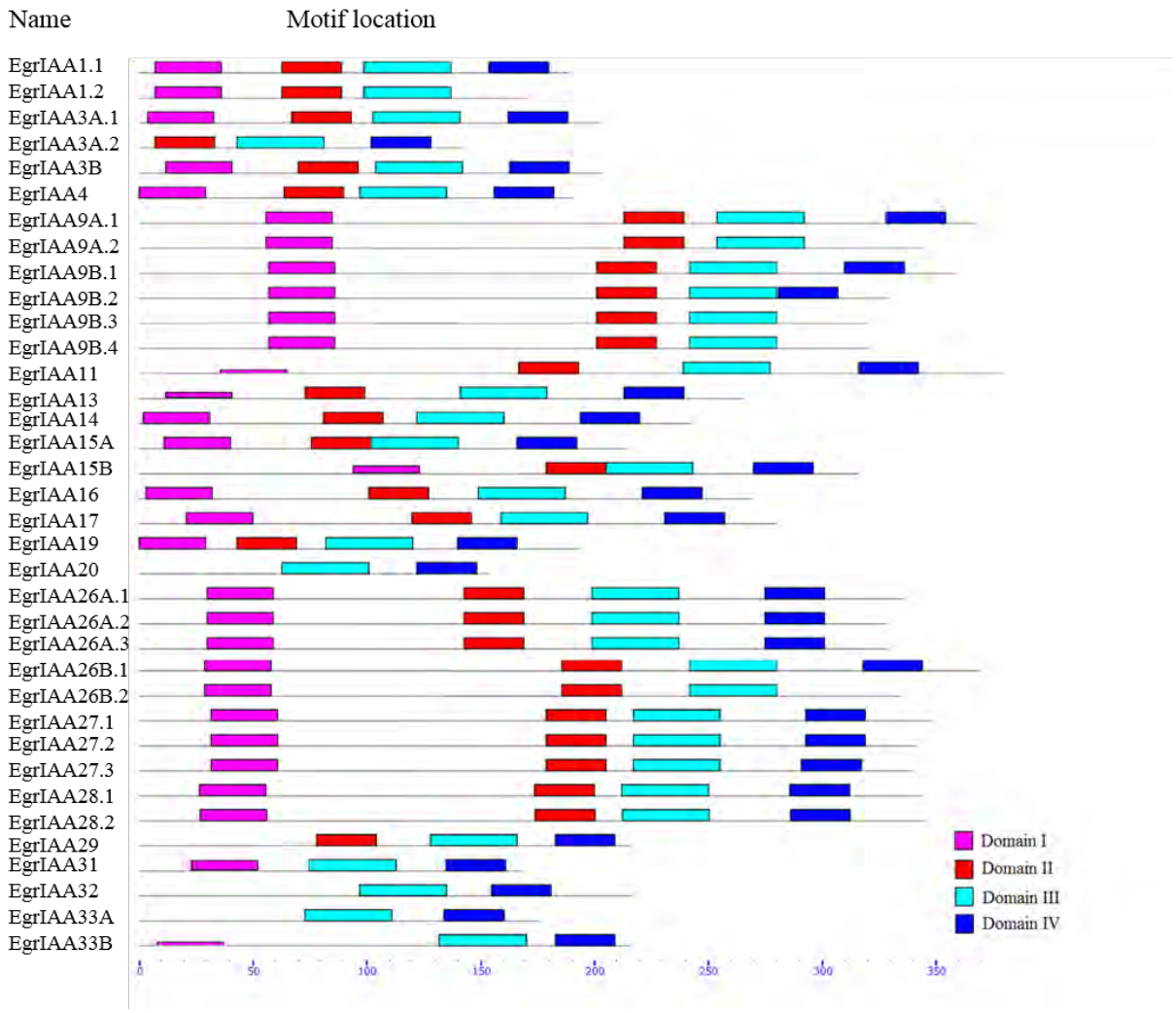
Supplementary Fig. SIII-4 Multiple sequence alignment of predicted amino acid sequences of EgrIAA proteins. The four highly conserved domains of Aux/IAA proteins were noted on the top of the alignment. Nuclear localization signals (NLSs) were indicated by filled circles. The amino acid position was given on the right of each sequence.



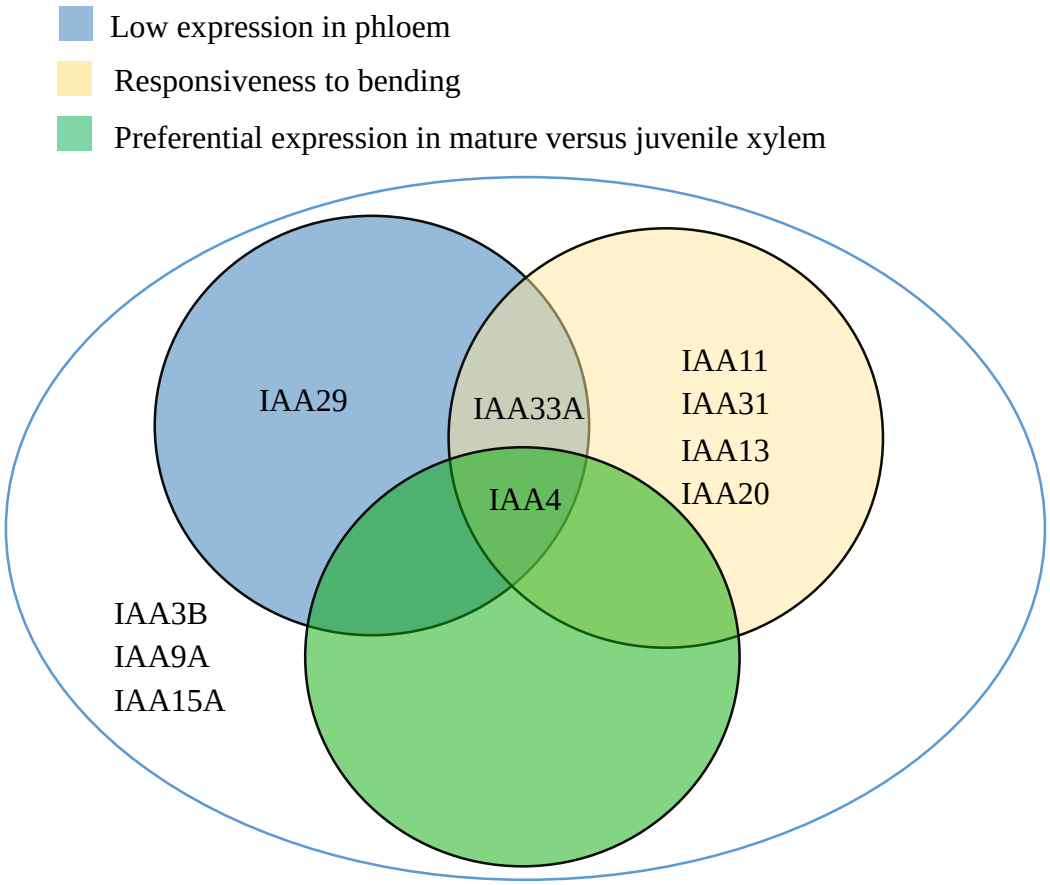
Supplementary Fig. SIII-5 Chromosomal positions of the *EgrIAA* genes. The tandem duplications (TD) and segmental duplication (SD) are written on the right side of the corresponding genes. Chromosome scaffolds (1–11) are indicated at the top end of each chromosome.



Supplementary Fig. SIII-6 Predicted proteins of all the transcripts of *EgrIAA* and their conserved domains. Conserved domains were predicted using the MEME web server with the following parameters: distribution of motif occurrences, zero or one per sequence; motif width ranges from 10 to 50 residues; maximum number of motifs to find, 4. The size and location of motifs can be estimated by the scale at bottom.



Supplementary Fig. SIII-7 Venn diagram of overall strategy to identify the best potential candidate(s) involved in wood formation. Best potential candidate(s) should be i) highly expressed in vascular tissues; ii) preferentially expressed in xylem/cambium compare to phloem; iii) showing a response to bending; iv) preferential expression in mature versus juvenile xylem. Only *EgrIAA4* is fulfilling these conditions.

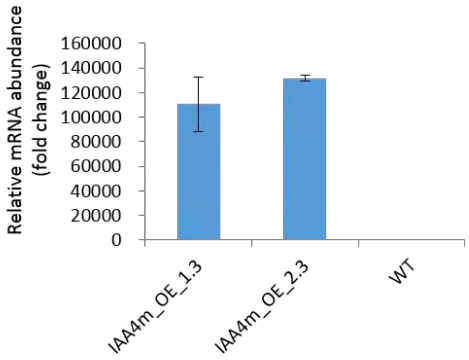


Supplementary Fig. SIII-8 Protein sequence of E*grIAA4* and the phenotype of gain-of-function transgenic *Arabidopsis*. (A) Amino acid sequence of mutated EgrIAA4m protein. In the mutated EgrIAA4m cDNA, the 80th codon encoding the conserved Pro residue (highlighted in red) was changed to a codon encoding a Ser residue. (B) Transcript levels of EgrIAA4m in transgenic *Arabidopsis* under the control of the 35SCaMV promoter. Three technical replicates were used for each qRT-PCR experiment. Bars shows averages ratio of the fragment intensity of interest relative to that of AtUBQ10. Error bars represent SE. (C) *EgrIAA4m* transgenic and wild type *Arabidopsis* plants grown in long/short days for 40 days. Scale bar: 1 cm.

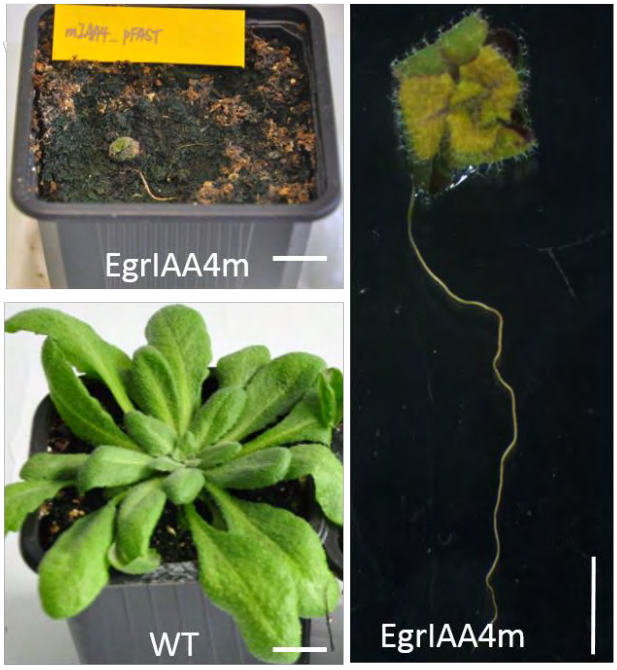
A

MAAQGEDLNLEATELRLGLPGTVEPEKQQAPLSGRSMKRNLIDVNNEYGSNEEESNGSSA
QKCDKQDVHRPSKAQVVGWPPVRSYRKNCFQKKAEGESTGVFIKVSMDGAPYLRKIDLKP
YKGYSDLLKDLQGMFKFKVGEYCEREGYNGSEFVPTYEDKDGDWMLVGDVPWNMFITSCK
RLRIMKGSSEV

B



C



Supplementary Table SIII-1. Protein identity matrix among EgrIAA

	IAA1	IAA3A	IAA3B	IAA4	IAA9A	IAA9B	IAA11	IAA13	IAA14	IAA15A	IAA15B	IAA16	IAA17	IAA19	IAA20	IAA26A	IAA26B	IAA27	IAA28	IAA29	IAA31	IAA32	IAA33A	IAA33B
IAA1	100%																							
IAA3A	78.04%	100%																						
IAA3B	76.62%	83.73%	100%																					
IAA4	78.25%	82.31%	79.06%	100%																				
IAA9A	44.51%	44.10%	44.30%	43.29%	100%																			
IAA9B	46%	45.93%	46.34%	44.51%	64.83%	100%																		
IAA11	50.60%	51.42%	51.62%	50.60%	28.04%	30.48%	100%																	
IAA13	56.09%	57.52%	57.11%	56.91%	29.47%	34.75%	60.56%	100%																
IAA14	66.26%	67.88%	67.07%	67.68%	45.52%	49.39%	50.20%	56.30%	100%															
IAA15A	67.27%	69.51%	71.74%	68.29%	41.66%	44.30%	52.23%	55.08%	67.07%	100%														
IAA15B	64.83%	65.04%	67.68%	65.65%	41.46%	44.51%	48.17%	54.26%	63.61%	69.30%	100%													
IAA16	61.58%	61.99%	61.38%	63.00%	44.30%	47.76%	48.17%	52.03%	78.65%	63.21%	61.17%	100%												
IAA17	64.83%	63.00%	63.41%	64.02%	45.52%	50.20%	46.34%	51.01%	76.21%	63.82%	61.38%	72.76%	100%											
IAA19	71.95%	72.96%	72.15%	71.54%	39.83%	42.27%	48.57%	59.34%	66.46%	69.30%	62.19%	62.80%	62.19%	100%										
IAA20	66.26%	64.83%	64.63%	65.24%	34.75%	34.34%	46.74%	52.43%	56.70%	61.78%	58.33%	52.23%	52.43%	63.82%	100%									
IAA26A	43.90%	42.68%	43.90%	42.47%	30.08%	33%	40.24%	43.49%	44.71%	45.73%	42.47%	44.51%	44.91%	43.29%	41.05%	100%								
IAA26B	37.60%	36.58%	37.39%	36.17%	27.23%	29.87%	33.53%	36.58%	38.41%	37.60%	36.17%	39.22%	38.41%	36.38%	33.13%	62.19%	100%							
IAA27	48.98%	51.62%	52.43%	49.18%	46.74%	48.37%	40.04%	41.05%	54.47%	52.03%	54.06%	54.87%	55.69%	48.17%	39.02%	38.00%	33.94%	100%						
IAA28	39.83%	39.43%	38.00%	39.43%	26.42%	28.25%	32.52%	34.95%	40.44%	38.82%	38.61%	40.44%	39.63%	39.63%	36.17%	48.78%	44.51%	34.14%	100%					
IAA29	62.80%	59.55%	60.56%	61.78%	33.94%	35.16%	48.98%	51.01%	54.47%	59.14%	54.47%	51.62%	52.03%	59.55%	63.61%	41.86%	35.16%	40.44%	38.61%	100%				
IAA31	62.80%	61.78%	61.58%	63.21%	35.97%	36.38%	49.18%	52.43%	54.67%	56.70%	55.08%	52.23%	51.42%	57.92%	70.12%	39.22%	34.75%	39.83%	36.58%	61.99%	100%			
IAA32	59.55%	57.52%	58.33%	57.11%	28.45%	31.30%	45.12%	52.03%	51.62%	56.09%	52.64%	47.56%	47.35%	62.39%	61.38%	39.83%	30.69%	35.56%	36.99%	56.91%	56.70%	100%		
IAA33A	65.85%	64.02%	62.80%	62.80%	31.09%	32.52%	44.71%	52.03%	55.28%	60.97%	57.11%	50.00%	51.82%	63.00%	65.24%	38.21%	30.28%	36.78%	33.94%	60.36%	58.33%	60%	100%	
IAA33B	60.36%	56.91%	57.92%	57.92%	31.30%	33.53%	46.34%	48.17%	51.42%	56.70%	52.03%	47.15%	48.37%	57.92%	61.38%	39.02%	31.09%	38.00%	35.36%	64.63%	57.52%	54.06%	58.53%	100%

Table SIII-2. Alternative transcript comparison between *Eucalyptus* and *Arabidopsis*

Gene name	Number of predicted alternative transcript	Gene name	Number of predicted alternative transcript
EgrIAA1	2	*AtIAA1*	1
EgrIAA3A	2	*AtIAA2*	1
EgrIAA3B	1	*AtIAA3*	1
EgrIAA4	1	*AtIAA4*	1
EgrIAA9A	2	*AtIAA5*	1
EgrIAA9B	4	*AtIAA6*	1
*EgrIAA11**	1	*AtIAA7*	2
EgrIAA13	1	*AtIAA8*	4
EgrIAA14	1	*AtIAA9*	2
EgrIAA15A	1	*AtIAA10*	1
*EgrIAA15B**	1	*AtIAA11*	3
EgrIAA16	1	*AtIAA12*	2
*EgrIAA17**	1	*AtIAA13*	3
EgrIAA19	1	*AtIAA14*	1
EgrIAA20	1	*AtIAA15*	1
EgrIAA26A	3	*AtIAA16*	1
EgrIAA26B	2	*AtIAA17*	1
EgrIAA27	3	*AtIAA18*	1
*EgrIAA28**	2	*AtIAA19*	1
EgrIAA29	1	*AtIAA20*	1
EgrIAA[P]29A	1	*AtIAA26*	1
EgrIAA[P]29B	1	*AtIAA27*	1
*EgrIAA31**	1	*AtIAA28*	1
EgrIAA32	1	*AtIAA29*	1
EgrIAA33A	1	*AtIAA30*	1
EgrIAA33B	1	*AtIAA31*	1
		AtIAA32	2
		AtIAA33	1
		AtIAA34	1
Total number of transcript	38		40

Table SIII-3. Bibliography of the genes which are potentially involved in wood formation in cluster III from the phylogenetic tree

Group in phylogenetic tree	Gene	Expression in Eucalyptus	Ortholog expression in Arabidopsis		Ortholog expression in Poplar	
			Expression in Genevestigator	Bibliography	Microarray Data (Kalluri et al. 2007)	Bibliography
B	*EgrIAA9A*	Vascular preferential expression	*AtIAA8* and *9* shows high expression in hypocotyl xylem cells	*AtIAA8* was up-regulated in xylem (Oh et al. 2003)	*PoptrIAA9* also showed high expression in xylem	*PttIAA2* showed high expression in cambium and xylem (Moyle et al. 2002)
	EgrIAA27	High expression in stem	*AtIAA27* shows high expression in hypocotyl xylem cells	*AtIAA27* was up-regulated in mature stem (Ko, et al. 2004)	*PoptrIAA27.2* also showed high expression in xylem	
G	*EgrIAA11*	Vascular preferential expression	*AtIAA11* showed preferential expression in root xylem cells	*IAA12/bdl* seeding showed reduced cotyledons vasculature (Hamann et al. 1999), *AtIAA12* and *13* was up-regulated in xylem (Oh et al. 2003), and up-regulated in mature stem (Ko, et al. 2004)	*PoptrIAA11* showed high expression in xylem	*PttIAA5* preferentially strong expressed in secondary wall forming xylem cells (Moyle et al. 2002)
	EgrIAA13	Vascular preferential expression	*AtIAA12* and *AtIAA13* display high/specific expression in hypocotyl xylem cells			
H	*EgrIAA29*	Vascular preferential expression	*AtIAA29* showed preferential expression in root xylem cells		*PoptrIAA29* showed high expression in xylem	
I	*EgrIAA20*	Vascular preferential expression		overexpressing these three genes (*AtIAA20, 30, 31*) result in reduced and discontinued vascular in cotyledons (Sato and Yamamoto 2008)	*PoptrIAA20.2* showed preferential expression in xylem	*PttIAA3* and *4* showed cambium and differentiating xylem preferencial expression (Moyle et al. 2002). *PttIAA3m* transgenic poplar showed reduced xylem development (Nilsson et al. 2008).
	EgrIAA31	Vascular preferential expression				

Supplementary Table SIII-4. The primers for *EgrARF* genes and reference genes used in qRT-PCR experiments

Name	Accession no	Forward 5'-3'	Reverse 5'-3'	Efficiency
EgrIAA1	Eucgr.K03314.1	AGGGTCTGATTATGCGCCTACC	TGTGGACGTAAACATGTTCCAAGG	2.01
EgrIAA3A	Eucgr.H03171.1	ACTCTGCTCCTTCTTCCAAAGCAC	GTTTCAGCCTCTGTCTTCTTCGTC	1.88
EgrIAA3B	Eucgr.H00216.1	TGGGCATCTGCCTGATGTCTTC	TATCTATCTGTGCGCCGTCCTC	2.06
EgrIAA4	Eucgr.H04336.1	GGTCCAATGAGGAGGAGTCCAATG	TTGTGCCTTAGATGGCCGATGC	2.05
EgrIAA9A	Eucgr.H02407.1	GCTTGAATTGCACCTAGAGCGTTG	AGGCAAGCATAGCAACCTGCAC	2.10
EgrIAA9B	Eucgr.F02172.1	TCAGTTGTTTCACCATAGGTGCTC	TCTTTCAGCTTGCTCTCACTTAGC	2.01
EgrIAA11	Eucgr.K01426.1	TCAACGCACCACGTTCAAATAGG	TGAGGCACTTGTTGCTCGTACC	1.93
EgrIAA13	Eucgr.H02914.1	AGAGGGAGACTGGATGCTCATTGG	TTGACGGCAGTGAGGAACATCC	1.96
EgrIAA14	Eucgr.H03170.1	TGCTCTGGGCAAGATGTTCAGC	ATCATTCCTTGCGCACCATAGC	2.10
EgrIAA15A	Eucgr.J03016.1	AGCTAAACCTCCTGCAGCCAAG	TCATCACGCTCTTCCTGAACGC	1.88
EgrIAA15B	Eucgr.C01083.1	TGCCATCATCGACGGAAACAGC	AAGGTCAGTTCCGTGTCCATGC	1.85
EgrIAA16	Eucgr.H04335.1	AGCGCCTGAGGATCATGAAAGG	TCTCTACTGCTCTCGGTGCAAG	1.91
EgrIAA17	Eucgr.K03313.1	AAGAGCTGTCTGAGGCCTTGTG	TCCTTCACTCCTTGGGATTCGC	1.98
EgrIAA19	Eucgr.F02578.1	ATACTGCAGACCAGGGACTTCCTC	AGAAGCGTTGCCCGTTCTCATC	1.99
EgrIAA20	Eucgr.K00561.1	AGATGGGCAAAGGAAGCAGCTC	TTCTTGAGACGCCGAGACATGG	2.03
EgrIAA26A	Eucgr.F03080.1	TCTGAGATTTCTGTGCTCAGTGC	GCTCATTGCATGGCAGAATCCG	1.73
EgrIAA26B	Eucgr.J02934.1	CCGAAGCAAAGTCGATTCAAGAGG	AACCAAGAGGATCTGCAACAAGAG	1.98
EgrIAA27	Eucgr.F03050.1	AGTTCCGAGGCAATTGGTCTAGC	CTAATCCCGGCTCTTGCACTTCTC	1.87
EgrIAA28	Eucgr.C02984.1	TAGAGGCCTTCTTGCTGCTCAC	GGCATTTGTGTCCTCTCTCTTTGC	1.87
EgrIAA29	Eucgr.C01734.1	TCAAGTCCTGGAGGAAGGAGTTCG	TCTCGTAATGATCCTGGTGTGGTG	2.11
EgrIAA31	Eucgr.H04141.1	TTGGATTGGCCCTTCGGTTTCG	TTGATGGGTGGCCAGTCTTGTC	1.83
EgrIAA32	Eucgr.B02853.1	GCGCATGCGGATCACAAGAAAG	TTGGTTGTTCGTCGGCTGATGC	2.43
EgrIAA33A	Eucgr.C02329.1	AGGCAGATGTTCGTGGACGATG	AAGGTGTCCGGGAACAGCATTG	1.69
EgrIAA33B	Eucgr.C01373.1	TCATCAACATGTTCGGCTTGTGC	CTCAGCAGTTTCAATCGCTGTGC	1.76
EF-1a	Eucgr.B02473.1	ATGCGTCAGACTGTGGCTGTTG	TTGGTCACCTTGGCTCCACTTG	1.84
SAND	Eucgr.B02502.1	TTGATCCACTTGCGGACAAGGC	TCACCCATTGACATACACGATTGC	2.11
PP2A_1	Eucgr.B03386.1	TCGAGCTTTGGACCGCATACAAG	ACCACAAGAGGTCACACATTGGC	1.98
*PP2A_3 5'*end	Eucgr.B03031.1	CGGAAGAACTGGGTGTGTTT	CACAGAGGGTCTCCAATGGT	2.03
*PP2A_3 3'*end	Eucgr.B03031.1	CAGCGGCAAACAACTTGAAGCG	ATTATGTGCTGCATTGCCCAGTC	2.02
*IDH 5'*end	Eucgr.F02901.1	AATCGACCTGCTTCGACCCTTC	TCGACCTTGATCTTCTCGAAACCC	1.92
*IDH 3'*end	Eucgr.F02901.1	TGCTGTGGCAGCTGAACTCAAG	ATGTTGTCCGCCAGTCACCTAC	1.86

Supplementary data for chapter IV

Fig. SIV-1 Maps and features of pENTR/D-TOPO which can cloning blunt-end PCR products into the empty used for the gateway system.

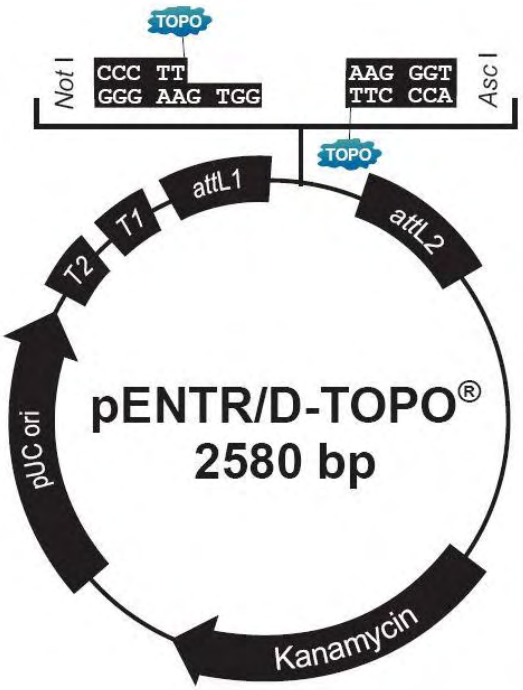


Fig. SIV-2 Maps and features of vectors pK7FWG2 (A) and pK7WGF2 (B) used for subcellular localization.

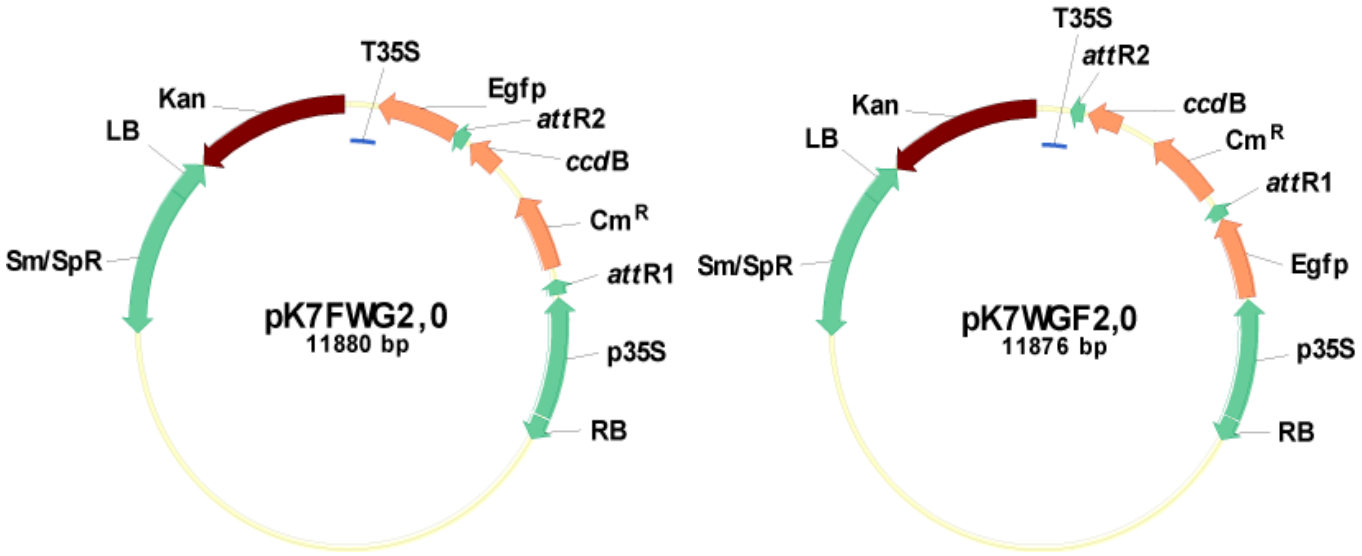


Fig. SIV-3 Maps and features of pGreen vector used for transcription activation analysis.

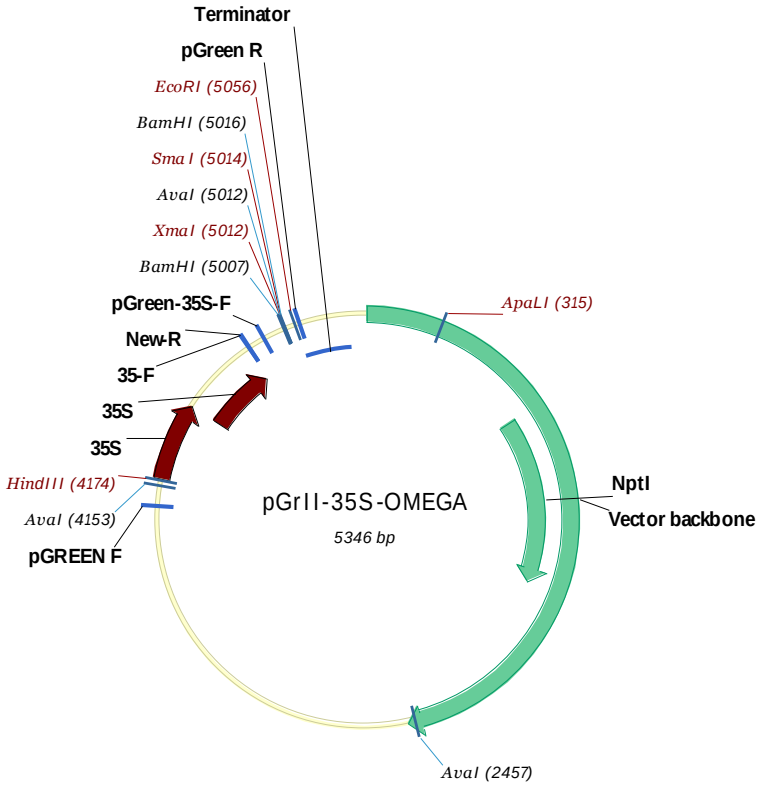


Fig. SIV-4 Maps and features of pFAST-G02 vector used for overexpression construction by gateway.

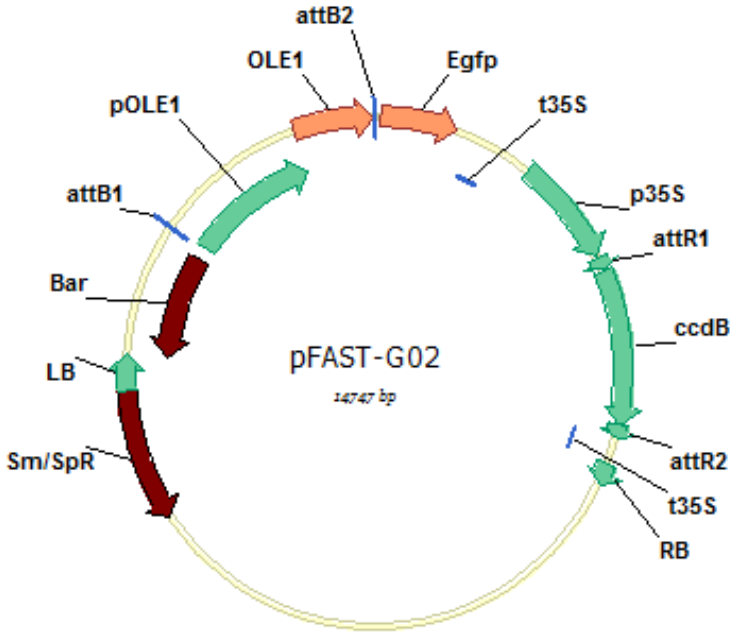


Fig. SIV-5 Maps and features of vector pH35GEAR used for dominant repression construction.

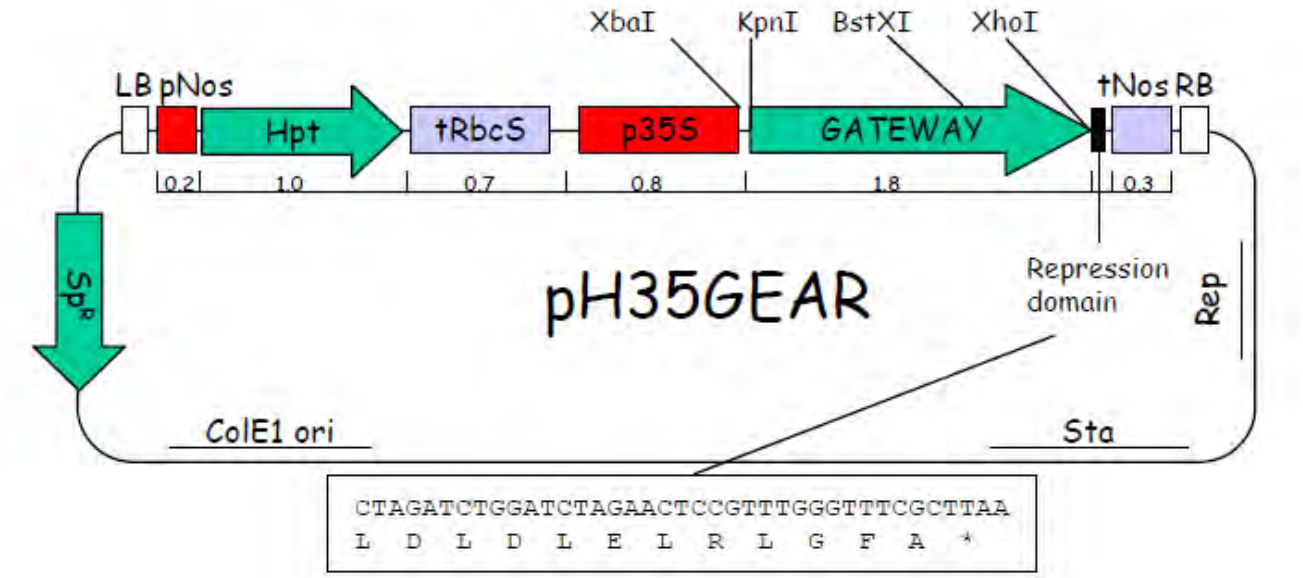


Table SIV-1. The Primers for wood associated CGs amplification

Name	Forward 5'-3'	Reverse 5'-3'
ARF4m_1	CACCATGGAAATTGATCTAAATCATGC	CTCCTGTCCCTGTAACACTTTAGAGGATCGTAAAGATTCCTCA
ARF4m_2	AAGTGTTACAGGGACAGGAGAAGTTGCATTTGGTGTCACCT	CTCCTGTCCCTGTAACACTCTTGGGAATCTATCGGATTCCA
ARF4m_3	AGAGTGTTACAGGGACAGGAGATATGCACACTGAAGTCTTTGACG	AAGCCTAACGACAGTAGGTGAA
ARF10m_1	CACCATGACGGTGTTGAAGGAGCCAGA	TTGTCTTGCACCTTGGATACCTGCAGGAGCATTATCAGATAGACAA
ARF10m_2	TGCAGGTATCCAAGGTGCAAGACAAGCTCAATTTGGAATACCCTTAAC	CGCGTAAGTGCTCAATGGAC
ARF19A	CACCATGAAGGTTCCTTCCAATGGATTTTTAGCT	TCTACTAATCAATATATCACCTCTCCATGCGT
IAA3Bm_1	ATGGAGTTTCGAGAAATGGAGAG	TCGGACAGGACTCCATCCCACGATCTGTGCTT
IAA3Bm_2	TGGGATGGAGTCCTGTCCGATCCTACCGAAA	AACACCATAACCCAAGCCCCTA
IAA4m_1	CACCATGGCAGCTCAAGGAGAGGAT	ATCGGACCGGACTCCACCCCACGACTTGTGCCTTA
IAA4m_2	CGTGGGGTGGAGTCCGGTCCGATCCTACCGAAA	AACCTCTGATGACCCTTTCATGATT
IAA9Am_1	CACCATGTCTCCACCACTTCTGG	CTAATAGGACTCCAACCAACGACCTGAGC
IAA9Am_2	GGTTGGAGTTCCTATTAGATCATACAGGAAGAATACA	ACTCTTCGATTTGTTCACTAGTCC
IAA13m_1	CACCATGGAAGCTCCACCTGCTC	TTATAGGACTCCATCCCACAACCTGACTTG
IAA13m_2	TTGTGGGATGGAGTCCTATAAGGGCATACAGGATGA	TATCGGCTTTCTCATTTGCCT
IAA20	CACCATGGGCAAAGGAAGCAGCTCAT	CTCTATTCTTGTAATCTTCAGTCTCTTCA
IAA29m_1	CACCATGGATCTTCAACTCGGTTTAGC	TTGATGGGACTCCACCCCAGAACCTGATCAT
IAA29m_2	TTCTGGGGTGGAGTCCCATCAAGTCCTGGAGGAAG	ATCTCTACTCCTTTGTATCTGCAGGC
IAA31	CACCATGGGAAGGGCTTCAGCTA	CTAACATCTGTCGGCTCTCG

Underline sequences (CACC) represent the extra sequence needed for the pENTR/D-TOPO cloning; sequences in red represent the overlap region for overlap PCR; sequences highlight in yellow represent the mutation base.

AUTHOR: Hong YU

TITLE: Characterization of *ARF* and *Aux/IAA* gene families in *Eucalyptus* and their roles in wood formation

PHD SUPERVISORS: Dr Jacqueline GRIMA-PETTENATI and Hua CASSAN-WANG

DEFENSE DATE AND PLACE: 2014, September 30th, conference room FRAIB

ABSTRACT

Auxin is a key regulator of cambium activity and wood formation. Auxin Response Factors (ARF) and Auxin/Indole-3-Acetic Acid (Aux/IAA) are important regulators of auxin responses in plants. Thanks to the recent *Eucalyptus grandis* genome sequence, we identified 17 ARF and 24 *Aux/IAA*, and further characterized their gene structure, protein motif architecture, chromosomal location. By a combination of comparative phylogeny and transcript profiling in a large panel of organs, tissues and environmental conditions, we identified promising candidates and carried out their functional characterization to study their potential roles in wood formation. The stabilized versions of *EgrIAA4m*, *9Am* and *20* were overexpressed in *Arabidopsis*. The transgenic lines exhibited auxin-related aberrant phenotypes, and obvious modifications in xylem and/or fiber cells differentiation and secondary cell wall composition. Altogether, this study provides a comprehensive characterization of the *Eucalyptus ARF* and *Aux/IAA* gene families and highlights the involvement of some members in the regulation of wood development.

KEY WORDS:

ARF, *Aux/IAA*, Auxin, *Eucalyptus*, Expression analysis, Wood formation, Xylem, Vascular cambium, Secondary cell walls, *Arabidopsis*, Functional genomics, Transcriptional regulation, Comparative phylogeny

---

ADMINISTRATIVE DISCIPLINE: Plant Biosciences

ADDRESS OF THE LABORATORY:

Laboratoire de Recherche en Sciences Végétales (LRSV), UMR5546, Université Toulouse 3, CNRS, BP 42617, Auzeville, 31326 Castanet Tolosan, France

AUTEUR: Hong YU

TITRE: Caractérisation des familles de facteurs de transcription *ARF* et *Aux/IAA* chez l'*Eucalyptus* et rôles dans la formation du bois

DIRECTEURS DE THESE: Dr Jacqueline GRIMA-PETTENATI et Hua CASSAN-WANG

LIEU ET DATE DE SOUTENANCE: 30 Septembre 2014, Salle de conférence FRAIB

RESUME

L'auxine est une hormone centrale du développement chez les végétaux et joue un rôle clé dans la régulation de la formation du bois. Les ARFs (Auxin-response factors) et les Aux/IAAs (Auxin/Indole acetic acid) sont des régulateurs importants de la voie de signalisation de l'auxine. La disponibilité de la séquence du génome *d'Eucalyptus grandis* nous a permis d'étudier les caractéristiques et l'histoire évolutive de ces deux familles de gènes. Nous avons identifié et caractérisé 17 *EgrARFs* et 24 *EgrIAAs*. Des analyses d'expression sur un large panel d'organes et de tissus et en réponse à des signaux environnementaux, ont mis en évidence des gènes candidats préférentiellement exprimés dans le xylème. La caractérisation fonctionnelle de *EgrIAA4*, *9* et *20* a été réalisée par surexpression chez *Arabidopsis*. Les analyses des lignées transgéniques ont montré des modifications évidentes dans la différentiation des cellules du xylème et/ou des fibres interfasciculaires. En conclusion, cette étude offre une caractérisation complète de deux familles de régulateurs importants de l'auxine chez l'*Eucalyptus* et met en évidence l'implication de certains membres dans la régulation de la formation du bois.

MOTS CLES:

Auxine, ARF, Aux/ IAA, *Eucalyptus*, Expression de gènes, Xylème, formation du bois, Cambium vasculaire, Paroi secondaire, *Arabidopsis*, Génomique fonctionnelle, Régulation transcriptionnelle, Phylogénie comparative

---

DISCIPLINE ADMINISTRATIVE: Biosciences végétales

INTITULE ET ADRESSE DU LABORATOIRE:

Laboratoire de Recherche en Sciences Végétales (LRSV), UMR5546, Université Toulouse 3, CNRS, BP 42617, Auzeville, 31326 Castanet Tolosan, France