



HAL
open science

Implementing the CRISPR-Cas9 technology in Eucalyptus hairy roots and functional characterization of auxin-dependent transcription factors involved in wood formation

Ying Dai

► **To cite this version:**

Ying Dai. Implementing the CRISPR-Cas9 technology in Eucalyptus hairy roots and functional characterization of auxin-dependent transcription factors involved in wood formation. *Vegetal Biology*. Université Paul Sabatier - Toulouse III, 2020. English. NNT : 2020TOU30220 . tel-03228121

HAL Id: tel-03228121

<https://theses.hal.science/tel-03228121>

Submitted on 17 May 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE

**En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE**

Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par

Ying DAI

Le 3 décembre 2020

Implementing the CRISPR/Cas9 technology in Eucalyptus hairy roots and functional characterization of auxin-dependent transcription factors involved in wood formation

Ecole doctorale : **SEVAB - Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingenieries**

Spécialité : **Développement des plantes**

Unité de recherche :

LRSV - Laboratoire de Recherche en Sciences Végétales

Thèse dirigée par

Jacqueline GRIMA-PETTENATI et Hua CASSAN-WANG

Jury

Mme Virginie LAUVERGEAT, Rapporteur

M. Richard SIBOUT, Rapporteur

M. Christian DUBOS, Examineur

M. Vincent BURLAT, Examineur

Mme Jacqueline GRIMA-PETTENATI, Directrice de thèse

Mme Hua CASSAN-WANG, Co-directrice de thèse

Acknowledgement

First and foremost, I would like to sincere thanks to my supervisors Dr. Jacqueline Grima-Pettenati and Dr. Hua Cassan-Wang, who gave me a lot of suggestions and constant support during my whole PhD study and life. Dr. Jacqueline gave me a lot of energy and courage to face difficulties, and also the precious criticism and guidance to push me make progresses and learn more knowledge of science, life philosophies, and logical mode of thinking and speaking. Thanks to Dr. Hua Cassan-Wang for her help all the time not only for study but also in everyday life. These experiences will be my valuable lifetime wealth.

I'm very grateful to my committee members: Dr. Christian Dubos, Dr. Andreas Niebel, and Dr. Nicolas Frei Dit Frey, for providing advice and patient guidance for my PhD research work.

I would also like to thank all members of GFE team, like a family to make me progress. Fabien MOUNET taught me how to present my work logically and confidently, and gave me encouragement. Nathalie LADOUCE and Annabelle DUPAS taught me technics that were important for my research, and also to have a rigorous and patient attitude to scientific stuffs. Thanks to Nathalie LADOUCE for sequencing service and DNA/RNA extraction teaching. Thanks to Annabelle DUPAS patient guidance and illustration of hairy root transformation, she helped me a lot to take care of my *in vitro*/hydroponic culture plant babies. Ines HADJ-BACHIR gave me the warm care in my daily life and be nice and patient for my demand. All these optimistic and active members really provided me help and support in both scientific area and emotions.

Thanks to Guojian HU (in GBF lab) for his experience sharing and guides to my research and thesis writing. Dr. Guojian HU helped me to troubleshooting especially in CRISPR parts during my experiments. Thanks to H  l  ne San Clemente for the support of bioinformatics. Thanks to Yves Martinez (FR3450) for assistance with microscopy (Plateforme Imagerie TRI). Thanks to Luciano Medina and Nils Blandels for their help of my research work, such as plants harvesting and FT-IR analysis. Thanks to Myriam Badawi, Guillermina Hernandez-Raquet for their contributions of the paper. Thanks to Dr. Zhangsheng ZHU to help me the 1st year experiment and gave me advice during processes.

Thanks to my friends accompany, encouragement and help (Jing AN, Luyang HU, Jie SHAO, Biao CAO, Tongming WANG), in both studies and life.

Thanks to China Scholarship Council (Beijing) provided this opportunity to me for my advanced study and live independently in France. Thanks to my country to provide this favorable chance to me for broaden my horzions and learnt more French culture.

And at last but not the least, I would like to thank to my parents to support and understand me all these years spiritually.

Abstract

Eucalyptus is the most planted hardwood worldwide for many industrial end-uses such as pulp and paper and emerging biofuel production. The analysis of the *Eucalyptus grandis* genome led to many candidate genes involved in wood formation including key mediators of auxin signaling (Auxin/Indole-3-Acetic Acid (Aux/IAA) and Auxin Response Factor (ARF)). The functional characterization of these candidate genes was hampered by the difficulty to general stable transgenic *Eucalyptus* and to knock out these genes. Taking advantage of rapid and efficient hairy root transformation mediated by *A. rhizogenes*, recently implemented by our team, the objectives of my work were to implement the powerful CRISPR/Cas9 gene editing tool and to use it to investigate the potential roles of three *Eucalyptus* auxin-dependent transcription factors (IAA9A, IAA20 and ARF5) in regulating wood formation.

First, as a proof-of-concept for implementing CRISPR/Cas9, We targeted *Cinnamoyl-CoA Reductase1 (CCR1)*, a key lignin biosynthetic gene whose down-regulation effects are well described in several plants. Almost all transgenic lines were edited but the allele-editing rates and profiles varied greatly depending on the genes targeted. Most edition events generated truncated proteins. The prevalent edition types were small deletions but large deletions were also observed. By using a combination of Fourier Transformed InfraRed (FT-IR) spectroscopy and multivariate analysis (partial least square analysis (PLS-DA)), we showed that the *CCR1*-edited lines, which were clearly separated from the controls. The most discriminant wave-numbers were attributed to lignin. Histochemical analyses further confirmed the decreased lignification and the presence of collapsed vessels in *CCR1*-edited lines, which are characteristics of CCR1 deficiency. Although the efficiency of editing could be improved, the method described here is already a useful tool to functionally characterize eucalypts genes.

In the second part of my work, we used this genome editing method to knock out two *Aux/IAAs (IAA9A and IAA20)* and one Auxin Response Factor (*ARF5*) in order to get more insights into the role of auxin in the regulation of wood formation in *Eucalyptus*. We generated transgenic *Eucalyptus* hairy root to overexpress and to knock out these genes. Unfortunately, all the transgenic plants overexpressing *IAA9A* and *IAA20* (under the control of 35S promoter) died during the Covid19 lockdown period and only three *IAA20*-CRISPR lines survived. Therefore, we could only analyze CRISPR/Cas9 edited transgenic plants for two candidates (*IAA9A* and *ARF5*). Editing events were detected either by subcloning and/or web-based tools (DSDdecode and ICE synthego). CRISPR/Cas9 generated *IAA9A*_lines had high knockout rates of 92.3% with 58.3% of biallelic mutations. In contrast, *ARF5* lines had quite low editing rates (43%) showing monoallelic and chimera mutations. In *IAA9A*_edited lines we observed precocious xylem development and increased xylem vessel diameters, while no obvious phenotype was detected in *ARF5*_edited lines. Finally, we screened a *Eucalyptus* developing xylem Yeast Two-Hybrid (Y2H) library to find potential partners of *IAA9A* and *IAA20*. For *IAA9A*, we found some potentially promising candidates such as Histone Linker (EgH1.3), CCoAOMT2 previously reported to be involved in xylem formation; for *IAA20* the main interactor revealed was *IAA9A*, suggesting that *IAA20* and *IAA9A* form dimers in developing xylem to regulate wood formation. In addition, we used the yeast two hybrid method to confirm protein-protein interactions of EgrIAA9A and EgrIAA20 with EgrARF5 and other candidates preferentially expressed in *Eucalyptus* wood-forming tissue.

Résumé

Les *Eucalyptus* sont les feuillus les plus plantés au monde pour les nombreuses utilisations industrielles de leurs bois telles que la pâte à papier et la production émergente de biocarburants. L'analyse du génome d'*Eucalyptus grandis* a conduit à l'identification de nombreux candidats impliqués dans la formation du bois, tels que des médiateurs clés de la signalisation de l'auxine (Aux/IAA et Auxin Response Factor (ARF)). La caractérisation fonctionnelle de ces gènes candidats a été retardée jusqu'à présent par la difficulté de supprimer leurs fonctions dans un système homologue. Pour pallier à cela, le premier objectif de mon travail a été de mettre au point le puissant outil d'édition de gènes "CRISPR / Cas9" en profitant de la transformation de "hairy roots" transgéniques médiée par *A. rhizogenes*, récemment développée dans l'équipe. Dans un deuxième temps, mon objectif était d'utiliser cette méthode d'édition de génome pour étudier les rôles potentiels de trois facteurs de transcription dépendant de l'auxine (*IAA9A*, *IAA20* et *ARF5*) dans la formation du bois d'*Eucalyptus*.

Premièrement, comme preuve de concept pour la mise en œuvre de la technologie CRISPR / Cas9, nous avons ciblé la *Cinnamoyl-CoA réductase1 (CCR1)*, un gène clé de la biosynthèse de la lignine dont les effets de "down-regulation" sont bien connus. Nous avons également utilisé le gène *IAA9A* comme cible. Presque toutes les lignées transgéniques ont été éditées, mais les taux et les profils d'édition alléliques variaient considérablement selon le gène ciblé. La plupart des événements d'édition ont généré des protéines tronquées. Les types d'édition les plus courants étaient de petites délétions, bien que de grandes délétions aient également été observées. En utilisant une combinaison de spectroscopie à infrarouge transformée de Fourier (FT-IR) et d'analyse multivariée (PLS-DA), j'ai pu montrer que les lignées éditées pour *CCR1*, étaient clairement séparées des témoins. Les nombres d'onde discriminant les deux groupes ont été attribués à la lignine. Les analyses histochimiques ont confirmé la diminution de la lignification et la présence de vaisseaux écrasés dans les lignées éditées pour *CCR1*, qui sont des caractéristiques de la déficience de ce gène. Bien que l'efficacité de l'édition puisse être améliorée, la méthode décrite ici est déjà un outil utile pour caractériser fonctionnellement des gènes chez l'*Eucalyptus*. Dans la deuxième partie de mon travail, j'ai utilisé cette méthode d'édition du génome pour muter deux *Aux/IAAs* (*IAA9A* et *IAA20*) ainsi que *ARF5* afin de mieux appréhender le rôle de l'auxine dans la régulation de la formation du bois chez l'*Eucalyptus*. J'ai généré des "hairy roots" soit pour surexprimer ces gènes, soit pour les muter par CRISPR/Cas9. Malheureusement, toutes les plantes transgéniques surexprimant *IAA9A* et *IAA20* (sous le contrôle du promoteur CaMV35S) sont mortes pendant la période de confinement liée au Covid19 et seules trois lignées CRISPR-*IAA20* ont survécu. Par conséquent, je n'ai pu analyser que des plantes transgéniques éditées pour deux candidats. Les lignées *IAA9A* générées par CRISPR / Cas9 présentaient des taux de knock-out élevés de 92,3% avec 58,3% de mutations bialléliques. En revanche, les lignées *ARF5* avaient des taux d'édition assez faibles (43%) et des mutations monoalléliques et/ou chimériques. Dans les lignées éditées pour *IAA9A*, nous avons observé un développement précoce du xylème et une augmentation du diamètre des vaisseaux du xylème, alors qu'aucun phénotype évident n'a été détecté dans les lignées éditées pour *ARF5*. Enfin, j'ai participé au criblage d'une banque double hybride de xylème d'*Eucalyptus* (Y2H) pour trouver des partenaires potentiels de *IAA9A* et *IAA20*. Pour *IAA9A*, des candidats prometteurs ont été obtenus tels que Histone Linker (EgH1.3) et CCoAOMT2, connus comme étant impliqués dans la formation du xylème; pour *IAA20*, l'interacteur principal est *IAA9A*, ce qui suggère que *IAA20* et *IAA9A* forment des dimères pour réguler la formation du bois. Nous avons également utilisé la méthode double hybride ciblée pour confirmer les interactions protéine-protéine d'*IAA9A* et d'*IAA20* avec *ARF5* ainsi qu'avec d'autres candidats préférentiellement exprimés dans le xylème d'*Eucalyptus*.

Table of Contents

ACKNOWLEDGEMENTS	I
ABSTRACT	III
Résumé	IV
TABLE OF CONTENTS	V
LIST OF FIGURES	IX
LIST OF TABLES	XI
ABBREVIATIONS	XIII
OBJECTIVES AND ORGANIZATION OF THE MANUSCRIPT	1
CHAPTER I: BIBLIOGRAPHIC REVIEW	3
PART 1 EUCALYPTUS AND WOOD FORMATION	5
1. Eucalyptus	5
2. The vascular system	6
2.1 The vascular system during primary growth	6
2.2 The vascular system during secondary growth	8
3. Wood formation (xylogenesis)	10
3.1 Secondary cell wall (SCW) structure and composition	12
<i>Cellulose</i>	13
<i>Hemicellulose</i>	13
<i>Lignin</i>	15
3.2 The transcriptional regulation of secondary cell wall formation	17
<i>Top-level NAC master regulators</i>	17
<i>SCW-associated MYB transcriptional factors</i>	19
<i>Lignin specific MYBs</i>	20
<i>MYB Negative regulators</i>	20
3.3 Hormone control of xylogenesis	21
PART 2 AUXIN REGULATES PLANT GROWTH AND DEVELOPMENT ESPECIALLY IN WOOD FORMATION	24
1. Auxin	24
2. Auxin metabolism	24
2.1 Auxin biosynthesis	24
2.2 Auxin homeostasis	26
3. Auxin signaling	27
3.1 Auxin signal perception and transduction	28

<i>SCF^{TIR1/AFB}-Aux/IAA-ARF signaling pathway</i> - - - - -	-28
<i>SKP2A signaling pathway</i> - - - - -	29
<i>ABP1 system</i> - - - - -	30
3.2 Early (Primary) auxin responsive genes induction - - - - -	-30
<i>SAUR and GH3</i> - - - - -	30
<i>AuxREs cis-element</i> - - - - -	31
4. Key mediators in auxin signaling - - - - -	31
4.1 Aux/IAA family (co-receptor, repressor) - - - - -	31
4.2 TOPLESS (TPL) family corepressors - - - - -	33
4.3 Auxin responsive factors (ARFs) family - - - - -	34
4.4 Protein-protein interactions in auxin signaling - - - - -	34
5. Roles of auxin and its cross-talk with other hormones in the control of wood formation - - - - -	35
5.1 Auxin influx- and efflux-carriers control vascular pattern and wood formation - - - - -	35
5.2 Reading the auxin gradient - - - - -	37
5.3 Identifying new Aux/IAAs and ARFs as mediators of auxin signaling in wood formation - - - - -	37
5.4 The cross-talk between auxin and cytokinins stimulates cambium activity - - - - -	38
5.5 The cross-talk between auxin and gibberellins promotes expansion of cambial derivatives - - - - -	-39
5.6 Interaction between auxin and other hormones (brassinosteroids and strigolactone)- - - - -	-40
PART 3 CRISPR/CAS9 CURRENT ADVANCES AND APPLICATIONS - - - - -	41
1. Introduction - - - - -	41
2. CRISPR/Cas system origin and mechanism - - - - -	44
3. CRISPR/Cas9 vector systems - - - - -	46
3.1 sgRNA expression cassettes- - - - -	51
3.2 Cas9 expression cassettes - - - - -	51
3.3 CRISPR/Cas9 constructs assembly - - - - -	-52
3.4 Delivery of sgRNA and Cas9 expression cassettes- - - - -	52
3.5 Methods to identify targeted mutations - - - - -	55
<i>Non-sequencing methods to detect mutations</i> - - - - -	56
<i>Sequencing methods to detect editing events</i> - - - - -	56
4. Factors affecting the editing efficiency and off-targets in tree - - - - -	57
4.1 Delivery methods and Cas9 optimization - - - - -	-57
4.2 sgRNA parameters affecting editing efficiency - - - - -	-57
<i>sgRNA components</i> - - - - -	57
<i>sgRNA positions and secondary structure</i> - - - - -	57
<i>sgRNA specificity and design</i> - - - - -	57
5. Examples of CRISPR/Cas9 applications in trees - - - - -	-59

5.1 CRISPR-Cas9 to modify Flowering-time	60
5.2 CRISPR-Cas9 to enhance resistance to biotic and abiotic stresses	60
5.3 CRISPR-Cas9 applications for gene functions in trees	60
Objectives of PhD thesis	62
CHAPTER II: IMPLEMENTING THE CRISPR/CAS9 TECHNOLOGY IN EUCALYPTUS HAIRY ROOTS USING WOOD-RELATED GENES	63
Abstract	65
1. Introduction	65
2. Results	67
2.1 Genotyping Revealed High Knock-Down Rate in CCR1 but High Knock-out Rate in IAA9A	67
<i>High Knock-Down Rate for CCR1 Editing</i>	67
<i>High Knock-Out Rate in IAA9A Lines</i>	70
<i>Mutation Spectra Vary Among sgRNA Targets</i>	70
2.2 Phenotyping Revealed Expected Alterations of Lignification in CCR1-Edited Lines	73
<i>Combination of FTIR Spectroscopy and Multivariate Analyses of CCR1-Edited Hairy Root Lines</i>	73
<i>Histochemical Characterization of CCR1-Edited Hairy Root Lines</i>	75
3. Discussion	75
4. Materials and Methods	77
4.1 Plant Material	77
4.2 CRISPR/Cas9 Targeted Mutagenesis System Selection and Pipeline	77
4.3 CRISPR/Cas9 Target Site Selection and sgRNAs Design	78
4.4 CRISPR/Cas9 Constructs Assembly	79
4.5 Agrobacterium Rhizogenes-Mediated Transformation	81
4.6 DNA Isolation, PCR Amplification and Mutation Identification	81
4.7 FTIR Analyses	81
4.8 Histochemical Analysis	82
Reference	82
CHAPTER III: WOOD FORMATION ASSOCIATED FUNCTIONAL CHARACTERIZATIONS OF EGRIAA9A, EGRIAA20 AND EGRARF5 IN HOMOLOGOUS ORGANISM	89
1. Introduction	91
1.1 <i>EgrIAA</i> genes functional characterizations in wood formation	91
1.2 <i>ARF5 (MONOPTEROS)</i> functional characterizations related to vascular tissues	93
1.3 Objective of Chapter III	94
2. Materials and Methods	94
2.1 Plant materials and culture conditions	94

2.2 CRISPR/Cas9 constructions to generate KO mutants of <i>EgrIAA9A</i> , <i>EgrIAA20</i> and <i>EgrARF5</i> - - - - -	95
2.3 Overexpression constructions of <i>EgrIAA9A</i> and <i>EgrIAA20</i> - - - - -	97
2.4 CRISPR/Cas9 editing events detection - - - - -	99
2.5 FT-IR and histology analyses of CRISPR/Cas9 mediated <i>IAA9A</i> , <i>IAA20</i> and <i>ARF5</i> lines - - - - -	100
2.6 Library screening and protein-protein interaction analysis by yeast two-hybrid (Y2H) system - - - - -	100
3. Results - - - - -	100
3.1 Mutations detection using web-based tools (ICE, DSDecode)- - - - -	100
<i>Mutations detection were only achieved by ICE for four ARF5 CRISPR-transgenic lines and one IAA20 line</i> - - - - -	103
<i>Comparison of web-based tools taking results of two lines ARF5_43, ARF5_44 as examples</i> - - - - -	104
<i>ICE analysis for IAA9A transgenic roots (compared with DSDecode and subcloning)</i> - - - - -	105
3.2 FTIR_PLSDA analysis for chemotypes discrimination - - - - -	105
<i>FT-IR_PLSDA analysis for CRISPR/Cas9 edited transgenic ARF5, IAA9A lines and controls</i> - - - - -	106
<i>Identification of wavelength numbers associated polymers</i> - - - - -	108
3.3 Histology analysis of <i>IAA9A</i> _edited lines and <i>ARF5</i> _edited lines - - - - -	110
3.4 Y2H results - - - - -	113
4. Discussion - - - - -	118
GENERAL DISCUSSIONS AND PERSPECTIVES - - - - -	121
REFERENCES - - - - -	127
SUPPLEMENTAL DATA - - - - -	157
SUPPLEMENTAL FIGURES AND TABLES IN CHAPTER I - - - - -	159
SUPPLEMENTAL FIGURES AND TABLES IN CHAPTER II - - - - -	163
SUPPLEMENTAL FIGURES AND TABLES IN CHAPTER II - - - - -	170

List of Figures

Figure I-1	Vascular development in dicot angiosperms.	8
Figure I-2	Xylem developing patterns in stems (endarch) and roots (exarch): xylem in green.	8
Figure I-3	Schematic representation of the radial growth of vascular cambium.	9
Figure I-4	Overview of procambial/cambial cell specification and xylem/phloem cell differentiation	11
Figure I-5	Transversal sections of poplar stem (stained with Calcofluor, auramine O, propidium iodide) showing wood cell formation processes.	12
Figure I-6	Chemical structures of cellulose (A), and building blocks of hemicelluloses (B)	14
Figure I-7	Lignin model structures of three major plant classes: gymnosperm/softwood (a), angiosperm/dicot/hardwood (b), and monocot (c)	15
Figure I-8	The lignin biosynthesis pathway in Eucalyptus adapted from (Carocha et al., 2015).	17
Figure I-9	Transcriptional network regulation of wood cell formation and SCW formation, involving in a battery of transcription factors such as MYB46, MYB83 master switches.	21
Figure I-10	Tryptophan dependent auxin biosynthesis pathway	25
Figure I-11	Homeostasis of IAA in plant cell (provided by PhD thesis of Hua Wang).	27
Figure I-12	The key components in auxin perception and signaling in Arabidopsis	29
Figure I-13	Protein structures of canonical Aux/IAA and ARF	31
Figure I-14	Different hormones distributions across cambial developmental zones	36
Figure I-15	Mechanisms of ZFN, TALEN, and CRISPR/Cas9 (A), generated double strand breaks (DSBs) (B) and possible types of mutations (C).	43
Figure I-16	Timeline of CRISPR/Cas9 genome editing key developments and progress in plants	45
Figure I-17	Pipelines of common-used Cloning methods for CRISPR/Cas9 applications in plants.	54
Figure 1	Genotyping of <i>CCR1</i> and <i>IAA9A</i> transformants and their corresponding prevalent edition types	69
Figure 2	Comparison between FT-IR spectra obtained from controls and <i>CCR1</i> edited lines	74
Figure S1	Wave numbers from FT-IR absorption spectra discriminating between controls and <i>CCR1</i> -edited lines.	165
Figure 3	Comparison of xylem development and lignification of xylem cells between control and <i>CCR1</i> edited lines' roots.	75
Figure S2	Comparison of xylem development and lignification of xylem cells between control and <i>CCR1</i> -edited lines' roots.	165
Figure 4	Pipeline of CRISPR/Cas9 implementation in <i>Eucalyptus grandis</i> hairy roots.	78
Figure 5	CRISPR/Cas9 sgRNA design and mutation detection in <i>CCR1</i> and <i>IAA9A</i> .	79
Figure 6	CRISPR/Cas9 binary vector targeting two loci simultaneously.	80
Figure III-1A	Histochemical analysis of cross sections of <i>OE_EgrIAA9Am</i> transgenic and wild-type <i>Arabidopsis</i> hypocotyls.	92
Figure III-1B	Histochemical analysis of cross sections of <i>OE_EgrIAA20</i> transgenic and wild-type <i>Arabidopsis</i> inflorescence stems.	93
Figure III-S1	Cross sections of inflorescence stems at the basal part showing dramatically inhibited fiber cells of secondary	170

	growth in <i>EgrIAA20</i> overexpressing lines.	
Figure III-S2	The procedure used in Chapter III	171
Figure III-2	CRISPR/Cas9 sgRNA design and mutation detection in <i>IAA9A</i> , <i>ARF5</i> and <i>IAA20</i> .	96
Figure III-3	Scheme illustrating the assembling of Cas9/sgRNA expressing constructs using Golden Gate Cloning.	96
Figure III-4	Scheme illustrating the assembling of overexpression constructs with CDS sequences of <i>EgrIAA9Amm</i> using Golden Gate Cloning (including three levels).	98
Figure III-5	Scheme illustrating the assembling of overexpression constructs with CDS sequences of <i>EgrIAA20m</i> using Golden Gate Cloning (including two levels).	99
Figure III-S3	Editing events of <i>ARF5_17</i> , <i>ARF5_29</i> , <i>ARF5_33</i> , <i>ARF5_43</i> , <i>ARF5_44</i> by ICE analysis	172
Figure III-6	Genotyping of <i>ARF5</i> and <i>IAA20</i> edited lines and their corresponding prevalent edition types and related amino acids changes deleted by ICE.	102
Figure III-7	Different edition types (A) and Indel contributions (B) of <i>ARF5_10</i> analyzed by ICE Synthego.	103
Figure III-S4	Allele mutations of <i>ARF5_43</i> and <i>ARF5_44</i> detected by DSDDecode.	173
Figure III-8	Different inferred different edited types of <i>ARF5_40</i> (A) and <i>IAA20_4</i> (B) obtained by ICE Synthego.	104
Figure III-S5	Combination of FT-IR and PCA (A) /PLS-DA (B) analysis for all <i>ARF5</i> lines	174
Figure III-S6	Combination of FT-IR and PCA (A) /PLS-DA (B) analysis for all <i>IAA9A</i> lines	175
Figure III-9	PLSDA analyses were performed using normalized values of FT-IR absorption spectrum (800-4000cm ⁻¹) obtained from controls and <i>ARF5</i> lines (A), <i>IAA9A</i> edited lines (B) transgenic hairy roots.	107
Figure III-10	Wavelength numbers (WN) from FT-IR absorption spectra discriminating between controls and <i>ARF5</i> lines (A)/ <i>IAA9A</i> edited lines (B).	109
Figure III-11	Comparison of xylem development and lignification of xylem cells especially vessels between transgenic roots of control and <i>IAA9A</i> edited lines.	111
Figure III-S7	The comparison of xylem vessel cells diameters between <i>IAA9A</i> edited lines and control in medium and old developing stages	176
Figure III-12	The median value comparison of xylem vessel cells diameters between all <i>IAA9A</i> edited lines and control	112
Figure III-13	Comparison of xylem development and lignification of xylem cells, especially vessels between transgenic roots of control and <i>ARF5</i> edited lines.	112
Figure III-S8	Comparison of <i>ARF5</i> lines and control under 40x magnification using UV light	177
Figure III-14	Protein-protein interactions of <i>EgrIAA9A</i> (A) and <i>EgrIAA20</i> (B) with potential candidates.	118
Figure III-S9	Vectors maps used in Chapter III	178

List of Tables

Table I-S1	ZFNs, TALENs, CRISPR/Cas9 pros and cons comparison	161
Table I-1	A summary of different optimized CRISPR/Cas9 systems applied in various tree species	47
Table I-S2	CRISPR/Cas9 applications in trees using different targets, associated phenotypes and mutated efficiency	162
Table S1	Details of mutations detected in 3 batches of <i>EgrCCR1</i> transgenic roots and control	166
Table S4	Details of mutations detected in <i>EgrIAA9A</i> transgenic roots and control	166
Table 1	Numbers of edited plants and type of mutations according to target genes	68
Table 2	Editing frequency and position (identified by PCR cloning and subsequent sequencing)	68
Table 3	Majority of altered alleles introduced significant modifications at the protein level	69
Table S2	The edition rate decrease along the transgenic plants age	166
Table S3	Mortality during the transfert from <i>in vitro</i> culture to hydroponic culture	166
Table S5	The edition type comparison of KOEgrIAA9A transgenic plants between DSDecode and sub-cloning methods	167
Table 4	Mutation types	72
Table S6	Wave numbers from FT-IR absorption spectra discriminating between controls and <i>CCR1</i> -edited lines and their related compounds	168
Table S7	The selected sgRNAs sequences designed using CRISPOR (http://crispor.tefor.net/)	169
Table S8	The primers and sgRNA scaffold used for generating sgRNA intermediary vectors.	170
Table S9	The primers used in CRISPR/Cas9 system, the primer names, sequences and the specific uses	171
Table III-S1A	Information of CRISPR/Cas9 mediated alived <i>ARF5</i> (30 fluorescenced roots / 50 in total), <i>IAA20</i> transgenic roots (three alive), and controls	179
Table III-S1B	Information of CRISPR/Cas9 mediated alived <i>IAA9A</i> , and controls	181
Table III-S2	The selected sgRNAs sequences of <i>EgrARF5</i> and <i>EgrIAA20</i> designed using CRISPOR (http://crispor.tefor.net/)	183
Table III-S3A	The primers used in CRISPR/Cas9 system, the primer names, sequences and the specific uses	184
Table III-S3B	The primers used in overexpressed constructions, the primer names, sequences and the specific uses	185
Table III-S4	Mutation types, occurrence, total Indel percentage and presumed knockout score of 11 <i>ARF5</i> _edited lines and one <i>IAA20</i> _edited line	186
Table III-1	Sequencing results of PCR products from Y2H library (secondary xylem) screening colonies using EgrIAA20 as a bait	114
Table III-2	Sequencing results of PCR products from Y2H library (secondary xylem) screening colonies using EgrIAA9 as a bait	116

ABREVIATION

2,4-D:	<u>2,4</u> - <u>d</u> ichlorophenoxyacetic acid
4CL:	<u>4</u> -hydroxycinnamoyl- <u>C</u> oA ligase
4-Cl-IAA:	<u>4</u> - <u>c</u> hloroindole-3- <u>a</u> cetic <u>a</u> cid
ABC:	<u>A</u> TP- <u>b</u> inding <u>c</u> assette
ABCB/D/G:	<u>A</u> TP- <u>b</u> inding <u>c</u> assette subfamily <u>B/D/G</u>
ABP1:	<u>A</u> XIN <u>B</u> INDING <u>P</u> ROTEIN <u>1</u>
ACC:	1- <u>a</u> mino- <u>c</u> yclopropane-1- <u>c</u> arboxylic acid
AD:	<u>A</u> ctivation <u>d</u> omain
ARF:	<u>A</u> uxin response <u>f</u> actor
AuxREs:	<u>A</u> uxin response <u>e</u> lements
Aux/IAA:	<u>A</u> uxin/Indole-3- <u>A</u> cetic <u>A</u> cid
AUX1/LAX:	<u>A</u> XIN RESISTANT <u>1</u> / <u>L</u> IKE- <u>A</u> X <u>1</u>
BD:	<u>B</u> inding <u>d</u> omain
BDL:	<u>B</u> ODEN <u>L</u> OS
BR:	<u>B</u> rassinosteroids
BSA:	<u>B</u> ovine <u>s</u> erum <u>a</u> lbumin
C4H:	<u>C</u> innamate <u>4</u> - <u>h</u> ydroxylase
C3H:	<i>p</i> - <u>c</u> oumarate <u>3</u> - <u>h</u> ydroxylase
CAD:	<u>C</u> innamyl <u>a</u> lcohol <u>d</u> ehydrogenase
CaMV:	<u>C</u> auliflower <u>M</u> osaic <u>V</u> irus
CCoAOMT:	<u>C</u> affeoyl- <u>C</u> oA <u>O</u> - <u>m</u> ethyltransferase
CCR:	<u>C</u> innamoyl <u>C</u> oA <u>r</u> eductase
cDNA:	<u>C</u> omplementary <u>d</u> eoxyribonucleic <u>a</u> cid
CDS:	<u>C</u> oding <u>s</u> equence
CesA:	<u>C</u> ellulose <u>s</u> ynthases
CK:	<u>C</u> ytokinin
CLE:	<u>C</u> LAVATA3/ <u>E</u> MBRYO SURROUNDING REGION
CML:	<u>C</u> ompound <u>m</u> iddle <u>l</u> amella
COMT:	<u>C</u> affeic acid <u>O</u> - <u>m</u> ethyltransferase
CRISPR:	<u>C</u> lustered <u>R</u> egularly <u>I</u> nterspaced <u>S</u> hort <u>P</u> alindromic <u>R</u> epeats
crRNA:	<u>C</u> RISPR <u>R</u> NA
CSC:	<u>C</u> ellulose synthase <u>c</u> omplex
CSE:	<u>C</u> affeoyl <u>s</u> hikimate <u>e</u> sterase
CTAB:	<u>C</u> etyl <u>t</u> rimethylammonium <u>b</u> romide
CTD:	<u>C</u> arboxy- <u>t</u> erminal dimerization <u>d</u> omain
DBD:	<u>D</u> NA <u>b</u> inding <u>d</u> omain

DSB:	<u>D</u> ouble <u>s</u> trand <u>b</u> reak
DSDecode:	<u>D</u> egenerate <u>s</u> equence <u>d</u> ecoding
EAR:	<u>E</u> RF-associated <u>a</u> mphiphilic <u>r</u> epression
EIN3D:	<u>E</u> thylene <u>i</u> nsensitive <u>3D</u>
EMNs:	<u>E</u> ngineered <u>m</u> eganuclease
ER:	<u>E</u> ndoplasmic <u>r</u> eticulum
ERF:	<u>E</u> thylene <u>r</u> esponse <u>f</u> actor
F5H:	<u>F</u> erulate <u>5</u> - <u>h</u> ydroxylase
FT:	<u>F</u> LOWERING <u>L</u> OCUS <u>T</u>
FT-IR:	<u>F</u> ourier <u>T</u> ransformed <u>I</u> nfra- <u>r</u> ed spectroscopy
G:	<u>G</u> uaiacyl
GA:	<u>G</u> ibberellin
GH3:	<u>G</u> RETCHEN <u>H</u> AGEN <u>3</u>
GT:	<u>G</u> lycosyl <u>t</u> ransferase
H:	<i>p</i> - <u>h</u> ydroxyphenyl
HDR:	<u>H</u> omology- <u>d</u> irected <u>r</u> epair
HD-ZIP III:	Class <u>III</u> <u>h</u> omeodomain-leucine <u>z</u> ipper
HRM:	<u>H</u> igh- <u>r</u> esolution <u>m</u> elting assay
IAA:	<u>I</u> ndole-3- <u>a</u> cetic <u>a</u> cid
IAM:	<u>I</u> ndole-3- <u>a</u> cetamide
IAOX:	<u>I</u> ndole-3- <u>a</u> ctaldoxime
IBA:	<u>I</u> ndole-3- <u>b</u> utyric <u>a</u> cid
ICE:	<u>I</u> nference of <u>C</u> RISPR <u>E</u> dits
IFL1/REV:	<u>I</u> NTERFASCICULAR <u>F</u> IBER <u>L</u> ESS/ <u>R</u> EVOLUTA
IPA:	<u>I</u> ndole-3- <u>p</u> yruvic <u>a</u> cid
LBD:	<u>L</u> ateral organ <u>b</u> oundaries <u>d</u> omain
MCTU:	<u>M</u> ulti- <u>c</u> omponent <u>t</u> ranscriptional <u>u</u> nit system
MP:	<u>ARF5</u> / <u>M</u> ONO <u>P</u> TEROS
NAA:	<u>N</u> aphthalene-1- <u>a</u> cetic <u>a</u> cid
NAC:	<u>N</u> AM/ <u>A</u> TAF/ <u>C</u> UC transcription factors
nCas9:	<u>N</u> ickase <u>C</u> as <u>9</u>
NHEJ:	<u>N</u> on- <u>h</u> omologous <u>e</u> nd <u>j</u> oining
NLS:	<u>N</u> uclear <u>l</u> ocalization <u>s</u> ignal
NST:	<u>N</u> AC <u>S</u> ECONDARY <u>W</u> ALL <u>T</u> HICKNING PROMOTING FACTOR
ORF:	<u>O</u> pen <u>r</u> eading <u>f</u> rame
PAA:	<u>P</u> henyl <u>a</u> cetic <u>a</u> cid
PAL:	<u>P</u> henylalanine <u>a</u> mmonia <u>l</u> yase
PAM:	<u>P</u> rotospacer <u>A</u> djacent <u>M</u> otif

PAT:	<u>A</u> uxin <u>p</u> olar <u>t</u> ransport
PCA:	<u>P</u> rincipal <u>c</u> omponent <u>a</u> nalysis
PCD:	<u>P</u> rogrammed <u>c</u> ell <u>d</u> eath
PCR:	<u>P</u> olymerase <u>c</u> hain <u>r</u> eaction
PDS:	<u>P</u> hytoene <u>d</u> esaturase
PIN:	<u>P</u> IN-FORMED
PILS:	<u>P</u> IN-Like transporters
PGP:	<u>P</u> - <u>G</u> LYCOPROTEIN
PLS-DA:	<u>P</u> artial <u>L</u> east <u>S</u> quare <u>D</u> iscriminant <u>A</u> nalysis
PXY:	<u>P</u> HLOEM INTERCALATED WITH <u>X</u> YLEM
RAM:	<u>R</u> oot <u>A</u> pical <u>M</u> eristem
RD:	<u>R</u> epression <u>d</u> omain
S:	<u>S</u> yringyl
SAUR:	<u>S</u> MALL <u>A</u> XIN <u>U</u> PREGULATED <u>R</u> NA
SAM:	<u>S</u> hoot <u>A</u> pical <u>M</u> eristem
SCF:	<u>S</u> KP1- <u>C</u> ullin- <u>F</u> -box
SCW:	<u>S</u> econdary <u>c</u> ell <u>w</u> all
sgRNA:	<u>S</u> ingle <u>g</u> uide <u>R</u> NA
SNP:	<u>S</u> ingle <u>N</u> ucleotide <u>P</u> olymorphism
SL:	<u>S</u> trigolactone
SSNs:	<u>S</u> equence-specific <u>n</u> ucleases
TAM:	<u>T</u> ryptamine
TALENs:	<u>T</u> ranscription <u>a</u> ctivator- <u>l</u> ike <u>e</u> ffector <u>n</u> ucleases
TE:	<u>T</u> racheary <u>E</u> lements
TF:	<u>T</u> ranscription <u>f</u> actor
TFL:	<u>T</u> ERMINAL <u>F</u> LOWER <u>1</u>
TIR1/AFBs:	<u>T</u> RANSPORT <u>I</u> NHIBITOR <u>R</u> ESPONSE1/ <u>A</u> XIN SIGNALING <u>E</u> - <u>B</u> OX PROTEINS
TPL:	<u>T</u> OPLESS
tracrRNA:	Trans-encoded small RNA
tRNA:	<u>T</u> ransfer <u>R</u> NA
VND:	<u>V</u> ascular-related <u>N</u> AC- <u>d</u> omain
WAT1:	<u>W</u> ALL <u>A</u> RE <u>T</u> HIN1
Y2H:	<u>Y</u> east <u>t</u> wo- <u>h</u> ybrid
ZFNs:	<u>Z</u> inc- <u>f</u> inger <u>n</u> ucleases

Objectives and Organization of the manuscript

The overall objective of my PhD was to implement the CRISPR genome editing tool to generate loss-of-function *Eucalyptus* mutants in transgenic hairy roots and to use this technology to functionally characterize auxin-signalling transcription factors Aux/IAs and ARFs potentially involved in the regulation of wood (secondary xylem) formation.

The thesis is organized in three Chapters. **Chapter I** consists in a bibliographic review divided in three main parts. The first one is dedicated to *Eucalyptus* and wood formation, as well as secondary cell walls and its transcriptional regulation. In the second part, the main features of auxin biology (synthesis, homeostasis and signaling), and its roles in the regulation of plant growth and development, with a special focus on wood formation are presented. The third part consists in an overview of the CRISPR/Cas9 technology, with a special emphasis of its implementation and applications in trees.

Chapter II is presented in the form of an article (**Implementing the CRISPR/Cas9 Technology in *Eucalyptus* Hairy Roots Using Wood-Related Genes**) published in “International Journal of Molecular Science” April 2020. As proofs-of-concept, we chose as target genes *Cinnamoyl-CoA Reductase1 (CCR1)*, a key lignin biosynthetic gene and *IAA9A* an auxin dependent transcription factor of *Aux/IAA* family. We designed two guide RNAs for each gene to generate CRISPR/Cas9-mediated mutations. Editions were detected in almost all transgenic lines with different allele-editing rates. The transgenic lines of *CCR1* were further analyzed by spectroscopic methods combined to multivariate analyses (FT-IR_PLSDA) as well as histochemical analyses, confirming the phenotypes induced by *CCR1*-deficiency, *i.e.* decreased lignification and irregular xylem vessels. Although the efficiency of editing could be improved, the method described here is already a useful tool to functionally characterize *Eucalyptus* genes.

Chapter III focus on the functional characterization of auxin-related candidate genes *EgrIAA9A*, *EgrIAA20* and *EgrARF5* in *E. grandis* transgenic hairy roots, using gain-of-function (overexpression under the constitutive 35S CamV promoter) and loss-of-function of mutants (CRISPR-cas9). We obtained CRISPR/Cas9 edited plants for all three targeted candidate genes. A fast-chemical screening by FT-IR_PLSDA analysis showed that the transgenic lines separated clearly from the wild-type control based on the global chemical composition analysis. Among them, bigger vessel cells and accelerated xylem development were identified in CRISPR/Cas9 generated *EgrIAA9A*_edited lines. However, no obvious phenotypes were detected for xylem formation in *EgrARF5* edited lines in our experimental conditions. We also generated overexpressing lines for *IAA9A* and *IAA20*, unfortunately we lost all of them during the covid lockdown, as well as most plants of CRISPR/Cas9 generated *EgrIAA20* lines. Finally, the potential partners of *EgrIAA9A* and *EgrIAA20* (as a bait) were screened out from *Eucalyptus* developing xylem cDNA library using Yeast Two Hybrid (Y2H). Interestingly, we found that *EgrIAA9A* is the main protein partner of *EgrIAA20*, suggesting that *EgrIAA20* forms complex with *EgrIAA9A* to regulate the transcriptional regulation during wood formation. We further using Y2H to validate the protein-protein interaction among our *Eucalyptus* Aux/IAA and ARF members potentially involved in wood formation.



Chapter I
Bibliographic review



Part 1 Eucalyptus and wood formation

1. Eucalyptus

The genus *Eucalyptus* belongs to a basal Rosid lineage (Myrtales order, Myrtaceae family) which evolved mostly in the isolation of the Australian continent, Tasmania and nearby islands. Therefore, they represent independent evolutionary experiments over 22 million years for studies of the woody perennial lifestyle (Myburg et al., 2014). The *Eucalyptus* genus is highly diverse and displays significant adaptability and phenotypic plasticity. These long-lived, sclerophyllous evergreen and flowering hardwood trees are well adapted to diverse climates from tropical rainfall to temperate semiarid zones, though only very few species are tolerant to coldness (Wiltshire, 2004) mainly due to the absence of dormancy. They can also grow on lean soils. Eucalypts come in a great range of shapes and sizes – from tall trees (Up to 100m for the tallest) to small multi-stemmed shrubs. Among its thirteen subgenera, the most important is *Symphyomyrtus*, comprising about 474 species including the ones of commercial importance (Grattapaglia et al., 2015). Only 20 or so of those *Eucalyptus* species have been extensively used in commercial forests plantation given they are fast and easy to grow, provide high forest productivities and deliver wood with high density, durability and interesting fibre properties (Grattapaglia et al., 2015). Because of their wide adaptability, fast growth and multipurpose uses of their wood, eucalypts are the most widely planted forest trees worldwide (Myburg et al., 2007), growing in over 100 countries such as Brazil, China, India, South Africa, Portugal, and covering more than 20 million hectares. China's eucalypt plantations make around 27% substantial contribution of a total of annual domestic timber output.

Global carbon cycle of forest contains carbon sequestration in tree biomass and soil through photosynthesis and respiration, and forestry timber sustainable harvesting contributes a significant pathway to carbon sequestration. In Brazil, along with the increased *Eucalyptus* plantation area and productivity, carbon contribution of *Eucalyptus* increased 13% (from 58% to 71%) in 25 years (Sanquetta et al., 2018). Because of the highly efficient C3 photosynthesis occurring in eucalypts, carbon sequestration is massive and will likely increase along with elevated ambient CO₂ and rising temperature (Ghannoum et al., 2010). This strong response to CO₂ contributes *Eucalyptus* to become a crucial tree for carbon sink and an environmental protector in the context of global warming.

Forests cover a third of world's total land area, and over a half of our forest conservation contributes to human activities, such as wood production, food and other forest products. The trend of global wood sustainable harvests will be up to 80% by planted forests expansion in 2030 (Jim Carle and Peter Homgren, 2008), comprised of softwood (60% supply by *Pinus*) and hardwood (prefer *Eucalyptus* and *Acacia*) (Ramage et al., 2017). *Eucalypt* native and domesticated forests are an important source of high-quality woody biomass for many wood products (pulp and paper, sawn timber, composites, fuelwood), which are dependent on variable key wood properties (wood density, hardness, color and shape, chemical composition, heartwood, and so on). Pulp and paper production is in nowadays the main purpose of industrial plantations (Greaves et al., 1997). Recently, the urgent need for fuels resources and applications gave rise to emerging biofuel as an alternative choice, hence bioethanol obtained from eucalyptus is a crucial feedstock (Shepherd et al., 2011; Verma et al., 2016). *Eucalyptus* have extending rotations such as high-value sawlogs with maximum biomass of clear wood, and bark

residues of *E.grandis* and *E.grandis* x *urophylla* for biofuel production (Lima et al., 2014). Mechanisms of lignin deposition and lignin changes along with tree maturation have been studied in *E.globulus* young and mature wood (Rencoret et al., 2011). In addition, non-destructive techniques for wood properties assessment (e.g., FT-IR spectroscopy) are required for solid wood products and researches of traits heritability (Raymond, 2000).

The *Eucalyptus grandis* genome 'BRASUZI' sequence has been released (Myburg et al., 2014), which provided insights into candidate genes and regulators involved in wood formation. Genome-wide analysis of lignin biosynthetic genes identified 17 candidate genes associated with *E. grandis* xylem lignification (Carocha et al., 2015). Also, members of several families of transcription factors, such as MYB, NAC, AFR and Aux/IAA, were identified to be involved in secondary cell wall (SCW) formation and wood formation in eucalypts (Yu et al., 2014, 2015; Soler et al., 2015; Hussey et al., 2015; Shinya et al., 2016). Further investigations of biosynthesis genes and transcription factors associated with regulation of wood formation in *Eucalyptus* will be described in the corresponding paragraphs, later in this chapter. Transformation using *Agrobacterium tumefaciens* has been implemented for several *Eucalyptus* species or hybrids (Tournier et al., 2003; Girijashankar, 2011; de la Torre et al., 2014; Plasencia et al., 2016) and recently freeze-resistant transgenic eucalyptus have been recently obtained in our lab (Cao et al., 2020). However, this is a very long and tedious process that do not allow easy functional characterization of candidate genes. To overcome this difficulty, a protocol using *Agrobacterium rhizogenes* has been implemented in our lab (Plasencia et al, 2016) and transgenic hairy roots are currently being used to functionally characterize candidate genes (Soler et al., 2017; Dai et al., 2020). For my PhD study, we utilize *Eucalyptus* as our experimental species to investigate the wood formation in tree.

2. The vascular system

From the numerous adaptations that land plants have developed during evolution, the acquisition of the vascular system some 400 million years ago have been a crucial event ensuring their successful earth colonization. The vascular system is thus an evolutionary innovation which enables delivering water and nutrients, as well as mechanical support to ensure plants growing tall to get more access of sun light for their photosynthesis and growth. It is composed of two main tissues, xylem and phloem. Xylem cells play crucial roles in water transport and mechanical support of the entire plant. Phloem is essential to the transport of photosynthate from photosynthetic tissues (leaves, source tissues) to developing tissues (sink tissues).

2.1 The vascular system during primary growth

During embryogenesis, at the globular stage, protoderm cell division generates vascular stem cells. Provascular tissue presents a vascular patterning geometry at the heart stage, and it is specified into Shoot Apical Meristem (SAM) and Root Apical Meristem (RAM) (Scheres et al., 1994; De Rybel et al., 2014) respectively, while no differentiation occurs yet (shown in Figure I-1a). The SAM and RAM are responsible for primary growth, i.e. extension of the shoots and roots, respectively. Later on, but still on embryo stage, vascular plants develop a lateral meristem called procambium. During a first phase of growth in the acropetal direction, these meristems produce the primary plant body including the primary vasculature composed of primary xylem and phloem. Procambium cells undergo two types of division: the periclinal

divisions parallel to the plant axis/surface, give rise to phloem and xylem precursor cells, whereas the anticlinal divisions perpetuate the procambium tissue along the plant axis giving rise the enlargement of procambium tissues themselves. In some species, primary vasculature, begin to differentiate from procambium in mature embryos, but in many others, vascular tissue differentiation only starts after the seed germinates (Evert, 2006).

In leaves, phloem and xylem generates towards abaxial and adaxial surfaces, respectively. In stems, primary vascular tissues are organized in discrete collateral vascular bundles separated by parenchyma cells, while in roots, the vascular tissue is organized in a bi- or multi-symmetric pattern (Figure I-1).

In both root and shoot vasculature, the most common organisation is, from the outside to the inside of the organ: phloem, procambium and primary xylem cells. Primary xylem is formed by differentiation of the procambium into protoxylem and metaxylem. During primary growth, strands of primary xylem are found in stems and roots, presenting narrow, small vessel cells of protoxylem and subsequent extended metaxylem which has larger size (Figure I-2). Among four patterns of protoxylem and metaxylem distribution, exarch development pattern (xylem develops centripetally) is found in vascular plant roots, showing protoxylem near pericycle and metaxylem close to pith. Another development pattern (xylem develops centrifugally) is the endarch one that exists in stems of seed plants, showing opposite distribution of exarch pattern of protoxylem and metaxylem (Figure I-2). Notably, secondary cell wall (SCW) is deposited in protoxylem in elongating organs and metaxylem in non-elongating organs by different patterns to provide mechanical support to xylem conduits (Ye and Zhong, 2015).

Roots differ from stems by the absence of pith, and the presence of two additional single-cell ring-shape layers, the endodermis and the pericycle. Endodermis will further form the Casparian strip, which is essential for impairing the diffusion of solutes from the inside to the outside of roots. The pericycle is a single layer of meristematic cells, the origin of lateral roots and have a role in secondary growth of roots. The vascular tissue in young roots presents a bisymmetric pattern, the primary xylem forms a central axis, and two flanking primary phloem poles (Figure I-1) (Nieminen et al., 2015).

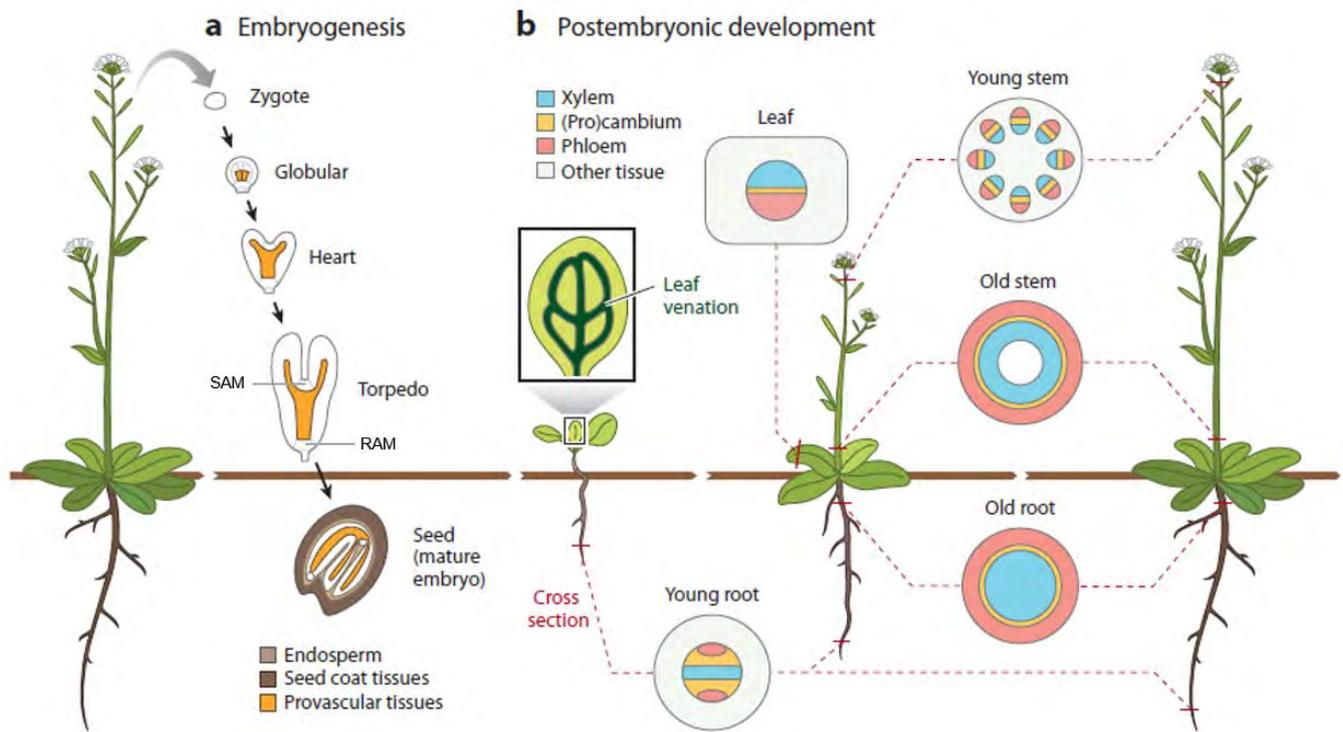


Figure I-1. Vascular development in dicot angiosperms (Adopted from (Ruonala et al., 2017)). (a) The provascular tissue (orange) is formed during embryogenesis. (b) The vasculature developing into different structures in organs and tissues (shoots, roots, leaves), starting from shoot apical meristem (SAM) and root apical meristem (RAM), and later develops into xylem (inner) and phloem (outer), throughout whole plant organs and tissues.

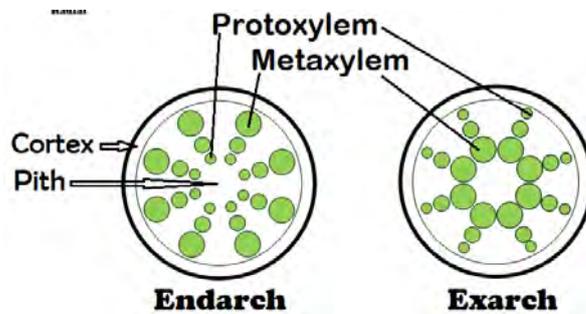


Figure I-2. Xylem developing patterns in stems (endarch) and roots (exarch): xylem in green.

2.2 The vascular system during secondary growth

During (primary growth), plants grow in the acropetal direction thanks to the activity of root and shoot apical meristems. With the notable exception of the monocotyledons, many vascular plants undergo a second phase of growth (secondary growth), which implies growth in diameter or radial growth. This secondary growth can be limited to the hypocotyls as for instance in *Arabidopsis* or can be particularly important like in trees where it produces large amount of secondary xylem (wood).

During secondary growth in angiosperm stems, procambium cells will give rise to the fascicular cambium whereas the interfascicular cambium is thought to arise through the de novo recruitment of interfascicular parenchyma cells (Schuetz

et al., 2013). In roots, the interfascicular cambium differentiates from pericycle cells. The junction of these two populations of cambial cells, makes the entire circular vascular cambium of mature tree trunks and roots (Figure I-1). The vascular cambium is a cylindrical meristem which contains both division and differentiation zones. Two types of divisions exist in cambial stem cells: anticlinal division (perpendicular to surface and axis) which extend the cambial rings laterally, and periclinal division (parallel to surface) which increases radial cell files for differentiation to secondary xylem inwards and to phloem outwards. So the vascular cambium not only proliferates to form new cambium cells to renew meristematic cells, but also differentiates to form secondary xylem inwards, or secondary phloem outwards (Groover and Robischon, 2006) in stems and roots/hypocotyls. There are two types of cambium stem cells: the larger fusiform initials and the small isodiametric ray initials. Fusiform initials are responsible for longitudinally aligned cells (such as vessels, fibres, sieve elements etc) with vacuolated cytoplasm generated by periclinal division, and ray initials are responsible for transversely aligned cells (ray parenchyma cells) connecting secondary phloem and xylem in radial cell system (Evert, 2006). The cambial cells undergo producing destined mother cells in differentiation zone, to form either xylem mother cell or phloem mother cell, and xylem and phloem elements afterwards (Figure I-3). Using lineage tracing analysis and sector analysis, vascular cambium was featured as a single layer of true cambial initials (being able to divide both anti- and periclinally) in *Arabidopsis* and poplar (Bossinger and Spokevicius, 2018; Shi et al., 2019).

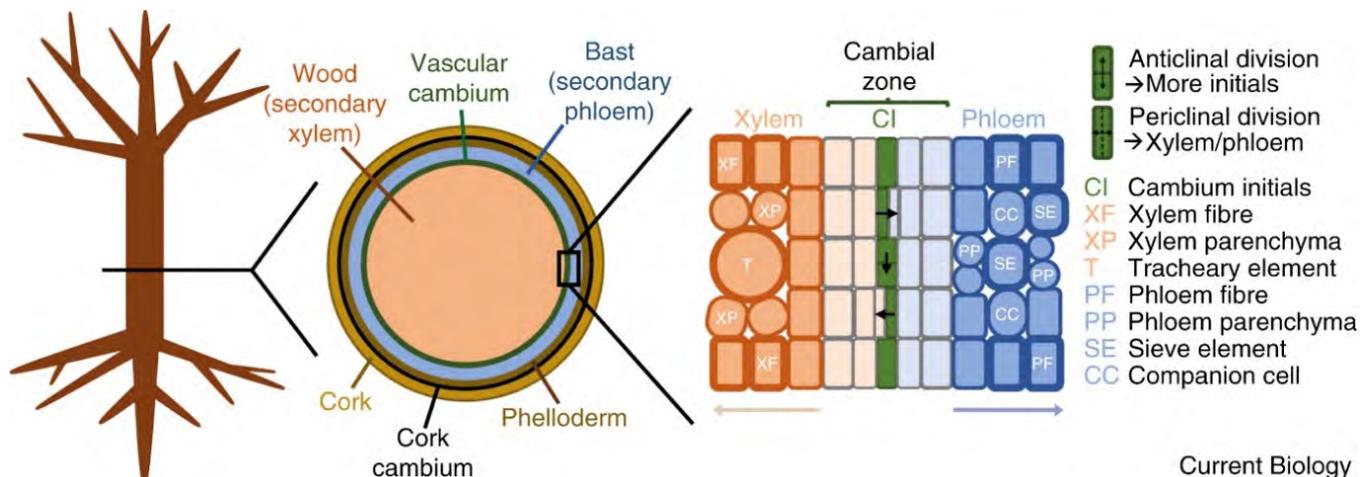


Figure I-3. Schematic representation of the radial growth of vascular cambium (Tonn and Greb, 2017). The secondary vasculature is produced by vascular cambium initials in central cambial zone. The cork cambium produces phelloderm inwards and cork outwards. Periclinal division of vascular cambium gives rise to differentiation into xylem or phloem. Anticlinal division produces more initials to increase the circumference of cambial zone.

The stem cell fates are determined by phytohormones and meristematic capacity, that will be maintained by xylem identity cells (initial xylem) directing adjacent vascular cambial cells to act as stem cells (Smetana et al., 2019), and will be complemented by pericycle division to form new vascular cambium (Chiatante et al., 2018). Except for peptide signaling pathways such as TDIF-PXY signaling which promotes cambium proliferation and inhibits xylem differentiation, the long-distance transported hormone signaling molecules involved in cambial cell proliferation and differentiation have been revealed in many researches, for example the pesipetal transported auxin. These genetic controls of vascular cambium

activity have been reviewed in (Nieminen et al., 2015; Ruonala et al., 2017). Vascular cambium activity is also affected by environmental factors (Fischer et al., 2019). In addition, hormone signal molecules are crucial for communication between primary and secondary meristems (Wang, 2020).

3. Wood formation (xylogenesis)

Wood (secondary xylem) is a sequential complex developmental process which starts in vascular cambium. Secondary xylem is produced after phloem, while has the larger quantity than phloem at the end of season in perennial tree stems. Fusiform initials of cambial cells divide periclinally to give rise to long narrow cells which differentiate into conducting cells, comprising tracheid elements and secondary xylem fibers towards inside and sieve elements (in phloem) towards outside (Figure I-4). Ray initials differentiate into small short ray parenchyma cells which provide radial transport of water and minerals between xylem and phloem.

Cambial cells periclinal division towards pith can be activated and subsequently undergo cell expansion to form specialized secondary xylem with final size, along with secondary cell wall deposition and programmed cell death (Figure I-5). There are three types of wood cells: tracheary elements (TEs, tracheids or vessels), xylem parenchyma cells and xylem fiber cells (Figure I-5). Of these, TEs (dead cells when matured) are the main type of secondary xylem acting as well-conducting tubes for transport throughout plants (Turner et al., 2007). The cell expansion of vessels is rapid to give large diameter and is limited in axial orientation as compared to fibers which have moderate expansion. Once vessels and fibers final size is reached, secondary cell wall (SCW) is deposited and then lignified, becoming inflexible and impermeable. Though SCW deposition is the dominant step of wood formation, programmed cell death (PCD) is also essential followed or simultaneously occurring with SCW lignification (Derbyshire et al., 2015). Programmed vessel cells death occurs rapidly starting by vacuolar integrity loss, while fiber cells death is a gradual degradative process in nucleus and cytoplasm (Courtois-Moreau et al., 2009). Parenchyma cells are dead at the end along with sapwood converts into heartwood. Due to complex tissues where TEs forming, TEs *in vitro* provides advantages of high percentage, easy accessibility, and molecular information about cell wall such as biochemical and gene expression changes (Devillard and Walter, 2014). Lignification is the biological process of lignin deposition in cell wall, which mainly includes cell-autonomous lignification such as vessel and fiber cells, where monolignols are produced and deposited by differentiating cells. Non-cell-autonomous lignification occurs in parenchyma cells, where monolignols produced by neighbored cells are transferred to non-autonomous-cells (Smith et al., 2017).

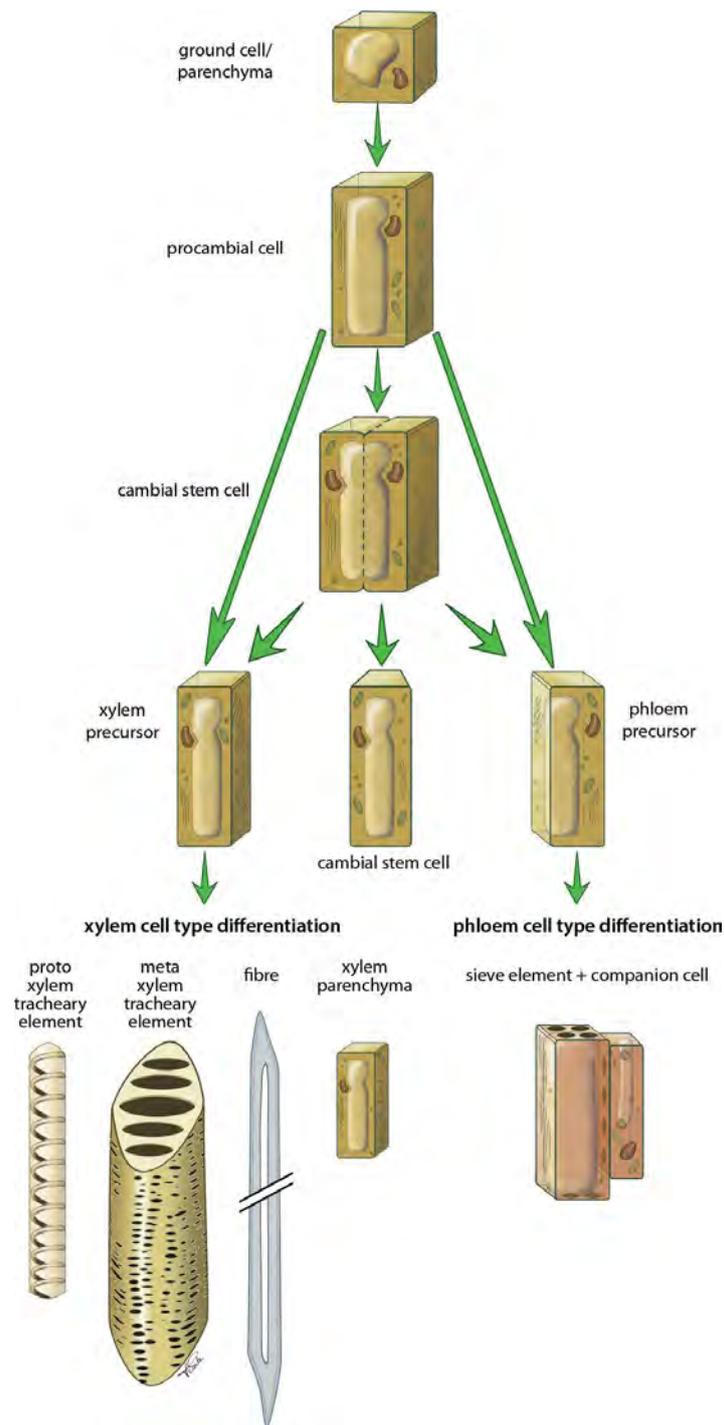


Figure I-4. Overview of procambial/cambial cell specification and xylem/phloem cell differentiation (Extracted from (Schuetz et al., 2013)).

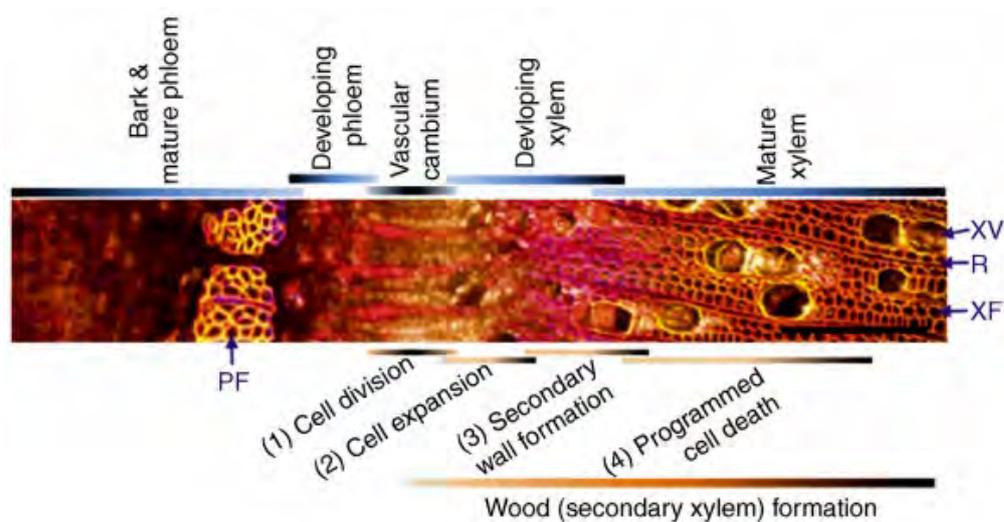


Figure I-5. Transversal sections of poplar stem (stained with Calcofluor, auramine O, propidium iodide) showing wood cell formation processes (Ko et al., 2016). The bars above show sequential wood formation stages. PF, phloem fiber; XV, xylem vessel; XF, xylary fiber; R, ray cell. Scale bars represent 200 μ m.

3.1 Secondary cell wall (SCW) structure and composition

After expansion of differentiated specialized cell types, secondary cell wall (SCW) is deposited between primary cell wall (PCW) and plasma membrane. Primary cell wall is thin, flexible and extensible layer of the cell wall composed of cellulose (50%), pectin (20-30%) and hemicellulose (20-30%). The PCW is a unique fabric that is strong but usually thin, flexible, and capable of both plastic and elastic extension. The primary wall is the cellulose-containing layer laid down by cells that are dividing and growing. To allow for cell wall expansion during growth, primary walls are thinner and less rigid than those of cells that have stopped growing. PCWs have functions of giving cells stability, determining shapes, protection, etc. Different from primary cell walls, SCWs function are to provide mechanical strength, water-proofing for water conduction, and barrier for protection against biotic and abiotic stresses. SCWs contain three major components: cellulose (40-50%), hemicelluloses (xylan and glucomannan (25%)), and lignin (25-35%). SCWs are mainly found in tracheary elements (TEs) and fibers, as well as in some other specialized tissues (Mauseth, 1988). The proportion of three components in SCWs is variable among different vascular plants in different developmental stages, in different cell types, and even changes when facing environmental stresses.

SCWs contain three distinct layers S1, S2 and S3 showing different cellulose content, polymerization, crystallinity and orientation (Timell, 1967; Müller et al., 2006; Mellerowicz and Sundberg, 2008). The S layers are abundant with xylan and cellulose while has less lignin than in compound middle lamella (CML) (Donaldson et al., 2001). The middle lamella serves as a cementing layer between the primary walls and adjacent cells. It is mainly composed of pectic polysaccharides, lignin, and a small amount of proteins. It is the first layer that is formed, which is deposited at the time of cytokinesis. Each S layer has special cellulose microfibril either be aligned irregularly, or in particular angle. In Eucalyptus bleached fibres, reproducible xylan (hemicellulose component) distribution pattern across SCW showed more xylan quantities in S1 and S3 than in S2 (Lekha et al., 2017).

Cellulose

Cellulose microfibril is the load bearing unit in both primary and secondary cell walls consisting of several β -1,4 linked chains of D-glucopyranosyl (Glu) residues (Doblin et al., 2002). It is a polysaccharide which is the most abundant component in SCWs and the first abundant biopolymer on earth, accounting for 40-50% of wood (structure in Figure I-6). In Arabidopsis, SCW cellulose synthases (CESAs) include three non-redundant proteins (CESA4/IRREGULAR XYLEM5b (IRX5), CESA7/IRX3 and CESA8/IRX1) belonging to the glycosyltransferase (GT) family 2 (Taylor et al., 1999, 2000, 2003). They are located at the plasma membrane, colocalize with cortical microtubules bands in older vessels and are functionally conserved in many vascular plants. Loss of function mutants of a single Cesa resulted in xylem morphology defects, showing collapsed xylem cells (Taylor et al., 2000). Cellulose synthases are integral plasma membrane proteins organized in hexameric rosette complexes which contain six-fold catalytic subunit trimers (Nixon et al., 2016) with different contributions of each CESA for cellulose synthesis (Kumar et al., 2018). Hampered cellulose synthesis genes led to little effect in the other two polymers (xylan and lignin) contents in SCWs (Turner and Somerville, 1997; Zhong et al., 2003). CESAs belong to multigene families, three PtrCesA1, PtrCesA2, PtrCesA3 in aspen (*Populus tremuloides*) are specifically highly expressed in xylem tissue during SCW deposition (Joshi et al., 2004), and more than 12 Cesa transcripts were identified in differentiated xylem undergoing SCW (Suzuki et al., 2006). Two types of cellulose synthase complexes (PtrCesA7A, PtrCesA3D) were further identified to influence crystalline multilaminar cellulose structure in wood SCW (Xi et al., 2017). Overexpression of PmCesA2 in hybrid poplar resulted in thickening SCW and increased cellulose and lignin content (Maleki et al., 2020). In addition, the membrane-bound β -1,4-endoglucanase KORRIGAN (KOR) is also required for proper cellulose-hemicellulose network synthesis, and the *kor* mutants of Arabidopsis exhibited decreased cellulose content in primary cell wall (Nicol, 1998; Sato et al., 2001), incomplete cell walls, defect of cell plate formation (Zuo et al., 2000), and irregular xylem vessel development (Szyjanowicz et al., 2004). The downregulation of *PdKOR* in poplar displayed increased crystalline cellulose but decreased polymerization of cellulose, reduced plant growth, and altered carbon allocation and biomass composition (Kalluri et al., 2016; Bali et al., 2016).

Hemicelluloses

Hemicelluloses account for around 20-35% of dry biomass and has amorphous structure with little strength (Figure I-6). Hemicelluloses consist of xylans, mannans, xyloglucans, and β -1,3;1,4-glucans, and have various structural features (as reviewed in (Zhou et al., 2017)), which are different from cellulose but prevents its flocculation. The β -(1,4)-linked xylose homopolymer xylan is the main hemicellulose of SCW in angiosperms and is one of the most abundant naturally occurring polymers (Ebringerová and Heinze, 2000; Scheller and Ulvskov, 2010). Using genetic approaches, genes associated with biosynthesis and structure of hemicellulose have influence growth and development in mutants as reviewed in (Pauly et al., 2013). Many glycosyltransferase (GT) genes located in Golgi membrane were identified to participate in xylan biosynthesis. A series of IRX, PARVUS, F8H genes in Arabidopsis and functional orthologs GT genes in poplar (Zhou et al., 2006; Lee et al., 2011; Ratke et al., 2018; Busse-Wicher et al., 2016; Ratke et al., 2018) were thought to play roles during SCW not only in linear β -(1,4)-linked xylose backbone formation, elongation and decoration pattern, but also

affected growth and lignocellulose saccharification. Moreover GT43B promoter was utilized specifically for secondary wall modification.

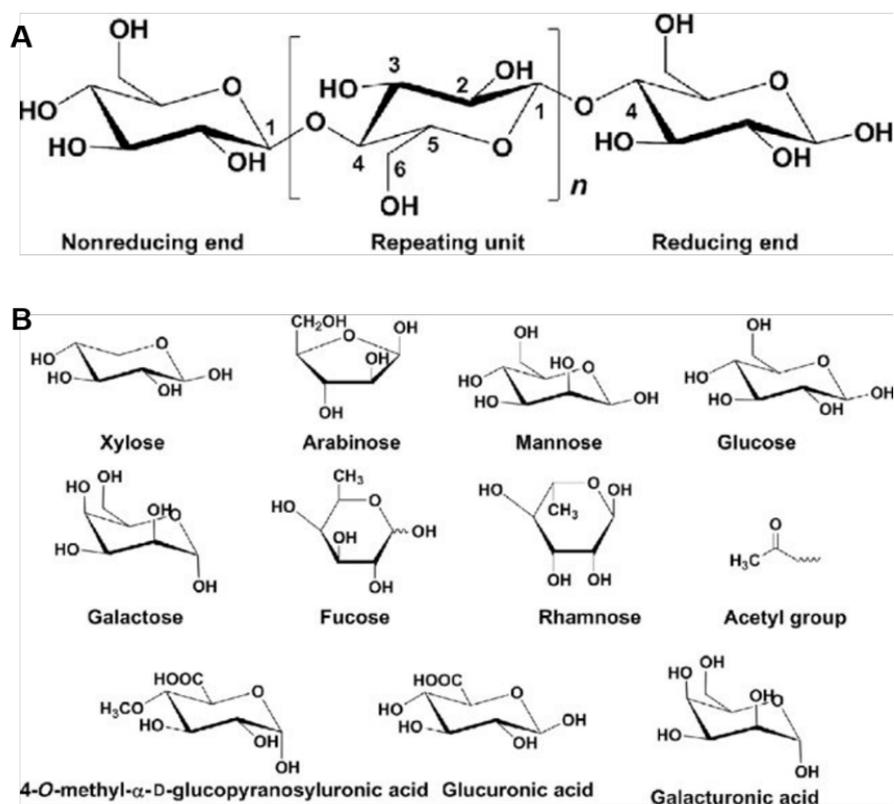


Figure I-6. Chemical structures of cellulose (A), and building blocks of hemicelluloses (B) (Zhou et al., 2016).

Lignin

The characteristic of xylem SCW is massive lignin deposition which provides rigidity and impervious to thickened cell walls. Lignin is a complex tridimensional aromatic polymer composed of phenyl propane units. It plays roles in water conduction, mechanical support and protect cell wall polysaccharides from pathogens. Lignin is mainly made of hydroxycinnamyl alcohols (or monolignols), coniferyl alcohol, sinapyl alcohol, and p-coumaryl alcohol, and this cross-linked complex gives rise to guaiacyl (G), syringyl (S) and p-hydroxyphenyl (H) units, as well as other aromatic monomers (Vanholme et al., 2019). Lignin model structures of gymnosperm, angiosperm (some monocots included) were shown in (Ralph et al., 2019) (Figure I-7). In different species and different wood cells, the proportions of G, H and S units are distinct, for instance, S lignin is prevalent in angiosperms but absent in gymnosperms. Lignin containing S units are less condensed compared to H and G lignin due to the absence of very strong [β -5', 5'-5', 4'-O-5'] C-C linkages (Ralph et al., 2004). Dicots lignin mainly contains G and S units (Weng and Chapple, 2010), and wood vessel cells are rich in G lignin while fibres contain S-G lignin (Lourenço et al., 2016). G lignin acts for tracheary elements (TEs) cell wall strength and conduction, while S lignin plays roles for derived traits such as defense (Renault et al., 2019). Lignin biosynthesis, monolignol contents and composition are affected by various factors such as plant growth and development, metabolic stresses, cell wall perturbation, wounding, and a series of biotic and abiotic stresses, (Cano-Delgado et al., 2003; Tronchet et al., 2010; Vanholme et al., 2019). Regulation of lignin biosynthesis genes results in altered H:G:S distribution.

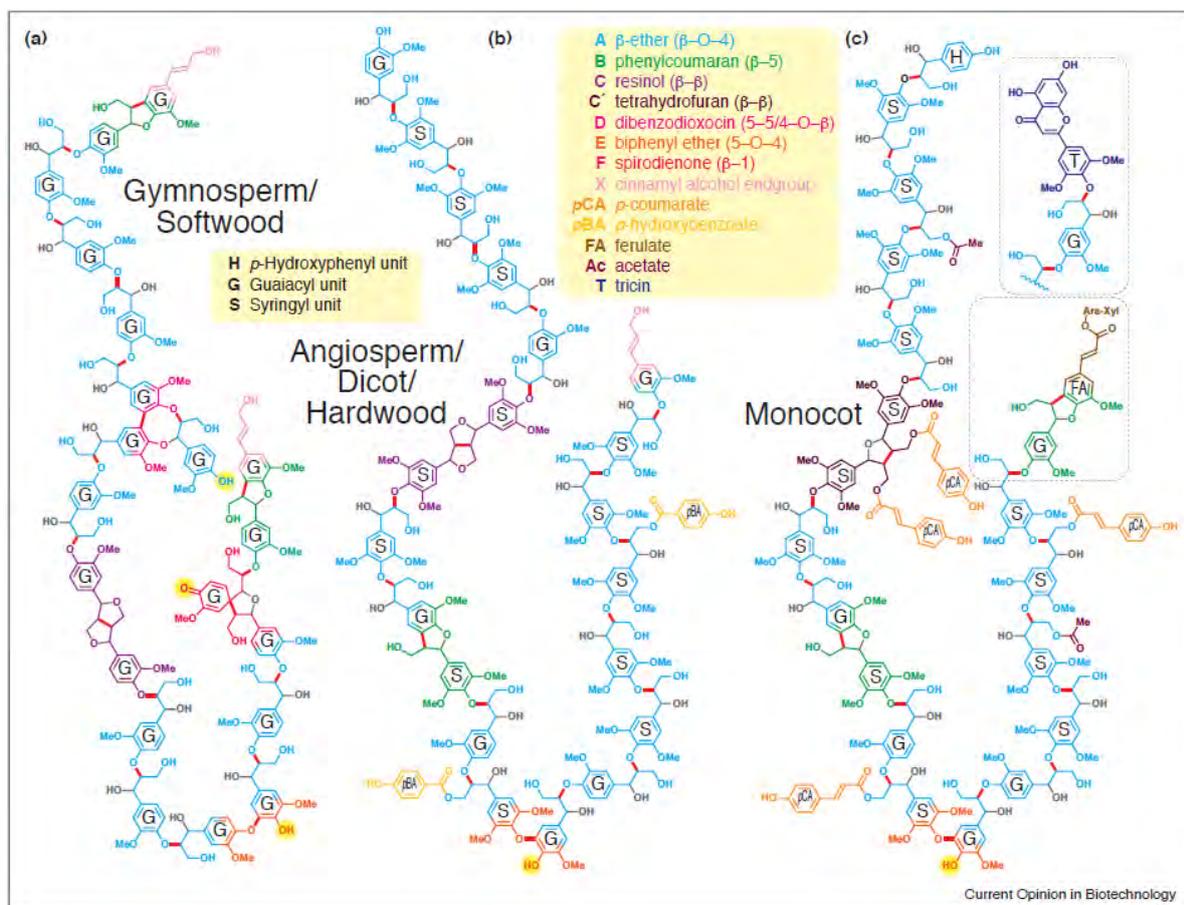


Figure I-7. Lignin model structures of three major plant classes: gymnosperm/softwood (a), angiosperm/dicot/hardwood (b), and monocot (c) (Ralph et al., 2019).

Lignin resistance to degradation is a major obstacle for industrial processing of wood such as during pulp and paper manufacturing where it requires harsh and costly chemical treatments. The huge economic importance of the pulp industry has been a driving force to decipher the lignin biosynthetic pathway, which has proven more complex and reticulate than initially thought. The topology of the pathway has been revised several times in the last decades [reviewed in (Humphreys and Chapple, 2002; Boerjan et al., 2003; Ralph et al., 2004; Vanholme et al., 2010)] and new alternative routes are still being discovered such as that involving the recently described caffeoyl shikimate esterase [CSE; (Vanholme et al., 2013)]. Altogether eleven enzymatic reactions (Figure I-8) are implicated in the synthesis of monolignols that involves the general phenylpropanoid pathway starting with the deamination of phenylalanine and leading to the production of hydroxycinnamoyl CoA esters. The enzymes involved in this short sequence of reactions are phenylalanine ammonia-lyase (PAL); cinnamate 4-hydroxylase (C4H) and 4-coumarate:CoA ligase (4CL). Hydroxycinnamoyl-CoA esters undergo successive hydroxylation and O- methylation of their aromatic rings (Boerjan et al., 2003) involving the following enzymatic activities: shikimate O-hydroxycinnamoyltransferase (HCT); caffeoyl shikimate esterase (CSE); p-coumarate 3-hydroxylase (C3'H); caffeoyl CoA 3-O-methyltransferase (CCoAOMT); ferulate 5-hydroxylase (F5H) and caffeate/5-hydroxyferulate O- methyltransferase (COMT). The conversion of the side-chain carboxyl to an alcohol group is catalyzed successively by cinnamoyl CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD), two enzymes considered to be the most specific of the monolignol biosynthesis pathway.

Notably, the CCR gene, was first cloned in *Eucalyptus gunnii* (*EguCCR*) and its identity was proven unambiguously by the enzymatic activity of the corresponding recombinant protein (Lacombe et al., 1997). Since the release of the *E. grandis* genome (Myburg et al., 2014), a genome-wide survey of the putative lignin biosynthesis genes, followed by comparative phylogenetic analyses, led to the identification of 38 genes of which 17 exhibit strong and preferential expression in highly lignified tissues. These 17 genes constitute the core set of a Eucalyptus lignification toolbox (Figure I-8).

The transport of monolignols from cytoplasm or near endoplasmic reticulum to developing cell wall occurs through mechanisms not yet elucidated. Polymerization is performed by oxidases such as laccases (LAC) and peroxidases (PRX) which can catalyze monolignol oxidation and polymerization (Bryan et al., 2016). In Arabidopsis stem, PRX64 localized in cell corners and fibers middle lamella, and LAC4 was in xylan-rich SCW layers of vessels and fibers, and both of them were highly expressed in lignifying tissue (Yi Chou et al., 2018).

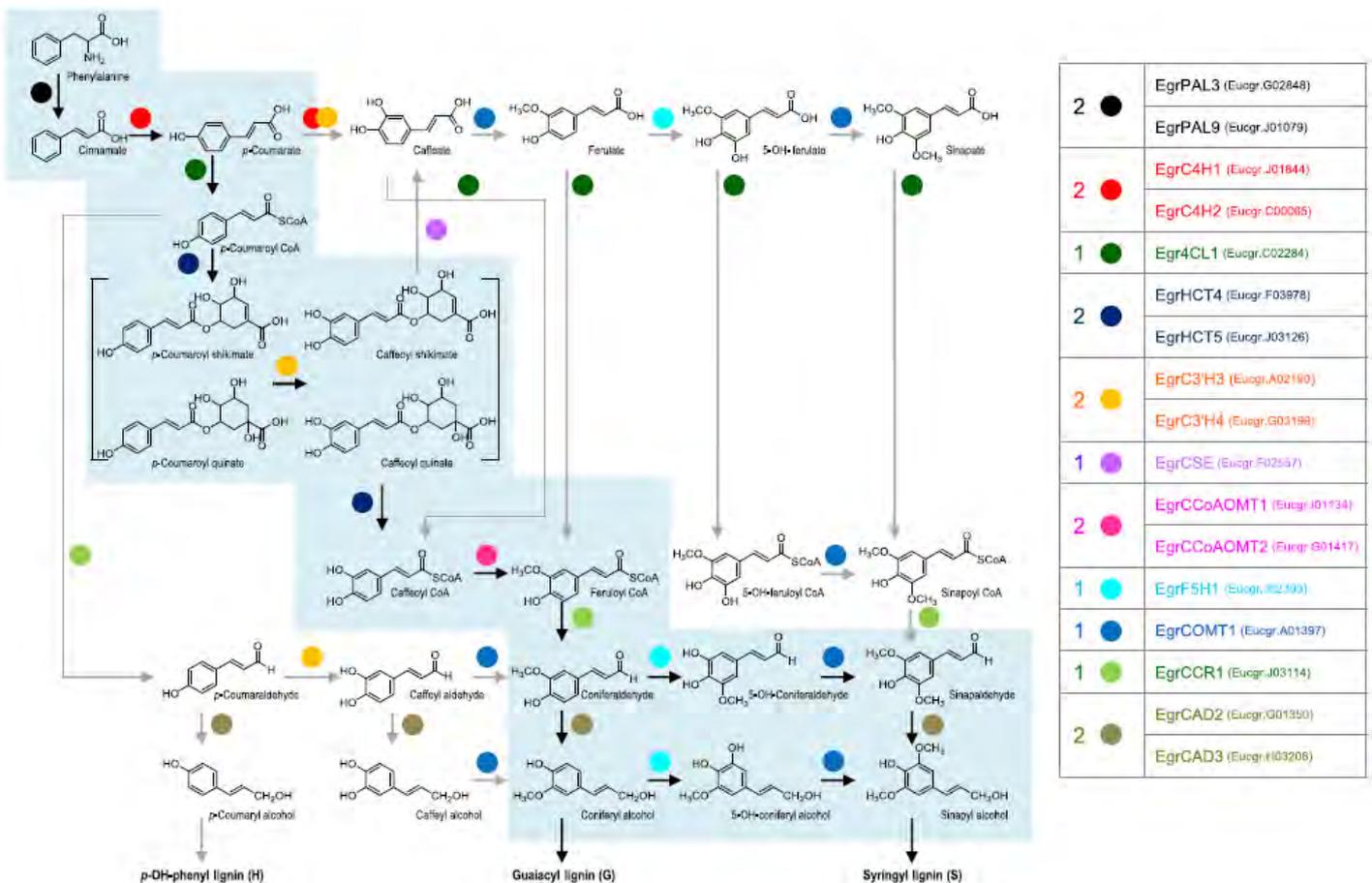


Figure I-8. The lignin biosynthesis pathway in Eucalyptus adapted from (Carocha et al., 2015). The 17 *E. grandis* genes encoding enzymes located in the bona fide clades constitute the core set of a Eucalyptus lignification toolbox. Enzymatic reactions thought to be key steps are indicated with black arrows.

3.2 The transcriptional regulation of secondary cell wall formation

The transcriptional network underlying SCW formation has been first elucidated (Figure I-9) in *Arabidopsis* and later in poplar as reviewed in (Hussey et al., 2013; Zhang et al., 2018a; Camargo et al., 2018). It involves mainly members of two transcription factors families: the NACs (NAM/ATAF/CUC) and the MYBs and (MYeloBlastosis) acting as first- and second-level master switches, respectively, to regulate a battery of downstream transcription factors and secondary cell wall biosynthesis genes (Wang et al., 2011; Schuetz et al., 2013). A recent comparison of the regulation of wood formation in angiosperm trees species *versus Arabidopsis* highlighted conserved and distinct mechanisms (Camargo et al., 2018).

Top-level NAC master regulators

Among the NACs, the VASCULAR-RELATED NAC-DOMAIN (VND) 1-7 and NAC SECONDARY WALL THICKENING PROMOTING FACTOR (NST) 1-3 are master switches of the entire SCW transcriptional regulation to

control xylem differentiation and lignification by regulating downstream partial overlapped target genes (Zhong et al., 2010d; Ohashi-Ito et al., 2010). In *Arabidopsis*, VNDs induce differentiation of vessels whereas NSTs/SNDs induce differentiation to be fibers (Mitsuda et al., 2005, 2007; Zhong et al., 2006; Yamaguchi et al., 2008).

Cellulose synthase complexes (CSC) are transcriptionally regulated during the initial of SCW synthesis by VNDs (VND1-7), especially VND6 and VND7 that were preferentially expressed in secondary xylem and have core functions for vessel formation in metaxylem and protoxylem (Kubo, 2005). In suspension culture cells system, VND6 directly regulated expression of cellulose synthesis genes *CESA4/IRX5*, *CESA7/IRX3*, *CESA8/IRX1* (Ohashi-Ito et al., 2010; Yamaguchi et al., 2010), whereas, VND6 and VND7 controlled lignin biosynthetic genes *CCoAOMT7*, *LAC4* and related PCD genes, *XCP1*, *XCP2*, *BFN1*, *RNS3* (Zhong et al., 2010d; Zhong and Ye, 2014). Overexpression of *AtVND6* or *VND7* was shown to induce xylem vessel transdifferentiation both in *Arabidopsis* and in poplar (Kubo et al., 2005; Yamaguchi et al., 2010), suggesting that the molecular mechanism of xylem vessel differentiation is, at least partially, conserved between these two species. VND6/7 were shown to be controlled by upstream regulators such as *Lateral organ boundaries domain 18 (LBD18)* and *LBD30* that control VND6/VND7, but are also targets of VND6/VND7 showing a feedback regulation (Soyano et al., 2008; Zhong et al., 2010d). E2Fc is also a regulator located upstream of VND6 and VND7 but has dual-function since it activates/repress by dose effect. E2Fc is also able to bind directly promoter regions of SCW components biosynthesis genes (*C4H*, *CCoAOMT*, *CAD*, *LAC4*) (Taylor-Teeple et al., 2015). VND-INTERACTING 2 (VNI2) is the negative regulator of VND, showing thickened SCW in xylem vessels of young overexpressing mutants (Yamaguchi et al., 2010). NST1-3 control fiber cells SCW formation and regulation; NST1 and NST3/SND1 are functionally redundant. Double *nst1/(snd1/nst3)* *Arabidopsis* mutant showed no thickness of SCW in fibres whereas SCW was normal in vessels. In this double mutant, the cellulose biosynthesis related genes *IRX3* and *IRX5*, lignin biosynthesis related *IRX4*, *IRX12* and *AtOMT1* were all down-regulated (Mitsuda et al., 2007). NST3 was shown to induce *PAL1*, *CCoAOMT*, *4CL3*, cellulose synthase-like B02 (*CSLB02*) and fasciclin-like arabinogalactan-protein 12 (*FLA12*) expression (Ohashi-Ito et al., 2010). The poplar orthologs of the *VNDs* and/or *NSTs/SNDs* genes are called *WNDs* (for wood-associated NAC domain transcription factors (Zhong et al., 2010b), or *VNSs* (for VND, NST/SND-, SOMBRERO-related proteins; (Ohtani et al., 2011) or *PtrSNDs/PtrVNDs* (Li et al., 2012; Johnsson et al., 2018). They could complement the SCW defects of the fibers in the double *nst1/(snd1/nst3)* *Arabidopsis* mutant (Zhong and Ye, 2010). However, surprisingly only *PtrWND2B* and *PtrWND6B* were able to induce ectopic deposition of SW when overexpressed in *Arabidopsis* (Zhong et al., 2010b). To explain the fact that other VNDs or NST/SNDs members could complement the *nst1/snd1* mutant but are not capable of inducing ectopic expression of SCWs, a likely hypothesis is that they need to cooperate with co-factors or other TFs which are only present in cells programmed to be sclerified.

It was initially reported that the fiber- or vessel-specific expression occurring in *Arabidopsis* was not occurring in poplar where all *WNDs/VNS* (both *VND* and *NST/SND*) were expressed in both developing vessels and fibers as well as in xylem ray parenchyma cells. The clear separation of the expression patterns of *VND* and *NST/SND* groups in *Arabidopsis* did not seem to be extensively shared with other plant species including poplar, rice, and maize (Zhong et al., 2010a, 2011a; Ohtani et al., 2011; Nakano et al., 2015). However, the recent high spatial-resolution RNA sequencing data spanning the secondary phloem, vascular cambium, and wood-forming tissues of *Populus tremula* (Sundell et al., 2017; Johnsson et al., 2018) provided new clues about the expression patterns of the genes of the VNDs and NSTs/SNDs clades which are more

complex and subtle than previously thought. Considering cell distance from cambium as a proxy for cell age, (Johnsson et al., 2018) showed that *VND6*-orthologs are induced in recently divided cambial xylem initials, while *SND1/NST1* orthologs are expressed in early xylem expansion zone. The expression patterns within the paralogous pairs of *PttSND1*, *PttNST1*, and *PttVND6* are highly similar whereas divergent expression profiles were observed within the *PttVND3* and *PttVND7* pairs. Both of the *PttVND7* paralogs are expressed in primary xylem. *PttVND7-1* is induced in the end of the maturation/programmed cell death zone indicating neofunctionalization of this paralog whereas *PttVND7-2* is not detectable in the secondary xylem (Johnsson et al., 2018). These authors further investigated the differences between the *PtrSND* and *PtrVND* clades by generating a co-expression network. They showed that *PtrVND6* orthologues together with *PtrVND3-2* formed a cluster connected to the bulk of pectin and xyloglucan biosynthetic genes as well as three primary wall-associated cellulose synthases *PtrCESAs*. *PtrSND1* and *PtrNST1* orthologs were most closely connected to secondary wall-associated *PtrCESAs*.

Using *PtrWNB2B/6B* as tools, (Zhong et al., 2011b) have uncovered a suite of up-regulated TFs, many of which were not reported previously. They further showed that *PtrWNBs/VNS* bind directly to SNBE sites in the promoters of their target genes like *Arabidopsis* *SND1* (Zhong et al., 2011b).

In Eucalyptus a genome-wide survey has identified in the *EgrNACs* members potentially involved in the control of SCW biosynthesis (Hussey et al, 2015).

SCW-associated MYB transcription factors

In *Arabidopsis*, *AtMYB46/83* are considered as the unique second-level master regulators since they are the direct targets of top-level *VND* and *NST/SND* master switches and are able to activate the promoters of all the three major SW polymers (i.e. cellulose, hemicelluloses and lignins). In poplar, the *AtMYB46/83* co-orthologs i.e. *PtrMYB2*, *PtrMYB3*, *PtrMYB20*, and *PtrMYB21* are direct targets of *PtWNB2B/6B* (Zhong et al., 2011b). As their *Arabidopsis* counterparts, they are also capable of activating the biosynthesis pathways of cellulose, xylan, and lignin, leading to ectopic SW deposition when overexpressed in *Arabidopsis* and poplar. Their dominant repression results in a reduction of SW thickening in transgenic poplar wood (McCarthy et al., 2010; Zhong et al., 2013). *EgMYB2* from Eucalyptus (Goicoechea et al., 2005), is also an ortholog of *AtMYB46/83* since it is able to activate the entire secondary wall biosynthesis program when overexpressed and can complement the *Arabidopsis* mutant *atmyb46/83* (Zhong et al., 2013).

One important difference between *Arabidopsis* and poplar is that other targets of *PtWNB2B/6B* than the orthologs of *AtMYB46/83* (*PtrMYB 2, 3, 20 & 21*), are capable of inducing the whole SW transcriptional program and may function as master switches (Zhong et al., 2011b). Among those, *PtrMYB18* (ortholog of *AtMYB20/43*), *PtrMYB75/92/125/199* (co-orthologs of *AtMYB42/85*), *PtrMYB10/128* (co-orthologs of *AtMYB103*), *PtrMYB74* & *PtrMYB121* can activate the promoter activities of several biosynthesis genes for cellulose, xylan, and lignin (Zhong et al., 2011b). *PtrMYB74* and *PtrMYB121* were among the 13 targets induced by *PtWNB2B*, for which no ortholog of *Arabidopsis* was previously shown to be involved in the regulation of SW biosynthesis. It is possible, as suggested by Zhong et al (2015), that additional master switches have been recruited in poplar to sustain a robust expression of secondary wall biosynthesis genes during wood formation, which requires the deposition of a massive amount of secondary wall components. Another hypothesis is

that these genes have specific cell localization and/or have diversified their functions for being involved in responses to stresses.

Among the poplar wood-associated MYB having diversified their transcriptional regulatory activities from those of their *Arabidopsis* counterparts is *PtrMYB128* (an ortholog of *AtMYB103*). *PtrMYB128* is capable to activate the promoters of the biosynthetic genes for all three secondary wall components in transient transactivation assays (Zhong et al., 2011b). *AtMYB103* was first shown to preferentially induce the expression of genes for the biosynthesis of cellulose but not xylan and lignin (Zhong et al., 2008). However, the characterization of two *myb103 Arabidopsis* T-DNA insertion mutants revealed that the main modification at the SW level was a change in lignin monomeric composition (decrease in the S/G ratio) in line with FERULATE 5 HYDROXYLASE (F5H) being the main target of *AtMYB103* (Öhman et al., 2013).

Lignin-specific MYBs

In *Arabidopsis*, three lignin specific MYB TFs have been identified: *AtMYB58*, *AtMYB63* and *AtMYB85* (Zhou et al., 2009). *PtrMYB28* is able to activate specifically lignin biosynthesis genes as its orthologs *AtMYB58/MYB63* (Zhou et al., 2009). Overexpression of *PtoMYB92*, an ortholog of *MYB85/MYB42* resulted in an increase in SW thickness in stems and ectopic deposition of lignin in leaves. *PtoMYB92* specifically activates the expression of lignin biosynthetic genes. Thus, *PtoMYB92* is involved in the regulation of SW formation in poplar by controlling the biosynthesis of monolignols as its *Arabidopsis* counterpart (Li et al., 2015).

MYB Negative regulators

SCW formation not only involves transcriptional activators but also entails transcriptional repressors. *Eucalyptus* *EgMYB1* represses the expression of secondary wall biosynthesis genes and inhibits secondary wall thickening in fibers when overexpressed in *Arabidopsis* and poplar, suggesting that it is a master transcriptional repressor of secondary wall formation (Legay et al., 2007, 2010). Notably, *EgMYB1* interacts specifically with a linker histone variant, *EgH1.3*. This interaction enhances the repression of *EgMYB1*'s target genes, strongly limiting the amount of lignin deposited in xylem cell walls. The expression profiles of *EgMYB1* and *EgH1.3* overlap in xylem cells at early stages of their differentiation as well as in mature parenchymatous xylem cells, which have no or only thin lignified secondary cell walls. This suggests that a complex between *EgMYB1* and *EgH1.3* integrates developmental signals to prevent premature or inappropriate lignification of secondary cell walls, providing a mechanism to fine-tune the differentiation of xylem cells in time and space (Soler et al., 2017).

Although, *EgMYB1* is a close ortholog of *AtMYB4*, their respective functions are different since *AtMYB4* has never been reported to regulate SW synthesis but instead it is known to regulate sinapate esters accumulation through its direct target C4H (cinnamate 4-hydroxylase) (Jin et al., 2000).

The two *EgMYB1*'s orthologs in *Populus* have been functionally characterized in independent studies. *PdMYB221* was overexpressed in *Arabidopsis* (Tang et al., 2015) and *PtoMYB156* was overexpressed and down-regulated by genome editing in poplar (Yang et al., 2017). Both genes were shown to negatively regulate the secondary wall thicknesses of xylem fibers and the content of cellulose, lignin and hemicelluloses but *PtoMYB156* was also shown to repress

phenylpropanoid biosynthesis genes, leading to a reduction in the amounts of total phenolic and flavonoid compounds (Yang et al., 2017). Whether PdMYB221 also regulate genes of the phenylpropanoid is still to be determined.

Thanks to the *E. grandis* genome availability, the whole R2R3-MYB TF family has been analyzed including the tissue/organ expression patterns of the members providing new candidates for the regulation of SCW as well as orthologs of regulators already characterized in *Arabidopsis* or other plants (Soler et al., 2015).

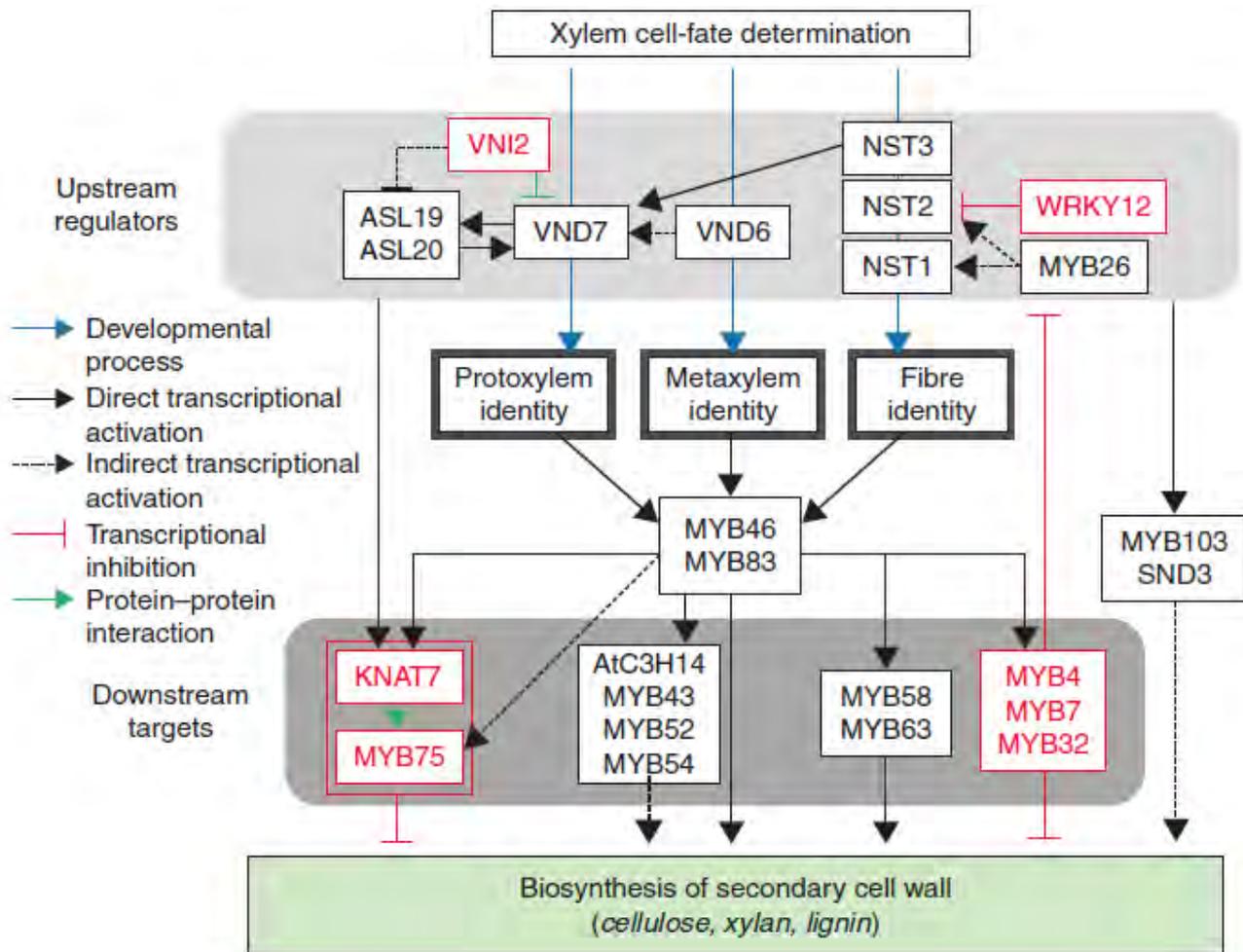


Figure I-9. Transcriptional network regulation of wood cell formation and SCW formation, involving in a battery of transcription factors such as MYB46, MYB83 master switches (Ko et al., 2014).

3.3 Hormonal control of xylogenesis

Secondary xylem formation is affected by endogenous factors (genetics, plant hormones) and exogenous factors (abiotic and biotic). Plant hormones levels in wood forming tissues are also changed under environmental stresses. The hormonal control of cambium activity and cell specification by auxin, cytokinin, gibberellins as well as thermospermine (Barra-Jiménez and Ragni, 2017), which were studied in models such as *Arabidopsis* and *Zinnia* and other annual species, and synthesized in many reviews (Mauriat et al., 2014; Nieminen et al., 2012; Milhinhos and Miguel, 2013; Sorce et al., 2013;

Smet and De Rybel, 2016). The central role of auxin and its cross talk with other hormones in initiating the transcriptional program of secondary xylem differentiation in trees will be developed in the next section (**Chapter I-Part 2-section 5**). Here, we will briefly comment on the role of ethylene and thermospermine in regulation of xylem cells development and cambium activity during wood formation.

Ethylene is produced during cell development especially in higher meristematic and ripen tissues, and it can also be induced by 1-amino-cyclopropane-1-carboxylic acid (ACC). Ethylene represses the ability of auxin polar transport across the developing zones, via decrease numbers of auxin efflux carriers (Junghans et al., 2004). Ethylene has a dual role for tree-ring formation, thus it promotes cell proliferation at low concentration, and it enables responses to environment stresses, such as drought. Ethylene promotes cell differentiation and arrives at peak in mature tracheid before programmed cell death (PCD) and lignification.

ACC precursors are synthesized in phloem, buds and needles, the ethylene synthesis occurs in cambium. The exogenous application of ethylene stimulates cambium activity but not in the ethylene-insensitive transgenic poplar obtained by overexpressing a dominant negative Arabidopsis ethylene receptor *ETR1* (Love et al., 2009). Consistent with these observations, the overexpression of ethylene biosynthesis gene *PttACO1* (ACC oxidase) stimulates cambium proliferation in *Populus*. The authors also showed that ethylene is at the origin of the eccentric cambial activity which gives rise to reaction wood formation in response to leaning. Applied exogenous ethylene and ACC to hybrid aspen enhanced cambial growth, decreased xylem vessels size and number, induced the gelatinous layers (G-layers) in tension wood, and changed the fiber cell wall cellulose microfibril angle (Seyfferth et al., 2018). (Seyfferth et al., 2018) also started to elucidate the ethylene signaling pathway in *Populus* trees and identified several putative downstream targets. *In silico* analysis of the AspWood transcriptome database (<http://aspwood.popgenie.org/>), which covers all stages of secondary growth in aspen stems, revealed that the ethylene precursor 1-aminocyclopropane-1-carboxylic acid (ACC) is synthesized during xylem expansion and cell maturation. More interestingly, ethylene-mediated transcriptional reprogramming occurs during all stages of secondary growth. The authors also identified new putative regulatory hub genes like *EIN3D* (*ETHYLENE INSENSITIVE 3D*) and 11 *ERFs* (*ETHYLENE RESPONSE FACTORS*) which, for most of them, were connected for the first time to wood formation (Seyfferth et al., 2018).

Thermospermine (Tspm) is one of the plant polyamines which has been studied of its regulations for vascular development especially xylem cells differentiation. *ACAULIS5* (*ACL5*) gene encodes thermospermine synthase which is specially expressed in xylem vessel element, and loss-of-function *acl5* Arabidopsis mutants displayed dwarf phenotype and accelerated vessel cells death (Hanzawa, 2000; Muñiz et al., 2008). Overexpressed *PtACL5* in poplar showed wider stem in dwarf plants, decreased number of metaxylem cells and no secondary growth (Milhinhos et al., 2011). In addition, *POPACL5* overexpression resulted in negative effect on auxin accumulation, while exogenous auxin promoted thermospermine accumulation and *ACL5* gene expression, demonstrating that Tspm level is controlled by the negative feedback loop mechanism (auxin is central component, also involved in *ACL5*, HD-ZIPIII, HB8) in stem secondary xylem differentiation (Milhinhos et al., 2013). Therefore, Overexpressed *ACL5* usually resulted in decreased *ACL5* expression. *ACL5* promotes SUPPRESSOR OF ACAULIS51 (*SACL*) translation which inhibits cytokinin biosynthesis and vascular cell division in Arabidopsis young hypocotyl and primary roots, while this regulatory is in contrast in secondary growth of poplar, showing that low *ACL5* expression and low cytokinin level led to increased cambial cells activity, which is the

evidence that ACL5 integrates auxin and cytokinin signaling to promote extensive secondary growth of tree stems (Milhinhos et al., 2020). However, thermospermine synthase is also activated in non-vascular plants and some aquatic plant, indicating its potential roles not only regulation of xylem cells development but also response to environmental stresses (Solé-Gil et al., 2019).

Part 2 Auxin regulates plant growth and development, especially in wood formation

1. Auxin

Auxin effect on plants was first discovered by Darwin's investigations of phototropic growth response (bending towards light) downward from coleoptile tip (Darwin, 1897), which may be caused by mobile signal transmission. This chemical substance (which named auxin Greek) was studied by Went through evidence of higher distribution on seedlings shaded side (Went, 1926), and then auxin was purified and characterized by Kenneth Thimann. Auxin represents a class compounds of plant hormones with an aromatic ring and a carboxylic acid group. To date, four naturally endogenous principal auxins was found in plants: indole-3-acetic acid (IAA), 4-chloroindole-3-acetic acid (4-Cl-IAA), phenylacetic acid (PAA) and indole-3-butyric acid (IBA) (Simon and Petrášek, 2011). The most important active auxin found in plants is IAA (indole-3-acetic acid), which is synthesized in young regions locally, such as shoot apical meristems and young leaves and less in root tips. Based on IAA structure analogical to amino acid tryptophan, synthetic compounds such as naphthalene-1-acetic acid (NAA) and 2,4-dichlorophenoxyacetic acid (2,4-D) also have been developed as plant chemical regulators by scientists and manufacturers.

Auxin regulates cell division, elongation and differentiation, and plays crucial roles in the control of plant growth and development. For instance, it controls primary root initiation, lateral root formation, apical dominance, phototropism, gravitropism, fruit development, embryogenesis, vascular patterning and wood formation. Auxin regulates many different plant developmental processes by auxin response, which is mainly mediated by dynamic, optimal auxin level (depending on its biosynthesis, distribution (transport) and homeostasis) on specific cells, and the auxin signalling triggers the regulation of genes expression precisely and rapidly.

2. Auxin metabolism

2.1 Auxin biosynthesis

Although virtually all plant tissues appear to be capable of synthesizing auxin, most is normally produced in young apical meristem and developing parts of plants, such as the shoot apex, emerging leaves and developing seeds (Ljung et al., 2002a, 2002b). Studies have also uncovered that auxin is also synthesized in roots, with the most prominent auxin source located in the meristematic zone of primary root tips and developing lateral roots (Ljung et al., 2005). The synthesized IAA in plants exists in active free form and conjugated states, which is synthesized mainly initial from Tryptophan precursor. De novo auxin biosynthesis is involved in multiple complex pathways, major containing the Trp-dependent and Trp-independent pathways (Mano and Nemoto, 2012), which both have distinct roles for development and environmental responses. However, the auxin biosynthesis in mainly achieved by Try-dependent pathway (Figure I-10). According to different intermediate types, Try-dependent IAA biosynthesis includes four postulated pathways: (i), the indole-3-

acetamide (IAM) pathway; (ii) the indole-3-pyruvic acid (IPA) pathway; (iii) the tryptamine (TAM) pathway; and (iv) the indole-3-acetaldoxime (IAOX) pathway. Of these, IPA is the most dominant pathway, which L-Tryptophan is converted by two steps using indole-3-pyruvate (IPyA) into IAA (Won et al., 2011; Zhao, 2012) with tandem operations of TRYPTOPHAN AMINOTRANSFERASE OF ARABIDOPSIS (TAA) (produces IPyA) and YUCCA (YUC) flavin monooxygenase-like enzymes (catalyzes the rate-limiting step) (Mashiguchi et al., 2011). TAA and YUCCA genes were discovered and functions were identified by genetic screens and analyzed mutants (Stepanova et al., 2008; Tao et al., 2008; Chen et al., 2014). The YUC genes family have been identified to investigate biological functions in more than 27 plant species, including twelve YUC genes in poplar (Ye et al., 2009) and twenty YUC genes in apple (Song et al., 2020). OsYUC1-auxin-OsWOX11 module controls crown root development in rice (Zhang et al., 2018b). This IPA two-step TAA/YUC pathway is conserved and essential for many developmental processes in planta. Try-independent pathway was found by isotope-labeling technology (Normanly et al., 1993) remaining ambiguous mechanisms, but recent genetic and biochemical evidence showed the specific component cytosol-localized indole synthase (INS) was involved in IAA Try-independent biosynthesis which spatiotemporal contributions in embryogenesis (apical-basal axis formation) in higher plants (Wang et al., 2015).

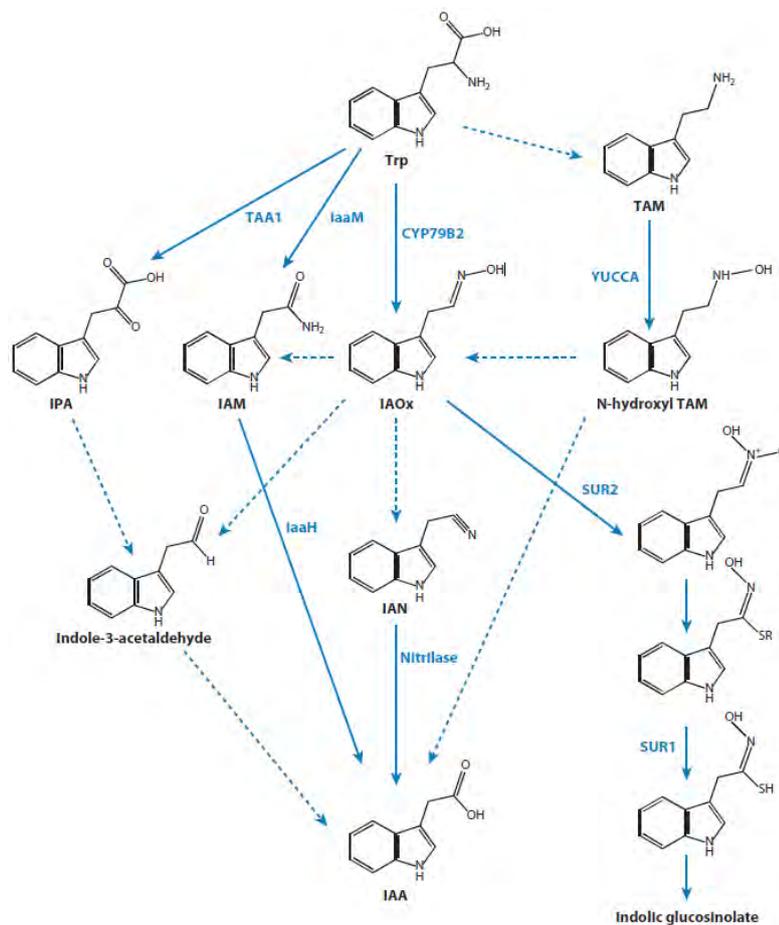


Figure I-10. Tryptophan dependent auxin biosynthesis pathway (Zhao, 2010). The identified (solid arrows) and proposed (dash arrows) genes in plants are responsible for steps. IAOx: indole-3-acetaldoxime; IPA: indole-3-pyruvate; IAM: indole-3-acetamide; IAN: indole-3-acetonitrile; TAM: tryptamine.

2.2 Auxin homeostasis

Cellular auxin homeostasis is under control of a series of mechanisms such as biosynthesis, degradation, conjugation, intra- and intercellular transport and compartmentation (Figure I-11). The active optimal auxin level modulated by homeostasis mechanisms is crucial for plant growth and development. The conjugated auxins are major and storage form in plants. There are three types of hydrolyzing auxin conjugates including (i), IAA-esters and IAA-saccharides, (ii), IAA-amino acids and (iii), amide-linked peptides or proteins (Ludwig-Müller, 2011), and the last conjugate storage form of IAA exists in most dicots (Woodward, 2005; Bajguz and Piotrowska, 2009). Genes related to IAA conjugate hydrolysis were identified, such as GH3, ILR, ILL, etc (reviewd in (Woodward, 2005; Fu et al., 2019)).

Auxin degradation includes decarboxylation on side chain or oxidation on indole ring (Normanly et al., 2010) by plant peroxidases, and irreversible conjugation which cannot be hydrolyzed back to free IAA once forming IAA-Asp and IAA-Glu amid conjugates (Ludwig-Müller, 2011) to form inactivate IAA. In addition, JA (jasmonic acid)-Trp and IAA-Trp inhibit auxin activity in Arabidopsis root growth (Staswick, 2009). Two major catabolic products of IAA are 2-oxoindole-3-acetic acid (OxIAA) and OxIAA glucopyranose (OxIAA-Glc) (Pěnčík et al., 2013), which are produced by Dioxygenase for Auxin Oxidation gene (DAO). DAO comprises a subfamily of 2-oxoglutarate and Fe(II) [2-OG Fe(II)] dependent dioxygenase superfamily (Zhang and Peer, 2017), and has functional redundancy with GRETCHEN HAGEN3 (GH3). The rice and Arabidopsis dao mutants presented increased IAA level inducing conversions from IAA to oxIAA (Zhao et al., 2013; Porco et al., 2016), and DAO1 enzyme acts a major role to catalyze auxin oxidation in Arabidopsis by correlated alternation of oxIAA levels and associated altered morphology in loss-and gain-of-function mutants (Zhang et al., 2016a). Auxin transport from the synthesis and storage sites throughout the entire plant over long and short distances in two pathways to cause spatial auxin distributions, (i), non-polar auxin transport from auxin sources sites and be distributed via phloem; (ii), cell-to-cell auxin polar transport (PAT) system, which is a directional flow through plant tissues. The PAT masters various of plant developments such as vascular differentiation, organogenesis, tropical growth, etc (Friml and Palme, 2002; Friml, 2003). Membrane permeability limits the auxin active processes movement, and the free auxin (weak acid form IAA-) in acidic environment can move across cell with passive processes (IAAH form, ~15%). While the rest majority of IAA remaining IAA- in cytoplasm (PH~7), requires the transporters to facilitate the IAA flow orientation by active uptake across the cells including influx and efflux carriers (Zazimalova et al., 2010), which lead auxin to be sensed by downstream regulators.

To date, experimental evidences of PAT associated carriers have been provided including one auxin influx carrier family AUXIN1/LIKE-AUX1 (AUX/LAX), whereas two main auxin efflux carriers families PIN-FORMED (PIN) and P-GLYCOPROTEIN (PGP). AUX/LAX homologs have been found in various plants species since 2000. Aux/LAX belongs to a group of proton-gradient-driven transporters. Four AUX/LAX members (AUX, LAX1, LAX2, LAX3) functions related to root development and vascular patterning were elucidated in Arabidopsis (Swarup and Péret, 2012), in addition, genetic evidences of AUX/LAX gene family roles in other model plants (Medicago, rice, etc) showed regulation of development processes such as root and vascular development, seed germination, and leaf morphogenesis (Swarup and Bhosale, 2019). PIN family of transmembrane proteins acting in auxin polar transport at intra- and intercellular levels, are mainly responsible to auxin asymmetric subcellular localizations. Conserved structure of PIN proteins contains an intracellular hydrophilic loop (HL) domain with various length (PIN with short length located on endoplasmic reticulum)

Bibliographic review---Part 2. Auxin regulates plant growth and development, especially in wood formation targeted by protein kinases for phosphorylation. The phosphorylated PIN proteins localized on plasma membranes (PMs) and/or endoplasmic reticulum (ER) are fine-tuned by a series of regulators (Adamowski and Friml, 2015). PIN family have been most well studied in various plant species playing crucial roles in plant developments and response to abiotic stresses (Zhou and Luo, 2018), and its structure-function connections and PIN-ABCB interactions were investigated and stimulated by supporting structural sequence motifs and amino acid strings (Zwiewka et al., 2019). Furthermore, the other carriers were divided into (i), located on PMs like nitrate transporter 1.1 (NRT1.1) involving in auxin homeostasis, and ATP-binding cassette (ABC) transporters containing three subfamilies (ABCB, ABCD, ABCG), which ABCB subfamily controls polar distribution of IAA whereas ABCG catalyzed IBA transport (Geisler et al., 2017); (ii), located on ER such as PIN-Like transporters (PILS) and WALLS ARE THIN 1 (WAT1). Expression analyses of a series of key genes related auxin synthesis, transport, and metabolism (MdYUCCA10a, MdPIN1b, MdPIN8a, MdGH3-5b, MdGH3-9a, MdIAR3c, MdILL6c) in different young dwarfing apple trees were shown that low free IAA contents and transport ability in phloem and roots resulted in dwarf growth (Song et al., 2016).

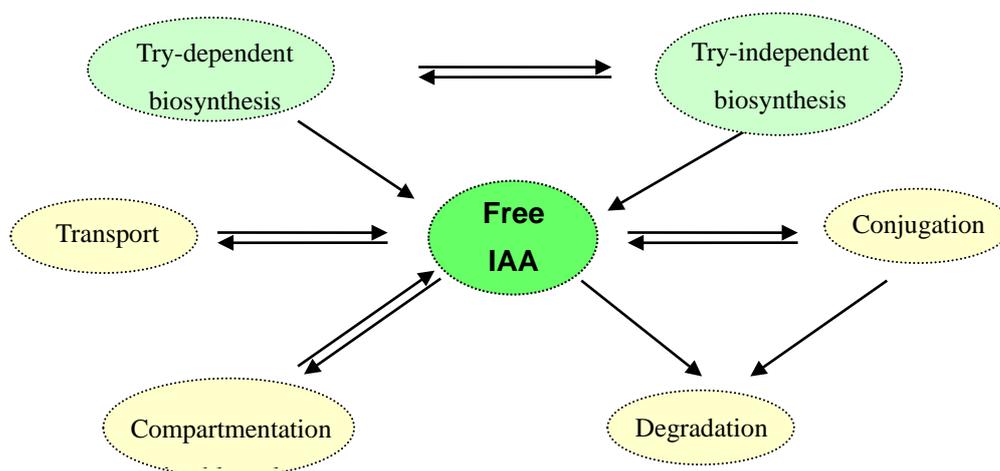


Figure I-11. Homeostasis of IAA in plant cell (provided by PhD thesis of Hua Wang). Biosynthesis by tryptophan-dependent and tryptophan-independent pathways can lead only to an increase in the concentration of free IAA. Degradation (either by non-decarboxylative oxidation or by decarboxylation) leads only to decrease in IAA concentration, while a part of conjugation is reversible and can therefore lead to either an increase or a decrease. Both transport and compartmentation can cause either an increase or a decrease in the cytosolic free IAA concentration, depending on the direction of hormone movement.

3. Auxin signaling

Auxin controls many plant growth and developmental processes by regulating genes expressions. In the last decades, studies in *Arabidopsis* and other plant species have identified auxin signaling pathways and associated mediators. Auxin specifically recognizes or binds with receptors which are located on membrane surface and/or in the cell nucleus to trigger signal transduction and lead to downstream biochemical events and diverse transcriptional responses.

3.1 Auxin signal perception and transduction

To date, there are three auxin perception and transduction systems in plants that have been studied. The major pathway is (a), SCF^{TIR1/AFB}-Aux/IAA auxin receptor complexes involved in transcriptional regulation, which has been well established and studied. As well, (b), S-PHASE KINASE ASSOCIATED PROTEIN 2A (SKP2A, F-box protein) signaling pathway that SKP2A binds with auxin to activate proteolysis of cell-cycle transcription factors, and (c), AUXIN BINDING PROTEIN1 (ABP1) system which auxin binds with ABP1 to form a complex with TMKs (receptor-like kinase) to regulate plant developmental processes in non-transcriptional genomic effects. The protein families involved in system (a) and related to auxin signaling transcription regulation will be further introduced in Part 2-section 4.

SCF^{TIR1/AFB}-Aux/IAA-ARF signaling pathway

In this pathway, auxin is directly perceived by a co-repressor complex consisting of an F-box protein from the TRANSPORT INHIBITOR RESPONSE1/AUXIN SIGNALING F-BOX PROTEINS (TIR1/AFBs) family and the AUXIN/INDOLE ACETIC ACID (Aux/IAA) family of transcription repressors. Transcription is directly regulated by a third family of transcription factor named AUXIN RESPONSE FACTORS (ARF). TIR1 belongs to F-box protein, that interacts with SKP1-like proteins and CULLIN-LIKE1 (CUL1) to form the ubiquitin-ligase (E3) complex within these three core subunits, called SCF^{TIR1} (Gray et al., 1999). SCF complex catalyzes substrate protein (for example, Aux/IAA) ubiquitination for 26S proteasomal degradation, and is required for nuclear auxin signaling. In Arabidopsis the TIR1/AFB auxin receptor family comprises six members: TIR1 and AFB1 through AFB5, each AFB member has distinct functions and unequal contributions to auxin response (Parry et al., 2009), drastically enhanced auxin response defect in triple mutant (*tir1afb2afb3*) indicates the existence of functional redundancy among the TIR1/AFB genes. The TIR1-Aux/IAA interaction was promoted by auxin (Gray et al., 2001; Kepinski and Leyser, 2004), which trigger Aux/IAA degradation (Figure I-12). When auxin levels are low, Aux/IAA proteins bind to ARFs and repress their transcriptional activity. The presence of auxin acts as a molecular glue to combine TIR1/AFB with domain II of Aux/IAA transcriptional repressors to mediate its degradation, and this process releases ARFs from Aux/IAA-ARF dimers for activating downstream auxin response transcription (Tan et al., 2007). Auxin co-receptor complexes consisting of TIR1 and Aux/IAA are various of auxin-binding affinities and auxin-sensing properties, which is determined by Aux/IAA (Calderón Villalobos et al., 2012). SCF^{TIR1/AFB} functions in auxin perception and transcriptional regulation in Arabidopsis were reviewed in (Salehin et al., 2015). Eight FBL genes (TIR1 homologs) were identified and characterized in *Populus* comprising four pairs, of which *PtrFBL1* and *PtrFBL7* were preferentially expressed in vascular and cambial tissues (Shu et al., 2015). As for adventitious roots (AR) formation, *PagFBL-IAA28* module was found to positively regulate AR formation and root biomass in poplar, sharing with mechanisms of lateral roots induction in Arabidopsis (Shu et al., 2019). However, root growth regulated by auxin in Arabidopsis was negatively correlated with canonical SCF^{TIR1/AFB}-Aux/IAA complex, which is probably regulated by of non-transcriptional auxin signaling branch (Fendrych et al., 2018).

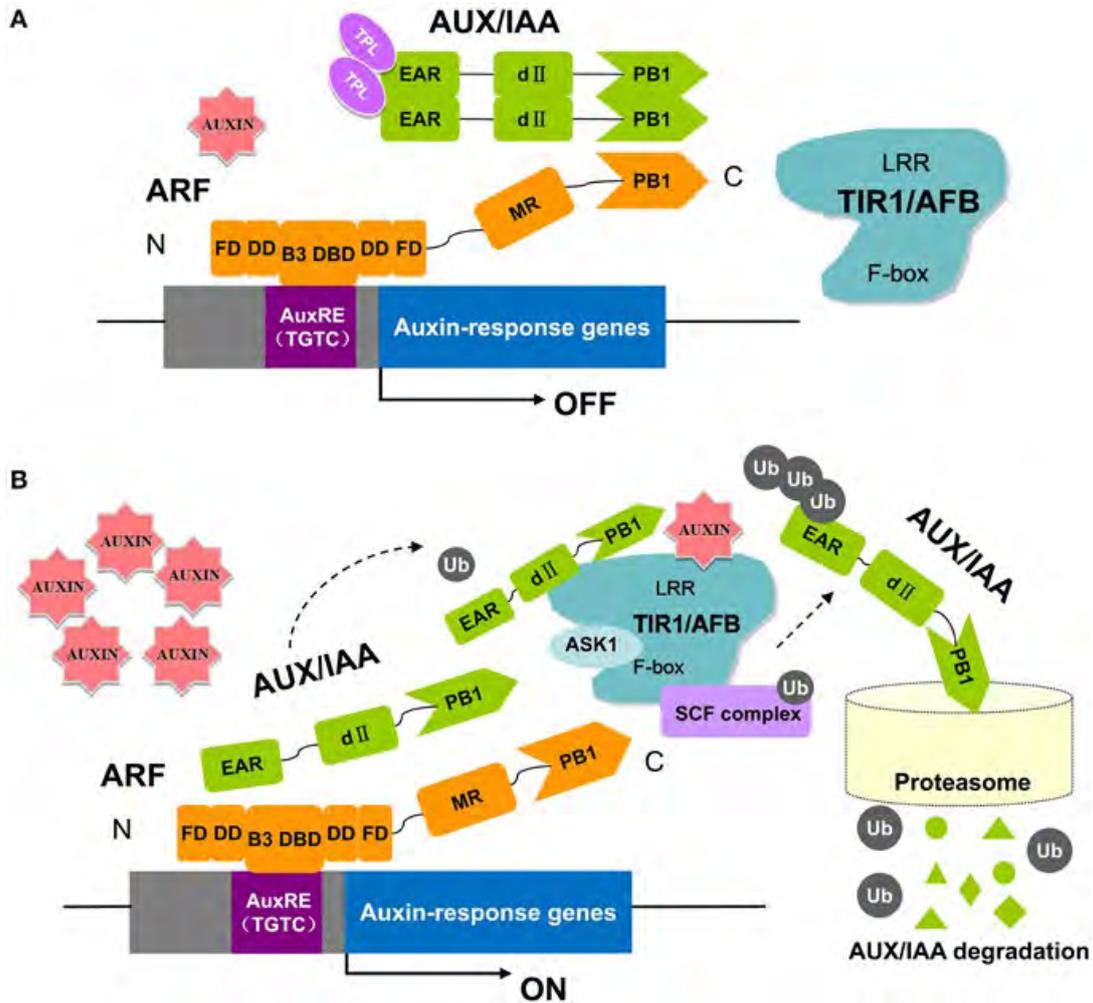


Figure I-12. The key components in auxin perception and signaling in *Arabidopsis* (Li et al., 2016). (A), at low auxin level, Aux/IAA proteins form dimers with ARF and recruit TPL to chromatin; (B), at high auxin level, Aux/IAA proteins are ubiquitinated and degraded through SCFTIR1/AFB complex and the 26S proteasome. The ubiquitin protein ligase complex SCFTIR1 is formed by F-box proteins (the TRANSPORT INHIBITOR RESPONSE 1/AUXIN SIGNALING F-BOX proteins (TIR1/AFBs), together with other proteins (ASK1, CUL1, RBX)). dII, domain II; DD, dimerization domain; FD, flanked domain.

SKP2A signaling pathway

SKP2A is an F-box protein which is diverse and the subunit of SCF complex belonging to E3 ubiquitin ligase (Jurado et al., 2008). Similar to SCFTIR1/AFB-Aux/IAA pathway, the auxin in SCFSKP2A pathway promotes the interactions of SKP2A and DPB. SKP2A is regulated by auxin signaling through Ub/26S degradation pathway and targets RETINOBLASTOMA/E2F/DP (E2FC/DPB) which regulates cell-cycle division and proliferation (del Pozo et al., 2006). Auxin binds SKP2A directly and specifically at identified binding site which is necessary for its stability and E2FC/DPB proteolysis, based on computational structure and SKP2A mutation form (Jurado et al., 2008; Mach, 2010), and phenotypes such as no-induced degradation and reduced cell division in root meristem were found in SKP2A mutated *Arabidopsis* (Jurado et al., 2010).

ABP1 system

In the past, ABP1(AUXIN-BINDING PROTEIN 1) was regarded as an auxin receptor to regulate cell expansion which is mainly located on ER (low amount on PM) (Jones and Herman, 1993) and mediates auxin binding nongenomic effects to induce downstream cascades events. ABP1 was well studied to play essential roles in many plant growth and developmental processes. ABP1-IAA complex forming and auxin binding site was identified by ABP1 crystal structure which is different from TIR/AFB (Woo, 2002). Overexpression of ABP1 (KDEL sequence mutation) resulted in phenotypes associated with auxin signaling such as decreased root length and apical dominance three cotyledons (Robert et al., 2010).

ABP1 was not easily accessible for molecular studies since the homozygous *abp1* Arabidopsis mutant (T-DNA insertion) is always embryo-lethal. Striking Gao et al. using CRISPR/Cas9 to successfully generated homozygous T2 mutants with null-alleles (five amino acids loss in first exon) which showed no difference between wild-type. Later the T-DNA insertion null mutant of *abp1* was discovered to be caused by the disruption of the adjacent gene BELAYA SMERT (BSM) (Michalko et al., 2015; Dai et al., 2015). And the *abp1* heterozygous mutant has been reported multiple defects in auxin physiology-related responses such as longer hypocotyls, agravitropic roots and hypocotyls, decreased apical dominance may also due to the disruption of adjacent gene BSM (Gao et al., 2015). In conclusion Gao et al demonstrated that ABP1 is not a key component in auxin signaling of Arabidopsis development.

3.2 Early (Primary) auxin responsive genes induction

Auxin leads to transcriptional response of a series of genes expressions which are called early (primary)-responsive genes with characteristics of rapid, transient, specific and precise induction by active auxin. Early responsive genes mainly contain three classes: Auxin/INDOLE-3-ACETIC ACID (Aux/IAA) family, SMALL AUXIN UPREGULATED RNA (SAUR) and auxin-responsive GRETCHEN HAGEN 3 (GH3) family. Both SAUR and GH3 expression are tissue- and organ-specific in soybean seedlings and flowers, which are involved in different auxin-mediated cellular responses (Gee et al., 1991). Furthermore, canonical auxin-response cis-elements (AuxREs) located in promoter regions of these auxin responsive genes have been identified with core sequence TGTCNN (TGTCTC most frequently) (Hagen and Guilfoyle, 2002). Identification and functional characterization of Aux/IAA will be introduced in next section 4.

SAUR and GH3

The first Small Auxin-Upregulated RNA (SAUR) was discovered in soybean hypocotyls, which was expressed in elongating tissues (McClure and Guilfoyle, 1987; Gil and Green, 1997) and is short half-lives induced by auxin. SAUR transcripts are unstable mRNAs and regulated post-transcriptionally, with presence of conserved downstream element (instability sequence found in plants) at 3'UTRs (Park et al., 2012). SAUR genes have been identified in various species such as Arabidopsis, rice, tomato, citrus, etc., while few functions were characterized. A total of 105 SAUR genes belonging to 10 subfamilies were identified in poplar, with different expression levels in tissues and the profile of stress-responses (Hu et al., 2018). Genome-wide analysis of SAUR genes in apple was carried out to obtain expression profile in different tissues and response to applied IAA (Wang et al., 2020). Activated by ABA, SAUR41 subfamily genes played

roles to modulate cell expansion and salt tolerance (Qiu et al., 2020).

The GRETCHEN HAGEN 3 (GH3) auxin-responsive gene was isolated in soybean and then identified in Arabidopsis (Takase et al., 2003). GH3 family is responsible for auxin feedback loop encoding IAA-amido synthetases, that is divided into three groups (Staswick et al., 2002) involving in auxin homeostasis and hormone response pathway. GH3 family modulated auxin level by formation of adenylated-IAA reaction (Chen et al., 2010). In Arabidopsis, GH3.11 (JAR1) acted as jasmonic acid-amido synthetase in conjugation to IAA or JA (Staswick et al., 2005). Since then, GH3 homologs have been identified in model plants, crops, moss and fruit woody trees. GH3 gene expression is regulated by plant hormones (auxin, ethylene, abscisic acid, salicylic acid) and environmental stresses to affect plant growth and development, as well as resistance to abiotic and biotic stresses (Domingo et al., 2009; Du et al., 2012; Chen et al., 2013a; Wei et al., 2019).

AuxREs cis-element

Canonical auxin response elements (AuxREs, TGTCNN) is a generic ARF-binding site which is functionally identified in the promoters of primary/early auxin response genes. AuxREs mainly mediate activation of transcription of auxin response rather than down-regulation, which is corresponding to Aux/IAA-ARF interactions analysis, showing that Aux/IAA usually form dimers with ARF activators (Vernoux et al., 2011). Auxin initiates signaling depend on nuclear auxin receptors, by increasing affinity between SCFTIR1/AFB complexes and Aux/IAA. Degradation of ubiquitin-modified Aux/IAA releases ARFs to regulate expressions of downstream auxin response genes by binding AuxREs (Wang and Estelle, 2014). Different from auxin presence which derepress ARFs, in absence or low auxin level, ARF transcription factor forms dimers with ARF repressors and/or Aux/IAAs, and its transcription activity is inhibited by Aux/IAA, through preventing contact with transcription initiation, and/or a repressive chromatin state mediated by TPL (Ito et al., 2016).

The widespread synthetic auxin response promoter called DR5 contains marks sites and 7-9 AuxREs repeats and acts as a reporter system by driving GUS (β -glucuronidase), fluorescent proteins (GFP or RFP) to investigate cellular level auxin distribution pattern and auxin-dependent processes in plants (Chen et al., 2013b). The AuxREs was identified in DR5 promoter firstly in soybean (Ulmasov et al., 1997), while AuxRE in DR5 is a less-affinity sites while TGTCGG site (DR5v2) is with higher affinity and functional in transformable plants, as visualized and quantified tools (Liao et al., 2015). Variants of AuxREs have been elucidated in (Mironova et al., 2014), and a series of hexamers (bind with bHLH and IZIP, or A/T-rich) were enriched more in upstream regions of auxin-responsive genes (Cherenkov et al., 2018). For putative cis-regulatory motif identification, finding tools have been developed for prediction of AuxREs (also AuxRE-like motifs) (Sghaier et al., 2018).

4. Key mediators of auxin signaling

4.1 Aux/IAAs family (co-receptor, repressor)

Aux/IAA is transcription factor which has been identified as short-lived nuclear proteins with distinct functions for plant growth and development, that plays crucial roles in auxin-mediated transcriptional regulation (Lavy and Estelle, 2016). In auxin signaling, Aux/IAA recruits TPL as co-repressors binding to the ARF to stop transcription in absence of auxin. Once auxin presents and acts as a glue to induce proteolysis of Aux/IAA, which release its partner protein ARFs to modulate

Chapter I:

Bibliographic review---Part 2. Auxin regulates plant growth and development, especially in wood formation

downstream auxin responsive genes to carry out auxin response in plants (Leyser, 2018).

There are four conserved domains of canonical Aux/IAA: Domain I-IV (Figure I-12, I-13). Domain I acts as an repression domain harboring EAR (ethylene response factor (ERF)-associated amphiphilic repression) motif (LxLxL, L is Leu and x is random) (Tiwari et al., 2004) which can recruit TPL-related (TPR) co-repressor and in turn recruit chromatin remodeling factors for stabilized repression (Szemenyei et al., 2008). The GWPPV/I degron motif in Domain II is conserved and required for interaction with TIR1/AFB. The gain-of-function mutation of Aux/IAA (mutated in Domain II) inhibited Aux/IAA degradation to present auxin-related phenotypes (Ramos et al., 2001). Domain II also has basic residues which is putative functional nuclear localization signal (NLS) (Abel et al., 1995). Two types of degenerated NLSs (SV40-like NLS located in Domain IV and bipartite NLS) were found in Aux/IAA until now targeting to nucleus (Wu et al., 2017). Domain III contains $\beta\alpha\alpha$ -fold and Domain IV contains an acidic region and an SV40 type NLS (PKKKRKY). As well, 'GDVP' motif located in Domain IV (between $\beta 1$ and $\alpha 2$) is responsible for electrostatic protein interactions (Guilfoyle and Hagen, 2012). Both Domain III and IV share the homologous region of carboxy-terminal dimerization domain (CTD) in ARF proteins, which serves for homo-and heterodimerizations with other Aux/IAAs and ARFs (dos Santos Maraschin et al., 2009), for further auxin responsive genes expression. Type I/II Phox and Bem1p (PB1) domain is formed by Aux/IAA domain III/IV for interactions with ARFs (Korasick et al., 2015). Aux/IAA plays crucial roles in SCF^{TIR1/AFB}-mediated auxin signaling not only as co-receptors, but also as repressors to auxin response.

Aux/IAA is a multiple gene family, auxin-binding affinities and sensing effects are dependent on Aux/IAA members in distinct tissues and developmental processes (Trenner et al., 2016). Aux/IAA was first discovered as *PS-IAA4/5* and *PS-IAA6* genes from pea (Oeller et al., 1993). Until now, Aux/IAA family has been identified in many plant species, within different numbers and corresponding molecular functions as reviewed in (Luo et al., 2018). Notably, comprehensive genome-wide analysis of Aux/IAA were carried out in poplar and Eucalyptus (Kalluri et al., 2007; Yu et al., 2015), in order to identify preferentially and highly expressed members functional characterizations associated wood formation. Some truncated Aux/IAA proteins were found in different species, for instance in Eucalyptus Aux/IAA family, domain I is partially conserved in EgrIAA11, 20 and 33B, and is absent in EgrIAA29, EgrIAA32, and EgrIAA33A. The conserved degron sequence VGWPP in domain II is missing in EgrIAA20, 32, 33A and 33B (Yu et al., 2015). While in *Carica papaya*, CpIAA11, CpIAA19, CpIAA27, CpIAA31 have been identified without domain III and IV (Liu et al., 2017). The variability of Aux/IAAs provides diverse functions in auxin signaling, resulting in developmental processes regulation and response to environmental stresses.

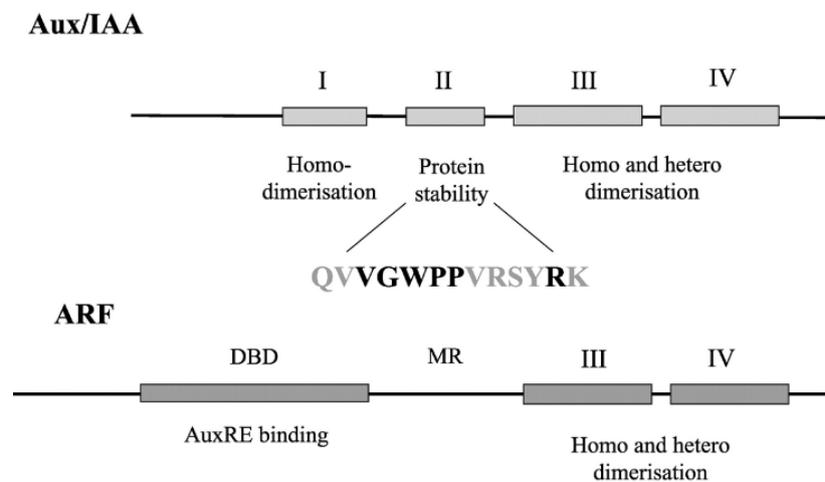


Figure I-13. Protein structures of canonical Aux/IAA and ARF (Kepinski and Leyser, 2002). The 13 amino acids serve to confer instability (conserved are bold). DBD, DNA binding domain; MR, middle region.

4.2 TOPLESS (TPL) family corepressors

When auxin levels are low, Aux/IAA proteins recruit co-repressor TOPLESS (TPL) bind to ARFs and repress their transcription. There are two main co-repressors of GROUCHO family in plants: LEUNIG/LEUNIG_HOMOLOG (LUG/LUH) and TOPLESS/TOPLESS-RELATED (TPL/TPR) groups (Liu and Karmarkar, 2008). Transcription factors within ERF-associated amphiphilic repression (EAR) domain (LxLxLx) was identified as the specific motif to recruit TPL/TPRs, for repression or activation of target genes expression (Szemenyei et al., 2008), as well as subsequently other identified repression domains (RDs, for example DLN_{xx}P, R/KLFGV, TLxLF) which were enriched in board range of TFs for TPL interaction and recruitment (Causier et al., 2012b). TPL/TPR family was identified first in Arabidopsis within five members, of which TPL1-4 interacted with WUSCHEL (WUS) (Kieffer et al., 2006). TPL/TPR proteins contain four canonical domains with conserved motifs: LisH domain at N-terminal region, and also at C-terminal region (CTLH), and two WD40-repeat domains at C-terminal. Of these, LisH domain promotes protein-protein interactions as demonstrated in (Cerna and Wilson, 2005).

The interaction partners and related mechanisms play crucial roles to modulate gene expression for multiple developmental and biological processes such as embryo development, plant immunity, and stress responses. Not all TFs which recruit TPL contain RDs, like JAZ proteins (involved in jasmonic acid signaling) that lack RDs but interact with TPL via adaptor NINJA (Pauwels et al., 2010). As for hormone signaling, a global regulatory role of TPL especially for hormone response (auxin, jasmonate, strigolactone, brassinosteroids, etc) was revealed in (Arabidopsis Interactome Mapping Consortium et al., 2011), through interactions with Aux/IAA, JAZ, SMLX, BES1 repressors (Szemenyei et al., 2008; Pauwels et al., 2010; Ryu et al., 2014; Espinosa-Ruiz et al., 2017). Furthermore, the crystallographic structure of Arabidopsis TPL gave insights into the common domain (LisH) of corepressors (TBL1) which have similar properties (Martin-Arevalillo et al., 2017). In auxin signaling, TPL interacts with Aux/IAA via EAR domain as the co-repressors binding to ARF to restrict auxin response. In addition, co-repressor complexes of BES1–TPL–HDA19 had integrating activations of BR-signaling pathway

(Kim et al., 2019), as well as the interactions of TPR-AtMYB44 to form the complex to recruit histone deacetylase to suppress target genes protein phosphatase 2C (PP2C) expression (Nguyen and Cheong, 2018).

4.3 Auxin responsive factors (ARFs) family

Auxin controls plant growth and development rapidly and precisely by regulating various genes expressions (such as *Aux/IAA*, *GH3* and *SAUR* primary auxin responsive genes), which is activated or repressed by transcription factors including Auxin response factor (ARF), through binding with auxin response cis-elements (AuxREs) located in promoter of these genes. Except for *Aux/IAA* binding with ARFs to control auxin responsive, ARF interacts with proteins by forming homodimers and/or heterodimers with *Aux/IAA* to modulate downstream target genes expression (Guilfoyle and Hagen, 2012). ARF family members contain highly conserved B3-type DNA binding domain (DBD) at N' terminal which recognizes AuxREs (TGTCNN, (Boer et al., 2014)), a variable middle region of transcriptional activation or repression domain (AD or RD), and C' terminal dimerization domain (CTD) called domain III and IV which interact with homologous sequences in *Aux/IAAs* and ARFs (Figure I-13).

ARF function is determined by amino biased acid sequences in middle region, performing distinction of transcriptional activation with glycine (Q)-abundance and transcriptional repression with serine (S)-abundance (Tiwari et al., 2003; Guilfoyle and Hagen, 2007). Arabidopsis ARF members ARF5, ARF6, ARF7, ARF8, ARF19 were identified to act as transcriptional activators and the other members were characterized as repressors based on protoplast transient transfection assay (Ulmasov et al., 1999; Tiwari et al., 2003). Not all ARF transcription factors harbor three relatively conserved domains. Recently, two additional domains associated with DBD domain, dimerization domain (DD) and Tudor-like ancillary domain were found in ARF1 and ARF5 according to crystal structures of ARFs (Boer et al., 2014). Evidence from X-ray crystal structure showed that domain III/IV in Arabidopsis ARF5 and ARF7 adopts Type I/II Phox/Bem1p (PB1) domain (Korasick et al., 2014; Nanao et al., 2014). To date, genome-wide analysis of ARF family has been investigated and characterized in some species besides model plants, and trees such as woody species (poplar, Eucalyptus) (Kalluri et al., 2007; Yu et al., 2014), fruit species (*Vitis*, Citrus, apple, papaya, lichi, pineapple, peach, etc) and tropical tree (physic nut).

4.4 Protein-protein interactions in auxin signaling

Auxin signaling is mainly regulated by interactions of *Aux/IAAs* and ARFs. Protein-protein interactions of *Aux/IAA*, ARF and TPL have crucial roles in transcriptional regulation in auxin signaling pathway. *Aux/IAA* domain I interacts with TPL via LxLxLx motif to form the complex to enhance repress activity (Szemenyei et al., 2008; Causier et al., 2012a). In addition, ARF repressors (such as ARF2 and ARF9) also directly interacted with TPL/TPR, implicating TPL/TPR co-repressors functions in both forms of ARF-mediated repression (Causier et al., 2012b).

The conserved domain III and IV in *Aux/IAA* shares homology with C-terminal domain (CTD) of ARF, which form the heterodimers of *Aux/IAA* and ARF via interactions between domain III/IV and CTD. Using yeast two hybrid system and split firefly luciferase complementation (SFLC) assay, a large scale analysis of *Aux/IAA* and ARF interactions was performed and the majority interactions were *Aux/IAA*-ARF activators, but with few repressors (Vernoux et al., 2011; Li

et al., 2011). Moreover, the comparative interactions showed ARF repressors were more likely to participate in AuxREs binding for regulation instead of involving in auxin signaling pathway with Aux/IAA (Weijers et al., 2005; Vernoux et al., 2011). Several pairs of Aux/IAA and ARF have been identified to regulate plant growth and development, for instance, AtIAA14/SLR interacted with AtARF7 and AtARF19 to regulate lateral root formation (Fukaki et al., 2006), AtIAA12 interacted AtARF5 to regulate embryonic root formation (Weijers et al., 2005), also playing roles for lateral roots formation after IAA14-ARF7-ARF19 module (De Smet, 2010; De Smet et al., 2010). Also, IAA28 was able to interact with ARFs (ARF5, 6, 7, 8, 19, with most prevalent ARF7 and 19) for root basal meristem (De Rybel et al., 2010). In apple, MdARF13 acted as a negative regulator of anthocyanin metabolic pathway via binding MdDFR promoter, while overexpressed MdIAA121 attenuated this inhibition, implicating MdARF13-MdIAA121 interactions participating in regulation of anthocyanin biosynthesis (Wang et al., 2018a). In *Populus*, PtoIAA9 interacted with PtoARF5 to regulate secondary xylem formation (Xu et al., 2019).

5. Roles of auxin and its cross-talk with other hormones in the control of wood formation

Since long, auxin was identified as a key regulator of wood formation. Landmark studies showed that an appropriate balance Auxin/cytokinin in suspension culture triggered mesophyll cells to differentiate into tracheary elements (TEs), lignified xylem-like cells (Fukuda and Komamine, 1980; Twumasi et al., 2009). In *Pinus*, a lack of auxin supply from shoot apex leads to a loss of fusiform shape of cambial derivatives (Savidge, 1983). Here, we focus on the role of auxin and its cross talk with other hormones in initiating the transcriptional program of secondary xylem differentiation. A recent publication described the spatial distribution of the plant hormones auxin, gibberellin and cytokinin in poplar sections encompassing the phloem, the vascular cambium and the xylem (Immanen et al., 2016). They showed that each hormone exhibits distinct concentration maxima with partially overlapping distribution profiles: auxin peaking in the cambial zone, cytokinins in the developing phloem cells and gibberellin in expanding xylem cells (Figure I-14). This suggests that auxin interacts with other hormones to contribute to subsequent cell fate decisions.

5.1 Auxin influx- and efflux- carriers control vascular pattern and wood formation

Auxin is synthesized in young developing parts (such as shoot apex) and be transported to other organs and tissues by cell-to-cell manner, relying on efflux (PINs) and influx carriers (AUX and LAX1, LAX2, LAX3). PIN protein is an auxin efflux carrier, controlling IAA flow and membrane permeability (Bennett et al., 2014). PIN1 localizes asymmetrically in plants basal and lateral plasma membranes, and is strongly expressed in *Arabidopsis* inflorescence stems. Auxin maintains canalization of polar flow by feedback regulation (Aux/IAA-ARF-dependent targeting PIN), which underlies vascular strands formation (Sauer et al., 2006). The *INTERFASCICULAR FIBERLESS/REVOLUTA (IFL1/REV)* gene is required for fiber differentiation, the *rev* mutants presented defection of fiber cells formation, and much lower expressions of *PIN3*, *PIN4* associated with much reduced polar auxin transport (Zhong and Ye, 2001). In contrast, *pin1pin2* mutants displayed increased xylem differentiation (Ibanes et al., 2009). In *Arabidopsis* quadruple *aux1lax1lax2lax3* and triple *aux1lax1lax2* mutants, less numbers of shoot vascular bundles and increased spaces in vascular patterns were observed, while the uneven vascular organization also showed increased xylem and procambial cells in shoots and roots, indicating their pervasive

Chapter I:

Bibliographic review---Part 2. Auxin regulates plant growth and development, especially in wood formation

roles xylem differentiation (Fàbregas et al., 2015). Auxin influx carriers play crucial roles in leaf vascular patterning. Except this, *lax2* mutants had similar phenotypes with overexpressed *HaHB4* (a sunflower HD-ZIP gene) in *Arabidopsis*, both displaying, higher xylem length and increased xylem cell rows. Therefore, LAX2 acts as a negative regulator for vascular patterning and xylem development (Moreno-Piovanò et al., 2017). Auxin polar transport through vascular cambium is a key regulator of xylem vessel cells development and patterning (Hacke et al., 2017). Auxin regulates transcription factors to produce cell wall components and affect cell wall loosening for cell expansion and enlargement (Majda and Robert, 2018). Although it is less traces that how IAA affects lignin deposition during late stages of wood formation, some experimental evidences showed the relationships with NAC transcription factors repression to regulate lignin deposition. It is also worth noting that polar auxin transport is important for vessel spatial patterning and size determination and by altering auxin transport, it is possible to shape the basic hydraulic properties of a woody stem (Johnsson et al., 2018).

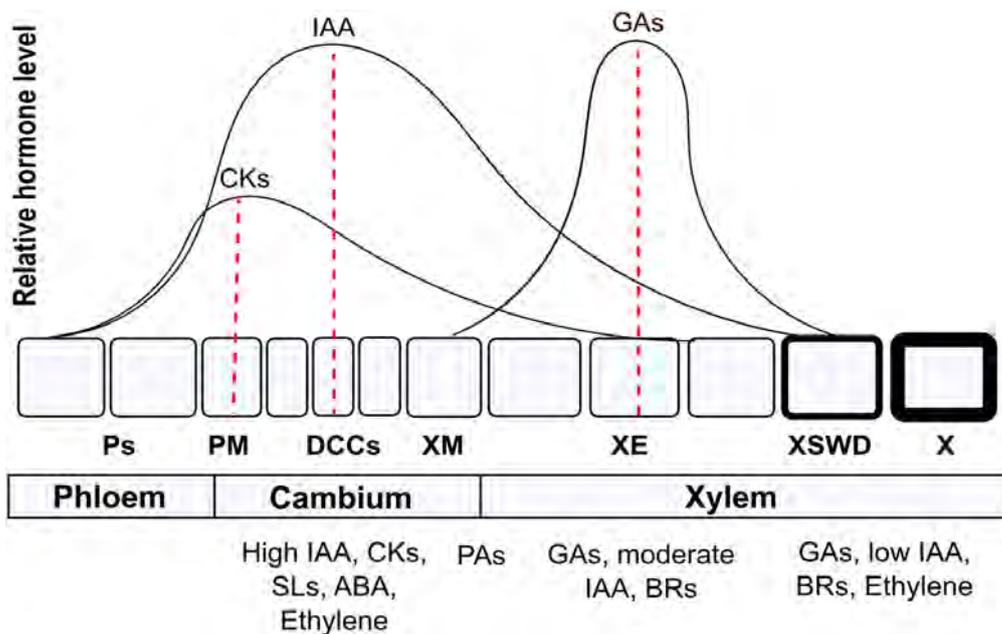


Figure I-14. Different hormones distributions across cambial developmental zones (Buttò et al., 2020). Each hormone maximum levels (IAA and ABA, auxin; CKs, cytokinins, BRs, brassinosteroids) are accumulated in different zones, which are indicated by red dashed lines. DCCs, developing cambial cells; PM, phloem mother cell; XM, xylem mother cell; Ps, Phloem cells; XE, enlarged xylem cells; XSWD, xylem secondary cell wall deposition; X, mature xylem cells.

ABCB (ATP-binding cassette (ABC) trans- porters of the B class)/PGP/MDR subfamily are involved in auxin transport (Geisler and Murphy, 2006). In *Arabidopsis*, four ABCB candidates (*ABCB11/MDR8*, *ABCB14/MDR12*, *ABCB15/MDR13*, and *ABCG33/PDR5*) were identified to express in developing stem vasculature, which may have functions in wood formation and SCW lignification. Similar with *abcb15* mutant, *abcb14* mutants showed reduced auxin polar transport, while smaller metaxylem vessels and altered vascular bundle organization (Kaneda et al., 2011). In *Populus*, expression patterns of 20 ABCB members were analyzed and their phylogenetic relationships and structures were compared with

Arabidopsis for selecting candidates. Moreover, *PttABC1.1* had highly and specifically expression in internodes and developing xylem, acting as an auxin transport protein in *Populus* (Carraro et al., 2012). WALL ARE THIN1 (WAT1) acts as an auxin efflux facilitator located in protoplast, and in *wat1* mutants significantly reduced SCW in xylary and early interfascicular fibers were identified, indicating WAT1 play roles specifically in SCW thickness of wood fiber cells (Ranocha et al., 2010, 2013).

5.2 Reading the auxin gradient

In wood-forming tissue, auxin concentrations peak in the cambium and decay rapidly toward the xylem and phloem (For a review see Bhalerao & Fischer, 2014 and references therein, Immanen et al., 2016). The mechanisms underlying the establishment of this gradient and how it could be interpreted have been deeply discussed, indicating that high auxin concentrations are 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 (Bhalerao & Fischer, 2014 and references therein). The landmark work of Nilsson et al. (2008) showed that auxin-responsive genes in wood-forming tissues of hybrid aspen respond dynamically to changes in cellular auxin levels but the expression patterns of most of them displayed limited correlation with the auxin concentration across this developmental zone.

5.3 Identifying Aux/IAAs and ARFs as mediators of auxin signaling in wood formation

Auxin regulates wood formation through the auxin responsiveness (Nilsson et al., 2008), which is carried out by a series of regulators such as Aux/IAA and ARF transcription factors (multigenic families) involved in auxin perception and signaling. Microarray analysis of xylem vessel formation using *Arabidopsis* suspension cells also identified key transcription factors involved in *Arabidopsis* xylem differentiation including also several members of Aux/IAA genes (Demura et al., 2002). *BDL* encodes IAA12 which represses ARF5-mediated transcription activation of auxin response, and the gain-of-function *bdl Arabidopsis* mutants showed reduced cotyledons vasculature, prevented xylem and SCW formation, and no primary roots formation (Hamann et al., 1999; Hamann, 2002). In *Arabidopsis* *AtIAA8* is highly expressed in vascular tissues and in tracheary elements (Oh et al., 2003; Groover et al., 2003). *AtIAA8/9* showed the similar expression patterns with *zIAA8* (IAA8 of zinnia) in developing vasculature, and regulated lateral roots formation (Groover et al., 2003; Arase et al., 2012). Plants overexpressing *AtIAA20*, *AtIAA30* or *AtIAA31* showed similar developing defects while *AtIAA20* overexpressed mutants exhibited most severe phenotypes like malformed vasculature in cotyledons and no primary roots formation (Sato and Yamamoto, 2008). In tomato, the *IAA9* knock-down *AS-IAA9* lines showed increased numbers of secondary veins in leaves and hypertrophic vascular xylem tissue in leaves and stem (Wang et al., 2005).

Several Aux/IAA genes were identified preferentially expressed in poplar differentiating xylem and cambium such as *PttIAA1*, 3, 4 and 8 (Moyle et al., 2002). Genome-wide analysis of Aux/IAA family showed members preferentially expressed in xylem and cambium tissues in Eucalyptus and poplars (Yu et al., 2015, Moyle et al., 2002; Kalluri et al., 2007; Nilsson et al., 2008; Sundell et al., 2017). In aspen the overexpression of a stabilized version of mutated *PttIAA3* (potential ortholog of *AtIAA20*) presented a reduced cambial cell division while radial extension of the cambial zone increases, as well as reduced wood formation due to decreased width and length of fiber cells and vessel elements and

other auxin-resistant phenotype (Nilsson et al., 2008). This suggests that auxin signaling not only promotes cambium proliferation but also spatially restricts stem cell characteristics within the cambium area (Bhalerao & Fischer, 2014). Overexpression stabilized version of *PtoIAA9m* (mutated version for stabilized protein) lead to significantly reduced radial growth as well as primary growth, inhibitions of cambium periclinal division and secondary xylem development (Xu et al., 2019). Poplar HB7 (*AtHB8* ortholog) is a dose-dependent regulator for cambium activity and xylem differentiation (Zhu et al., 2013). It was proved that auxin mediated module *PtoIAA9-ARF5-HB7/8* could regulate secondary xylem formation (Xu et al., 2019).

Some ARFs are also reported to have potential functions associated with vascular patterning (especially in leaves) and cambium activity mediated by auxin. In *Arabidopsis*, ARF 3 & 4 were recently pointed out as general promoters of cambial activity whereas ARF5/MP has a more specific role in attenuating the activity of *WOX4* to stimulate differentiation (Brackmann et al., 2018). *MP*, *ARF7* and *ARF19* showed strongly and consistently expressed in vascular cambium, implicated by vascular pattern defects in *arf7arf19* mutants, and no initiation of roots in *mp* mutants (Smetana et al., 2019). Genome-wide analysis of *ARF* family showed members preferentially expressed in xylem and cambium tissues in Eucalyptus and poplar (Yu et al 2014, Kalluri et al., 2007, Sundell et al., 2017). Most functional study of ARF for vascular tissue formation concentrated in the ARF member ARF5/MP, which regulate directly nearly half of Aux/IAA members expression, whereas Aux/IAA proteins had negative feed-back effects on ARF5/MP protein activity by forming homo-/heterodimers (Krogan et al., 2014; Ramachandran et al., 2017). To date, ARF5/MP was demonstrated regulating meristems activity either embryogenesis or post-embryo organs developments, which are mainly involved in root formation, vascular patterning, cotyledon separation, shoot apical meristem (SAM) and floral formation (more *ARF5/MP* functional characterizations were shown in Chapter III-Introduction part).

5.4 The cross-talk between auxin and cytokinins stimulates cambium activity

Cytokinin (CK) is produced at root tips, developing seeds, leaves or other aerial parts, and can promote cell division and cytokinesis in plant shoot and root growth and development. Cytokinins (CK) signaling is known to be important in the maintenance and proliferation of cambial cells and in cambium cell specification (Milhinhos and Miguel, 2013; Bhalerao and Fischer, 2014 and references therein). A reduction in cytokinin level by overexpression of a CK catabolic gene, *Arabidopsis CYTOKININ OXIDASE 2 (CKX2)*, in transgenic poplar leads to a decrease in the number of cambium cells and concomitantly a reduced stem diameter (Nieminen et al., 2008). CK regulates auxin polar transport to influent vascular development. Cytokinins promote the bisymmetric distribution of the PIN-FORMED (PIN) auxin efflux proteins, which channel auxin toward a central domain. High auxin promotes transcription of the cytokinin signaling inhibitor AHP6, which closes the interaction loop (Bishopp et al., 2011). On the other hand, aspen trees overexpressing the *AtIPT7* gene, encoding one key enzyme in the biosynthesis of major bioactive CKs, displayed stimulated cambial cell division activity resulting in dramatically increased production of the lignocellulosic trunk biomass. The elevation of the CK content (reached a peak in developing phloem zone) led to an increase in cambial auxin concentration, highlighting the interconnected nature of these two hormonal gradients to stimulate cambial activity (Immanen et al., 2016).

5.5 The cross-talk between auxin and gibberellins promotes expansion of cambial derivatives

The role of gibberellin (GA) in wood development in trees is supported by several lines of evidence (Mauriat et al., 2014; Milhinhos and Miguel, 2013 and references therein). For instance, the overexpression of *GIBBERELLIN 20-OXIDASE1* (*GA20ox*), a GA biosynthetic gene, in poplar results in increased growth and xylem fiber length (Eriksson et al., 2000). Overexpression of poplar orthologues of *AtGID1* (*GIBBERELLIN INSENSITIVE DWARF1*) resulted in phenotypes similar to those of the constitutive overexpressors of *GA20ox* i.e. increases in xylogenesis and cell elongation with the exception of increased fiber length (Mauriat and Moritz, 2009). Noteworthy, the increase in xylogenesis in decapitated aspen trees is stronger when adding both IAA and GA compared to what is observed with adding only one hormone at a time (Björklund et al., 2007). Consistent with the alteration of cambium proliferation observed in poplars affected in GA metabolism, GA stimulates IAA transport and in turn, IAA stimulates GA biosynthesis genes likely through Aux/IAA-ARF signaling elements. Strikingly, both hormones shared a common transcriptomic signature including many transcripts related to cell growth. These findings strongly support that these two hormones play important roles in the post-meristematic expansion of cambial derivatives (Björklund et al., 2007).

The gradual reduction in auxin concentration with increasing distance from the cambium and concurrent increase concentration of active GA may constitute a signal for cells to transition from an expansive phase to one of maturation (Immanen et al., 2016). Building on the hypothesis that the interplay between GA and IAA may act as a signal to initiate fiber differentiation, (Johnson et al., 2018) tested if these hormones were able to regulate the expression of the top level regulators of xylem cell fate and SW deposition, the secondary wall NAC (NAM/ATAF/CUC)-domain transcription factors. Based on high spatial resolution profiling in *Populus* (Sundell et al., 2017) and comparative phylogeny, they firstly identified two clades of secondary wall NACs presenting distinct expression profiles and containing in their promoters both GA and IAA response elements. Interestingly, they reported that those associated with fiber and SCW formation, the so-called “SND/NST”, were induced by GA treatment and tended to be repressed by auxin. On the other hand, both GA and auxin could induce vessel-specific “VND” genes. These findings were reflected at the anatomical level. In decapitated poplar stems, exogenous auxin treatment reduced cell wall thickness while GA promotes SCW deposition. High concentration of auxin in the cambial and expansion zones represses fiber-specific NAC TFs and allows differentiation of xylem vessels. When auxin concentration decreases and GA levels are high, all wood-associated NACs are induced (Johnsson et al., 2018). This supports a central role of the IAA/GA cross-talk in determining cell fate possibly through the control of SND/NST and VND master regulators and/or direct regulation of SCW biosynthesis (Johnson et al., 2018). However, the underlying mechanisms of how auxin and GA coordinate expression of wood-associated NAC TFs remains unclear although the presence of overlapping GA and auxin responses elements in the promoter of some wood-associated NAC genes could provide a way for the suggested IAA/GA cross-talk. It is also worth noting that deregulation of a SND/NST member, can in turn, affect auxin homeostasis. For instance, transformation of the PtSND2 activator into a strong repressor affects auxin biosynthesis, transport and signaling and, as a result, repress the normal growth and vascular development of transgenic poplar plants (Wang et al., 2014). The promoters of genes involved in auxin biosynthesis, transport and signaling contain SNBE sites suggesting that they could be targets of NAC TFs (Johnson et al., 2018).

5.6 Interaction between auxin and other hormones (brassinosteroids and strigolactone)

Pioneer works on *Zinnia* mesophyll cell cultures demonstrated that brassinosteroids (BRs) are able to regulate xylem differentiation (Yamamoto et al., 2006). In *Arabidopsis* shoots, Ibanes (Ibanes et al., 2009) showed that auxin polar transport coupled to BR signaling is required to determine the radial pattern of vascular bundles. Recently, Jin have overexpressed one poplar ortholog of AtCYP85A2, known to encode a bifunctional cytochrome P450 monooxygenase, catalyzing a final rate-limiting step in the BR-biosynthetic pathway. Overexpression of PtCYP85A3 in poplar increased the endogenous BR levels and significantly promoted growth, particularly xylem formation without apparent alteration of SCW properties (Jin et al., 2017).

The work of Agusti revealed a role for strigolactone (SL) signaling in the regulation of secondary growth conserved among species. Exogenous application of artificial SL (GR24) on stems of *Eucalyptus globulus* induced cambium division and *Arabidopsis* SL deficient mutants exhibited less radial growth while presenting high IAA levels and signaling (Agusti et al., 2011). The use of double mutants SL and IAA deficient suggested that SL function predominantly downstream of auxin signaling that positively regulates secondary growth.

Part 3. CRISPR/Cas9 current advances and applications

1. Introduction

Genome editing is a technic which modifies specific DNA sequences in cells and organisms using sequence-specific nucleases (SSNs). Compared with random mutagenesis (induced by physical, chemical or biological mutagenesis) and interruptions of gene functions by repressing the corresponding mRNA (antisense RNA, virus-induced gene silencing, RNA interference), genome editing simulates the natural mutation and generates site specific mutagenesis for more precise and less time- and cost-consuming genetic evolution.

SSNs-induced genome editing technics include several generations: 1) engineered meganuclease (EMNs)/homing endonuclease, 2) zinc-finger nucleases (ZFNs), 3) transcription activator-like effector nucleases (TALENs) and 4) clustered regularly interspaced short palindromic repeats (CRISPR/Cas9), which accomplish the precise alteration of DNA sequence *in vivo* (Figure I-15A). SSNs induce double strand breaks (DSBs) in specific chromosomal sites leading to various DNA modifications, such as indels (insertions and/or deletions) and substitutions produced by non-homologous end joining (NHEJ) pathway, and insertions and substitutions of homologous donor templates generated by homology-directed repair (HDR) pathway, which produces the precise repair (Figure I-15B). If the mutation sites occurred at the gene coding region, reading frame shift, amino acid alteration and/or premature stop codon will be generated, leading to truncated and/or inactive proteins.

Meganucleases existing nearly in all microbes, recognize and cleave typically 12-40bp long DNA sequence to generate DSBs in eukaryote genomes. The mega nuclease, I-SecI belonging to LAGLIDADG family, was first applied in tobacco plants to enhance homology-directed repair (HDR) (Daboussi et al., 2015). Although engineered meganucleases have been used in plants such as Arabidopsis, cotton and maize, this genome editing tool is restricted because of the overlap between the DNA binding domain and the cleavage regions. Zinc-finger nucleases (ZFNs) contain two domains: the DNA-binding domain (with zinc fingers binding motifs) and the FokI restriction endonuclease domain (for DNA cleavage). The binding specificity of the designed zinc-finger domain directs the ZFN dimers to a specific genomic site. Over the last decade, ZFNs have been successfully used in some plant species (Arabidopsis, maize, tobacco, soybean), and the first use in trees was to develop apple and fig tree transgenic lines with GUS gene repair by heat-inducible ZFN expression system (Peer et al., 2015). ZFN was then used to mutagenize two essential floral genes in poplar, LEAFY (LFY) and AGAMOUS (AG) in 2016, showing very low mutated efficiencies (Lu et al., 2016). ZFN has fewer off-target effects but its construction is difficult and costly, with relative low efficiency, limiting its use in various organisms. Transcription activator-like effector nucleases (TALENs) also contain FokI nuclease domain like ZFNs fused with transcription activator-like effector (TALE) repeats, which is secreted by plant pathogenic *Xanthomonas* bacterium via Type III secretion system. TALEN emerged in 2009 (Moscou and Bogdanove, 2009) and was first applied in Arabidopsis (Cermak et al., 2011) and tobacco (Boch et al., 2009). The site specificity depends on the repeat variable di-residue (RVD) located in TALE repeats (position 12 and 13), which make easier engineering gene targeting. However, TALENs has relative low efficiency and is sensitive to target's DNA methylation. The difficulty of constructs assembly was addressed by the Golden Gate Cloning strategy.

The clustered regularly interspaced short palindromic repeats (CRISPR)-associated protein (Cas) type II system was first identified as an adaptive immune system in *Streptococcus pyogenes* (Jinek et al., 2012). The CRISPR Type II system has been widely applied in many animal and plant organisms as an efficient genome editing technology, just using the necessary components of Cas9 nuclease and a single guide RNA (sgRNA). In addition, it is easier to manipulate because of the small size of constructs and is less time-consuming and cost than ZFNs and TALENs, and even feasible for multiple targeting (Gaj et al., 2013; Ain et al., 2015; Yamamoto, 2015). The mechanisms of action of the ZFNs, TALENs and CRISPR/Cas9 are reported in Figure I-15-A. The three methods generate double strand breaks (DSB) which need repair either Non-homologous end joining (NHEJ) pathway or Homology-directed repair (Figure I-15B). The possible types of mutations, include monoallelic, bi-allelic (in diploid species) and chimeric mutations (Figure I-15C). The pros and cons of ZFNs, TALENs and CRISPR/Cas9 are reported in Table I-S1.

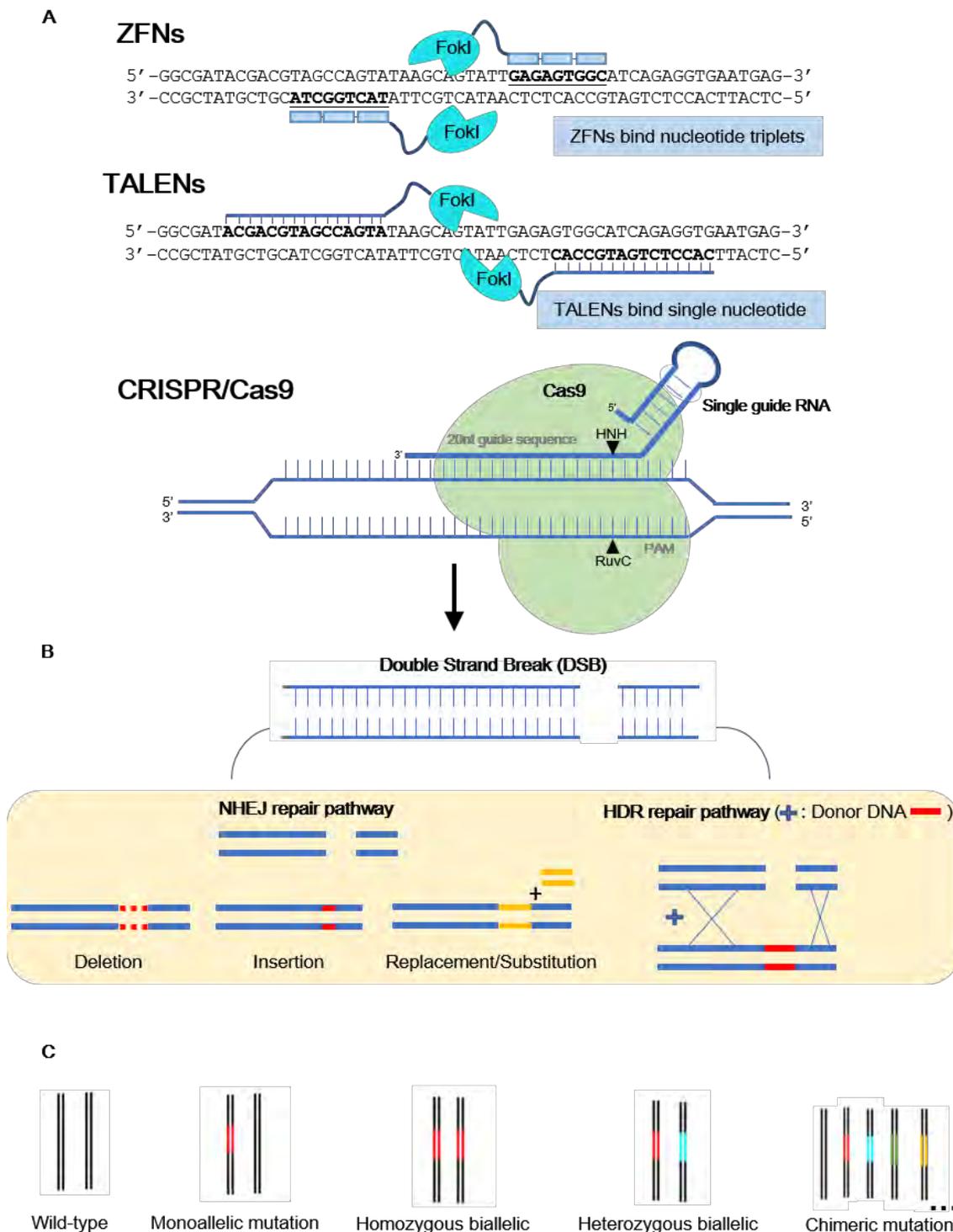


Figure I-15. Mechanisms of ZFN, TALEN, and CRISPR/Cas9 (A), site-specific endonucleases generate double strand breaks (DSBs) (B) and possible types of mutations in diploid organisms (C). Double strands break (DSB) are repaired either by Non-homologous end joining (NHEJ) pathway or by Homology-directed repair; Possible types of mutations in diploid species, including monoallelic, biallelic and chimeric mutations.

2. CRISPR/Cas system origin and mechanism

Here we will review the CRISPR/Cas origin and mechanism, survey its applications in plants with an emphasis on tree species. We will also present the optimization of (i) the single guide RNAs (sgRNA) and of the Cas9 expression cassettes, (ii) the delivery methods (iii) the methods of mutation detection.

CRISPR-Cas system was identified early as an adaptive immune system from invading virus in bacteria (Mojica et al., 2005). Briefly, in the first-step of infection, the bacteria keep partial fragment of viral DNA in its genome (stored in CRISPR spacers) as a memory, which works for defense for the next virus attack. CRISPR locus is transcribed into mature CRISPR RNA (crRNA) which is the basis for resistance, binding with Cas proteins, to form the protein-RNA complex for degradation of the invading DNA. A short palindromic repeated sequence was discovered in *E.coli* in 1980s (Ishino et al., 1987). The CRISPR loci includes spacers (from foreign DNA) and repeats (20-50bp) interval, and a leader sequence located in upstream of repeats. The associated cas genes (located adjacent to CRISPR locus) were identified in prokaryotes (bacteria and archaea) (Jansen et al., 2002). A specific short DNA sequence (2-6bp) adjacent to proto-spacers is called “protospacer adjacent motif” (PAM). It is associated with proto-spacers acquisition from extrachromosomal elements, and exists in diverse CRISPR-Cas systems, its sequence for spCas9 is NGG (Mojica et al., 2009). The timeline of CRISPR/Cas history is presented in Figure I-16.

In Type II CRISPR-Cas9 system, a trans-encoded small RNA (tracrRNA), complementarity to crRNA precursor transcript, serves as a guide for ribonuclease 3-aided (RNaseIII) processing of mature pre-crRNA (Deltcheva et al., 2011). The chimeric tracrRNA:crRNA (together with a loop sequence to generate single guide RNA) recruits CRISPR-associated Cas9 protein (formerly named Cas5 or Csn1) which belongs to large families in bacteria species (Hsu et al., 2013). The engineered sgRNA and Cas9 form a complex to cleave cognate DNA adjacent to PAM both *in vitro* and *in vivo*. In contrast to ZFNs and TALENs which depend on DNA-protein domain for target recognition, CRISPR/Cas9 only needs to design the sgRNA sequence within the specific target sequence generally at the 5'-ter ends. Cas9 protein contains HNH and RucV domains to generate site specific blunt-end DSBs. HNH domain cleaves the target strand (complementary to the guide RNA sequence), whereas the RucV domain cleaves the non-target strand (Figure I-15A) (Gasiunas et al., 2012). The main steps for CRISPR/Cas9 applications are the followings: 1) CRISPR/sgRNAs construction, 2) delivery transformation, 3) Mutations detection and corresponding phenotypes observations.

Like CRISPR/Cas9 system, recently, new Cas efficient enzyme Cpf1 (also called Cas12a) has also been used for genome editing in many organisms including plants, which recognizes T-rich PAM sequence for staggered DSB cleavage at distal and downstream of PAM, with the 5-nucleotides overhang starting at 18 nucleotides 3' of the PAM (Zetsche et al., 2015). Cas12a requires only one single crRNA for cleavage which is shorter than engineered single guide RNA for Cas9, and the smaller size of Cas12a has advantages of vector construction and delivery (Ledford, 2015). In order to expand Cas12a application, variants of Cas12a (AsCas12a, FnCas12a, LbCas12a) have been used for genome editing in many plants including crops and model species.

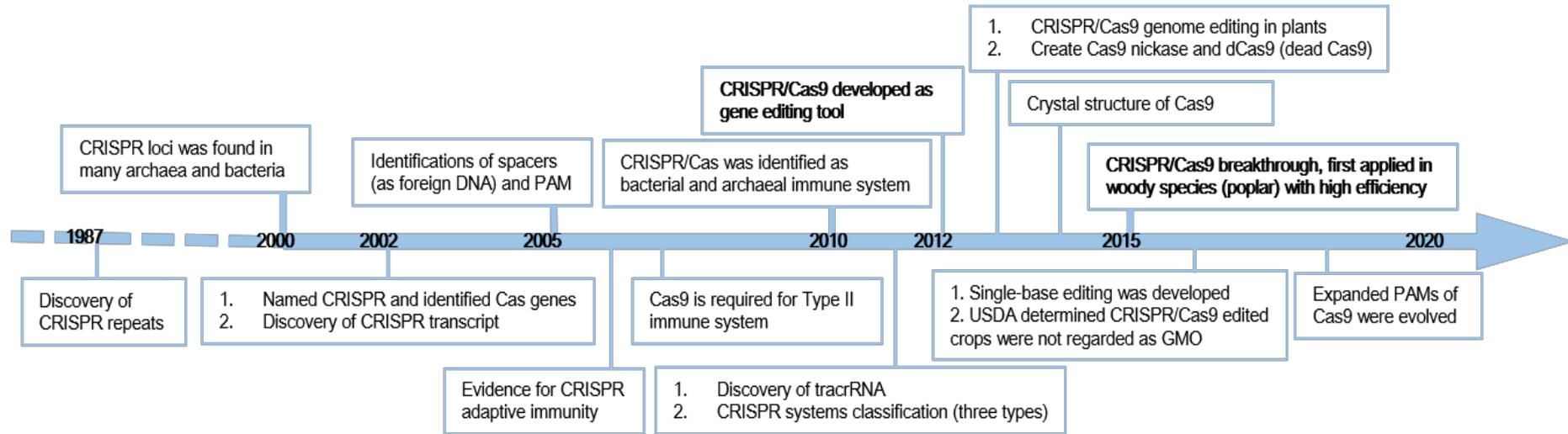


Figure I-16. Timeline of CRISPR/Cas9 genome editing key developments and progress in plants (Adapted from (Han and She, 2017; Mushtaq et al., 2018; Wu et al., 2020)).

3. CRISPR/Cas9 vector systems

CRISPR/Cas9 has been applied in many plant species since 2013, as attested by continuous reports in Nature Biotechnology (Li, 2013; Nekrasov et al., 2013; Shan et al., 2013).

At the beginning of CRISPR/Cas9 technique's development, very low mutation efficiencies were obtained but since that various optimized vectors have been emerged for more efficient editing (Ma et al., 2015; Ren et al., 2016), CRISPR/Cas9 system has been widely used in monocots such as wheat, rice, maize with stable vectors and considerable efficiency, while it is still a challenge to implement high edition efficiency for polyploidy plant species due to genetic redundancy (Ryder et al., 2017). The earliest implementation in tree was in citrus with the target gene *PHYTOENE DESATURASE (PDS)*, and its disruption leading to an easy identification of visible albino and often death of knockout mutants (Jia and Wang, 2014). The greatest progress of gene editing in heterozygous tree species has been made in poplar with high efficiency targeting of two lignin biosynthesis genes *4CL1*, *4CL2* (Zhou et al., 2015; Tsai and Xue, 2015). However, there still exists some technical obstacles in most gymnosperms (non-flowering) mainly due to the lack of efficient transformation methods, although several transformation protocols were carried out in a few conifer species such as Korean fir (Lee, Hyoshin et al., 2014) and pines (Richard Wenck et al., 1999). Therefore, most CRISPR/Cas9 system applications in trees are dicots species, mainly including fruit trees, tropical trees and woody tree species (Table I-1).

With the acceleration by CRISPR/Cas9, T0 generation trees with knockout mutations will be obtained in tremendously shorter reproductive cycles (Fan et al., 2015; Tsai and Xue, 2015; Igarashi et al., 2016). However, genome editing in outcrossing trees faces the challenge of heterozygosity, mainly due to the single nucleotide polymorphism (SNP) existing at target sites, which increases off-targets and can be harnessed by CRISPR-specific allele editing (Bewg et al., 2018). In general, by DNA repair of NHEJ and HDR, more than half of mutations in T0 trees are bi-allelic (homozygous and/or heterozygous, Figure I-15C), indicative of important inheritance in vegetative propagation (such as in poplar) (Bewg et al., 2018). In addition, considerable chimeric mutation rates occurred in T0 always results in ectopic and unstable phenotypes (Ding et al., 2020). The success of CRISPR/Cas9-mediated editing in trees with considerable edited efficiency depends on the delivery methods and the optimized, stable and adoptable CRISPR/sgRNA vectors used in different species.

Table I-1. A summary of different optimized CRISPR/Cas9 systems applied in various tree species

tree species	Cas9 type ¹	promoter for sgRNA	delivery method	sgRNA ²	assembly method	methods for genotyping	mutation efficiency	types of mutations	reference
citrus	spCas9, CaMV 35S	CaMV 35S	Agrobacterium, epicotyl	1, 1	three-way ligation	subcloning	3.2-89.4%	biallelic	Jia and Wang, 2014; Jia et al., 2016; Jia et al., 2017b
	saCas9, CaMV 35S	CaMV 35S, AtU6-1	Agrobacterium, epicotyl	3, 1	three-way ligation	Next Generation Sequencing	15.55-79.67%	chimeric	Jia et al., 2017a
	spCas9, <i>A. thaliana</i> YAO, human	AtU6-26	Agrobacterium, epicotyl	3, 1	conventional cloning (pYAO:hSpCas9 modified)	subcloning	45.5(10/22)-75(21/28)%	homozygous, monoallelic, biallelic	Zhang et al., 2017
	spCas9, 35S, plant	AtU6-1	Agrobacterium, epicotyl	1 or 5, 1	conventional cloning	PCR amplicons sequencing, subcloning	11.5-85.7%	biallelic, chimeric, homozygous, heterozygous	Peng et al., 2017; Wang et al., 2019
apple	spCas9, 2x35S, plant	AtU6-1	Agrobacterium, leaf disc	4, 1	Golden Gate Cloning	subcloning, PCR amplicons sequencing	13.6(6/44)%	biallelic, chimeric	Nishitani et al., 2016
	spCas9, PcUbi4-2, <i>A. thaliana</i>	MdU3, MdU6	Agrobacterium, leaf disc, vacuum infiltration	2, 2	Gateway LR recombination	subcloning	85.1-93.0%	homozygous, heterozygous, biallelic, chimeric	Charrier et al., 2019
	CRISPR RNPs	/	Protoplast Transformation	6-8, 1	<i>in vitro</i> cleavage assay	Targeted deep sequencing	0.5-6.9%	/	Malnoy et al., 2016
grape	spCas9, 35S, plant	AtU6	Agrobacterium, suspension-cell	2, 1	PCR, homologous recombination(HR)	subcloning, PCR amplicons sequencing	100.0%	heterozygous, chimeric	Ren et al., 2016
	spCas9, Pcub, <i>A. thaliana</i> , plant	AtU6-26	Agrobacterium, embryogenic callus	4, 1	conventional cloning	subcloning	3.2-72.2%	chimeric	Nakajima et al., 2017

Chapter I:

Bibliographic review----Part 3. CRISPR/Cas9 current advances and applications

	spCas9, 2x35S, plant	AtU3b, AtU3d, AtU6-1, AtU6-29	Agrobacterium, proembryonal masses (PEM)	4, 4	Golden Gate Cloning	sequence PCR amplicons	31.0%	monoallelic, biallelic	Wang et al., 2018a
	spCas9, CaMV 35S	AtU6	Agrobacterium, Grape Suspension Cells	4, 1	PCR, homologous recombination(HR)	subcloning, T7EI, PCR/RE	1.7-86.6%	chimeric	Ren et al., 2019
	CRISPR RNPs	/	Protoplast Transformation	4, 1	<i>in vitro</i> cleavage assay	Targeted deep sequencing	0.1%	/	Malnoy et al., 2016
poplar	spCas9, 35S, plant	AtU3b, AtU3d, AtU6-1, AtU6-29	Agrobacterium, leaf disc	4, 1-4	Golden Gate Cloning	subcloning	50-86.4%	chimeric	Liu et al., 2015
	spCas9, CaMV 35S, human	MtU6.6	Agrobacterium, callus	1, 1	blunt ended In Fusion Cloning, conventional cloning	PCR amplicons sequencing	100.0%	biallelic	Zhou et al., 2015, Tsai and Xue, 2015
	spCas9, 2x35S, plant	AtU3b, AtU6-1, AtU6-29	Agrobacterium, leaf disc	3, 3	Golden Gate Cloning	subcloning	45.4-50.9%	homozygous, heterozygous, chimeric	Fan et al., 2015, Yang et al., 2017, Wan et al., 2017, Wang et al., 2017a, Xu et al., 2017, Shen et al., 2018, Gui et al., 2019
	spCas9, 2x35S, human	AtU6-26	Agrobacterium, callus	2, 1 or 2	conventional cloning	PCR amplicons sequencing, sbucloning	62.3- 81.3%	homozygous, biallelic, chimeric	Elorriaga et al., 2018

	spCas9, PcUbi4-2 (<i>Petroselinum crispum ubiquitin4- 2</i>), <i>A. thaliana</i>	AtU6-26	Agrobacterium	1, 2	GATEWAY cloning	Amplicon deep sequencing	12-87%	chimeric	Takata et al., 2019
	spCas9, CaMV 35S	AtU6	Agrobacterium, leaf disc	1, 1	conventional cloning	subcloning	30-73.4%	homozygous, biallelic, chimeric	Bruegmann et al., 2019
	spCas9, 35S	AtU3b, AtU3d	Agrobacterium, explants	2, 1	conventional cloning	PCR amplicons sequencing	69.6-89.7%	homozygous, biallelic, heterozgous, chimeric	Muhr et al., 2018
cassava	spCas9, CaMV 35S	AtU6-26	Agrobacterium, embryogenic callus	2, 1	conventional cloning	subcloning	90%-100%	monoallelic, biallelic, chimeric	Odipio et al., 2017
kiwifruit	spCas9, 35S, plant	AtU6-1	Agrobacterium, leaf discs	4, 1-2	Golden Gate Cloning	T7 endonuclease I, subcloning	0-8.3% for CRISPR/Cas9, 65.4-91.7% for PTG/Cas9	biallelic, chimeric	Wang et al., 2018
<i>Parasponia andersonii</i>	spCas9, 35S, <i>A.thaliana</i>	AtU6	Agrobacterium, tissue explants	1 or 3, 1	Golden Gate Cloning	PCR amplicons sequencing	48.3-88.9%	heterozygous	van Zeijl et al., 2018
cacao	spCas9, CaMV 35S	AtU6-26	transient transformation, Agrobacterium infiltration, leaf	2, 2	Golden Gate Cloning	subcloning	27.0%	chimeric, heterozygous	Fister et al., 2018
rubber tree	CRISPR RNPs	T7 RNA polymerase	PEG-mediated rubber tree protoplast transformation	2, 1-2	T7 High Efficiency Transcription Kit	Targeted deep sequencing	2.7-5.6%	/	Fan et al., 2020

Chapter I:

Bibliographic review----Part 3. CRISPR/Cas9 current advances and applications

	CRISPR RNPs	T7 RNA polymerase	PEG-mediated rubber tree protoplast transformation	3, 1-3	T8 High Efficiency Transcription Kit	Targeted deep sequencing	1.3-20.1%	/	Fan et al., 2020
coffee	spCas9, 2x35S, rice/plant	CcU6	Agrobacterium, embryogenic callus	3, 1-2	overlapped PCR	subcloning	30.4(28/92)%	heterozygous, homozygous	Breitler et al., 2018
Eucalyptus	spCas9, 2x35S, human	AtU6	Agrobacterium, hairy roots	2, 2	Golden Gate Cloning	PCR amplicons sequencing, sbucloning	93.3-100%	monoallelic, biallelic, chimeric	Dai et al., 2020

Cas9 type¹: Cas9 with protein tag, promoter for Cas9, and codon optimized;
 sgRNA²: total sgRNAs for each target, and sgRNA number in each vector.

3.1 sgRNA expression cassettes

The typical synthetic gRNA which is simplified from tracrRNA and crRNA, acts as a guide for Cas9/sgRNA nuclease complex. The promoter chosen to drive the sgRNA determines its expression. The sgRNA's length and specificity affects off-target and on-target editing efficiencies. In general, U3 or U6 small nuclear RNA gene promoters are selected as sgRNA's promoters (Jiang et al., 2013), because of their small sizes and suitability to drive sgRNA transcription by RNA polymerase III. The U3/U6 promoter-sgRNA cassettes can be generated by either target-adaptor ligation or by overlapping PCR. Arabidopsis U3 (AtU3b, AtU3d) and U6 (AtU6-1, AtU6-26, AtU6-29) are the most used in many tree species (refer to pYLCRISPR/Cas9 intermediate construct created by (Ma et al., 2016). In apple, the homologous MdU3/MdU6 were used for driving sgRNAs (Charrier et al., 2019). Medicago U6 promoters (MtU6) were selected in citrus (Jia and Wang, 2014; Jia et al., 2017a, 2017b, 2016) and poplar (Zhou et al., 2015), respectively, as sgRNA's promoters (Table I-1). Polycistronic tRNA (transfer RNA) precursor sequences (pre-tRNA) which are recognized by RNase P and RNase Z, also can promote sgRNA expression for multiplex genome editing. Since the limitations of multiple sgRNA editing via CRISPR/Cas9, like the assembled capacity of a large-size vector within multiple-modules and specific start ribonucleotide due to Pol III-transcription, endogenous RNA-processing system was performed using a single synthetic gene with tRNA-gRNA architecture, which simulates precise transcript *in vivo* using endogenous RNases (Xie et al., 2015) (Figure I-17B). The PTG/Cas9 system was developed since 2015 based on plant endogenous glycine tRNA (77bp) processing, and this system utilized a single RNA Polymerase III promoter to drive multiple sgRNAs, performing up to 100% mutation rate in rice (Xie et al., 2015). Pair-sgRNAs cloning was carried out in kiwifruit using polycistronic tRNA-sgRNA cassette (PTG/Cas9), showing 10-fold higher efficiency than CRISPR/Cas9 (Wang et al., 2018b) (Table I-1).

3.2 Cas9 expression cassettes

Cas9 is a dual RNA-guided endonuclease that recognizes and cleaves DNA based on complementary base pairing. The optimization of the heavily utilized Cas9-mediated genome editing in various organisms includes: 1) the Cas9 type, 2) the Cas9 codon optimization, and 3) the promoters driving Cas9 expression. In eukaryotes, Cas9 requires single nuclear localization signal (NLS) or two NLS flanked to function in cell nuclei.

In plants including trees, most used Cas9 are from *Streptococcus pyogenes*'s origin and plant- or Arabidopsis-codon optimized versions of Cas9 genes were designed to guarantee high expression (Table I-1). In poplar, citrus and Eucalyptus, the Cas9 used was human-codon optimized (Zhou et al., 2015; Elorriaga et al., 2018; Zhang et al., 2017; Dai et al., 2020), exhibiting high editing efficiency reaching 100% in poplar. *Staphylococcus aureus* Cas9 (SaCas9) has a smaller size than spCas9 and was tried to facilitate efficient genome editing in citrus, with higher editing efficiency (up to 79.67%) as compared with spCas9 (Jia et al., 2017a). SaCas9 /sgRNA was employed as a functional alternative tool in several plants including Duncan grapefruit to knockout *PDS*, targeting 21-23 nucleotides with variable (15.6-79.7%) editing efficiency. As an expanded technic, CRISPR/LbCas12a (LbCpf1, PAM recognition is TTTV, V=A/C/G) generates staggered ends distal from PAM using single crRNA. It has been applied in targeting *CsPDS* (by agroinfiltration) for implementation and CsLOBP (*CsLOB* promoter, by *Agrobacterium*-mediated constitutive expression) in citrus to obtain reduced canker symptoms (Jia et al., 2017b).

Dead Cas9 (dCas9) is the Cas9 point modified version resulting in loss of endonuclease activity. The dCas9 activation system (CRISPRa) employs the transcription factors or proteins mainly applied for regulation of gene expression and epigenome editing (Brocken et al., 2018). Base-editing systems are accomplished by nCas9 (Cas9 nicknase), which can introduce the specific base changes, including cytosine base-editor (CBE) system and adenine base editors (ABEs) system (Komor et al., 2016; Gaudelli et al., 2017). For widely used in plant species, Liu group developed the efficient base-editor 3 (BE3) system and Gao group optimized ABE systems with enhanced sgRNA to increase conversion efficiencies (Hess et al., 2017). To date, no applications of dCas9, nCas9 in trees yet.

Constitutive promoters such as Ubiquitin (Ubi) and Cauliflower mosaic virus (CaMV) 35S promoters are used for driving Cas9 in plants including tree species (Table I-1). In citrus, the Cas9 was driven by the Arabidopsis YAO promoter, which is preferentially expressed in undergoing meristem or embryo organisms, inducing editing efficiency of 45.5-75% (Zhang et al., 2017). In addition, N-terminal FLAG-protein tag (hydrophilic peptide: DYKDDDDK) is fused with Cas9 in some cases, to increasing expression efficiency in eukaryotes, for instance in citrus and apple.

However, Cas9 optimized codon sequence neither protein tags seem to be necessary *in planta*, because comparable high editing efficiencies were detected with no optimized cassettes (Mao et al., 2013; Jiang et al., 2013; Nekrasov et al., 2013; Xie and Yang, 2013; Lawrenson et al., 2015).

3.3 CRISPR/Cas9 constructs assembly

Simultaneous expression of engineered cassettes, including programmed Cas9 cassette, single or multiple sgRNA cassette(s) as well as the selectable marker, is necessary to introduce CRISPR/Cas9-mediated precise genome editing. Several recombinational cloning methods have been developed to generate single CRISPR/Cas9 binary vector containing multiple modules, for most used *Agrobacterium*-mediated transformation in plants: 1) Conventional cloning; 2) Golden Gate Cloning; 3) Gateway Cloning; 4) Gibson Assembly (Gibson et al., 2009). Different construction methods of CRISPR/Cas9 used in trees are shown in Figure I-17, and the first two methods are commonly used in trees. Multiple sgRNAs were constructed by Golden Gate Cloning (or Gibson assembly) in the same cassette, aiming to target homologous genes/multiple gene families, or to target the same gene for increasing editing efficiency.

Conventional PCR cloning, Golden Gate and Gateway cloning are all based on restriction enzyme sites for assembling one or more sgRNA expression cassettes. Until now, both regular PCR cloning and Golden Gate Cloning are widely used for CRISPR/Cas9-sgRNA vector construction. Gateway cloning was used in apple, pear and poplar, it needs BP and LR reaction to introduce engineered target regions into the destination vector, using attB and attR specific sites (Charrier et al., 2019; Takata et al., 2019). Homologous recombination (HR) by overlapped PCR was used to connect AtU6-sgRNA fragments and CRISPR/Cas9 vector in grapes (Ren et al., 2016) (Table I-1).

One distinction from animals and a challenge of CRISPR/Cas9 in plants is to resolve multiple targeting, facing the problem of assembling multiple cassettes including Cas9, selectable marker, sgRNAs into one plasmid for stable *Agrobacterium*-mediated transformation. Therefore, the strategy of multiple targeting in plants is imperative, as an emerged typical multi-component transcriptional unit system (MCTU) (Figure I-17C).

Golden Gate Cloning has the advantage of simultaneously cut and ligate multiple (eight at most in plants) well PCR-prepared and arranged intermediate vectors in two steps cloning, based on four Type II restriction enzymes: BsaI, BbsI,

and BsmBI (Esp3I) (Engler et al., 2008). Nowadays by means of Golden Gate Cloning, there are three CRISPR/Cas9 binary vector systems adoptable and simplified for monocots and dicots:

1) one toolkit requires only BsaI for final construct harboring maize-codon optimized Cas9 and two sgRNA cassettes (Xing et al., 2014); 2) A versatile toolbox is pYLCRISPR/Cas9 system which is mostly applied in plant species, with plant codon optimized Cas9 and allowing to target up to eight genes (e.g., 8 FTL homologous gens in rice) (Ma et al., 2016) (Table I-1); 3) A similar MCTU frame (up to six simultaneous targeting, using AtU6-26, AtU3b, and At7SL-2 as sgRNA promoters) for Arabidopsis created in 2016 (Zhang et al., 2016b).

All three systems worked as first generation of MCTU and succeeded to realize multiple targets knockout in plants.

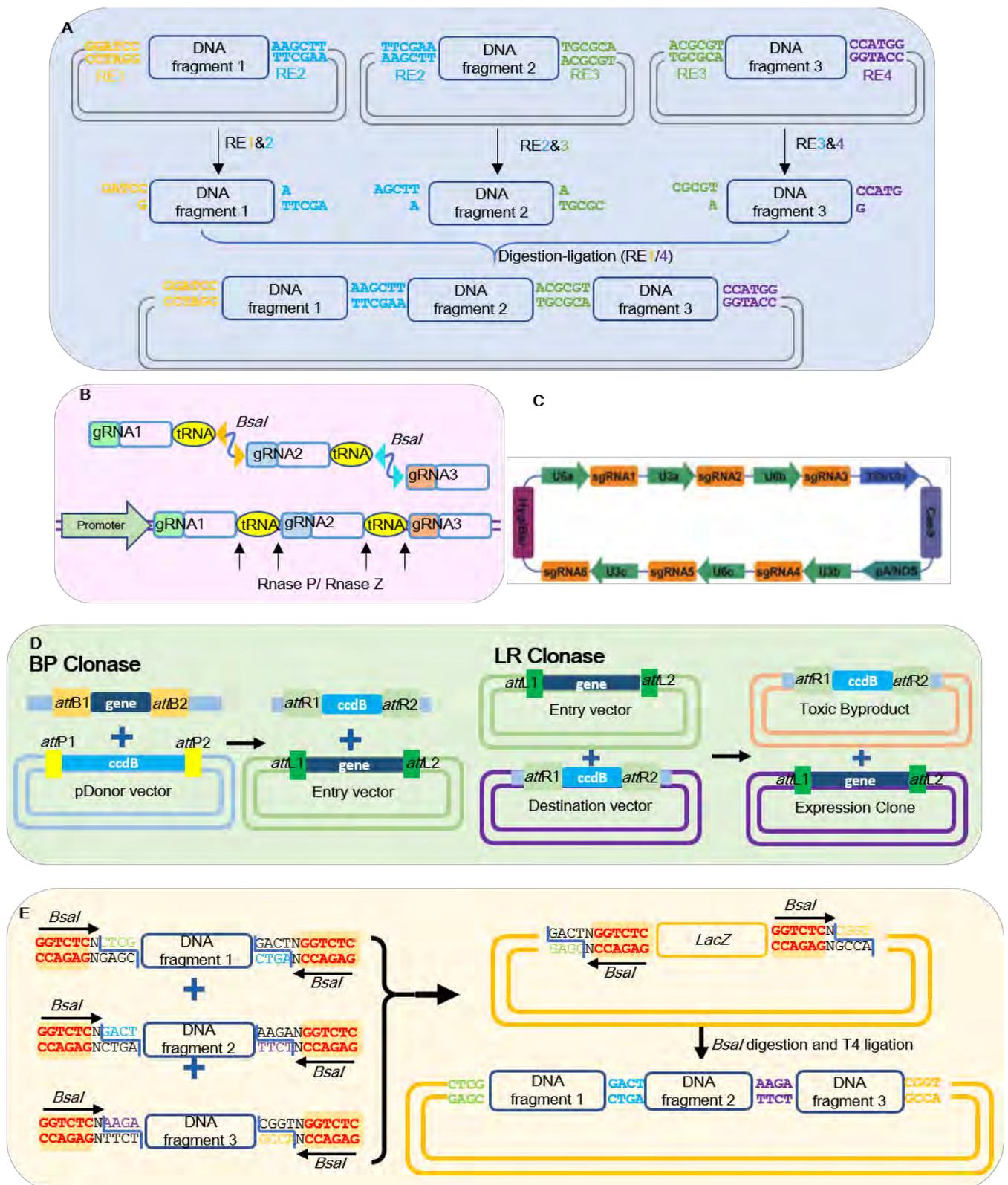


Figure I-17. Pipelines of common-used Cloning methods for CRISPR/Cas9 applications in plants. A, Conventional Cloning; B, PTG/Cas9 system indicates polycistronic tRNA-sgRNA driven by one promoter; C, extending and simplification for multiple assembly, the versatile and simplified MCTU system; D, Gateway Cloning; E, Golden Gate Cloning. (Zhang et al., 2016b).

3.4 Delivery of Cas9 and sgRNA expression cassettes

CRISPR/Cas9-sgRNA constitutive construct and its recombinant protein needs to be delivered and expressed in plant cells through plant transformation methods. *Agrobacterium*-mediated stable transformation is a routine method to produce transgenic plants, integrating exogenous DNA (carrying Cas9 and sgRNA expression cassettes) into plant genome and expressing it for editing. It is the most effective method to obtain stable transgenic mutants. Leaf discs transformation is generally used in most plants, while in citrus is epicotyl, and in grape is suspension cells (Table I-1). In *Eucalyptus*, *A. rhizogenes* mediated hairy roots are used as in tomato, soybean, chicory and *Medicago* (Ron et al., 2014; Sun et al., 2015; Cai et al., 2015; Jacobs et al., 2015; Bernard et al., 2019; Zhang et al., 2020), etc. In order to verify transgenic mutants carrying T-DNA insertions, either antibiotics or fluorescent reporter genes are used for screening. For instance, Green fluorescent protein (GFP) was introduced into a binary vector together with Cas9 and sgRNA in cacao and coffee (Fister et al., 2018; Breitler et al., 2018), and fused in frame with resistance genes (NPTII or blasticidin) serving as dual selectable marker in citrus, apple (Zhang et al., 2017; Nishitani et al., 2016). CRISPR/sgRNA expression binary vectors harboring β -glucuronidase (GUS) was first transformed into Wanjincheng orange epicotyl, and the identified GUS-positive shoots were grafted on Troyer citrange to recover (Peng et al., 2017; Wang et al., 2019). DsRed1-fluorescence was observed in transgenics *Parasponia andersonii* explants (van Zeijl et al., 2018) and *Eucalyptus* hairy roots (Dai et al., 2020) indicating a successful transfer of T-DNA in the corresponding plant genomes.

The versatile and efficient transient transformation methods by directly delivering foreign DNA, (including PEG protoplast transformation, agroinfiltration and particle bombardment), usually worked for early trials to test Cas9/sgRNA constructs efficiency. In cacao, agroinfiltration was used to introduce CRISPR/Cas9 plasmid in leaf tissue (GFP as selectable marker), showing efficacy as premier test (Fister et al., 2018). In apple, transient agroinfiltration was adopted to obtain T-DNA-free transgenic lines (CRISPR-PDS), with 0.4% mutation efficiency (3/747) presenting albino buds (Malnoy et al., 2016). In citrus, *Xanthomonas citri* subsp. *citri* (Xcc) pre-treatment followed by agroinfiltration enhanced efficiency of transient expression in leaves.

In order to avoid the transgene integration into organisms, DNA-free genome editing system was carried out in various plants (potato, wheat, apple, *petunia hybrida*, maize) using Cas9-sgRNA ribonucleoprotein (Cas9/sgRNA RNPs), which is degraded rapidly after cleaving targets without transcription and translation (Malnoy et al., 2016; Subburaj et al., 2016; Svitashv et al., 2016; Liang et al., 2018). CRISPR RNPs technic avoids integration of foreign DNA and produced less off-targets. In maize and wheat, CRISPR/Cas9 RNA reagents were delivered by particle bombardment into embryos. RNPs has also been used with Cpf1 in soybean and tobacco (Kim et al., 2017). In general, protoplast transformation was selected for direct delivery of CRISPR/Cas9 RNPs in trees (grape, apple, rubber tree), though it is limited in plants because of difficulty to acquire heritable mutations.

3.5 Method to identify targeted mutations

To identify genotype mutations induced by CRISPR/Cas9 system, several methods of verification and detection were developed based on different principles. Methods to verify the success of editing use reporter genes, PCR/restriction enzymes (PCR-RE), T7 endonucleases I (T7EI) assay, PAGE, High-resolution melting assay (HRM). The methods to

characterize editing events use Sanger sequencing (of either PCR amplicons or after subcloning of the latter) and high-throughput sequencing methods. These steps must be carried out before further analysis of CRISPR/Cas9-mediated transgenic individuals.

Non-sequencing methods to detect mutations

In order to simply and rapidly verify occurrence of editing events and screen mutations, non-sequencing methods were developed, mainly consisting of reporter genes, PCR/restriction enzymes (PCR-RE), PAGE, and High-resolution melting assay (HRM). Reporter genes used in trees aim to verify transgenic plants if carrying constitutive CRISPR/Cas9 plasmids. PCR/restriction enzymes (PCR-RE) method is based on the enzyme sites located in target sequence. Once the target mutations are induced by CRISPR/Cas9, the enzyme site is destroyed and cannot be amplified by PCR. The limitation of this strategy is that target sequences contain enzyme sites. MfeI restriction enzyme site was designed in citrus target regions (CsPDS) for further mutation detection (Jia and Wang, 2014) (Table I-1).

The mixed PCR amplicons within target regions from transgenics and controls generate unpaired nucleotides if mutations exist, which can be evidenced using SURVEYOR nuclease and T7 Endonuclease I assay. Though it has less limitation than PCR-RE, the sensitivity is also lower (Voytas, 2013). PCR amplicons were denatured and reannealed, using DNA of transgenic kiwifruit and control, for the subsequent T7 endonuclease I (T7E1) assay (Wang et al., 2018b) (Table I-1).

Polyacrylamide gel electrophoresis (PAGE) and high-resolution melting (HRM) methods for mutation verification have not been applied in trees yet, which rely on single-strand conformation polymorphism (SSCP) and different melting temperatures between control and mutants, respectively, but this strategy is not very sensitive to detect mutations.

Sequencing methods to detect editing events

CRISPR/Cas9 applied in diploid plants generates several types of mutations including monoallelic, biallelic (homozygous, heterozygous) and chimeric mutations (Figure I-15C). Sanger sequencing detects both simple and complicated mutations patterns. Two main strategies can be carried out: 1) sequencing multiple clones, which obtained by PCR products subcloning; 2) direct sequencing of PCR amplicons. The classical method in which the amplified PCR products containing target sites are cloned into T-vectors, and multiple clones are selected for sequencing, is precise and reliable for mutations detection, though it is time- and cost-consuming. Direct sequencing PCR amplicons works well for monoallelic and biallelic-homozygous mutations, however, when faced with complicated mutations (heterozygous and chimeric) in transgenic plants, the superimposed sequencing chromatograms lead to failed decode of various mutation types. In order to further analysis this bimodal pattern, web-based tools are explored, such as DSDecode (Degenerate sequence decoding) (<http://skl.scau.edu.cn/dsdecode/>) (Liu et al., 2015; Ma et al., 2015) and ICE (Inference of CRISPR Edits) Synthego (<https://ice.synthego.com/>) (Roginsky, 2018), both of which boost analysis single or batch of sequencing files.

High-throughput Next Generation sequencing (target deep sequencing) is suitable for identification of complex or rare mutations or editing events occurred in polyploid plants, and also used for screening off-targets in whole genome, though it is expensive and time-consuming (Fauser et al., 2014; Feng et al., 2014). Until now, target deep sequencing was applied in grape and rubber tree, together with CRISPR RNP. And the high-throughput tracking of mutations was analyzed by

online tools or platforms including AGEseq, Cas-analyzer, CRISPR-GA, CIRSPRESSO, Hi-TOM (Xue and Tsai, 2015; Park et al., 2017; Guell et al., 2014; Pinello et al., 2016; Liu et al., 2019), etc.

4. Factors affecting the editing efficiency and off-targets in trees

4.1 Delivery methods and Cas9 optimization

CRISPR/Cas9 system has been widely used in plants and in trees with various editing efficiencies. Editing efficiency is affected by a series of factors such as- 1) proper transformation methods selection, 2) optimized Cas9 and its promoter, 3) sgRNA positions, length, specific sequences and its promoter. Transient transformation (protoplast or agroinfiltration) led to lower editing efficiency in apple, grape and kiwifruit, compared with *Agrobacterium*-mediated stable transformation, although there still exists difficulties of plant regeneration from various tissue cultures. What Farboud and coll. found was that the delivery methods of Cas9 also influence the editing efficiency, CRISPR RNP for instance, boosts more Cas9 editing than constitutive DNA expression of Cas9 and produces less off-targets (Farboud et al., 2019). Though Cas9 function cannot be affected by DNA methylation, the tolerance of Cas9 to mismatches (up to 5 nucleotides) within gRNA:DNA (Fu et al., 2013) or PAM (Hsu et al., 2013) leads to off-targets, which may disrupt mutation analysis and related phenotypes variations. These mismatches showed only a few unpaired nucleotides with original sequence even close to PAM. Hence, off-targets detection is needed, which often requires whole genome sequencing which is time- and cost-consuming for minimization and elimination, or web-based tools predicted analysis and evaluation. According to off-targets detecting results in some plant species, its occurrence rate is low, however, restraint of off-targets is also needed because of inheritance or uncontrollable negative effects in traits improvement. Cas9 from *Streptococcus pyogenes* has been widely used because its associated PAM sequence (NGG) is highly abundant throughout the genome, which supports more options to select target sequences. Expression level of spCas9 is determined by driving promoter, terminator and codon optimization. In general, RNA polymerase II type of promoters 35S or Ubi are widely used, with Poly A or NOS terminator. In grape, spCas9 expression level affected less editing efficiency than the GC content in sgRNA (Ren et al., 2019).

Cas9 protein contains conserved RuvC domain and HNH domain, of each is responsible for one strand cleavage at target regions (3bp close to PAM sequence) to produce blunt-ended DSB. Mutation of either Cas9 domain result in single-strand break rather than DSB, and it is called Cas9 nickases (nCas9, D10A or H840A for spCas9). Together with paired-gRNAs in PAM-out orientation, nCas9-gRNAs complex can make robust staggered cuts to create a double overhang-nicks for more specificity (Ran et al., 2013).

4.2 sgRNA parameters affecting editing efficiency

sgRNA components

Initial sequence 'A' or 'G' is required for sgRNA promoter (U3 or U6) and its homologous types from *Arabidopsis* or other organisms have been widely applied in plants and affected editing efficiency. The composition of sgRNAs determines its activity. Depending on comparative research for 1841 sgRNAs by Doench (Doench et al., 2014) and Farboud (Farboud

and Meyer, 2015), achievable high editing efficiency needs 'A', 'C', 'G' at sgRNA positions 3, 16 and 20 from 5' terminal, or 'GG' located at sgRNA distal. Higher GC content (ranging from 50% to 70%) in sgRNA sequence leads to higher editing efficiency, as feature in Gramineae genomes harboring high GC content at 5' terminal of a gene. Though PAM sequence 'NGG' is canonical for spCas9, the comparison results of different PAMs by Zhang (Zhang et al., 2015) showed ranks of editing efficiencies is NGG>NGA>NAG. The specificity of sgRNA lies on 'seed sequence' which is located 8-12 nucleotides near PAM. Restrictions and no mismatch of seed sequences give less possible off-targets. Some authors used truncated sgRNA (18bp) to decrease off-target effects like in mammal cells (Fu et al., 2014). Truncated sgRNA (17-18nt) maintains its specificity and activity along with less off-targets. It was applied in apple to successfully induce mutations (Nishitani et al., 2016).

sgRNA positions and secondary structure

T-DNA insertion positions, depending on PAM sequence recognition and targeting sites locations in genome, also impacts the targeting efficiency, performing predicted mutations occurred close to 5' terminal of open reading frame (ORF) which results in heavy alternation of gene associated protein. Secondary structure of sgRNA is crucial for editing efficiency and there are three types of stem-loop structures related to sgRNA activity and effectiveness, one stem loop (GAAA) participates in gRNA processing, and the other two loops (GAAA and AGU) act for stable complex formation (Liang et al., 2016). By high-throughput chemical screening, small molecules such as a β 3-adrenergic receptor or Brefeldin A (lactone antiviral) improve 2-3 folds editing efficiency of HDR template insertions and mutations (Yu et al., 2015).

sgRNA specificity and design

Gene duplication occurs randomly in cells and is essential for plants evolution, especially segmental and tandem duplications. Duplicated or highly homologous genes induce minor variability in plant genome, which gives risks of off-targets cleavage by CRISPR/Cas9. Functional redundancy of homologous genes is the obstacle of CRISPR/Cas9-mediated knockout investigation. Single sequence polymorphisms (SNP) is another concern to be addressed in outcrossing species especially trees, which are highly heterozygous and harbor high frequency of SNPs in whole genome. Potential SNPs are identified using whole-genome resequencing and RNA-seq data. SNPs existing in coding or non-coding regions result in mismatches of sgRNA reducing editing efficiency. Because of code degeneracy, some coding SNPs (cSNP) can lead to no alteration of amino acids.

Using web-based tools to design specific target sequence for CRISPR/Cas9, the potential off-targets can be limited and avoided, and the editing efficiency will be improved, considering the parameters mentioned above.

CRISPR-P (<http://crispr.hzau.edu.cn/CRISPR2/>) provides three kinds of methods to design sgRNA, based on 33 plant species genome such as Arabidopsis, rice, maize (Lei et al., 2014). Once upload the Locus tag, gene Positions or Sequences, 'G(N)20GG' or 'A(N)20GG' sgRNAs are obtained after submission. The evaluation and risk of off-targets are analyzed with service 'BLASTn' in whole genome. Using CRISPR-P, efficient sgRNAs are designed and worked for CRISPR/Cas9 induced mutations in citrus, grape, poplar and cassava.

ZiFiT Targeter (<http://zifit.partners.org/ZiFiT/Introduction.aspx>) identifies potential binding sites for zinc finger proteins

(ZFPs), and its Version 4.2 aims to design sgRNA and evaluate off-targets for both wild type (WT) -Cas9 and Cas9 nickases (design paired-gRNA) (Sander et al., 2010). In trees, ZiFiT has been utilized in poplar for WT-Cas9, targeting PDS and MYBs.

CRISPOR (<http://crispore.tefor.net/>) combines more species genomes and predictive scoring models, presenting a list sgRNAs by ranking potential off-targets and predicted on-target efficiency after submitting single genomic sequence (< 2000 bp) (Haeussler et al., 2016; Concordet and Haeussler, 2018). The latest version tries to provide a complete and detailed experimental workflow. Until now, CRISPOR based sgRNA design was carried out in apple and Eucalyptus, with variability of editing efficiencies.

5. Examples of CRISPR/Cas9 applications in trees

Woody plants have economic and ecological benefits especially fruit trees and fast-growing trees. They are characterized by long reproductive cycles. Whole genome sequences were recently released in many tree species, which make it possible to implement and apply CRISPR/Cas9 to: 1) create novel germplasm resources (genetically improved); 2) establish mutants' libraries; and 3) perform gene functional characterization.

The first successful CRISPR/Cas9 -mediated genome editing was carried out in populus (*Populus tomentosa* Carr) with relative high efficiency, accomplishing lignin biosynthesis-related targets editing (Zhou et al., 2015). Until now, CRISPR/Cas9 knockout has been widely used not only in poplar, but also in citrus, apple, grape, pear, kiwifruit, cassava, cacao, rubber tree, *Parasponia andersonii* and Eucalyptus, targeting series of genes besides phytoene desaturase (PDS) (up to now, CRISPR-induced pds mutants has been obtained in citrus, poplar, apple, grape, cassava, kiwifruit and coffee, with various mutated efficiencies from 3.2% to 100%), such as reduce diseases susceptibility related (MLO, WRKY, LOB, DIPM, NPR), flowering time-related (FT, TFL), hormonal pathways related (NSP, ARF, HK4, EIN, IAA), and transcription factors (MYBs, NST/SND, TCP), though existing large variability of editing efficiency (CRISPR/Cas9 application summary in trees and related KO phenotypes were shown in Table I-S2).

In some cases, constructions for CRISPR/Cas9 systems have been described for use in further research. For instance, a CRISPR/Cas9 knockout binary vector targeting the cellulose synthase (CES-A) from Chinese tulip tree was assembled (Xu et al., 2017). Two sgRNAs designed for tea caffeine synthase were introduced into a CRISPR/Cas9-mediated binary vector for further research. Cytokinin oxidase/dehydrogenase (CTX) gene family (CTXA, CTXB, CTXC) and miR156 family (miR156A, miR156B, miR156C) of *Eucalyptus grandis* were assembled into the pHDE-Cas9 system (mcherry as selectable marker), laying foundation for point editing even large deletions (Li et al., 2018).

Providing a DNA fragment as a repair template, CRISPR/Cas9 knock-in technology can realize more precise editing by HR repair (insert donor sequences matched into target sites). CRISPR/Cas9 knock-in technic have been carried out in a few plant species such as *Arabidopsis*, maize, rice and soybean. However, CRISPR knock-in presented low efficiencies of both HR repair and donor template transfection (Puchta and Fauser, 2014), which limited its application in plants even tree species.

5.1 CRISPR-Cas9 to modify Flowering-time

The breeding cycle of fruit trees is under the control of vegetative growth and reproductive development, thus manipulating precisely the trait of flowering time (delayed- or early- flowering) is interesting both for yield and fruit quality. FLOWERING LOCUS T (FT) is a florigen signal for flowering initiation, and TERMINAL FLOWER1 (TFL1) is a floral repressor. CRISPR-induced *tfl* mutants of apple and pear presented early flowering (Charrier et al., 2019). CRISPR RNP was used together with two FT genes (HbFT1 and HbFT2) and introduced in rubber tree protoplasts (Bi et al., 2019; Fan et al., 2020).

In poplar, essential floral identity genes LFY (LEAFY), AGAMOUS (AGs) and the floral integrator SOC1 (suppressor of overexpression of CONSTANS1) and their paralogs which functions were previously known, were selected as target genes to generate CRISPR/Cas9-mediated knockout mutants, with 56.4 to 81.3% targeting efficiencies (Elorriaga et al., 2018; Bruegmann et al., 2019).

5.2 CRISPR-Cas9 to enhance resistance to biotic and abiotic stresses

Recently, CRISPR/Cas was used to create mutants with enhanced tolerance to biotic and abiotic stresses. LATERAL ORGAN BOUNDARIES 1 (LOB1) is the susceptibility gene for citrus canker caused by *Xanthomonas citri* subsp. *citri* (Xcc). CsLOB1 citrus knockout mutants were generated by CRISPR/Cas9, which led to modifications of the coding region resulting in a range of resistance degrees (Jia et al., 2017b). Cas9/sgRNA targeting to CsLOB1 promoter region disrupting cis-element binding sites affected gene expression, resulting in alleviated canker symptoms in citrus.

In order to increase resistance to powdery mildew, Malnoy and collaborators (Malnoy et al., 2016) targeted MLO-7 (MLO, Mildew Locus O) susceptibility gene to powdery mildew] in grape creating knockout mutants (protoplasts) by CRISPR RNP. Similarly, they targeted DIPM-1, DIPM-2, DIPM-4 (DIPM, DspE-interacting proteins of *Malus*) known to interact with *Erwinia amylovora* carrying disease-specific (*dsp*) gene *dspA/E* creating knockout Apple protoplast mutants with increased resistance to fire blight disease.

WRKY proteins belongs to a large transcription family, involved in response to abiotic and biotic stresses as well as in many plant developmental physiological processes (Rushton et al., 2010; Jiang et al., 2017). In grape, CRISPR/Cas9-mediated knockout mutants of VvWRKY52 presented increased defense against *Botrytis cinerea* (Wang et al., 2017). CRISPR/Csa9 induced loss-of-function of CsWRKY22 in Citrus mutants showed severe decrease susceptibility to Citrus canker (Wang et al., 2019). In *Theobroma cacao*, a suppressor of the defense response, Non-Expressor of Pathogenesis-Related 3 (TcNPR3) was targeted efficiently by CRISPR/Cas9 using transient transformation.

The optimization of plant regeneration methods and the availability of target genes provide settlements for tree improvement and increased resistance (Fister et al., 2018).

5.3 CRISPR-Cas9 applications for gene functions in trees

CRISPR/Cas9 also serves as a robust tool for investigating functions of traits-related genes and/or transcription factors (TF) in trees. The L-idonate dehydrogenase gene (*IdnDH*) in grape regulates tartaric acid (TA) biosynthesis, which is important for the organoleptic properties of grape berries and essential to kill bacteria (DeBolt et al., 2006). *IdnDH* mutants

induced via a CRISPR plasmid harboring one sgRNA (located in the first exon) and spCas9 showed expected indel mutations (Ren et al., 2016).

Parasponia andersonii a fast-growing tropical tree, is one of the important non-leguminous plants can perform nitrogen fixation through symbiosis with *Rhizobium*. Components of plant hormonal pathways (strigolactone biosynthesis, cytokinin, and ethylene) related DWARF27, HK4, and EIN2, as well as NSP associated with nodule organogenesis, were targeted by CRISPR/Cas9, with KO phenotypes such as inhibition of root nodules formation and reduced procambium activity.

Poplar was the first stable transformed woody species edited by CRISPR, mainly focusing on lignin biosynthesis, wood formation and secondary cell wall (SCW) traits. Two 4-coumarate: CoA ligases (4CL) 4CL1 and 4CL2 were selected as target genes accounting to SNPs frequent occurrence in heterozygous poplar genome to implement CRISPR/cas9 in poplar. Knockout-induced 4cl and 4cl mutants indicated the primary roles of two genes in lignin biosynthesis and in flavonoid biosynthesis, respectively (Zhou et al., 2015).

DWF4 encodes a cytochrome P450 protein which catalyzes the rate-limiting step in brassinosteroid biosynthesis in *Arabidopsis* (Si et al., 2016). CRISPR/Cas9 generated PtoDWF4-KO lines showed reduced biomass and down regulation of SCW-related biosynthesis genes, suggesting that DWF4 could act as a positive regulator of wood development in poplar (Shen et al., 2018).

CRISPR-induced knockout Pto-myb156 and Ptr-myb57 mutants showed increased proanthocyanidin (PA), indicating that PtoMYB156 and PtrMYB57 play negative roles in the regulation of anthocyanin biosynthesis (Yang et al., 2017; Wan et al., 2017).

The MYB transcription factor (LTF1) binds to the 4CL promoter. The phosphorylation status of LTF1 affects the protein function, interactions and stability. CRISPR/Cas9-generated knockout LTF1 mutants identified increased lignin deposition in developing xylem of *Populus* (Gui et al., 2019).

In the CRISPR knockout mutants of the four NST/SND using, SCW formation was inhibited in xylem ray parenchyma cells and almost irregular fibres, indicating that four NST/SND orthologs (belonging to the NAC transcription factors family) are master switches for SCW formation in wood fibers in *Populus* (Takata et al., 2019).

BRC belongs to the TB1 CYCLOIDEA PCF (TCP) family of transcription factors, which controls a series of plant developmental processes. Phenotypes of CRISPR/Cas9-mediated knockout populus mutants of PcBRC2 showed two ectopic leaves at each node, and PcBRC1 mutants retained even enhanced bud outgrowth (Muhr et al., 2018).

Objectives of PhD thesis

The overall objective of my PhD was to Implementing the **CRISPR/Cas9** technology in *Eucalyptus hairy roots* and Functional characterization of **auxin-dependent transcription factors** involved in **wood formation**. Wood is the most abundant biomass on earth, fulfils key roles in trees, and is also a raw material for multiple end-uses by mankind. Wood utilization is impacted by structure and composition of secondary cell wall (SCW), which is characteristic of wood cells. SCW deposition requires coordinated expression of biosynthetic genes, which is under the control of transcriptional network (NACs and MYBs master switches). Wood formation is controlled by exogenous signals, the environmental stimuli, and endogenous factors, plant hormones (auxin, gibberellin, cytokinin, etc).

Based on the previous researches in *Arabidopsis*, auxin was identified as a key regulator in wood formation through auxin response, which is achieved by auxin signaling (including two key mediators: Aux/IAA and ARF). Here we investigate the regulatory role of auxin signaling in wood formation, using *Eucalyptus* as our experimental system. The genome of *Eucalyptus grandis* was released in 2014. Our team identified in *Eucalyptus* some candidates potentially involved in regulation of wood formation such as transcriptional factors MYB, NAC, Aux/IAA and ARF. Twenty-four Aux/IAA members were identified in *E. grandis* genome and the relative expression profiling in 13 different tissues and organs pointed out that 11 Aux/IAA members are highly expressed in wood forming tissues. Among them, *EgrIAA4*, *EgrIAA9A*, *EgrIAA13* and *EgrIAA20* were chosen to be transformed into heterologous system, *Arabidopsis*, and overexpressed *EgrIAA9A* and *EgrIAA20* in *Arabidopsis* transgenic lines showed modified wood cells phenotypes, such as reduced radial growth and greatly inhibited fiber cells development. Besides, ARF5/MONOPTEROS (MP) was identified as a key regulator for vascular tissue specifications in embryo, leaf and root, while it was unclear in ‘tree’ cambium specification, which need further investigations to know which Aux/IAA and ARF members (especially IAA9, IAA20 and ARF5) are involved. Therefore, we selected *EgrIAA9A*, *EgrIAA20* and *EgrARF5* as our candidates to validate their functions on wood formation in homologous system (*Eucalyptus*).

The establishment of of *Eucalyptus* hairy roots in 2016 by our team provided a stable transformation system which is proved to be suitable for functional characterizations of SCW genes, using *EgrCCR1* as a proof-of-concept. We obtained the strong phenotypes of down regulated *EgrCCR1* in *Eucalyptus* hairy roots, such as decreased lignin content and collapsed vessels in CCR1 antisense lines.

Above all, my first objective is **Implementing the CRISPR/Cas9 in *Eucalyptus* Hairy Roots** (choose *EgrCCR1* and *EgrIAA9A* as proofs of concept genes) to generate loss-of-function mutants. My second objective is **Functional Characterization of *EgrIAA9A*, *EgrIAA20* and *EgrARF5* in *Eucalyptus* hairy roots** using reverse genetics approaches by generating loss-of-function (CRISPR/Cas9) and gain-of-function (35S promoter, Golden Gate Cloning). My third objective is **seeking for the interacting partners of our candidates in developing secondary xylem of *Eucalyptus*** (using Yeast two Hybrid xylem library screening).

Chapter II

Implementing the CRISPR/Cas9 Technology in Eucalyptus Hairy Roots Using Wood-Related Genes

In Chapter II, we implemented CRISPR/Cas9 genome editing to obtain loss-of-function genes in Eucalyptus hairy roots, using *EgrCCR1* and *EgrIAA9A* as proofs of concept (Dai et al., 2020).

Specific difficulties were due to frequent SNPs (single nucleotide polymorphism) due to the heterozygous genome of *Eucalyptus grandis*. We obtained various types of editions for *ccr1* and *iaa9* knockout/knockdown mutants,.

We further characterized cell wall related phenotypes by FT-IR chemical composition profile screening and histochemical analysis for *CCR1*_edited lines.



Article

Implementing the CRISPR/Cas9 Technology in *Eucalyptus* Hairy Roots Using Wood-Related Genes

Ying Dai ¹, Guojian Hu ², Annabelle Dupas ¹, Luciano Medina ¹, Nils Blandels ¹, H el ene San Clemente ¹, Nathalie Ladouce ¹, Myriam Badawi ^{3,4}, Guillermina Hernandez-Raquet ³, Fabien Mounet ¹, Jacqueline Grima-Pettenati ¹ and Hua Cassan-Wang ^{1,*}

¹ Laboratoire de Recherche en Sciences V eg etales, Universit e de Toulouse III, CNRS, UPS, UMR 5546, 24 Chemin de Borde Rouge, 31320 Castanet-Tolosan, France; ying.dai@lrsv.ups-tlse.fr (Y.D.); annabelle.dupas@lrsv.ups-tlse.fr (A.D.); luciano.medina@lrsv.ups-tlse.fr (L.M.);

nils.blandel@lrsv.ups-tlse.fr (N.B.); san-clemente@lrsv.ups-tlse.fr (H.S.C.); ladouce@lrsv.ups-tlse.fr (N.L.); mounet@lrsv.ups-tlse.fr (F.M.); grima@lrsv.ups-tlse.fr (J.G.-P.)

² UMR 990, G enomique et Biotechnologie des Fruits, Universit e de Toulouse, INP-ENSA Toulouse, Avenue de l'Agrobiopole, 31326 Castanet-Tolosan, France; guojian.hu@etu.ensat.fr

³ TBI, Universit e de Toulouse, CNRS, INRAE, INSA, 31400 Toulouse, France; Myriam.Badawi@univ-lemans.fr (M.B.); hernandg@insa-toulouse.fr (G.H.-R.)

⁴ Laboratoire Mer Mol ecules Sant e, MMS EA2160 Le Mans Universit e, 72085 Le Mans, France

* Correspondence: wang@lrsv.ups-tlse.fr

Received: 14 April 2020; Accepted: 8 May 2020; Published: 12 May 2020

Abstract: Eucalypts are the most planted hardwoods worldwide. The availability of the *Eucalyptus grandis* genome highlighted many genes awaiting functional characterization, lagging behind because of the lack of efficient genetic transformation protocols. In order to efficiently generate knock-out mutants to study the function of eucalypts genes, we implemented the powerful CRISPR/Cas9 gene editing technology with the hairy roots transformation system. As proofs-of-concept, we targeted two wood-related genes: *Cinnamoyl-CoA Reductase1* (*CCR1*), a key lignin biosynthetic gene and *IAA9A* an auxin dependent transcription factor of *Aux/IAA* family. Almost all transgenic hairy roots were edited but the allele-editing rates and spectra varied greatly depending on the gene targeted. Most edition events generated truncated proteins, the prevalent edition types were small deletions but large deletions were also quite frequent. By using a combination of FT-IR spectroscopy and multivariate analysis (partial least square analysis (PLS-DA)), we showed that the *CCR1*-edited lines, which were clearly separated from the controls. The most discriminant wave-numbers were attributed to lignin. Histochemical analyses further confirmed the decreased lignification and the presence of collapsed vessels in *CCR1*-edited lines, which are characteristics of *CCR1* deficiency. Although the efficiency of editing could be improved, the method described here is already a powerful tool to functionally characterize eucalypts genes for both basic research and industry purposes.

Keywords: CRISPR/Cas9; genome editing; cinnamoyl-CoA reductase; Aux/IAA; wood; secondary cell walls; *Eucalyptus*; lignin; hairy roots; FT-IR spectroscopy

1. Introduction

Eucalypts are among the leading sources of woody biomass worldwide [1,2]. Due to their rapid growth rate, broad adaptability to diverse edaphoclimatic conditions and their multipurpose wood properties, they are the most planted trees worldwide. Eucalypts wood is currently used in the emerging areas of biofuels and biomaterials in addition to more traditional uses such as pulp and paper production, thereby extending the already considerable economic importance of these trees. Wood is mainly composed of secondary cell walls (SCWs), which contain three major polymers:

cellulose, hemicelluloses and lignins. The proportions of each of these polymers and the interactions between them underlie the composition and the structure of the SCWs, which are major determinants of industrial processing efficiencies [3].

The availability of the *Eucalyptus grandis* genome [4] has allowed genome-wide characterization of many gene families, notably those involved in the lignin biosynthetic pathway [5] as well as transcription factor families containing members known to regulate SCW formation such as the R2R3-MYB [6], NAC [7], ARF [8] and Aux/IAA [9] among others. These studies have underscored many new candidates potentially regulating wood formation that need to be functionally characterized. The bottleneck to functionally characterize genes in *Eucalyptus* has always been stable transformation. Although stable transformation protocols have been established for several *Eucalyptus* species [10–13], they are very time-consuming and present low efficiencies explaining why only very few functional studies have been performed in transgenic eucalypts (reviewed in [11]). To overcome these limitations, we have set up an alternative stable transformation system for *E. grandis* using *Agrobacterium rhizogenes* that allows the development of composite plants with wild-type shoots and transgenic roots easily detectable by fluorescent markers [14]. We have further shown that this system is suitable to elucidate the function of genes involved in xylem or SCW formation particularly important for woody species. As a proof-of-concept, we used the down-regulation of *Cinnamoyl CoA reductase1* (*EgrCCR1*), the penultimate step of the lignin branch pathway through antisense strategy [14].

During the last three decades, antisense RNA, virus-induced gene silencing and RNA interference were the most used methods for gene silencing in plants and provided very useful insights in the function of many genes especially in plants for which no mutant collection was available. However, silencing was not as effective as in mutants since gene function was interrupted indirectly by repressing the corresponding mRNA, often leading to partial repressive effect and in some cases to unpredicted effects. The Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-associated protein 9 (CRISPR/Cas9) based genome editing that can induce efficiently targeted mutations, revolutionized reverse genetics in all systems and was considered as the breakthrough of year 2015 [15,16]. The CRISPR/Cas9 method is based on the ability of Cas9, a RNA-guided endonuclease from *Streptococcus pyogenes*, to cut a DNA double strand at a specific region [17]. A complementary single-guide RNA (sgRNA) forming a complex with Cas9 will specifically recognize a target DNA region by base-pairing. The Cas9 will then cut directly upstream of a DNA motif of a 2–6 base pair long, called Protospacer Adjacent Motif (PAM). The sequence of the PAM depends from the origin of the Cas9 endonuclease [18].

Currently, CRISPR/Cas9 is the system of choice to targeted mutagenesis in a growing number of plants including woody plants [19]. The possibility to obtain null mutations in the T0 generation is especially important for trees that have very long reproductive cycles [20,21] and like poplar or eucalypts are propagated vegetatively. Another characteristic of outcrossing trees is their high degree of genome heterozygosity and the presence of sequence polymorphisms at the target sites can render CRISPR editing ineffective [19]. The greatest progresses have been made with poplar, the first tree to be genome-edited by CRISPR with high efficiency [20,22] and for which allele-sensitive bioinformatic resources facilitating genome editing in heterozygous species have been developed [21,23]. Due to the importance of wood properties for industrial applications, most of CRISPR gene editing studies in poplar have targeted SCW composition and phenylpropanoid metabolism including lignin [24–28]. A large study encompassing more than 500 transgenic events has also reported successful mutations of essential flowering genes to prevent bisexual fertility [29]. The goal was to produce infertile trees thereby solving potential seed or pollen dispersal, which are of concern for trees that are vegetatively propagated for industrial plantations.

Eucalypts are diploid species ($2n = 22$), each homologous chromosome contains one allele, so each gene has two alleles. Theoretically for each sgRNA, the CRISPR/Cas9 system can introduce up to two types of mutations in the targeted gene except the chimer: (i) if the two alleles are mutated, it is a biallelic mutation, which can be either homozygote mutation (if the two alleles share exactly the same mutation), or biallelic (if the mutations are different in the two alleles) and (ii) if only one of the

alleles is mutated, it is a monoallelic mutation, which does not theoretically produce full knock-out mutations because a wild-type allele is remaining, but it may produce knock-down mutations. However, in the case of chimera, three or more mutations are detected simultaneously in the same plant. There are two types of chimera; (i) the ones with two or more mutated alleles that also present a wild type allele (may introduce knock-down but never complete knock-out) and (ii) the chimera with all alleles mutated that may trigger complete knock-out.

Here, we tested the potential of the CRISPR/Cas9 system to induce gene knock-out in *E. grandis* hairy roots. We first targeted the *EgrCCR1* (*CCR1*) gene because the phenotypes induced by the down regulation of this gene in *E. grandis* hairy roots have been well characterized [14]. We also selected the candidate *Aux/IAA* gene, *EgrIAA9A* (*IAA9A*), known to be preferentially expressed in vascular cambium and developing secondary xylem [9]. For CRISPR/Cas9 editing, we used two single guide RNAs (*sgRNAs*) under the control of *Arabidopsis* *U6* promoter and *Cas9* under the control of the Cauliflower Mosaic Virus *CaMV* 35S promoter in a single vector [30]. We report here that the CRISPR/Cas9 was efficient in generating mutations in both *CCR1* and *IAA9A* in *E. grandis* hairy roots but with very different editing efficiency rates. The phenotyping of *CCR1*-edited lines by FTIR spectroscopy and histochemical analyses confirmed the decreased lignin phenotypes expected in response to *CCR*-downregulation.

2. Results

To implement the CRISPR/Cas9 technology in the *Eucalyptus* hairy roots transformation system, we firstly generated the constructs comprising *Cas9* protein expression cassette and two single guide RNAs targets for the same gene in a single vector as described in the Materials and Methods section. We then transferred the constructs into *E. grandis* hairy roots using *Agrobacterium rhizogenes*. The transgenic roots were selected by the selection marker of DsRed fluorescence. We then extracted the DNA of transgenic roots and sequenced the corresponding region to reveal the gene editions in the target genes in order to determine whether or not generated the expected knock-out mutants.

2.1. Genotyping Revealed High Knock-Down Rate in *CCR1* but High Knock-out Rate in *IAA9A*

To reveal the CRISPR/Cas9 system genome editing and characterize the mutation types, we extracted genomic DNA from *E. grandis* transgenic roots. We then amplified by PCR the target gene regions including the two *sgRNA* sequences. The PCR amplicons were directly sequenced and analyzed by the web-based tool “Degenerate Sequence Decoding (DSDecode, <http://dsdecode.scgene.com>). For *CCR1*-lines the majority of the direct sequencing of PCR amplicons could not be read by the DSDecode program. We then subcloned PCR amplicons into pGEM-T vectors and several subclones were sequenced by Sanger sequencing.

2.1.1. High Knock-Down Rate for *CCR1* Editing

The rates of edited plants (hairy roots) were calculated on the basis of the PCR subcloning sequencing results. First, we verified that mutations in either *CCR1* or *IAA9A* genes were totally absent in the control plants transformed with an empty vector harboring the *Cas9* cassette without guide RNA sequences (Supplementary Table S1 and Supplementary Table S4). Then, we separated the CRISPR-*CCR1* and *IAA9A* hairy roots into two groups: (i) putative knock-out if all alleles were altered and no WT allele was present and (ii) putative knock-down if a WT allele sequence was still present (Table 1). For instance, a chimera plant present three and more alleles simultaneously with all alleles edited (Altered allele 1/Altered allele 2/Altered allele3, A1/A2/A3, etc.) was considered as a putative knock-out whereas a chimera plant with two or more mutated alleles and one WT allele simultaneous in the same plant (WT/A1/A2, etc.) was considered as a putative knock-down.

For *CCR1*, 100% of the 24 independent transgenic hairy roots (obtained from three independent transformation batches) were edited. All exhibited mutations in at least one allele (Table 1). Ten out of these 24 (41.2%) exhibited monoallelic mutation with a mutated allele and a WT allele (A1/WT). Thirteen out of 24 (54.2%) were chimeric with a WT allele and two or more mutated alleles

(WT/A1/A2, etc.; Table 1). Neither homozygote (A1/A1) nor biallelic mutation (A1/A2) was detected. Only one plant was determined as a chimera having editing in all tested alleles (no WT allele) comprising of seven different edited alleles (Table 1, *CCR1_22* in Supplementary Table S1). One out of the seven editions was a 1-bp substitution generating no change in the *CCR1* amino acid sequence and we also could not totally rule out the possibility that this 1-bp substitution was introduced during the PCR step. Whatever the case, this should not lead to a complete *CCR1* knock-out but most likely to a knock-down.

Table 1. Numbers of edited plants and type of mutations according to target genes.

Gene	Edited Plants/Total Transgenic Plants	Plants with All Alleles Altered (no WT)			Plants with one WT Allele		WT/WT
		Homoz. (A1/A1)	Biallelic (A1/A2)	Chimera (A1/A2/A3...)	Monoallelic (WT/A1)	Chimera (WT/A1/A2...)	
<i>CCR1</i>	24/24 (100%)	0	0	1 (4.2%)	10 (41.2%)	13 (54.2%)	0
<i>IAA9A</i>	12/13 (92.3%)	0	7 (53.8%)	4 (30.8%)	1 (7.7%)	0	1 (7.7%)

A1, edited allele 1; A2, edited allele 2, WT, wild-type allele, Homoz. Homozygote mutation. Different numbers in the alleles stand for distinct alleles. The number and percentage of the most prevalent genotypes for each CRISPR/Cas9 edited gene is in bold, highlighted in green for *CCR1*-transformants and in yellow for *IAA9A*-transformants.

For *CCR1*, we obtained an allele editing rate of 32.0%, i.e., only 89 among the 278 PCR subclones sequenced exhibited mutated alleles (Table 2). Forty six subclones (51.7%) had mutations in the first sgRNA position, 65 subclones (73.0%) in the second sgRNA position and 22 subclones (24.7%) contained mutations at both sgRNA1 and sgRNA2 positions (Table 2). Noteworthy 19 subclones (21.3%) showed large deletions between sgRNA1 and sgRNA2 as expected when using two guide RNAs separated from approximately 100 bp [30].

Table 2. Editing frequency and position (identified by PCR subcloning and subsequent sequencing).

Gene	Total Sequenced Subclones	Edited Subclones	Editions in sgRNA1	Editions in sgRNA2	Editions in sgRNA1&2	Large Deletions
<i>CCR1</i>	278	89 (32.0%)	46 (51.7%)	65 (73.0%)	22 (24.7%)	19 (21.3%)
<i>IAA9A</i>	95	88 (92.6%)	84 (95.5%)	79 (89.9%)	75 (85.2%)	27 (30.7%)

The detailed mutations for each edited allele can be found in Supplementary Table S1. The two most prevalent edition types were 82 bp deletion (14 alleles, 16%) and 1 bp deletion (11 alleles, 12%), both generating a reading frame shift and premature stop codon as illustrated in Figure 1. Among the 89 edited alleles, 87 generated amino acid changes, potentially modifying the activity of the corresponding enzyme (Supplementary Table S1). Noteworthy, the majority (69.7%) of the *CCR1*-targeted edition introduced significant modifications among which 67.4% were reading frame shifts (Table 3) and 24.7% were 15 bp or more insertion/deletions. The very highly complex pattern of *CCR1* edition explained why the DSDcode online tool failed to run, thereby confirming that this analytic tool is not suitable for chimeric or multiple mutations.

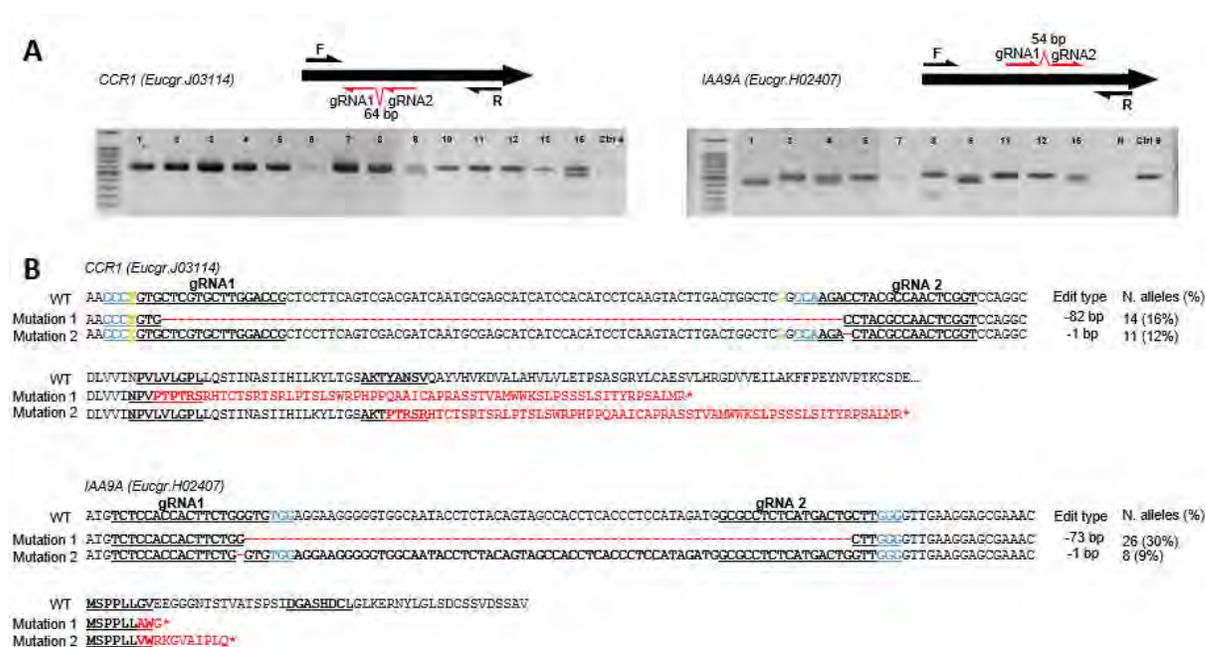


Figure 1. Genotyping of *CCR1* and *IAA9A* transformants and their corresponding prevalent edition types. (A) Electrophoresis gels showing PCR amplicons for *CCR1* (left panel) and *IAA9A* (right panel) transformants; the positions of the guide RNAs and of the primers used for PCR are indicated on the schematized ORF (open reading frame) sequences of the genes. The symbols above each lane indicate the transformant lines. In the panel of *CCR1* transformants: 1 for *CCR1_1*, 2 for *CCR1_2* and so on; Ctrl-4, control line 4. In the panel of *IAA9A* transformants: 1 for *IAA9A_1*, 2 for *IAA9A_2* and so on; N, PCR negative control. (B) Prevalent mutation types in alleles from transformants *CCR1* and *IAA9A*, respectively. The top alignments show mutations in DNA sequences compared to the WT sequence (red dashes indicate deleted base pairs) and the consequences on the protein sequences are displayed below (altered amino acid sequences are in red). The sequences of sgRNA1 and sgRNA2 are underlined, Protospacer Adjacent Motif (PAM) sequences are indicated in blue. Two SNP positions in *CCR1* are indicated in yellow. The mutation types, number of altered alleles detected and their corresponding occurrence in percentage in brackets are indicated on the right part of the figure.

Table 3. Majority of altered alleles introduced significant modifications at the protein level.

Gene	Edited Clones	Presumed Significant Modifications		Presumed Less Significant Modifications (15 bp Indel, Substitution without Shift)
		Reading Frame Shift ¹	≥15 bp Indel	
<i>CCR1</i>	89	60 (67.4%)	22 (24.7%)	27 (30.3%)
		62 (69.7%)		
<i>IAA9A</i>	88	75 (85.2%)	50 (56.8%)	12 (13.6%)
		76 (86.4%)		

¹ Reading frame shifts all induced premature stop codons. Indel: insertion and deletion; sub: substitution.

We performed three independent transformation batches for *CCR1*. For the first batch, we collected our samples on 153-day-old roots (109 days of in vitro culture + 44 days in hydroponic culture) in order to obtain roots containing secondary xylem. Although, all the roots were edited, the allele edition rate was low (23.5%; only 50 alleles altered among 213 alleles sequenced; Supplementary Table S2). A likely explanation is that some *CCR1* editing events may have impacted too severely roots development as observed previously when down-regulating *CCR1* in hairy roots [14]. Supporting this hypothesis was the dramatic rate of mortality (59.2%) for this first batch of *CCR1* transgenics right after the transfer from in vitro culture to hydroponic medium (Supplementary Table S3). This prompted us to verify if the edition rate could be higher if we harvested younger roots

cultivated only in vitro without adding the stress of transferring them to hydroponic culture. For the second batch and third batch of transformation CRISPR *CCR1*, we collected the samples at 73 days and 54 days, respectively. The genotyping results clearly showed that the younger the transgenic roots, the higher the edition rate. The editing rate reached 80.8% for the roots collected at 54 days (around 1–2 cm long) and dropped to 46.2% for those sampled at 73 days (Supplementary Table S2).

2.1.2. High Knock-Out Rate in *IAA9A* Lines

For *IAA9A*, we obtained 13 independent transgenic hairy roots plants. Twelve out of 13 had editions/mutations in at least one allele, leading to a very high edited plants rate (92.3%; Table 1). In contrast to *CCR1*, most of mutations profiles suggested *IAA9A* knock-out (91.7%; 11 plants out of 12 edited plants). Among them seven plants (58.3%) were biallelic mutations, four plants (33.3%) were chimera with only mutated alleles and no WT allele. One plant (8.3%) had a monoallelic mutation (A1/W). Unexpectedly, one transgenic hairy root had only WT alleles although it exhibited DsRed florescence indicating the presence of the T-DNA. Neither homozygote mutations nor chimera with the WT allele were detected (Table 1).

In total, we sequenced 95 subclones, among which 88 presented mutations, leading to an allele edition rate as high as 92.6%, which was in sharp contrast to the allele edition rate of *CCR1* (32.0%). In the first guide RNA and the second sgRNA positions, 84 (95.5%) and 79 clones (89.8%) had mutations, respectively; whereas 75 clones (85.2%) contained mutations in both sgRNA1 and sgRNA2 positions (Table 2). Thus, eleven out of 13 plants had all alleles altered leading to a putative knock out rate of 84.6%.

The detailed mutations for each altered allele detected can be found in Supplementary Table S4. The two most prevalent edition types were 73 bp deletion (26 alleles, 30%) and 1 bp deletion (8 alleles, 9%), both generated reading frame shifts and premature stop codons (Figure 1). Among the 88 edited alleles, 87 generated amino acid sequence changes, thus potentially modifying the properties of the encoded protein (Supplementary Table S4). The majority (86.4%) of *IAA9A* targeted edition introduced significant protein modifications (Table 3) including 85.2% of reading frame shifts with premature stop codons, and 56.8% of 15 bp or more indels (insertion/deletions).

In parallel to subcloning, we used the DSDDecode online tool on the 13 PCR amplicon sequences (13 transgenic hairy roots). Nine were successfully treated by DSDDecode (Supplementary Table S5). For three of them, we got the same results as our subcloning and sequencing data (highlighted in green, Supplementary Table S5). They included one monoallelic mutation line (*IAA9A_5*), one biallelic line (*IAA9A_15*) and a line without any mutation (*IAA9A_10*). For two of them the results obtained with the two methods were different (highlighted in yellow, Supplementary Table S5). For the line *IAA9A_3*, for instance, the analysis of the subcloning results showed that it was a biallelic mutation whereas the DSDDecode concluded that it was a homozygote mutation with the allele of 1 bp deletion; for the line *IAA9A_11*, the two edition types detected by subcloning were different from the one edition type and WT allele detected by DSDDecode. For the four others, we observed only partial overlapping of the results between the two methods (Supplementary Table S5). For example, for the line *IAA9A_1*, the subcloning results showed biallelic mutations (eight out of nine subclones showed a 73 bp deletion and one showed a 6 bp deletion) whereas the DSDDecode online tool concluded that it was a homozygote mutation consisting of the allele with a 73 bp deletion. For the four lines failed to be analyzed by DSDDecode, the subcloning sequencing revealed one chimeric line (no WT, A1/A2/A3, etc.) and three biallelic mutations (A1/A2).

2.1.3. Mutation Spectra Vary Among sgRNA Targets

Various types of editing were detected in both *CCR1* (43 types) and *IAA9A* (20 types) lines (Supplementary Table S1 and Supplementary S4), including deletions, insertions and substitutions (Table 4). The most prevalent edition type was deletion (Table 4, Figure 1). For *CCR1*, 52 and 21 alleles exhibited small and large deletions, respectively. In total, 73 out of 89 alleles (82.0%) had deletions. For *IAA9A*, 49 and 38 alleles had small (smaller than 15bp) and large (larger than 15bp) deletions, respectively, leading to an overwhelmed majority (87 out of 88 alleles, 98.9%) of the deletion editing

type. As expected, using two guide RNAs separated around 100 bp, large deletions occurring between the two sgRNAs were frequently observed, representing 21.3% and 30.7% of the total edition types for *CCR1* and *IAA9A*, respectively. The second most frequent edition type was substitution. For *CCR1* small substitutions represented the second most prevalent type, scoring as high as 32.6% while no large substitutions happened in *CCR1* or *IAA9A* alleles. Insertions were not often seen in *CCR1* edited alleles; only 4 small insertions were detected among 89 altered alleles and no large insertions were observed. In contrast, in *IAA9A* lines, 16 and 12 alleles had small and large insertions, respectively, representing 31.8% (28 alleles out of 88 alleles) of the editing events.

Table 4. Mutation types.

Gene	Total Edited Clones	Total Edition Types	Deletion			Insertion			Substitution			Expected Large Deletion									
			Small Deletion (≤15bp)			Large Deletion (>15bp)			Small Substitution (≤15bp)				Large Substitution (>15bp)								
			sg1	sg2	sg1&2	sg1	sg2	sg1&2	sg1	sg2	sg1&2		sg1	sg2	sg1&2	sg1	sg2	sg1&2			
CCR1	89	43	12	40	0	19	21	19	3	1	0	0	0	0	24	5	0	0	0	0	19 (21.3%)
			52 (58.4%)			21 (23.6%)			4 (4.5%)			0			29 (32.6%)			0			
IAA9A	88	20	29	39	19	38	27	27	5	11	0	12	0	0	12	2	0	0	0	0	27 (30.7%)
			49 (55.7%)			38 (43.2%)			16 (18.2%)			12 (13.6%)			14 (15.9%)			0			

sg1, sgRNA1; sg2, sgRNA2; sg1&2, sgRNA1 and sgRNA2. The most prevalent edition type is highlighted in green and the second most prevalent types are highlighted in yellow for both *CCR1* and *IAA9A* alleles.

The comparison of the editing rates between the two different sgRNAs for *CCR1*, revealed that 46 and 65 mutations occurred in sgRNA1 and sgRNA2, respectively (Table 2). Using a larger data set of RNAseq from *E. grandis* than the one used for designing the sgRNAs, we detected the presence of a SNP at position 20 flanked by PAM GGG of sgRNA1 (Figure 1, indicated in a yellow character; Supplementary Table S1, footnote) that likely explains the different editing rates between the two sgRNAs. Since we used *E. grandis* seeds and not a clone for hairy root transformation, we assumed that those containing a SNP in *CCR1* were not edited at all as shown for 4-coumarate: CoA ligase by [22]. The editing types were also different between the two *CCR1* sgRNAs. Small substitutions as well as insertions were more frequently observed in the sgRNA1 position than in the sgRNA2 position (Table 4). The two sgRNAs of *IAA9A* displayed equivalent editing rates: 84 and 79 alleles had mutations in sgRNA1 and sgRNA2 positions, respectively.

2.2. Phenotyping Revealed Expected Alterations of Lignification in *CCR1*-Edited Lines

The phenotyping was focused on *CCR1*-edited line because we knew the phenotypes of knock-down mutants (previously generated by antisense techniques in our team [14]) to validate our CRISPR/Cas9 technology was successfully implemented in *Eucalyptus* hairy roots. However, we are uncertain of the phenotypes of *IAA9A*-edited lines, and the phenotypic analyses of *IAA9A*-edited lines are ongoing, so it will not be presented here.

2.2.1. Combination of FTIR Spectroscopy and Multivariate Analyses of *CCR1*-Edited Hairy Root Lines

In order to discriminate rapidly and efficiently between the chemotypes of the CRISPR/Cas9 edited-*CCR1* roots (batch 1) and the control ones, we used the Fourier transformed infra-red (FT-IR) spectroscopy, a fast, cheap and non-destructive technique that provides information about the structure of secondary xylem constituents and chemical changes in wood samples [31]. We analyzed all the FTIR spectra obtained by a multivariate statistical tool (here partial least square analysis (PLS-DA)) because such combinations were shown to be powerful to characterize differences between complex biological samples and provide clues concerning the chemical nature of their divergence [32]. As shown in Figure 2A, the two first components of the PLS-DA explained more than 50% of total variability of the samples. The first component (PC1 axis), which explained 42% of the variability clearly separated *CCR1*-edited lines from controls. The second component (PC2 axis) explained different patterns among the *CCR1* transgenic lines.

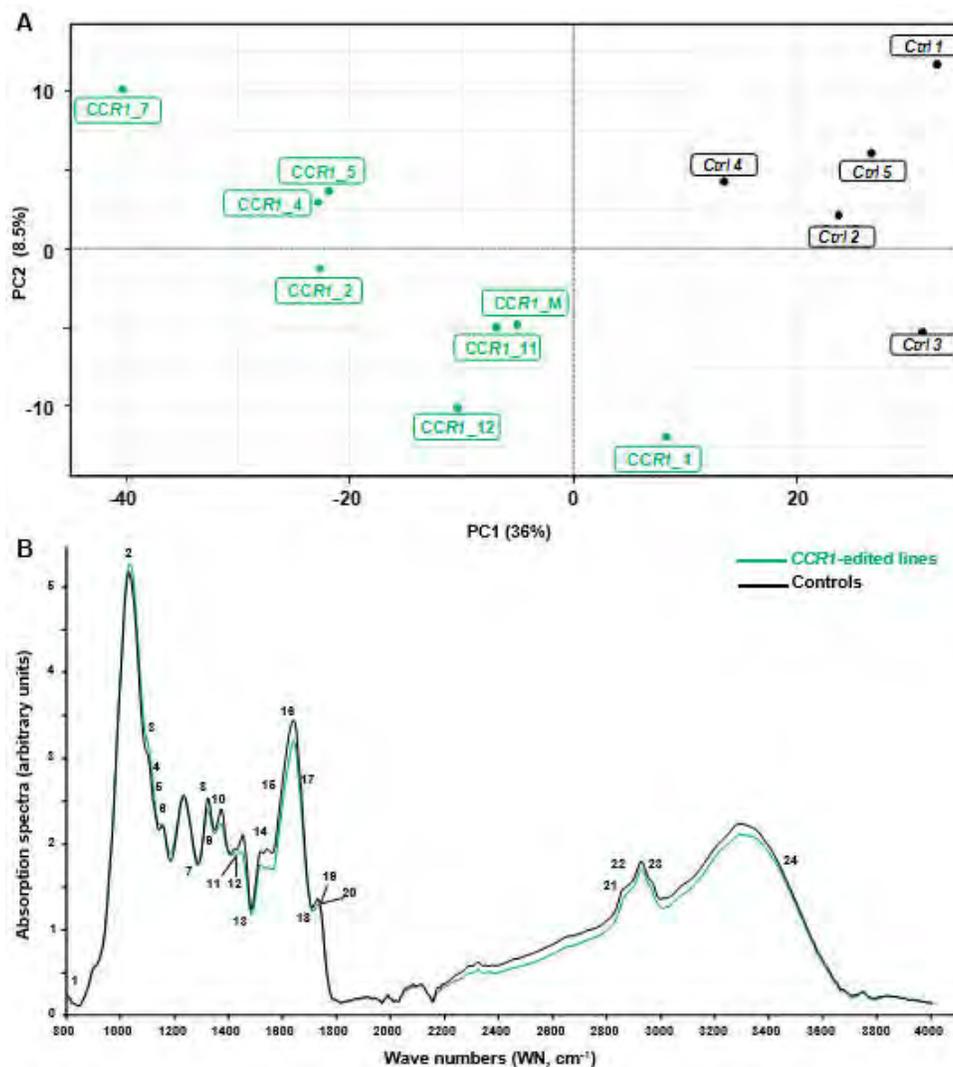


Figure 2. Comparison between FT-IR spectra obtained from controls and *CCR1* edited lines. (A) partial least square analysis (PLS-DA) analysis was performed using the normalized values of Fourier-transformed infra-red (FTIR) absorption spectra (800–4000 cm⁻¹), acquired from hairy roots samples. The first principal component (PC) separates controls from *CCR1*-edited lines and explains more than 36% of total variability. PC2 axis (8%) mostly explains the separation of different groups within *CCR1*-edited lines. *CCR1_M* represent a mixture of five *CCR1*-edited samples (*CCR1_6*, *CCR1_8*, *CCR1_9*, *CCR1_10* and *CCR1_13*) due to not enough materials if proceeded individually. (B) FT-IR absorption spectra of controls (black) and *CCR1*-edited lines (green). The curves were drawn using the median of controls and *CCR1*-edited lines absorption values (except *CCR1-1* and *CCR1-14*). Numbers 1–24 are the most significant wave numbers related to secondary cell wall polymers involved in the separation between controls and *CCR1*-edited lines (see Supplementary Figure S1 and Supplementary Table S6) [33–43].

We used the median of the FT-IR absorption spectra (Figure 2B) and the loadings contribution values to PC1 and/or PC2 to further identify the most discriminant wavenumbers explaining the separation between these contrasting samples. In total, we found 29 discriminant wavenumbers explaining the separation between *CCR1*-edited lines from controls that belong to three main regions of the spectra: 1000–1200cm⁻¹, 1300–1800cm⁻¹ and 2200–3600cm⁻¹ (see Supplementary Figure S1 and Supplementary Table S6). Among them 24 reported on FT-IR spectra (Figure 2B) were already described in the literature as bounds related to SCW composition. Notably, the majority of these

bands were related to lignin structure and composition (Supplementary Table S6). For most of them, the absorption of *CCR1*-edited lines was lower than in controls (Figure 2B) as expected from the down-regulation of a step-limiting enzyme of the lignin biosynthesis pathway.

2.2.2. Histochemical Characterization of *CCR1*-Edited Hairy Root Lines

We further examined the vascular tissues of the transgenic roots by performing histological analyses using either the phloroglucinol-HCl, which stains lignin polymers in red-purple or the natural auto fluorescence of phenolic compounds (including lignin) under UV-light. In root sections performed at around 10 cm from the root apex, most CRISPR/Cas9 edited *CCR1*-lines displayed a clear *CCR1* down-regulation phenotype: xylem cell walls (vessels and fibers) stained faintly with phloroglucinol-HCl in comparison to roots transformed with control vectors, which appeared strongly stained in red (Figure 3, Supplementary Figure S2). As the intensity of the phloroglucinol-HCl staining is indicative of the lignin content, the faint staining in the CRISPR/Cas9 edited lines strongly suggest a reduced lignin content. A frequent collapsed xylem vessels and irregular shapes for both xylem vessels and fibers were also detected in those edited lines (Figure 3, indicated by arrows), due to lower lignification, as well as cells with greatly decreased lignification (non-phloroglucinol staining and/or no auto fluorescence under UV light; Figure 3, indicated by *). These observations were consistently obtained from the several lines (Figure 3, Supplementary Figure S2), especially the strongest phenotypes were more distinct in those such as the lines comprised big deletions (e.g., *CCR1_5*, Figure 3. Supplementary Table S1) and the lines had significant mutated alleles (e.g., *CCR1_7*, all edited alleles had shifted reading frame, Supplementary Table S1). Under UV-light the intensity of auto-fluorescence was also lower in the CRISPR/Cas9 edited *CCR1*-lines as compared to controls, further supporting the hypolignified phenotypes of the *CCR1*-edited lines.

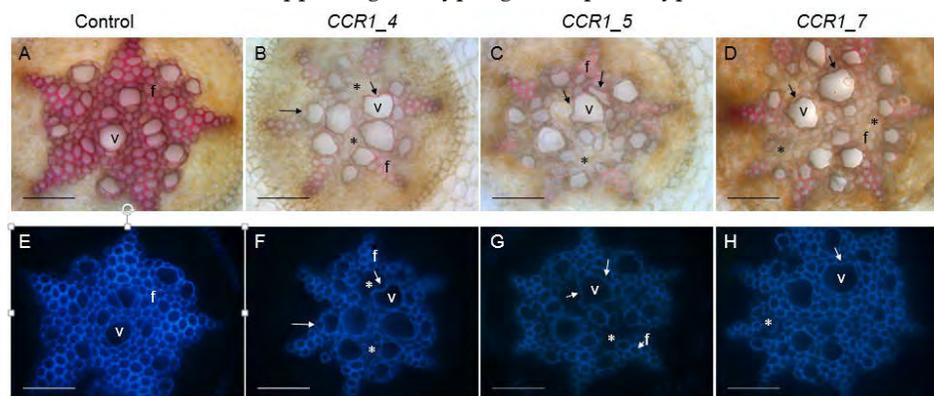


Figure 3. Comparison of xylem development and lignification of xylem cells between control and *CCR1* edited lines' roots. Transversal root sections made at around 10 cm from the root apex of control (A,E) and edited lines (B–D,F–H). Lignified cell walls are visualized in red/purple by phloroglucinol-HCl (A–D) and in blue by UV auto fluorescence (E–H). Collapsed vessels in *CCR1*-edited lines are indicated by arrows. Cells with greatly decreased lignification (non-phloroglucinol staining and/or no auto fluorescence under UV light) are indicated by *. V, vessels; f, fibers. Scale bar = 50 μ m.

3. Discussion

Whereas some studies have implemented gene editing by CRISPR/Cas9 in woody species including fruit trees [44–47] and forestry species like poplar [20,22,24,25,29,48] none was yet reported in eucalypts probably because they are particularly recalcitrant to genetic transformation [11,14]. The purpose of this study was to combine the CRISPR/Cas9 gene editing system with the efficient *Eucalyptus grandis* hairy roots transformation system in order to obtain a powerful knock-out system for gene functional studies. We investigated the mutagenesis efficiencies and patterns produced by the CRISPR/Cas9 nuclease directed against two distinct target genes: (i) the lignin biosynthetic gene *CCR1* used previously as a proof-of-concept to show that *Eucalyptus* hairy roots were adapted to

functionally characterize SCW-related genes [14] and (ii) the *Aux/IAA* gene *IAA9A*, a potential regulator of xylem formation [9].

For both genes, we obtained very high percentages of edited hairy roots amongst the cotransformed ones, i.e., 100% and 92% for *CCR1* and *IAA9A*, respectively (Table 1). However, the allele-editing rates varied considerably between these two targets. While the allele-editing rate was very high (92.6%) in *IAA9A* transgenic roots, it was low (32.0%) in *CCR1* transgenic roots. Indeed, in *CCR1*-edited transgenic roots, we were surprised by the absence of biallelic editing and by the very high percentage of chimera. In strong contrast, for *IAA9A*-edited transgenic roots, the level of biallelic mutations was high (58.3%) and much less chimera was detected. Moreover, the percentage of potential knock-out among the *IAA9A*-edited lines (11 out of 12 lines) was extremely high, indicating that the CRISPR/Cas9 nuclease editing system used in this study could be highly efficient in *E. grandis* hairy roots. In addition, based on the differences observed between the *IAA9A*- and the *CCR1*-edited hairy roots, the editing efficiency appeared gene- and sgRNA-dependent. The results obtained with *IAA9A* are closer to those obtained in rice [18,49] and poplar [22,29,49,50] where high proportions of biallelic mutations were reported in the T0 transgenic plants but in comparison, we got more chimeric mutations in *IAA9A*-edited hairy roots. Obtaining biallelic mutations in T0 transformants is important for trees, since in poplar for instance, it was shown that biallelic mutations were stably inherited through clonal propagation [22,29,51]. This is in agreement with the fact that CRISPR-induced biallelic DNA modifications lead to permanent mutations in edited cells that are inherited mitotically and no more editions are possible.

The phenotypes of the *CCR1*-edited lines analyzed by the combination of FTIR/PLS-DA on one hand, and by histochemistry on the other hand, shared features characteristics of *CCR1*-deficiency reported in other plants such as a low lignin content and collapsed vessels [14,41,42]. We analyzed more in depth the *CCR1*-editing case to understand why we obtained so many chimera transgenic hairy roots likely leading to *CCR1*-knock-down and no biallelic mutations that would lead to *CCR1* knock-out. Indeed, many examples in the literature revealed that too strong *CCR1*-down regulation leads to deleterious effects. In transgenic tobacco plants, the antisense line with the most severely depressed *CCR1* activity exhibited dramatic development alterations with reduced size, abnormal morphology of the leaves, collapsed vessels [52]. In poplar down-regulated for *CCR* by sense or antisense strategies, 5% of the transformed plants were dwarf and unable to be acclimatized [53]. Our hypothesis thus is that biallelic *CCR1* editing would be either lethal or would lead to too severely impaired *CCR1*-lines with very poor development that would die prematurely as it was the case for the more severely down regulated *CCR1* transgenic hairy roots obtained by antisense strategy [14]. Since the hairy roots derive from the transformation of one single cell, the chimera may result from monoallelic edition. In this case, as only one allele is edited, the second allele (wild-type) still contains intact sgRNA target sites. During cell division, the Cas9 protein is able to edit the wild-type targets generating a second type of edition and so on and so forth. In the case of CRISPR-edited *CCR1*, we found two types of chimera (i) those with only edited alleles (A1/A2/A3, etc.) and (ii) those still having wild type alleles (WT/A1/A2/A3, etc.). In the former case, the chimera could be considered as stable because all the alleles are altered and no more target sequences are available for Cas9. In the latter case, the transformants are potentially “not stable” since the target sequences (contained in the wild-type allele(s)) are still available for Cas9 editing, especially because *Cas9* is under the control of the constitutive *CaMV35S* promoter. Taking advantage of the three different transformation batches made for *CCR1*, we compared the allele-editing rates at different times after transformation. Indeed, the allele-editing rate was quite high for the younger hairy roots recently transformed with *A. rhizogenes* and decreased rapidly along with the age of the *CCR1*- transformants (Supplementary Table S2). One possible explanation is that in chimera hairy roots comprising three or more transformed cell lineages, the severely *CCR1*-impaired cell lineages were constantly facing the concurrence from surrounding wild-type cell lineages with normal *CCR1* function and/or *CCR1* edited cell lineages in which *CCR1* function was less impaired. We believed that the severely *CCR1*-deficient cell lineages would be less competitive and could gradually disappear, leading to a lower allele edition rate in the older transformants. This indirect argument also supports the lethality of a

too severe *CCR1*-down regulation and a fortiori of a *CCR1*-knock out and the fact that we did not find any biallelic mutations.

Since the *E. grandis* genome is available [4], we chose the on-line tool 'CRISPOR' (<http://crispor.tefor.net/>) to design the sgRNAs, because it can directly evaluate and score the genome scale off-target risk as well as the editing efficiency [54]. We also selected very low off-target risk sgRNA targets (Supplementary Table S4). Although we cannot completely rule out the possibility of having off-targets in our system, most reports in various plants reported a lack of or low percentage of off-target mutagenesis (reviewed by [29]).

In this study, we implemented the CRISPR/Cas9 gene editing system in *Eucalyptus* hairy roots. Among the mutations generated, the majority introduced frame shifts with premature stop codons and thus truncated proteins for both *CCR1* and *IAA9A* genes. In addition, as expected using two gRNAs, large deletions were frequently seen most likely leading to non-functional proteins. The low level of biallelic mutations and the high level of chimera were unexpected especially when compared to poplar [20,22,29], or even to hairy roots system in soybean, tomato or chicory [55–59], where high percentage of biallelic mutations and low number of chimera were reported in general. Although in some studies, the percentage of chimera could have been underestimated by the use of the DSDecode software, which is not able to detect complex patterns of mutations such as chimeric or multiple mutations and is less accurate than subcloning to identify chimera as we showed here. There is room in our system to increase the percentage of biallelic knockout and reduce the mono allelic mutation and chimera percentages. Indeed, for many species, it has been observed that monoallelic/chimera mutations predominate in the first-generation (T0) when CRISPR editing efficiencies were low [19,23]. Multiple parameters were reported impact the edition efficiency such as the use of the native U6 or U3 promoters to drive the expression of the sgRNAs [56,60], the Cas9 expression cassettes. Bruegmann T. and his colleagues [51] also reported that the structure of sgRNA impacts gene editing efficiency, in particular the GC content, the presence of purine residues at the sgRNA end and the free accessibility of the seed region seemed to be highly important for genome editing in poplar. Further studies are needed to explore the impact of the use of (i) *Eucalyptus* native U6 or U3 promoters, (ii) different Cas9 expression cassettes and (iii) optimized sgRNA structures, to achieve higher biallelic edition rate in the T0 generation. This is particularly important for eucalypts that are like poplars, clonally propagated. Although some progress are also needed to improve eucalypts transformation, further research may also be guided to generate CRISPR/Cas9 editing without any transgenic DNA integration by transferring just ribonucleic-protein complexes into plant cells [61] to overcome the persistent societal hostility to transgenic trees.

4. Materials and Methods

4.1. Plant Material

Commercial *E. grandis* seeds (W. Hill ex Maiden, cultivar LCFA001) purchased at Instituto de Pesquisas e Estudos Florestais (IPEF, Piracicaba, Brazil) were surface-sterilized by 30-min treatment in a 1% sodium hypochlorite solution containing Twin-20. Germination was carried out on 1/4 strength Murashige and Skoog medium (MS medium; Sigma-Aldrich, St. Louis, MO, USA) solidified with 8 g/L (Sigma-Aldrich) at 25 °C in the dark for 3 days. To obtain in vitro plantlets around 1cm long with the hypocotyls fully expanded and the first two leaves just appearing, we germinated seeds inside plates at normal position at 25 °C for 12 days in light (12 $\mu\text{mol}/\text{m}^2/\text{s}$, 8–16-h photoperiod, 50% humidity).

4.2. CRISPR/Cas9 Targeted Mutagenesis System Selection and Pipeline

To implement the CRISPR/Cas9 system in *Eucalyptus*, we selected the method introducing selective marker, 35S-Cas9-Nos expression cassette and two sgRNAs under the corresponding promoter (here *Arabidopsis* U6 promoter) in a single construct that had previously been proven highly active in various plants such as tomato and *Nicotiana benthamiana* [30,62,63] using golden gate cloning [64,65]. The scheme illustrating the pipeline of the targeted mutagenesis in plants is shown in Figure

4, including the main steps as (1) sgRNAs design, (2) construct assembly, (3) hairy roots transformation, (4) genotyping and (5) phenotyping.

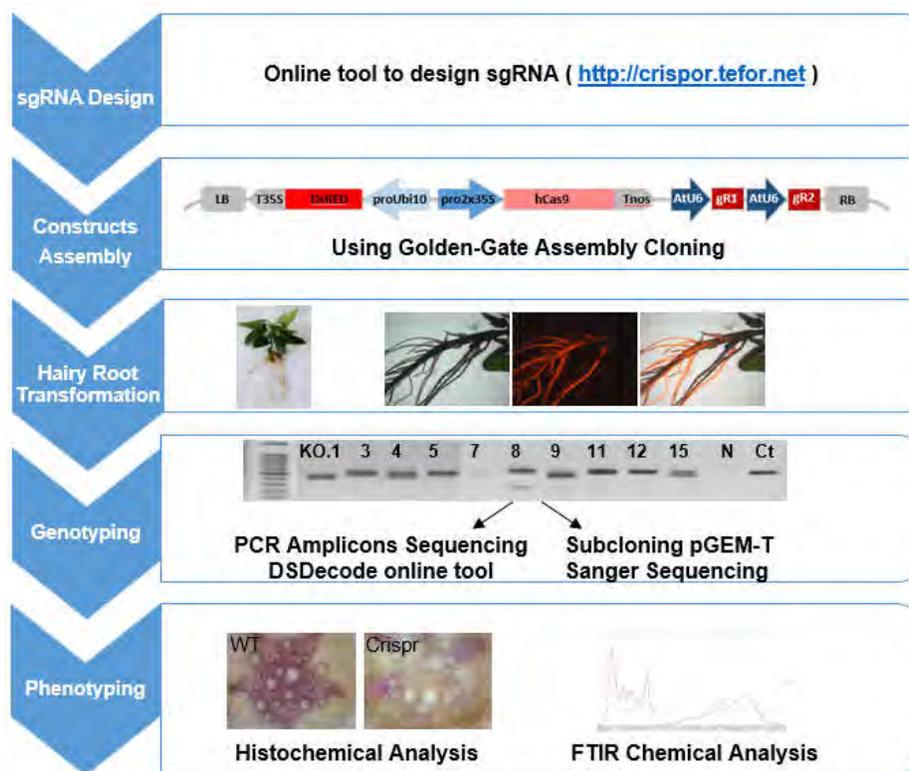


Figure 4. Pipeline of CRISPR/Cas9 implementation in *Eucalyptus grandis* hairy roots.

4.3. CRISPR/Cas9 Target Site Selection and sgRNAs Design

We selected two different target sites for each gene with the help of the sgRNA design online tool ‘CRISPOR’ (<http://crispor.tefor.net/>), which enables to evaluate the genome wide off-targets and to score on-target efficiency [54]. We first retrieved the genomic sequences of our two target genes obtained from *Eucalyptus grandis* genome sequencing database of Phytozome (<https://phytozome.jgi.doe.gov>). All possible sgRNAs on the genomic sequence were ranked by CRISPOR based on the off-target risk and on target efficiency. Among those, we selected pairs of sgRNAs at around 100 bp interval to increase editing efficiency and to possibly create a large deletion between the two sgRNAs [62]. The target sites selected had a ‘G’ as their first base to function as the RNA polymerase III start site (Guide RNA prefix for U6 promoter) and were followed by the Protospacer Adjacent Motif (PAM) sequence ‘NGG’ given the *Streptococcus pyogenes* Cas9 as PAM selection preference. The selected sgRNAs and the off-target risk and editing efficiency prediction were presented in Supplementary Table S7 [66].

Since eucalypts are highly heterozygous, we verified the absence of SNPs in the guide RNAs by blasting them against a large RNAseq data set of *E. grandis* [67] registered at the NCBI SRA database (PRJNA514408). Unexpectedly, the alignment between the RNAseq and the genomic sequence of *CCR1* retrieved from Phytozome (*Eucgr. J03114*) allowed us to detect a deletion of 14 bp (5'-gcttctcctcgcagc-3') at position 33649876 (Chr. J; Figure 5) in the latter and consequently the predicted exon/intron structure and protein sequence were wrong. We corrected manually the sequence and both gene structure and proteic sequence were in perfect agreement with our previous work [2,68].

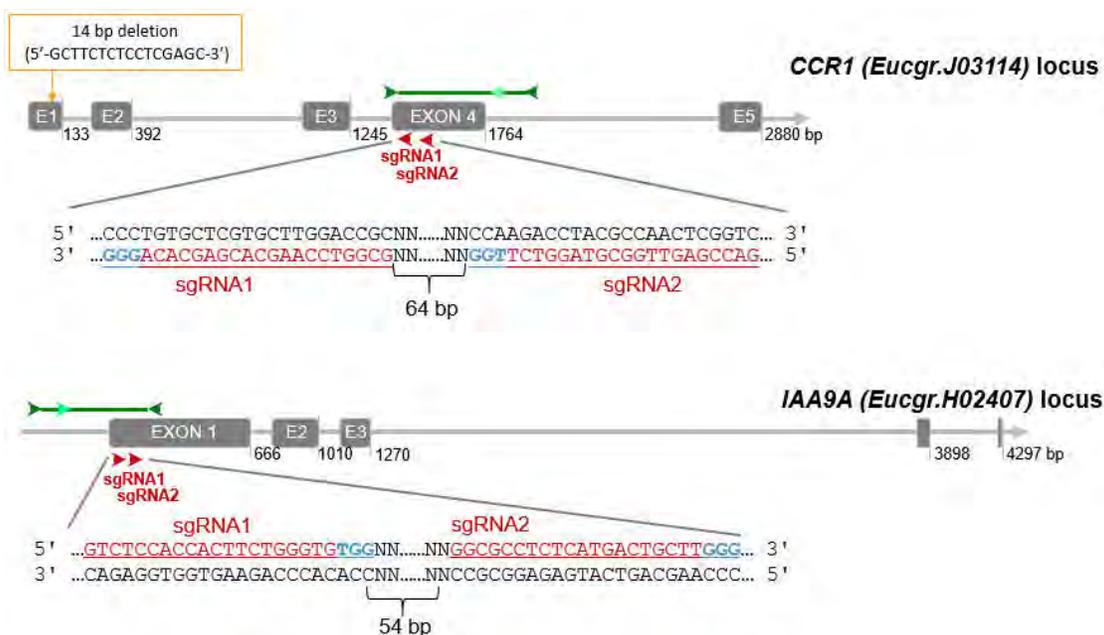


Figure 5. CRISPR/Cas9 sgRNA design and mutation detection in *CCR1* and *IAA9A*. Schematic representation of the target sites and the PCR assay for Sanger sequencing. Exons and introns are represented by gray boxes and gray lines, respectively. The target sites for each CRISPR/Cas9 nuclease are indicated by red arrows, sgRNA target sequence are indicated in underlined red characters, PAM sequences in blue. The dark green arrows indicate approximately the location of the primers for PCR amplification, the light green arrows indicate the nested primers designed for DSDecode mutation identification.

Finally, for *CCR1* (*Eucgr.J03114*), we chose two sgRNAs (sgRNA1_ *CCR1* (5'-GCGGTCCAAGCAGAGCACA-3') and sgRNA2_ *CCR1* (5'-GACCGAGTTGGCGTAGGTCT-3')) separated by 64 bp and both located on the antisense strand in exon 4 (Figure 5). This exon contains a highly conserved region among different *Eucalyptus* species [2]. For *IAA9A* (*Eucgr.H02407*) we selected a pair of sgRNAs (sgRNA1_ *IAA9A* (5'-GTCTCCACCACTTCTGGGTG-3') and sgRNA2_ *IAA9A* (5'-GGCGCCTCTCATGACTGCTT-3')) located at 54 bp interval in exon 1 (Figure 5). The sgRNA1 and sgRNA2 of *IAA9A* locate 23 bp and 96 bp after ATG start codon respectively.

4.4. CRISPR/Cas9 Constructs Assembly

The construct assembly was achieved by two steps using golden gate cloning as described in [62]. In brief, for the first step we generated two intermediary vectors (AtU6p::sgRNA1_ *CCR1* and AtU6p::sgRNA2_ *CCR1*) by cut-ligation using *Bsa*I endonuclease and T4 ligase. Each vector harbored *Arabidopsis* U6 promoter, corresponding *CCR1*-sgRNA1 or 2 target sequence, sgRNA scaffold and U6 terminator. The primers and sgRNA scaffold used for generating sgRNA intermediary vectors were described in Supplementary Table S8. sgRNA scaffold template (pICH86966) and level 0 construct (pICSL01009::AtU6 SpecR), level 1 destination vector (pICH47751 (CarbR) and pICH47761 (CarbR) were provided by Dr. G HU (UMR990 GBF, Toulouse France), which can be ordered from Addgene. The second step in one reaction we cut-ligated all intermediary vectors into one final binary vector pICSL4723 (LB-DsRed-Cas9-sgRNA1-sgRNA2-RB) using *Bbs*I endonuclease and T4 ligase (Figure 6), which includes all the CRISPR/Cas9 components: the plant selective marker DsRed expression cassette (AtUbi10p::DsRed::T35S-terminator) at position 1 flanked to the left border of binary vector, domesticated human codon optimized Cas9 expression cassette (2x35Sp-5'UTR::Cas9::NOS Terminator) at position 2, and two sgRNA expression cassettes (AtU6p::sgRNA::U6terminator) at position 3 and position 4 (Figure 6). The DsRed selection marker vector (AtUbi10p::DsRed::T35S-terminator) was provided by Dr. PM Delaux, UMR5546 LRSV, France. The Cas9 expression cassette

vector (pICH47742::35S::Cas9-NOST) and the linker vector (pICH41780 Linker) can be obtained from Addgene plasmids (<https://www.addgene.org/>) The destination vector pICSL4723 was provided by M. Youles (Sainsbury Laboratory, Norwich, UK). For the empty vector control construct, we cut-ligated the DsRed and 35S-Cas9-NOST expression cassettes but without any sgRNA as described following (Figure 6).

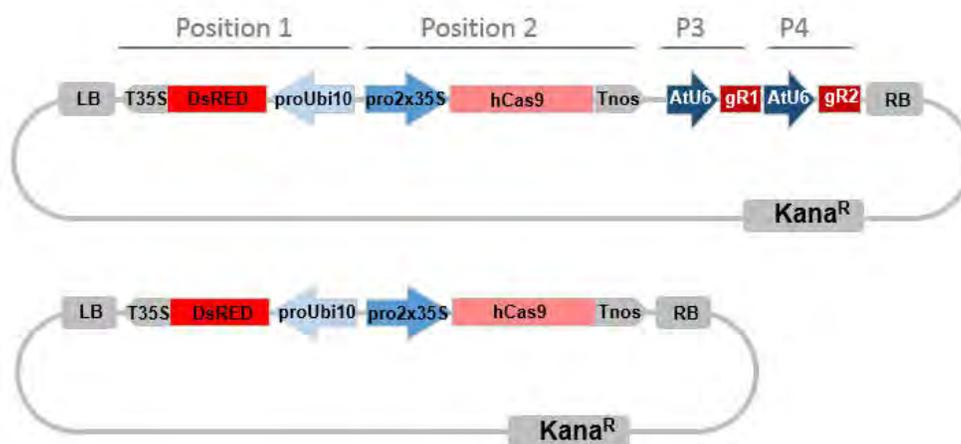


Figure 6. CRISPR/Cas9 binary vector targeting two loci simultaneously. Promoters are indicated in blue arrows, terminators are indicated in gray arrows. *Pro2x35*, double Cauliflower mosaic virus *CaMV* 35S promoter; *AtU6*, *Arabidopsis thaliana* U6 gene promoter; *hCas9*, human codon-optimized *Cas9* gene sequence from *Streptococcus pyogenes*; LB, left T-DNA border; *Kana^R*, kanamycin resistance gene sequence; DsRed, DsRed fluorescent marker gene sequence, *ProUbi10* *Arabidopsis thaliana* Ubiquitin 10 gene promoter; *T35S*, terminator region of *CaMV35S* gene; *TNos*, terminator region of the *nopaline synthetase* gene from *Agrobacterium tumefaciens*; RB, right T-DNA border, P3, Position 3; P4, Position 4, miss description of the gRNA. The bottom construct is the Cas9 control plasmid without sgRNA.

For the first step, the cut-ligation (Type II restriction endonucleases–T4 ligation) reaction (20 μ L) was prepared to contain 2 μ L BsaI 10 \times reaction buffer (NEB), 20 U of BsaI (NEB), 20 U of T4 DNA ligase (using high concentration ligase, 20 U/ μ L, Promega (Charbonnières-les-Bains, France)), approximately 40 fmol insert (100 ng of DNA for a 4 kb plasmid) of pICSL01009::AtU6p vector, sgRNA PCR amplicons harboring targeted sequences (15 ng) and 20 fmol destination vector pICH47751 or pICH47761 (with molar ratio 2:2:1). The reactions were incubated in a thermocycler (marque) for 13 cycles (37 $^{\circ}$ C, 10 min; 16 $^{\circ}$ C, 10 min), followed by 20 min of digestion at 37 $^{\circ}$ C, and 10 min of denaturation at 55 $^{\circ}$ C. The ligations were then transformed into DH5 α *E. coli* (Thermo Fisher (Illkirch-Graffenstaden, France)). White colonies were selected on agar with X-Gal and carbenicillin. *E. coli* PCR, plasmid Miniprep (Promega (Charbonnières-les-Bains, France)), restriction endonuclease digestion and Sanger sequencing were carried out to screen and obtain the assembled level 1 vector (AtU6 promoter and gRNA sequence).

For the second step, a restriction-ligation reaction (20 μ L) was set up using T4 ligase buffer (Promega (Charbonnières-les-Bains, France)) plus 2 ng BSA (Promega (Charbonnières-les-Bains, France)), 15 U of BpiI (Thermo Fisher (Illkirch-Graffenstaden, France)), 20 U of T4 DNA ligase, 40 fmol of each insert elements (Vector for DsRed, Cas9, sgRNA1, sgRNA2 and linker) and 20 fmol of destination vector (pICSL4723). The reactions were incubated in a thermo-cycler for 30 cycles (37 $^{\circ}$ C, 5 min; 16 $^{\circ}$ C, 5 min), 20 min digestion at 37 $^{\circ}$ C and 10 min denaturation at 55 $^{\circ}$ C. The cut-ligation products were transformed in One Shot™ TOP10 Chemically *E. coli* Competent cells (Thermo Fisher (Illkirch-Graffenstaden, France)), and a single white colony was selected (white/orange selection) on LB plates supplemented with kanamycin (100 mg/mL). The colony was grown in liquid culture for 12–16 h 37 $^{\circ}$ C and the “DsRed-Cas9-sgRNA1-sgRNA2” vector was extracted using a Miniprep kit, ready to transform *Agrobacterium rhizogenes*.

4.5. *Agrobacterium Rhizogenes*-Mediated Transformation

The binary vectors were transferred into *E. grandis* hairy roots using *A. rhizogenes* strain A4RS as described by [14]. *Eucalyptus* composite plants harboring transgenic hairy roots and wild-type shoots were grown in vitro culture (MS with ½ strength of macro elements) for a period of three to ten weeks (7–12 h light/16 h dark, 40% humidity, 22/20 °C). In order to obtain roots containing enough xylem, for the first batch of *CCR1* transformation, 109-day-old in vitro culture composite plants were transferred in hydroponic culture using MS with ½ strength of macro elements solution in a phytotron (130 µmol/m²/s, 8–16-h photoperiod, 80% humidity, 25/22 °C). After 4 weeks DsRed fluorescence was verified again. For the first batch of *CCR1* transformation, 153 days old hydroponic cultured fluorescent roots expressing DsRed were collected for DNA extraction, chemical analysis (FT-IR) and histochemical analysis. For the second and third batch of *CCR1* transformation, 54 days old and 73 days old in vitro culture fluorescent roots were sampled for genotyping. For *IAA9A* transformation 165 days old (77 days in vitro culture + 88 days in hydroponic culture) fluorescent roots were samples for genotyping.

4.6. DNA Isolation, PCR Amplification and Mutation Identification

Transformed *Eucalyptus* hairy roots (2 cm from the apex) expressing DsRed fluorescence were harvested for genomic DNA extraction by the CTAB method according to [69]. The quality and concentration of the genomic DNA were measured by a Nanodrop (DS-11 Spectrophotometer).

We used PCR to amplify the genomic region flanking the target sites. For *CCR1*, the forward and reverse primers were 102 bp upstream of sgRNA1 and 376 bp downstream of sgRNA2, respectively (*CCR1_edit check_NCBI_F* and *CCR1_edit check_NCBI_R*, amplicon size 579 bp, Supplementary Table S9); For *IAA9A* the forward and reverse primers were 372 bp upstream of sgRNA1 and 131 bp downstream of sgRNA2, respectively (*IAA9A_edit check_NCBI_F*, *IAA9A_edit check_NCBI_R*; amplicon size 595 bp; Supplementary Table S9). The PCR amplifications were performed using the High-Fidelity Phusion DNA Polymerase (Thermo Fisher (Illkirch-Graffenstaden, France)).

Two methods were used to identify mutations. The PCR amplicons were either directly sequenced and analyzed by the web-based tool “Degenerate Sequence Decoding (DSDDecode, <http://skl.scau.edu.cn/sadsdecode/#>), and/or were subcloned into pGEM-T vectors after adding ‘A’ tail by GoTaq polymerase (Promega (Charbonnières-les-Bains, France)) and up to 22 colonies were sequenced by Sanger sequencing. For the first method, instead of using the primers for previous amplification, two nested primers were used for PCR amplicon directly sequencing followed the instruction of the DSDDecode online tool: nested primer *CCR1_R* (located 220 bp downstream of gRNA2) and nested primer *IAA9A_F* (located 191 bp upstream of gRNA1; Supplementary Table S9) to avoid the noise signals of sequencing results.

4.7. FTIR Analyses

Hairy roots were harvested, frozen in liquid nitrogen and kept at -80 °C until use. Samples were freeze-dried during 48 h and milled with a Mixer Mill MM 400 (Retsch). Fourier transform infrared spectroscopy (FT-IR) analysis was performed on 100–200 mg lyophilized roots dried powder samples using an attenuated total reflection (ATR) Nicolet 6700 FT-IR spectrometer (Thermo Fisher (Illkirch-Graffenstaden, France)) equipped with a deuterated-triglycine sulfate (DTGS) detector. Some *CCR*-edited roots were too small to generate enough material to be analyzed by FT-IR.

Spectra were recorded in the range 400–4000 cm⁻¹ with a 4 cm⁻¹ resolution and 32 scans per spectrum. We used hyperspectr v0.99 [70], prospect v0.1.3 [71] and base v3.6.2 packages [72] to perform baseline correction, normalization and offset correction, respectively. All packages were compiled with R version i386 3.5.2. Analyses were performed using the mean spectra resulting from ten individual replicates. Partial least square-Discriminant analysis (PLS-DA) was performed using mixOmics R package [73] to compare samples spectra and identify wavenumbers responsible for samples discrimination.

4.8. Histochemical Analysis

DsRed fluorescence indicating co-transformed roots was detected using a stereomicroscope Axiozoom V16 (Zeiss, Marly le Roi, France) equipped with a color CCD camera (ICC5; Zeiss) and with filter sets for DsRed (607/80 nm). Transverse sections (60 µm thick) of roots embedded in 6% low gelling point agarose (Sigma-Aldrich) were obtained using a Vibratome (VT 100S; Leica) and observed using an inverted microscope (DM IRBE; Leica) equipped with a CDD color camera (DFC300 FX; Leica). Lignified secondary cell walls were visualized either in red/purple by phloroglucinol-HCl staining or in blue due to auto fluorescence under UV light.

Supplementary Materials: Supplementary materials can be found at www.mdpi.com/1422-0067/21/10/3408/s1.

Author Contributions: “Conceptualization, J.G.P. and H.C.W.; methodology, H.C.W., Y.D., F.M., N.L., A.D., L.M. N.B.; investigation, Y.D., H.C.W., F.M.; F.T.I.R. script, G.H. writing—original draft preparation, Y.D., H.C.W., F.M. and J.G.P.; writing—review and editing, H.C.W., and J.G.P.; funding acquisition, J.G.P. All authors have read and agreed to the published version of the manuscript

Funding: This research was funded 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 China Scholarship Council (a PhD grant to Y.D.), the National Council for Scientific and Technological Development of Brazil (CNPq; a PostDoc grant to L.M.).

Acknowledgments: The authors acknowledge the help of Yves Martinez (FR3450) for assistance with microscopy (Plateforme Imagerie TRI; <http://tri.ups-tlse.fr/>).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

BSA	Bovine serum albumin
CaMV	Cauliflower mosaic virus
CCR	Cinnamoyl CoA Reductase
CRISPR	The Clustered Regularly Interspaced Short Palindromic Repeats
CTAB	Cetyl Trimethylammonium Bromide
DSDcode	Degenerate Sequence Decoding
FT-IR	Fourier Transformed Infra-red spectroscopy
Kana	Kanamycin
ORF	Open reading frame
PAM	Protospacer Adjacent Motif
PLS-DA	Partial Least Square Analysis
sgRNA	Single guide RNA
SNP	Single Nucleotide Polymorphism

References

1. Myburg, A.A.; Potts, B.M.; Marques, C.M.; Kirst, M.; Gion, J.-M.; Grattapaglia, D.; Grima-Pettenatti, J. *Eucalypts*. In *Forest Trees*; Springer: Berlin, Germany, 2007; pp. 115–160.
2. Paiva, J.A.; Rodrigues, J.C.; Fevereiro, P.; Neves, L.; Araújo, C.; Marques, C.; Freitas, A.T.; Bergès, H.; Grima-Pettenatti, J. Building up resources and knowledge to unravel transcriptomics dynamics underlying *Eucalyptus globulus* xylogenesis. *BMC Proc.* **2011**, *5*, O52.
3. Holladay, J.E.; White, J.F.; Bozell, J.J.; Johnson, D. *Top Value-Added Chemicals from Biomass-Volume II—Results of Screening for Potential Candidates from Biorefinery Lignin*; Prepared for the U.S. Department of Energy: Richland, WA, USA, 2007; pp. PNNL-16983, 921839.
4. Myburg, A.A.; Grattapaglia, D.; Tuskan, G.A.; Hellsten, U.; Hayes, R.D.; Grimwood, J.; Jenkins, J.; Lindquist, E.; Tice, H.; Bauer, D.; et al. The genome of *Eucalyptus grandis*. *Nature* **2014**, *510*, 356–362.
5. Carocha, V.; Soler, M.; Hefer, C.; Cassan-Wang, H.; Fevereiro, P.; Myburg, A.A.; Paiva, J.A.P.; Grima-Pettenatti, J. Genome-wide analysis of the lignin toolbox of *Eucalyptus grandis*. *New Phytol.* **2015**, *206*, 1297–1313.

6. Soler, M.; Camargo, E.L.O.; Carocha, V.; Cassan-Wang, H.; San Clemente, H.; Savelli, B.; Hefer, C.A.; Paiva, J.A.P.; Myburg, A.A.; Grima-Pettenati, J. The *Eucalyptus grandis* R2R3-MYB transcription factor family: Evidence for woody growth-related evolution and function. *New Phytol.* **2015**, *206*, 1364–1377.
7. Hussey, S.G.; Saïdi, M.N.; Hefer, C.A.; Myburg, A.A.; Grima-Pettenati, J. Structural, evolutionary and functional analysis of the NAC domain protein family in *Eucalyptus*. *New Phytol.* **2015**, *206*, 1337–1350.
8. Yu, H.; Soler, M.; Mila, I.; San Clemente, H.; Savelli, B.; Dunand, C.; Paiva, J.A.P.; Myburg, A.A.; Bouzayen, M.; Grima-Pettenati, J.; et al. Genome-Wide Characterization and Expression Profiling of the AUXIN RESPONSE FACTOR (ARF) Gene Family in *Eucalyptus grandis*. *PLoS ONE* **2014**, *9*, e108906.
9. Yu, H.; Soler, M.; San Clemente, H.; Mila, I.; Paiva, J.A.P.; Myburg, A.A.; Bouzayen, M.; Grima-Pettenati, J.; Cassan-Wang, H. Comprehensive Genome-Wide Analysis of the Aux/IAA Gene Family in *Eucalyptus*: Evidence for the Role of EgrIAA4 in Wood Formation. *Plant Cell Physiol.* **2015**, *56*, 700–714.
10. Tournier, V.; Grat, S.; Marque, C.; El Kayal, W.; Penchel, R.; de Andrade, G.; Boudet, A.-M.; Teulière, C. An Efficient Procedure to Stably Introduce Genes into an Economically Important Pulp Tree (*Eucalyptus grandis* × *Eucalyptus urophylla*). *Transgenic Res.* **2003**, *12*, 403–411.
11. Girijashankar, V. Genetic transformation of *Eucalyptus*. *Physiol. Mol. Biol. Plants* **2011**, *17*, 9–23.
12. de la Torre, F.; Rodríguez, R.; Jorge, G.; Villar, B.; Álvarez-Otero, R.; Grima-Pettenati, J.; Gallego, P.P. Genetic transformation of *Eucalyptus globulus* using the vascular-specific EgCCR as an alternative to the constitutive CaMV35S promoter. *Plant Cell Tissue Organ Cult.* **2014**, *117*, 77–84.
13. Cao, P.B.; Ployet, R.; Nguyen, C.; Dupas, A.; Ladouce, N.; Martinez, Y.; Grima-Pettenati, J.; Marque, C.; Mounet, F.; Teulière, C. Wood architecture and composition are deeply remodeled in frost sensitive *Eucalyptus* overexpressing CBF/DREB1 transcription factors. *Int. J. Mol. Sci.* **2020**, *21*, 3019.
14. Plasencia, A.; Soler, M.; Dupas, A.; Ladouce, N.; Silva-Martins, G.; Martinez, Y.; Lapierre, C.; Franche, C.; Truchet, I.; Grima-Pettenati, J. *Eucalyptus* hairy roots, a fast, efficient and versatile tool to explore function and expression of genes involved in wood formation. *Plant Biotechnol. J.* **2016**, *14*, 1381–1393.
15. Travis, J. Making the cut. *Science* **2015**, *350*, 1456–1457.
16. Elorriaga, E.; Klocko, A.L.; Ma, C.; Strauss, S.H. CRISPR-Cas nuclease mutagenesis for genetic containment of genetically engineered forest trees. *Mosaic* **2015**, *42*, 46.
17. Bortesi, L.; Fischer, R. The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnol. Adv.* **2015**, *33*, 41–52.
18. Ma, X.; Zhu, Q.; Chen, Y.; Liu, Y.-G. CRISPR/Cas9 Platforms for Genome Editing in Plants: Developments and Applications. *Mol. Plant* **2016**, *9*, 961–974.
19. Bewg, W.P.; Ci, D.; Tsai, C.-J. Genome Editing in Trees: From Multiple Repair Pathways to Long-Term Stability. *Front. Plant Sci.* **2018**, *9*, 1732.
20. Fan, D.; Liu, T.; Li, C.; Jiao, B.; Li, S.; Hou, Y.; Luo, K. Efficient CRISPR/Cas9-mediated Targeted Mutagenesis in *Populus* in the First Generation. *Sci. Rep.* **2015**, *5*, 12217.
21. Tsai, C.-J.; Xue, L.-J. CRISPRing into the woods. *GM Crops Food* **2015**, *6*, 206–215.
22. Zhou, X.; Jacobs, T.B.; Xue, L.-J.; Harding, S.A.; Tsai, C.-J. Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial *Populus* reveals 4-coumarate: CoA ligase specificity and redundancy. *New Phytol.* **2015**, *208*, 298–301.
23. Xu, R.-F.; Li, H.; Qin, R.-Y.; Li, J.; Qiu, C.-H.; Yang, Y.-C.; Ma, H.; Li, L.; Wei, P.-C.; Yang, J.-B. Generation of inheritable and “transgene clean” targeted genome-modified rice in later generations using the CRISPR/Cas9 system. *Sci. Rep.* **2015**, *5*, 11491.
24. Wang, L.; Ran, L.; Hou, Y.; Tian, Q.; Li, C.; Liu, R.; Fan, D.; Luo, K. The transcription factor MYB115 contributes to the regulation of proanthocyanidin biosynthesis and enhances fungal resistance in poplar. *New Phytol.* **2017**, *215*, 351–367.
25. Wan, S.; Li, C.; Ma, X.; Luo, K. PtrMYB57 contributes to the negative regulation of anthocyanin and proanthocyanidin biosynthesis in poplar. *Plant Cell Rep.* **2017**, *36*, 1263–1276.
26. Xu, C.; Fu, X.; Liu, R.; Guo, L.; Ran, L.; Li, C.; Tian, Q.; Jiao, B.; Wang, B.; Luo, K. PtoMYB170 positively regulates lignin deposition during wood formation in poplar and confers drought tolerance in transgenic *Arabidopsis*. *Tree Physiol.* **2017**, *37*, 1713–1726.
27. Yang, L.; Zhao, X.; Ran, L.; Li, C.; Fan, D.; Luo, K. PtoMYB156 is involved in negative regulation of phenylpropanoid metabolism and secondary cell wall biosynthesis during wood formation in poplar. *Sci. Rep.* **2017**, *7*, 41209.

28. Shen, Y.; Li, Y.; Xu, D.; Yang, C.; Li, C.; Luo, K. Molecular cloning and characterization of a brassinosteroid biosynthesis-related gene PtoDWF4 from *Populus tomentosa*. *Tree Physiol.* **2018**, *38*, 1424–1436.
29. Elorriaga, E.; Klocko, A.L.; Ma, C.; Strauss, S.H. Variation in Mutation Spectra among CRISPR/Cas9 Mutagenized Poplars. *Front. Plant Sci.* **2018**, *9*, 594.
30. Brooks, C.; Nekrasov, V.; Lippman, Z.B.; Van Eck, J. Efficient Gene Editing in Tomato in the First Generation Using the Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-Associated9 System. *Plant Physiol.* **2014**, *166*, 1292–1297.
31. Reyes-Rivera, J.; Terrazas, T. Lignin Analysis by HPLC and FTIR. In *Xylem*; de Lucas, M., Etchells, J.P., Eds.; Springer: New York, NY, USA, 2017; Volume 1544, pp. 193–211, ISBN 978-1-4939-6720-9.
32. Bjarnestad, S.; Dahlman, O. Chemical Compositions of Hardwood and Softwood Pulps Employing Photoacoustic Fourier Transform Infrared Spectroscopy in Combination with Partial Least-Squares Analysis. *Anal. Chem.* **2002**, *74*, 5851–5858.
33. Sammons, R.J.; Harper, D.P.; Labbé, N.; Bozell, J.J.; Elder, T.; Rials, T.G. Characterization of Organosolv Lignins using Thermal and FT-IR Spectroscopic Analysis. *Bioresources* **2013**, *8*, 2752–2767.
34. Szymanska-Chargot, M.; Zdunek, A. Use of FT-IR Spectra and PCA to the Bulk Characterization of Cell Wall Residues of Fruits and Vegetables Along a Fraction Process. *Food Biophys.* **2013**, *8*, 29–42.
35. Casas, A.; Alonso, M.V.; Oliet, M.; Rojo, E.; Rodríguez, F. FTIR analysis of lignin regenerated from *Pinus radiata* and *Eucalyptus globulus* woods dissolved in imidazolium-based ionic liquids. *J. Chem. Technol. Biotechnol.* **2012**, *87*, 472–480.
36. Popescu, M.; Zanoaga, M.; Mamunya, Y.; Myshak, V.; Vasile, C. Two dimensional infrared correlation spectroscopy studies of wood-plastic composites with a copolyamide as matrix. *J. Optoelectron. Adv. Mater.* **2007**, *9*, 923.
37. Faix, O. Classification of lignins from different botanical origins by FT-IR spectroscopy. *Holzforsch. Int. J. Biol. Chem. Phys. Technol. Wood* **1991**, *45*, 21–28.
38. Largo-Gosens, A.; Hernández-Altamirano, M.; García-Calvo, L.; Alonso-Simón, A.; Álvarez, J.; Acebes, J.L. Fourier transform mid infrared spectroscopy applications for monitoring the structural plasticity of plant cell walls. *Front. Plant Sci.* **2014**, *5*, 303.
39. Adamafio, N.; Kyeremeh, K.; Datsomor, A.; Osei-Owusu, J. Others Cocoa pod ash pre-treatment of wawa (*Triplochiton scleroxylon*) and sapele (*Entandrophragma cylindricum*) sawdust: Fourier transform infrared spectroscopic characterization of lignin. *Asian J. Sci. Res.* **2013**, *6*, 812–818.
40. Shi, J.; Xing, D.; Lia, J. FTIR Studies of the Changes in Wood Chemistry from Wood Forming Tissue under Inclined Treatment. *Energy Proced.* **2012**, *16*, 758–762.
41. Stark, N.M.; Yelle, D.J.; Agarwal, U.P. Techniques for Characterizing Lignin. In *Lignin in Polymer Composites*; Elsevier: Amsterdam, The Netherlands, 2016; pp. 49–66, ISBN 978-0-323-35565-0.
42. Mouille, G.; Robin, S.; Lecomte, M.; Pagant, S.; Höfte, H. Classification and identification of *Arabidopsis* cell wall mutants using Fourier-Transform InfraRed (FT-IR) microspectroscopy. *Plant J.* **2003**, *35*, 393–404.
43. Li, X.; Sun, C.; Zhou, B.; He, Y. Determination of Hemicellulose, Cellulose and Lignin in Moso Bamboo by Near Infrared Spectroscopy. *Sci. Rep.* **2015**, *5*, 17210.
44. Jia, H.; Wang, N. Targeted Genome Editing of Sweet Orange Using Cas9/sgRNA. *PLoS ONE* **2014**, *9*, e93806.
45. Jia, H.; Xu, J.; Orbović, V.; Zhang, Y.; Wang, N. Editing Citrus Genome via SaCas9/sgRNA System. *Front. Plant Sci.* **2017**, *8*, 2135.
46. Wang, Z.; Wang, S.; Li, D.; Zhang, Q.; Li, L.; Zhong, C.; Liu, Y.; Huang, H. Optimized paired-sgRNA/Cas9 cloning and expression cassette triggers high-efficiency multiplex genome editing in kiwifruit. *Plant Biotechnol. J.* **2018**, *16*, 1424–1433.
47. Breitler, J.-C.; Dechamp, E.; Campa, C.; Zebal Rodrigues, L.A.; Guyot, R.; Marraccini, P.; Etienne, H. CRISPR/Cas9-mediated efficient targeted mutagenesis has the potential to accelerate the domestication of *Coffea canephora*. *Plant Cell Tissue Organ Cult.* **2018**, *134*, 383–394.
48. Jiang, Y.; Guo, L.; Ma, X.; Zhao, X.; Jiao, B.; Li, C.; Luo, K. The WRKY transcription factors PtrWRKY18 and PtrWRKY35 promote Melampsora resistance in Populus. *Tree Physiol.* **2017**, *37*, 665–675.
49. Ma, X.; Zhang, Q.; Zhu, Q.; Liu, W.; Chen, Y.; Qiu, R.; Wang, B.; Yang, Z.; Li, H.; Lin, Y.; et al. A Robust CRISPR/Cas9 System for Convenient, High-Efficiency Multiplex Genome Editing in Monocot and Dicot Plants. *Mol. Plant* **2015**, *8*, 1274–1284.

50. Zhang, H.; Zhang, J.; Wei, P.; Zhang, B.; Gou, F.; Feng, Z.; Mao, Y.; Yang, L.; Zhang, H.; Xu, N.; et al. The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. *Plant Biotechnol. J.* **2014**, *12*, 797–807.
51. Bruegmann, T.; Deecke, K.; Fladung, M. Evaluating the Efficiency of gRNAs in CRISPR/Cas9 Mediated Genome Editing in Poplars. *IJMS* **2019**, *20*, 3623.
52. Piquemal, J.; Lapierre, C.; Myton, K.; O'connell, A.; Schuch, W.; Grima-pettenati, J.; Boudet, A.-M. Down-regulation of Cinnamoyl-CoA Reductase induces significant changes of lignin profiles in transgenic tobacco plants. *Plant J.* **1998**, *13*, 71–83.
53. Leplé, J.-C.; Dauwe, R.; Morreel, K.; Storme, V.; Lapierre, C.; Pollet, B.; Naumann, A.; Kang, K.-Y.; Kim, H.; Ruel, K.; et al. Downregulation of Cinnamoyl-Coenzyme A Reductase in Poplar: Multiple-Level Phenotyping Reveals Effects on Cell Wall Polymer Metabolism and Structure. *Plant Cell* **2007**, *19*, 3669–3691.
54. Haeussler, M.; Schönig, K.; Eckert, H.; Eschstruth, A.; Mianné, J.; Renaud, J.-B.; Schneider-Maunoury, S.; Shkumatava, A.; Teboul, L.; Kent, J.; et al. Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol.* **2016**, *17*, 148.
55. Ron, M.; Kajala, K.; Pauluzzi, G.; Wang, D.; Reynoso, M.A.; Zumstein, K.; Garcha, J.; Winte, S.; Masson, H.; Inagaki, S.; et al. Hairy Root Transformation Using *Agrobacterium rhizogenes* as a Tool for Exploring Cell Type-Specific Gene Expression and Function Using Tomato as a Model. *Plant Physiol.* **2014**, *166*, 455–469.
56. Sun, X.; Hu, Z.; Chen, R.; Jiang, Q.; Song, G.; Zhang, H.; Xi, Y. Targeted mutagenesis in soybean using the CRISPR-Cas9 system. *Sci. Rep.* **2015**, *5*, 10342.
57. Cai, Y.; Chen, L.; Liu, X.; Sun, S.; Wu, C.; Jiang, B.; Han, T.; Hou, W. CRISPR/Cas9-Mediated Genome Editing in Soybean Hairy Roots. *PLoS ONE* **2015**, *10*, e0136064.
58. Jacobs, T.B.; LaFayette, P.R.; Schmitz, R.J.; Parrott, W.A. Targeted genome modifications in soybean with CRISPR/Cas9. *BMC Biotechnol.* **2015**, *15*, 16.
59. Bernard, G.; Gagneul, D.; Alves Dos Santos, H.; Etienne, A.; Hilbert, J.-L.; Rambaud, C. Efficient Genome Editing Using CRISPR/Cas9 Technology in Chicory. *IJMS* **2019**, *20*, 1155.
60. Shan, Q.; Wang, Y.; Li, J.; Zhang, Y.; Chen, K.; Liang, Z.; Zhang, K.; Liu, J.; Xi, J.J.; Qiu, J.-L.; et al. Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat. Biotechnol.* **2013**, *31*, 686–688.
61. Chen, K.; Wang, Y.; Zhang, R.; Zhang, H.; Gao, C. CRISPR/Cas Genome Editing and Precision Plant Breeding in Agriculture. *Annu. Rev. Plant Biol.* **2019**, *70*, 667–697.
62. Belhaj, K.; Chaparro-Garcia, A.; Kamoun, S.; Nekrasov, V. Plant genome editing made easy: Targeted mutagenesis in model and crop plants using the CRISPR/Cas system. *Plant Methods* **2013**, *9*, 39.
63. Nekrasov, V.; Staskawicz, B.; Weigel, D.; Jones, J.D.G.; Kamoun, S. Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* **2013**, *31*, 691–693.
64. Weber, E.; Engler, C.; Gruetzner, R.; Werner, S.; Marillonnet, S. A Modular Cloning System for Standardized Assembly of Multigene Constructs. *PLoS ONE* **2011**, *6*, e16765.
65. Patron, N.J.; Orzaez, D.; Marillonnet, S.; Warzecha, H.; Matthewman, C.; Youles, M.; Raitskin, O.; Leveau, A.; Farré, G.; Rogers, C. Standards for plant synthetic biology: A common syntax for exchange of DNA parts. *New Phytol.* **2015**, *208*, 13–19.
66. Hsu, P.D.; Scott, D.A.; Weinstein, J.A.; Ran, F.A.; Konermann, S.; Agarwala, V.; Li, Y.; Fine, E.J.; Wu, X.; Shalem, O.; et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **2013**, *31*, 827–832.
67. Ployet, R.; Veneziano Labate, M.T.; Regiani Cataldi, T.; Christina, M.; Morel, M.; San Clemente, H.; Denis, M.; Favreau, B.; Tomazello Filho, M.; Laclau, J.; et al. A systems biology view of wood formation in *Eucalyptus grandis* trees submitted to different potassium and water regimes. *New Phytol.* **2019**, *223*, 766–782.
68. Lacombe, E.; Hawkins, S.; Van Doorselaere, J.; Piquemal, J.; Goffner, D.; Poeydomenge, O.; Boudet, A.-M.; Grima-Pettenati, J. Cinnamoyl CoA reductase, the first committed enzyme of the lignin branch biosynthetic pathway: Cloning, expression and phylogenetic relationships. *Plant J.* **1997**, *11*, 429–441.
69. Verbylaitė, R.; Beiðys, P.; Rimas, V.; Kuusienė, S. Comparison of Ten DNA Extraction Protocols from Wood of European Aspen (*Populus tremula* L.). *Balt. For.* **2010**, *16*, 8.
70. Beleites, C., Sergio V (2020). *HyperSpec: a package to handle hyperspectral data sets in R*. R package version 0.99-20200213.1, <https://github.com/cbeleites/hyperSpec>.

71. Barnes, R.J.; Dhanoa, M.S.; Lister, S.J. Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra. *Appl. Spectrosc.* **1989**, *43*, 772–777.
72. Becker, R.A.; Chambers, J.M.; Wilks, A.R. *The New S Language*; Wadsworth & Brooks; Chapman & Hall: London, UK, 1988.
73. Rohart, F.; Gautier, B.; Singh, A.; Lê Cao, K.-A. mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput. Biol.* **2017**, *13*, e1005752.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).





Chapter III:
Functional Characterization of Wood-Associated *EgrIAA9A*, *EgrIAA20*
and *EgrARF5* in Eucalyptus hairy roots

Chapter III:

Functional Characterization of Wood-Associated *EgrIAA9A*, *EgrIAA20* and *EgrARF5* in Eucalyptus hairy roots

1. Introduction

With the objective of identifying auxin signaling mediators involved in wood formation, a genome-wide identification of the *ARFs* and *Aux/IAAs* gene families was performed in *Eucalyptus grandis* (Yu et al., 2014, 2015). While the *E. grandis* has one of the largest proportion of tandem duplicated genes (34%; (Myburg et al., 2014)), these two families which contain 17 *ARFs* and 24 *Aux/IAAs* members, respectively, are not affected at all by tandem duplication. They contain much less genes than poplar (35 *ARFs*, 39 *Aux/IAAs*) and strikingly even less than *Arabidopsis* (23 *ARFs* and 29 *Aux/IAAs*). A combination of comparative phylogenetic analysis and large-scale gene expression profiling enabled the identification of several candidates. Interestingly, in both poplar and *Eucalyptus*, analyses of xylem gene networks revealed several *Aux/IAAs* as being tightly correlated with secondary wall TFs (Johnson et al., 2018, Ployet et al, 2018).

1.1 *EgrIAA* genes functional characterizations in wood formation

Previous studies from our team identified several *Aux/IAA* members preferentially expressed during wood formation (Paux et al., 2004) as well as some members deregulated during tension wood formation (Paux et al., 2005). The genome-wide analysis of *Aux/IAA* family in *E. grandis* identified members preferentially expressed in xylem and cambium tissues (Yu et al., 2015). To functionally characterize these candidate genes, they were first stabilized by mutations and then overexpressed in *Arabidopsis* under the control of a constitutive promoter. The overexpression of a stabilized version of *Eucalyptus IAA4* (*EgrIAA4m*, mutated at domain II for stabilized version) driven by 35S CaMV promoter in *Arabidopsis* led to dramatically reduced secondary xylem formation (Yu et al., 2015). The formation of xylary fibres and vessels in vascular bundles and fiber cells in the interfascicular region were inhibited. In addition, primary roots elongation and lateral roots emergence were impaired (Yu et al., 2015). The transgenic lines overexpressing two other *Eucalyptus Aux/IAA* members (*EgrIAA9Am*, *EgrIAA20*) exhibited obvious phenotypes of wood cells (unpublished data): both OE_*EgrIAA9Am* and OE_*EgrIAA20* lines showed significant reduction of radial growth presenting thinner and floppy stems. Histochemical analysis demonstrated that the formation of secondary fiber cells was either completely abolished or dramatically inhibited in the transgenic lines whereas it was not the case for primary fiber cells in both vascular bundles and interfascicular regions (Figure III-1A & B, Figure III-S1). The development of vessels was promoted in OE_*EgrIAA9Am* lines (Figure III-1A) but not in OE_*EgrIAA20* lines (Figure III-1B).

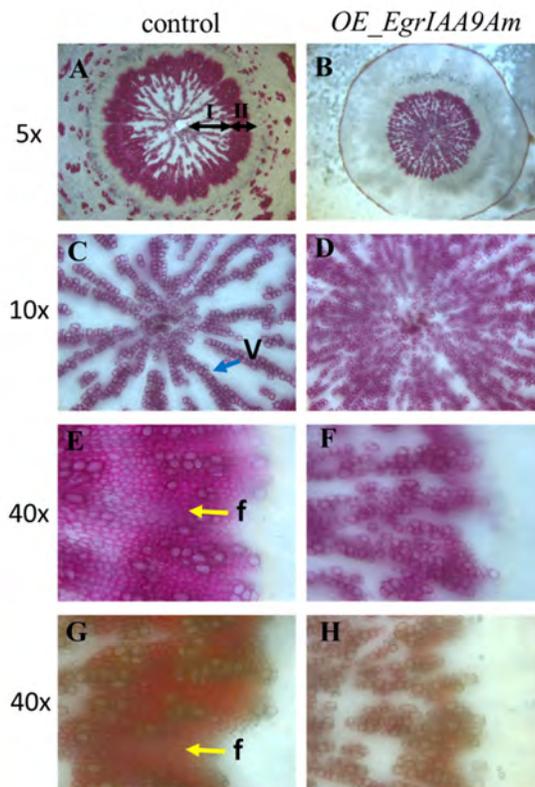


Figure III-1A. Histochemical analysis of cross sections of *OE_EgrIAA9Am* transgenic and wild-type *Arabidopsis* hypocotyls (from internship report of Liu Mingjun (2015)). Sections from A-F were stained by phloroglucinol-HCl; sections G & H were stained by Mäule coloration (fiber cells stained in red, and vessel cells in brown); left panel (A, C, E and G): wild-type plants; right panel (B, D, F and H): *OE_EgrIAA9Am* plants. *OE_EgrIAA9Am* lines showed more vessels in phase I growth (compare B and D to A and C, respectively). During phase II growth, no red stained secondary fiber cells were observed in *OE_EgrIAA9Am* lines in contrast to intensive red stained fiber cells in wild-type (H compared to G), indicating fiber cells formation was inhibited during secondary growth in *OE_EgrIAA9Am*. F: fiber cell, v: vessel cell. I: phase I growth of hypocotyl (before flowering), II: phase II growth of hypocotyl (after flowering).

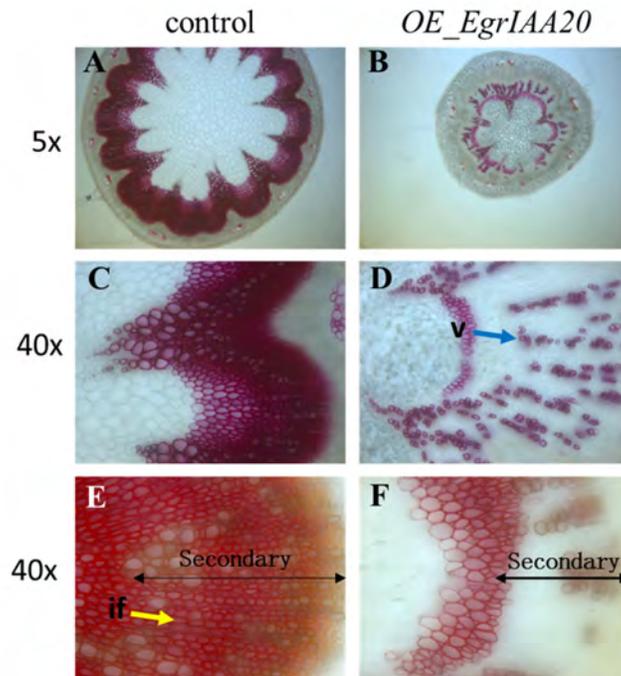


Figure III-1B. Histochemical analysis of cross sections of *OE_EgrIAA20* transgenic lines and wild-type *Arabidopsis* inflorescence stems. Sections from A to D were stained by phloroglucinol-HCl; sections E & F were stained by Maüle reagent that stains in red fiber cells rich in S lignin and vessel cells in brown (rich in lignin units G). Left panel is for wild-type plants (A, C and E) and right panel is for *OE_EgrIAA20* transgenic plants (B, D and F) were. Focus on secondary growth region in E and F. v: vessels, if: interfascicular fibers.

1.2 *ARF5* (*MONOPTEROS*) functional characterizations related to vascular tissues

In contrast to the *Aux/IAA* family where several members were shown to regulate wood formation, only one member of the *ARF* gene family, the *ARF5* (*MONOPTEROS*, *MP*) gene was shown to be directly involved in wood formation. In *mp* *Arabidopsis* mutants, reduced xylem and phloem strands were found in vascular bundles of inflorescence axes, and less secondary veins were present in rosette leaves (Przemeck et al., 1996). Both *mp* and *pin* mutants had defects in the leaves vascular patterning, which is regulated probably by auxin-MP-PIN feedback loop module (Wenzel et al., 2007). The *ARF5/MP* activates downstream *DOF5.8*, which is expressed in pro-vascular tissue in leaf primordia. The *DOF5.8* mutations enhanced *mp* defected phenotypes of root formation and vein patterning (Konishi et al., 2015). Mutations in different positions of *ARF5/MP* resulted in degrees of vascular defects. If disruption occurred in activating domain, vascular strands in cotyledons of seedlings were ramified with even a lack of mid-vein. The reduction of the vascular system was more severe when a stop codon was found close to N' terminal (Hardtke, 2004). The *BDL*(*IAA12/BODENLOS*)-*MP* module controls provascular specification and patterning during embryogenesis (Hamann, 2002). The *bdl* is a gain-of-function mutation of *IAA12*, thus *BDL* constitutively inhibit *MP* activity, mimicking the loss-of-function of *mp* mutation. *IAA20/30-ARF5/MP* module was found to affect root development and root vascular patterning by forming a loop with HD-ZIPIII (Müller et al., 2016). *ARF3*, *ARF4*, *ARF5* were identified as cambium regulators, especially *ARF5* plays roles in cambium stem cells, and *ARF5-WOX4* module (*ARF5* attenuates *WOX4*

expression) is crucial in auxin signaling especially for cambium regulation (Brackmann et al., 2018).

Genome-wide analyses of *ARF* families in *Eucalyptus* and Poplar identified members preferentially expressed in xylem and cambium tissues (Kalluri et al., 2007; Yu et al., 2014). In *Populus* there is two potential orthologs of *MP* named *PoptrMP1* and *PoptrMP2*, *PoptrMP1* is specially expressed in developing secondary xylem. Furthermore, the MP-HB8 module known to regulate vascular development in *Arabidopsis* (Mattsson et al., 2003) *AtHB8* being a target of *MP* is conserved in *Populus* (Johnson and Douglas, 2007). Constitutive expression of truncated *PtoARF5.1* without C terminal protein binding domain preventing binding to Aux/IAA inhibitors in OE_ *PtoIAA9m* poplar lines let to partial recovery of lateral growth and showed increased numbers and layers of xylem cells in stem (Xu et al., 2019). Expression profiles of *EgrARFs* provided potential candidates involved in auxin regulation in wood formation, such as *EgrARF10*, *EgrARF19* which were preferentially expressed in vascular cambium (Yu et al., 2014), as well as their orthologs in *Arabidopsis* *AtARF10* and *AtARF19* were also expressed in vascular tissue (Liu et al., 2007; Wilmoth et al., 2005). Overexpressing *EgrARF10m* (stabilized version) in *Arabidopsis* generated some transgenic plants showing severe defects in leaf development with needle-like leaves, whereas no obvious phenotype was found in dominant repression of *EgrARF19A* in *Arabidopsis* (Yu, 2014).

1.3 Objective of Chapter III

Previous work from our team used overexpression in *Arabidopsis* as a strategy to functionally characterize *EgrAux/IAA* and *EgrARF* candidate genes. Here, we took the opportunity the *E. grandis* hairy roots stable transformation system and of the powerful CRISPR/Cas9 gene editing technology that I implemented during my PhD, to generate both loss-of-function and gain-of-function in a homologous system (The strategy is shown in Figure III-S2). Besides this first objective of making an in-depth functional characterization of three candidates [*EgrIAA9A*, *EgrIAA20*, and *EgrARF5*], we decided to seek for protein partners of *EgrIAA9A*, *EgrIAA20* by screening of a eucalypts xylem Yeast two hybrid library (Soler et al 2017).

We generated putative knockout (KO) mutants of *EgrIAA9A*, *EgrIAA20*, and *EgrARF5* using CRISPR/Cas9 genome edition tools in *Eucalyptus* hairy roots. We also generated gain-of-function by overexpressing (OE) *EgrIAA9A*, *EgrIAA20* under the control of the 35S CaMV promoter in *Eucalyptus* transgenic hairy roots. Although the OE constructs of *EgrIAA9A*, *EgrIAA20* and hairy root transformation were successfully obtained using the golden gate cloning strategy, unfortunately the newly generated *E.grandis* hairy roots (413 Plants) all died during the Covid-19 lockdown period.

2. Material and Methods

2.1 Plant material and culture conditions

The *E.grandis* plant material, *in vitro* culture and hydroponic culture conditions were the same as described in Chapter II. After A4RS *Agrobacterium Rhizogenes*-mediated transformation, 70-day-old *ARF5*_lines and *IAA20*_lines *in vitro* composite plants with DsRED fluorescence were transferred in hydroponic culture. Before harvesting, we verified DsRED fluorescence again to discard the non transformed roots exhibiting no fluorescence. Finally, the 213-day-old transformed

hairy roots were harvested for genotyping (delayed harvesting due to the lockdown period of Covid-19), FT-IR and histology analyses (Table III-S1A). Of these, FT-IR and histology analyses were already described in Chapter II.

2.2 CRISPR/Cas9 constructions to generate KO mutants of *EgrIAA9A*, *EgrIAA20* and *EgrARF5*

The sgRNAs design and CRISPR/Cas9-mediated construction of *EgrIAA9A* using Golden Gate cloning were described in Chapter II ‘Material and Methods’ part, and the transgenic roots phenotypes generated by CRISPR/Cas9 will be introduced in this Chapter (*IAA9A*_lines information shown in Table III-S1B). Similarly, the selected sgRNAs for *EgrIAA20* and *EgrARF5* designed by CRISPOR online, the predicted off-target risk and the editing efficiency are represented in Table III-S2. For *EgrIAA20*, we chose two sgRNAs (sgRNA1_IAA20 (5’-GAGACGCCGAGACATGGACG-3’) and sgRNA2 (5’-GCCATCTTGGTGAGAAGCCG-3’)) separated by 46bp and located on the antisense strand in exon 1. For *EgrARF5*, we chose two sgRNAs (sgRNA1_ARF5 (5’-GTATTTCCAGTACCTGACTT-3’) located on the sense strand and sgRNA2 (5’-GTCTTGCAAAAGAAGTCACT-3’) located on antisense strand) separated by 52bp in exon 3 (short exon 1 (30bp) and exon 2 (57bp)) (Figure III-2). CRISPR/Cas9 constructs assembly of *EgrIAA20* and *EgrARF5* were carried out as described in Chapter II. The pipeline of two levels of Golden Gate Cloning and the restriction enzymes are shown in Figure III-3, and the primers used are shown in Table III-S3A.

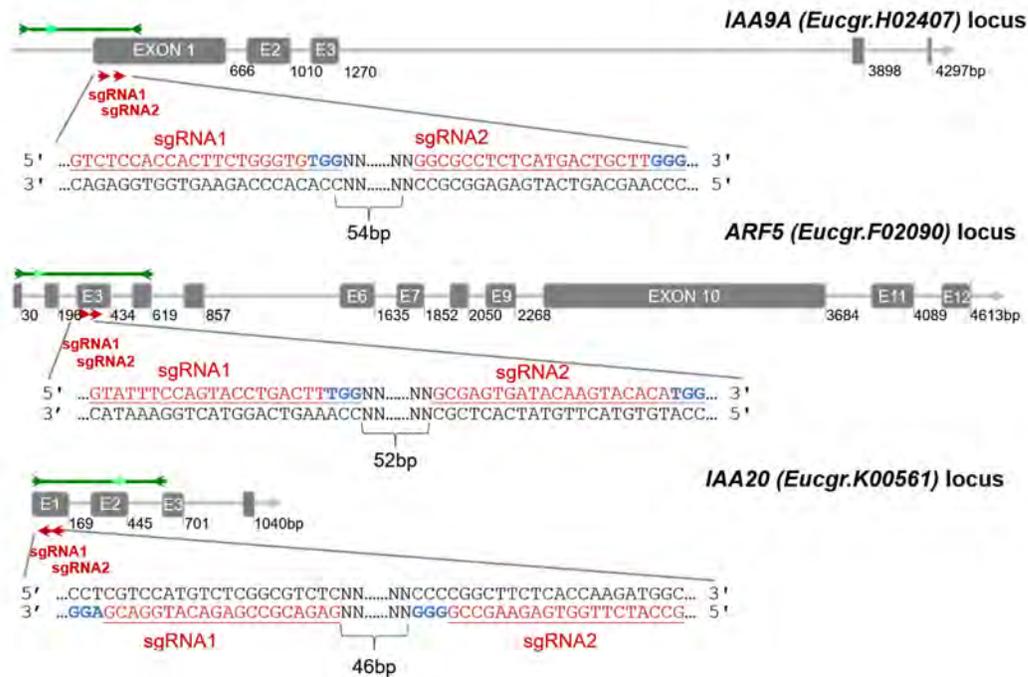


Figure III-2. CRISPR/Cas9 sgRNA design and mutation detection in *IAA9A*, *ARF5* and *IAA20*. Schematic representation of the target sites and the PCR assay for Sanger sequencing. Exons and introns are represented by grey boxes and grey lines, respectively. Red arrows represented target sites for each CRISPR/Cas9 nuclease, and target sequences are indicated with underlined red characters, while blue characters are PAM. The dark green arrows indicate the location of the primers for PCR amplification, the light green arrows indicate the nested primers designed for DSDecode mutation identification.

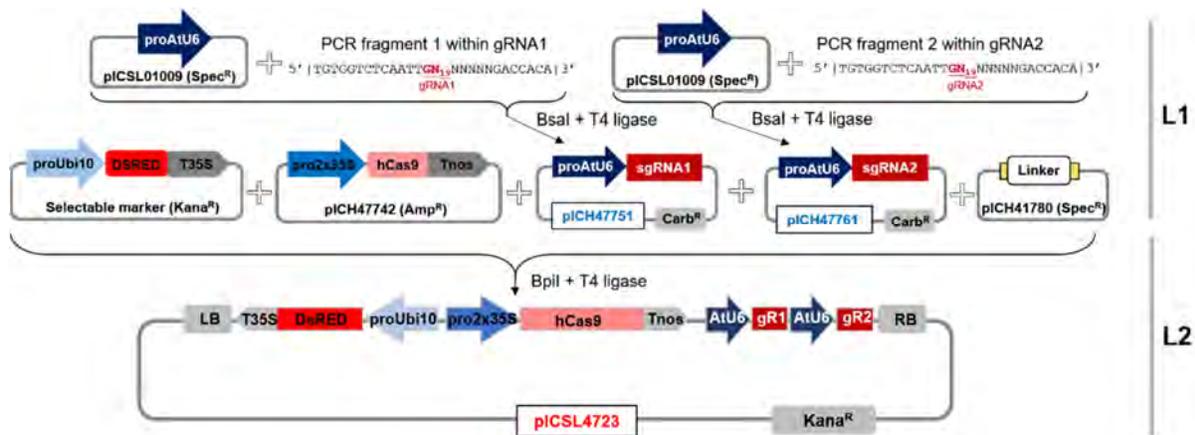


Figure III-3. Scheme illustrating the assembling of Cas9/sgRNA expressing constructs using Golden Gate Cloning. Promoters are indicated in blue arrows, and terminators are in gray arrows. In level 1, sgRNA is placed under the *Arabidopsis* U6 promoter, and two sgRNAs are separately cloned into L1 Golden Gate vector. DsRED cassette is a selectable marker. The hCas9 is human codon-optimized Cas9 gene of *Streptococcus pyogenes*. proUbi10, *Arabidopsis* Ubiquitin 10 promoter; pro2x35S, double *Cauliflower mosaic virus* CaMV 35S promoter. All L1 intermediary vectors were assembled into L2 acceptor (ratio DNA amount is module: acceptor=1:2). The final binary vector was targeting two loci simultaneously. *Kana^R*, kanamycin resistance gene sequence; *Carb^R*, carbenicillin resistance gene sequence. LB, left T-DNA border; RB, right T-DNA border.

2.3 Overexpression constructions of *EgrIAA9A* and *EgrIAA20*

The CDS sequences (from Phytozome *E. grandis* genome v12.1) of *EgrIAA9A* (*Eucgr.H02407*) and *EgrIAA20* (*Eucgr.K00561*) genes were amplified by PCR using High-Fidelity Phusion DNA Polymerase (Thermo Scientific), cDNA strand (reverse transcribed from total RNA) as template, and the pairs of gene-specific primers (Table III-S3B). In order to obtain gain-of-function mutant of *EgrIAA9A*, we generated a more stabilized version of Aux/IAA protein named *EgrIAA9Am*, by introducing a mutant in degron motif 'VGWPP' of domain II which prevent auxin-mediated Aux/IAA protein degradation. The *EgrIAA9Am* fragment was cloned into pENTRY-TOPO vector (Invitrogen, France). Since *EgrIAA20* does not contain domain II there was no need for stabilization.

The risks linked when generating overexpressing cassettes by Golden Gate cloning system is the presence of Type II restriction enzyme sites (BsaI, BpiI/BbsI, BsmBI/Esp3I) in the CDS of the gene. This can be solved by domestication (Golden Braid 4.0 <https://gbccloning.upv.es/do/domestication/>) based on codon degeneracy. There are one BsaI site in the CDS of *EgrIAA9Am*, two BsmBI sites and one BpiI/BbsI site in *EgrIAA20*, and in the three steps of Golden Gate cloning, BsmBI, BsaI and BpiI are required for level 0, level 1 and level 2 assembly, respectively. We first made virtual assembly in Benchling (<https://benchling.com/>). As for *EgrIAA9Am*, we amplified two PCR fragments splitting at BsaI site, we domesticated two fragments of coding sequence by performing overlapping PCR with specific pairs of oligos (Table III-S3B) to obtain PCR product (named *EgrIAA9Amm*) presenting two BsmBI sites added overhang at 5' and 3' terminals, and one BsaI site loss (GGTCTC was mutated into GGTCTG without changing the protein sequence). *EgrIAA9Amm* was prepared to be cloned into the pUPD2 vector as Level 0 vector, which was assembled into the Level 1 pICH47742 acceptor (with lacZ, blue/white colonies selection), together with CaMV 35S promoter (pICH51266) and Nos terminator (pICH41421). DsRED expression cassette (AtUbi10p::DsRed::T35S-terminator) at position 1, Level 1 intermediary vector at position 2, and corresponding Linker (pICH41744) were assembled in Level 2 into the pICSL4723 destination vector (LB-DsRed-CaMV 35S-CDS *EgrIAA9Amm*-Nos Terminator-RB), which included all components. The final binary vector was used for *Agrobacterium* mediated transformation (Figure III-4).

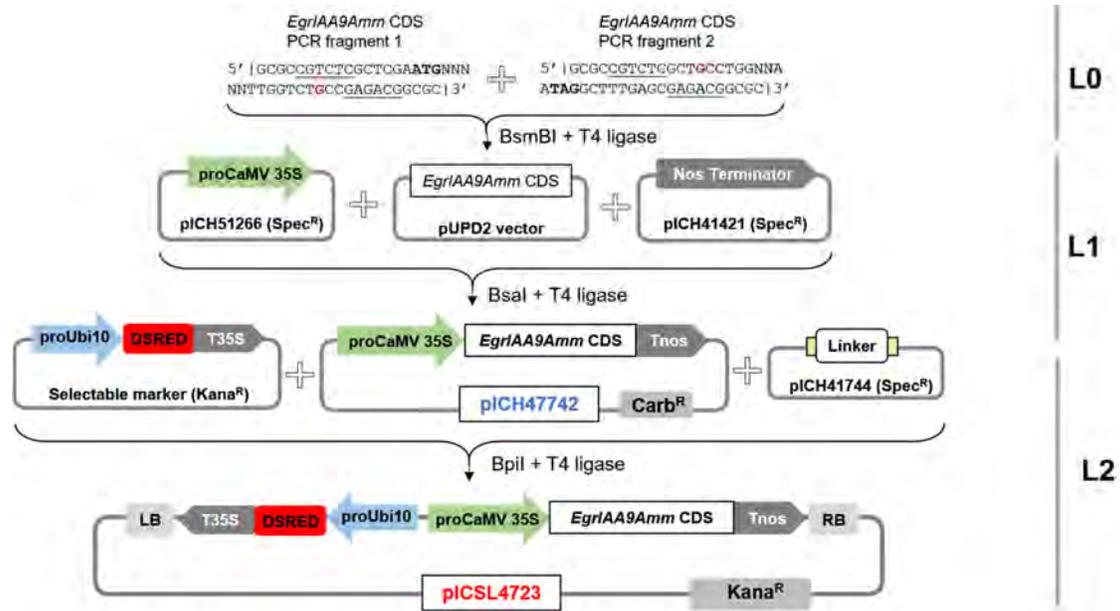


Figure III-4. Scheme illustrating the assembling of overexpression constructs with CDS sequences of *EgrIAA9Amm* using Golden Gate Cloning (including three levels). Two mutated CDS PCR fragments of *EgrIAA9Amm* were cloned into pUPD2 vector in Level 0 (L0). L1 intermediary vector was under the Cauliflower mosaic virus CaMV 35S promoter (green arrow) in Level 1 (L1). DsRED (acting as a selectable marker) cassette (driven by *Arabidopsis* Ubiquitin 10 gene promoter) and intermediary L1 vector were assembled together in Level 2 (L2) Golden Gate vector. In two PCR fragments, the red characters are mutated BsaI restriction enzyme site, the underlined characters are BsmBI enzyme sites added, and the bold characters are start/stop codons. Kana^R, kanamycin resistance gene sequence; Carb^R, carbenicillin resistance gene sequence. LB, left T-DNA border; RB, right T-DNA border.

As for *EgrIAA20*, due to two BsmBI sites which also exist in pUPD2 in Level 0 assembly, we domesticated CDS to lose the BpiI site (GAAGAC was mutated to GAAGAT) using overlapping PCR and directly cloned the PCR product (renamed as *EgrIAA20m*) into the Level 1 acceptor, and then assembled in Level 2 to generate the binary vector similar to what was described for *EgrIAA9Amm* (Figure III-5).

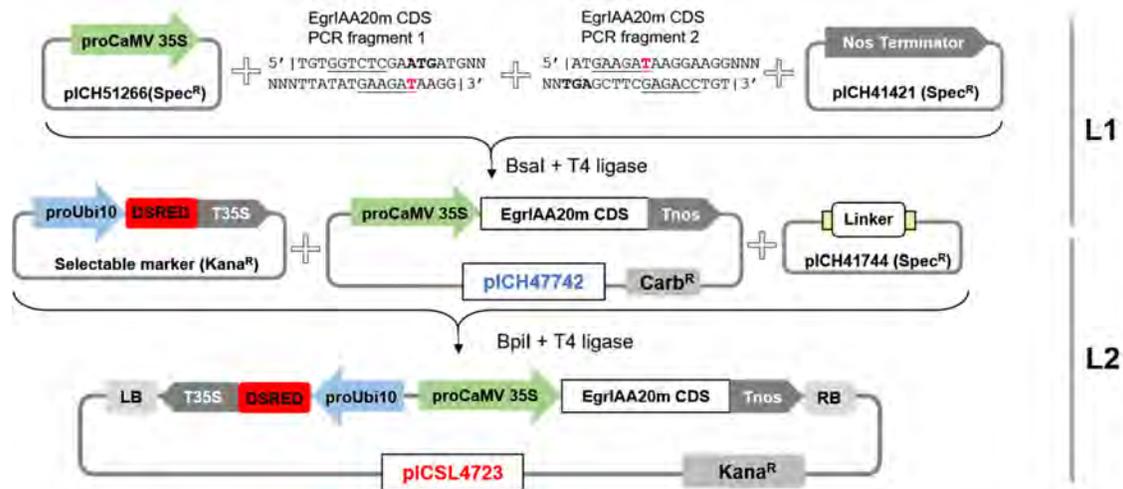


Figure III-5. Scheme illustrating the assembling of overexpression constructs with CDS sequences of *EgrIAA20m* using Golden Gate Cloning (including two levels). Two mutated CDS PCR fragments of *EgrIAA20m* were cloned directly into L1 acceptor, under the driven of Cauliflower mosaic virus CaMV 35S promoter (green arrow) and stopped by Nos terminator (Tnos in gray) in Level 1 (L1). DsRED (a selectable marker) cassette (driven by *Arabidopsis* Ubiquitin 10 gene promoter) and intermediary L1 vector were assembled together in Level 2 (L2) Golden Gate vector. The red characters are mutated BpiI restriction enzyme site, and the underlined sequences are BsaI enzyme sites added. Kana^R, kanamycin resistance gene sequence; Carb^R, carbenicillin resistance gene sequence. LB, left T-DNA border; RB, right T-DNA border.

2.4 CRISPR/Cas9 editing events detection

DNA Isolation and PCR Amplification were carried out as described in Chapter II 'Material and Methods' part. The PCR amplicons comprising the two sgRNAs sites were amplified by High-Fidelity Phusion Polymerase (Thermo Fisher, Illkirch-Graffenstaden, France) using genomic DNA as templates and primers flanking the two target sites. For *ARF5*, the forward primer is 265bp upstream of the sgRNA1, and the reverse primer 235bp downstream of the sgRNA2 (*ARF5_edit check_NCBI_F* and *ARF5_edit check_NCBI_R*, amplicon size was 635bp, Table III-S3A); For *IAA20*, the forward and reverse primers are 35bp upstream of the sgRNA1 and 441bp downstream sgRNA2, respectively (*IAA20_edit check_NCBI_F* and *IAA20_edit check_NCBI_R*, amplicon size was 602bp, Table III-S3B).

In addition of the subcloning method and DSDcode software mentioned in Chapter II, we used another online software ICE (Inference of CRISPR Edits) Synthego (<https://ice.synthego.com/#/>) (Roginsky, 2018) for mutations detection. ICE analysis generates CRISPR editing results with an accuracy highly comparable to that of Next Generation Sequencing (NGS) data. ICE tool cannot distinguish less than 5% between inherent noise of Sanger Sequencing and samples' signal, which depends on the qualities of PCR and sequencing results. When using multiplex gRNAs (no more than three for ICE) in CRISPR/Cas9 knockout, ICE manages to analyze the deletions up to 150bp. For genome editing detection using DSDcode, the PCR amplicons were directly sequenced using nested primer *ARF5_F* (located 133bp upstream of the sgRNA1) or *IAA20_R* (located 174bp downstream of sgRNA2) to avoid the noise signals of sequencing results.

2.5 FT-IR and histology analyses of CRISPR/Cas9 mediated *IAA9A*, *IAA20* and *ARF5* lines

As for FT-IR analysis, the harvested hairy roots from hydroponic culture were stored in -80°C and further freeze-dried in 48-72h with Freeze Dryer machine (Alpha 1-4 LSCbasic, Martin Christ, France) and ground around 1.5min with Mixer Mill MM 400 (Retsch). The FT-IR and histology analyses were carried out as described in Chapter II ‘Material and Methods’.

2.6 Library screening and protein-protein interaction analysis by yeast two-hybrid (Y2H) system

The haploid yeast strains (Y187 and Y2HGold) are able to mate with each other naturally to form a diploid cell, and this property is to detect protein-protein interaction. The Yeast two hybrid library was constructed starting with mixed total RNAs isolated from developing xylem samples (Soler et al, 2017). ‘Make Your Own Mate & Plate™ Library System’ was then used to generate corresponding cDNA pool, which was cloned into pGADT7 vector and transformed into Y187 yeast competent cells, with the proper final density ($>1 \times 10^7$ cells) (Soler et al., 2017). *EgrIAA9A* and *EgrIAA20* CDS were cloned in frame into pGBKT7 vector, and then transferred into Y2HGold competent cells, to be used as a bait to screen the eucalypts xylem Yeast two hybrid cDNA library. The mating protocol of Y2H library screening was carried out as described in Matchmaker® Gold Yeast Two-Hybrid System User Manual). The three-lobed structure zygotes were observed under microscope. The mating colonies were grown on QDO (Quadruple dropout medium: SD/-Ade/-His/-Leu/-Trp) plates. PCR on colonies and on plasmids were performed using primers located on pGADT7 vector. The purified PCR products were then sequenced and the results were blasted to identify the partners of *EgrIAA9A* and *EgrIAA20*, respectively.

For the targeted yeast two hybrid, the cds of *EgrIAA9A* or *EgrIAA20* were each cloned in frame into the BD-bait vector pGBKT7 containing the Trp gene. The candidates obtained by the screening of the library were cloned into the AD-prey vector pGADT7 vector containing the Leu gene. Once the two constructions were co-transformed into Y2HGold yeast strain, the transformants harboring the two constructs will grow on a media without tryptophane and leucine. If there is an interaction between the two proteins (prey and bait), the yeast will grow on media lacking Trp, Leu, His and Ade. All processes related to Yeast two hybrid followed the instructions of the (Matchmaker® Gold Yeast Two-Hybrid System User Manual).

3. Results

After screening and selection of fluorescent roots, genome editing events were detected using PCR amplicon direct sequencing. In order to rapidly see if the chemotypes of the edited roots differed from control plants, a combination of FTIR spectroscopy and multivariate analyses (PCA/PLSDA) analyses were carried out. Finally, histochemical analyses and microscopic observations were performed focusing on the xylem architecture of transgenic roots and controls. The details of all transgenic plants were shown in Table III-S1.

3.1 Mutations detection using web-based tools (ICE, DSDecode)

At the beginning, we used the web-based tool “DSDecode” to detect the genome editing events, but we found that it was

not able to detect chimeric mutations or multiple mutations and thus we adopted the sub-cloning and sequencing strategy which is reliable but costly and time consuming. Here, we tried another web-based analysis tool “ICE” which was reported to give accurate results highly comparable (with $R^2=0.96$ or better) to that of NGS sequencing data (Roginsky, 2018). We first compared the results obtained by the two web-based editing detection methods. In general, the simple mutations such as homozygous biallelic mutations and monoallelic mutations were easily detected by both DSDecode and ICE, but for complex mutations such as chimeric mutations, DSDecode failed to detect them just mentioning “complicated variant” as output. ICE gives the occurrence percentage of each type of mutation. Even when both methods provide successful results they very often present some discrepancies. ICE seemed to give more accurate results, so we decided to use preferentially ICE as mutation detection tool, and to include the sub-cloning method whenever needed. In practice for the ICE method, the targeted mutations regions including the two sgRNA sequences were amplified by PCR, then the PCR amplicons were directly sequenced by Sanger method using nested primers; the sequencing results were proceeded by the web-based tool ICE, the putative inferred mutations types and occurrence percentage were reported as editing detection results. In total, for the *ARF5_* and *IAA20_CRISPR* lines, using ICE we detected editing events in 11 *ARF5_CRISPR* lines and one *IAA20_CRISPR*. Seven of these (one *IAA20_* line and six *ARF5_* lines) had over 50% mutation percentage (ICE parameters were shown in Table III-S4). Here we report the mutations detection results in three classes: (a) lines with chimeric mutations which can be decoded only by ICE (not by DSDecode); (b), lines with up to three types of mutations (monoallelic mutations) in which the edition detection can be obtained from both DSDecode and ICE showing the same results in most of the cases; (c), lines with low occurrence of editing event (< 50% editing rate).

As for *IAA9A_* edited lines, the comparison between DSDecode and subcloning was performed as described in Chapter II, providing three classes of CRISPR/Cas9-generated *IAA9A* lines, (a) same (fully or partially) edition types detected by the two methods; (b) totally different edition types using two methods (c) edited lines that failed to be analyzed by DSDecode. We further proceeded the data with ICE and compared the three methods for detecting mutations (ICE, DSDecode and sub-cloning) on the same set of editing plants.

We did PCR amplifications on all CRISPR/Cas9 generated transgenic hairy roots and all the amplicons were directly sequenced by Sanger Sequencing. Three lines *ARF5_29*, *ARF5_44* and *IAA20_4* presented two PCR bands of different sizes compared with that of control (Figure III-6A), suggesting the presence of large deletions in these lines. The prevalent edition types of *ARF5_* lines and *IAA20_* lines are showed in Figure III-6B. Among the various mutation types, the prevalent edition types detected were deletions as compared to insertions and substitutions. For example, one bp deletion at sgRNA1 of *ARF5* (occurrence as high as 62%) led to a reading frame shift with a premature stop codon near the N-terminal and thus to a severely truncated protein. A large mutation type [78bp deletion in *ARF5_44*] is predicted to generate a 26 amino acid deletion. In *IAA20_4*, a 66bp deletion (55%) is the major edition type leading to a 22 amino acid deletion (as illustrated in Figure III-6).

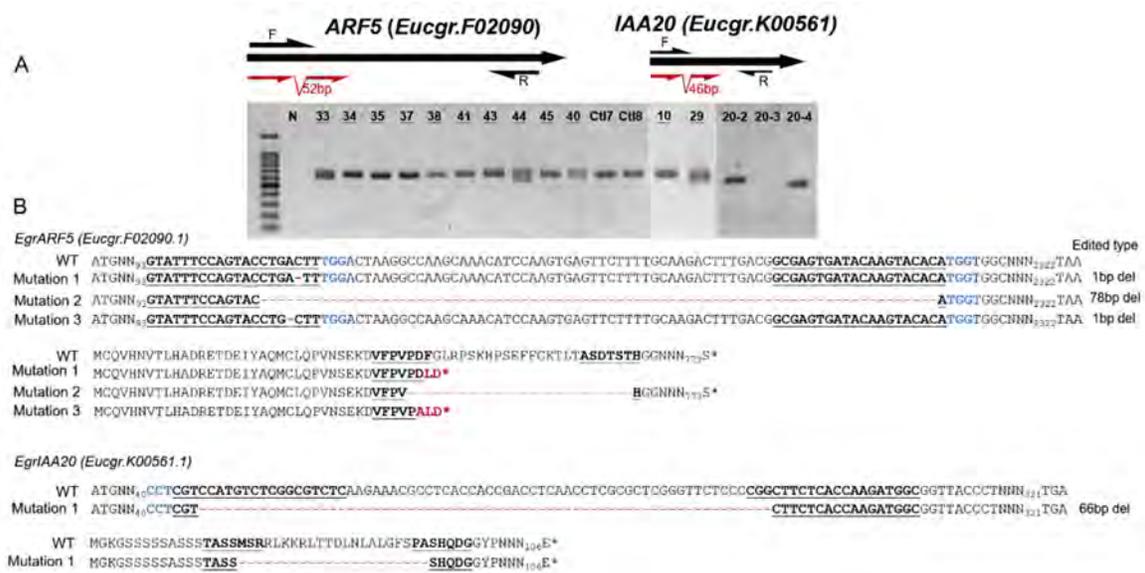


Figure III-6. Genotyping of *ARF5* and *IAA20* edited lines and their corresponding prevalent edition types and related amino acids changes detected by ICE. (A) Electrophoresis gels showing PCR amplicons for *ARF5* (left panel) and *IAA20_4* (right panel) transformants. The positions of the guide RNAs and the primers used for PCR amplification are located in the schematized ORF (open reading frame) sequences of genes. The symbols above each lane indicate the transformant lines. Ctrl7, control_7. N, negative control. (B) Prevalent mutation types for *ARF5* edited lines and *IAA20_4*, respectively. The top alignments showed DNA sequence mutations compared with controls (read dashes present deleted base pairs) and the consequences on the proteins are shown below (altered or deleted amino acids were in red). The sequences of sgRNA1 and sgRNA2 and corresponding amino acids are bold and underlined, and the protospacer adjacent motif (PAM) sequences are indicated in blue. The edited types detected by ICE are shown on the right part of the figure.

Taking *ARF5_10* as an example, ICE analysis revealed chimeric mutations (Figure III-7A). The output results listed all different mutation types detected with the corresponding occurrence percentage. The knockout-score was reported as high as 46, indicating that 46% of the mutations were presumed knock-out lines (frameshift or more than 21bp Indel according to the user's instruction). In addition, there is also 3bp insertions (4%) and of 6bp insertion (3%) without frameshift, reaching 53% of total editing. Among these mutations, 41% of indels occurred at sgRNA1 position, and 32% of mutations occurred in sgRNA2 with more frequent insertions. Twenty percent of the indels can be found at both sgRNA1 and sgRNA2. The major edition type in *ARF5_10* was two bp deletion (31% out of 53%) as shown in distribution of Indels (Figure III-7B). The difference between control and edited samples usually starts around the cut site generated by CRISPR/Cas9. However, for *ARF5_10*, the predicted cut site position located upstream of sgRNA1 was decoded by DSDecode and was not as the main cleavage site in the target region (3bp after NGG PAM). Unfortunately, we obtained only three *IAA20* transgenic roots. Line *IAA20_3* failed to be amplified by PCR even after several attempts (changed primers and new DNA extraction) (Figure III-6A).



Figure III-7. Different edition types (A) and Indel distributions (B) of *ARF5_10* analyzed by ICE Synthego. (A) The “Contribution” tab shows the inferred sequences and the relative representation within the edited population. The black vertical dotted line represents the cut sites, located 3bp upstream of sgRNAs (forward) distal end. The mutation types and positions are marked on the left. (B) The “Indel Distribution” tab shows an Indel plot indicating the percentages of the different indel sizes in summary in the whole edited population. On each blue bar are given details of indel size, along with the related percentage. The discordance plot highlights the level of disagreement between control (orange) and edited line (green).

a. Mutations detection was only achieved by ICE for four *ARF5* CRISPR-transgenic lines and one *IAA20* line

Four ARF lines (*ARF5_17*, *ARF5_29*, *ARF5_33*, *ARF5_40*) (Figure III-S3) and one *IAA20* line could not be analyzed by DSDecode, which only indicated ‘complicated variant’ as output. Both *ARF5_40* and *IAA20_4* had expected large deletions between two sgRNAs which contributed to 7% and 45% of the PCR amplicons, respectively. For line *ARF5_40*, mutations at sgRNA1 position were found in all detected mutation sequences, except for 4% still having a wild type sequence. Obviously, 2bp deletion at sgRNA1 position was the prevalent mutation type in *ARF5_40* (58%). There was no insertion found in ICE analysis of *ARF5_40*, showing all edited types were deletions in this line (Figure III-8A). *ARF5_33* had most edited Indel types among all detected *ARF5* edited lines, meaning its high level of chimera. Small size fragments (1-12bp) were simultaneously deleted at both sgRNA1 and sgRNA2 positions, which was the prevalent mutation type (20% out of 42%) in this line. The large deletion (23bp) at sgRNA1 contributed only to 1% (Figure III-S3).

For two *IAA20* edited lines (*IAA20_2* and *IAA20_4*) PCR amplicons were analyzed by ICE but also directly sequenced by Sanger Sequencing. ICE detected mutations in *IAA20_4* (Figure III-8B), but not in *IAA20_2*. In line *IAA20_4*, the main mutation types were deletions ranging from 12bp to 65bp. The large deletion between two sgRNAs contributed for 55%. Sixty-five % of the mutations were found at sgRNA2 position whereas only 4% were detected at sgRNA1 position. Large deletions (≥ 15 bp) represented 92.5% of all mutated sequences, explaining the smaller PCR band I compared to control (Figure III-6A). We performed the subcloning in order to further validate the ICE result, A-tailed PCR amplicons were cloned into pGEM-T vector and five *IAA20_2* clones and two *IAA20_4* clones were sequenced. Large deletion was detected in *IAA20_4* which was in accordance with the ICE results, and no mutation was detected in *IAA20_2* which also confirmed the result obtained from ICE.

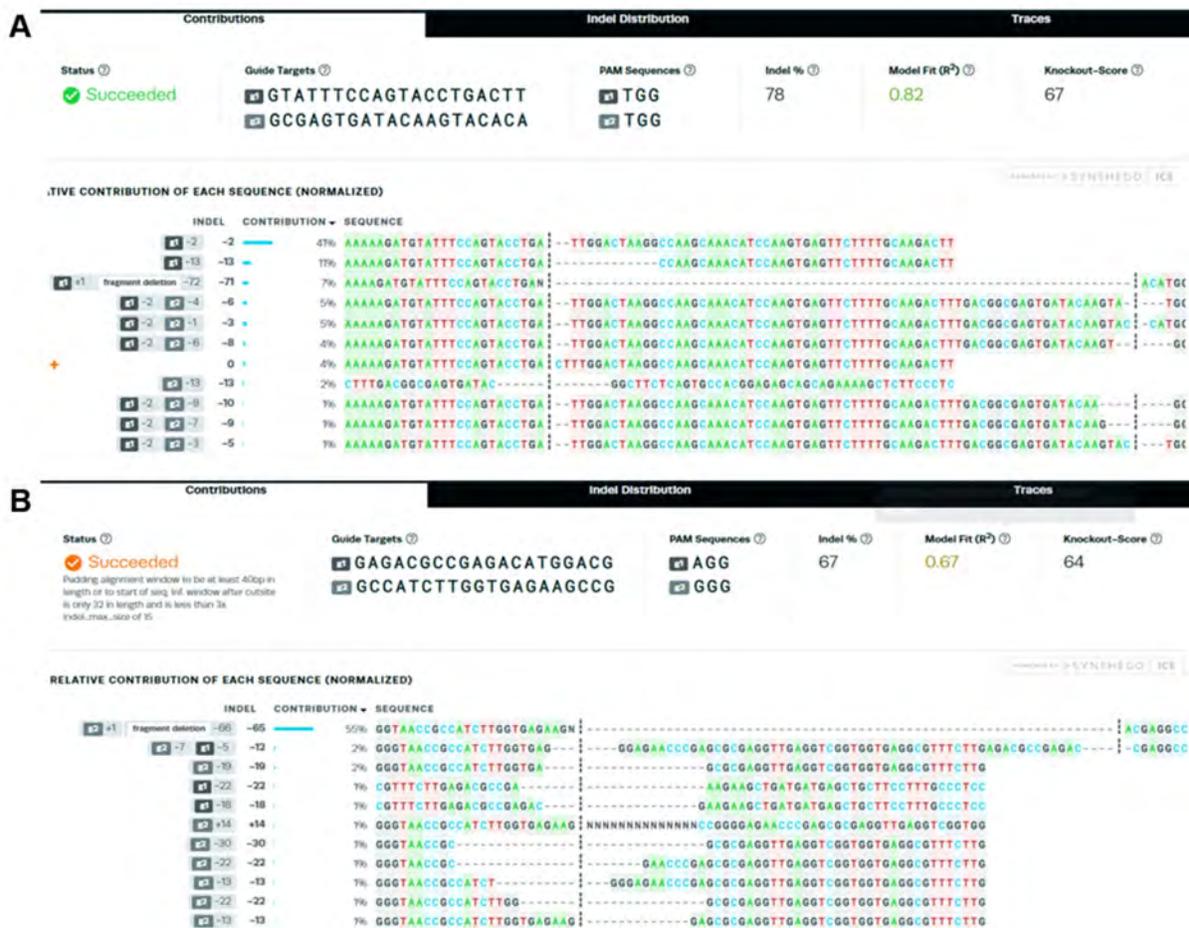


Figure III-8. Different inferred edition types of *ARF5_40* (A) and *IAA20_4* (B) obtained by ICE Synthego. Same as *ARF5_10* above, both two edited lines were chimeric mutations. The symbol “+” on the left in *ARF5_40* represents wild type allele. The insertion (14bp) at sgRNA2 is shown in *IAA20_4*.

b. Comparison of web-based tools taking results of two lines *ARF5_43*, *ARF5_44* as examples

The results of edition detection were in many cases in agreement between the two online tools (ICE vs. DSDDecode), but

some discrepancies existed. DSDecode is able to detect mutations with low complexity (up to 3 different PCR sequences) corresponding to the cases of monoallelic or biallelic mutations. For instance, monoallelic mutation (78bp deletion in one allele and a WT allele) was detected in line *ARF5_44* by both ICE and DSDecode. However, ICE gave more information about mutation allele percentage (57%) (Figure III-S3). For line *ARF5_43*, both tools ICE and DSDecode detected one WT allele and one mutated allele. However, ICE detected also 2% of 1bp deletion at sgRNA2 position, revealing that it was a chimera. This deletion was not detected by DSDecode probably due to its very low frequency (2%) and only a monoallelic mutation was reported by DSDecode (Figure III-S4).

c. ICE analysis for IAA9A transgenic roots (compared with DSDecode and subcloning)

In total, 9 PCR amplicons of *IAA9A* lines were successfully treated by DSDecode. We then performed the ICE mutation detection method and compared the results among the three different methods. In general, ICE results showed more edition types covering 70% (7/10) of 9 lines edition types by DSDecode, and 42.9% (12/28) of 13 lines by subcloning. The mutation types of lines *IAA9A_5*, *IAA9A_10* and *IAA9A_15* detected by DSDecode were the same as corresponding subcloning. One WT allele and one mutated (1bp del at sgRNA1) allele were found in line *IAA9A_5*. However, ICE reported one more edited type occurred at sgRNA2 (1bp del), which contributed to 4%. No mutation was detected in the line of *IAA9A_10* by subcloning, and it was also confirmed by both DSDecode and ICE as wild type alleles. As for *IAA9A_15*, ICE detected the main mutation type was large deletion (73bp del) between two sgRNAs with the occurrence as high as 38%, which was detected by both DSDecode and subcloning as well. In general ICE revealed more chimeric mutation than DSDecode.

Based on mutation detection results using two methods in Chapter II, the second group of *IAA9A* lines (*IAA9A_1*, *IAA9A_8*, *IAA9A_12*, *IAA9A_20*) were partial same between DSDecode and subcloning. As for these four *IAA9A* edited lines, the edited type with highest contribution (even second high percentage edited types of *IAA9A_8* and *IAA9A_12*) of each line in ICE was as the same performing in subcloning and/or DSDecode. For example, in ICE results, *IAA9A_1* and *IAA9A_20* had allele with 73bp del (95%) and 2bp del (simultaneously at two sgRNAs, 94%), respectively, and the two edited alleles were the major edition types in subcloning as well (occupied 88.9% and 66.7%). Different edited results of *IAA9A_3* and *IAA9A_11* were found in DSDecode and subcloning; furthermore, ICE results obtained chimeric mutations of two lines, with only minor editions the same as DSDecode, showing one allele (1bp del at sgRNA1, 4%) in *IAA9A_3*, and WT allele (93%) in *IAA9A_11*.

In brief, the comparison of the three mutation detection methods showed that ICE reported more edition types with low occurrence frequency, which was not always detected by DSDecode and even by subcloning.

3.2 FTIR_PLSDA analysis for chemotypes discrimination

In order to rapidly discriminate the chemotypes of CRISPR/Cas9 generated transgenic roots from controls, we used FT-IR. For each sample, ten absorption spectra were measured as technical replicates, and the results were analysed using PLS-DA (partial least square analysis). The median value of all absorption spectra (transgenic lines and controls) were compared, and then, main contributions of the wavelength numbers to the two major Principal Components (PC) axes were

identified.

a. FT-IR_PLSDA analysis for CRISPR/Cas9 edited transgenic *ARF5*, *IAA9A* lines and controls

Absorption spectra value of all *ARF5* samples were analysed using PCA and PLS-DA (Figure III-S5). After discarding outliers (control_8, control_17, *ARF5_2*) which were far apart from the other similar samples, FT-IR_PLSDA analysis was again carried out using remaining samples. As shown in Figure III-9A, together the two major components of PLSDA explained 56.2% of the total variability of all the samples. First component (PC1 axis) contributed to 32.1% variability which mainly separates controls from CRISPR- edited-*ARF5*_lines.

Interestingly, the few *ARF5*_lines close to the control group, were lines for which no edition was detected by ICE. In contrast, the remaining eleven *ARF5*_CRISPR edited lines (detected by ICE) were clearly separated from controls. Among those, three edited lines *ARF5_29*, *ARF5_40*, *ARF5_44* exhibiting large deletions, were located relatively far from the control group.

For *IAA9A* transgenic lines, we checked the editions using both sub-cloning and web-based tools and we obtained more reliable mutations characterization of each line. We selected only the true *IAA9A*_edited lines and controls (the outlier control_8 was discarded) for FT-IR_PLSDA analysis to obtain a better separation (all *IAA9A*_lines and control were analysed using PCA and PLSDA and the results were shown in Figure III-S6). Together, the PC1 and PC2 components of PLSDA explained 28.6% of the total variability (PC1 axis explained 17.8 %). In fourteen *IAA9A*_edited lines, mutation frequencies inferred from ICE and/or subcloning were noted along with red stars. *IAA9A_4*, *IAA9A_9*, *IAA9A_11* had low editing rates, which were not consistent with sub-cloning results (10 sequenced clones showed 100% mutation in *IAA9A_4* and *IAA9A_9* lines, 3 sequenced clones were all edited in *IAA9A_11* with 100% mutation) (Figure III-9B).

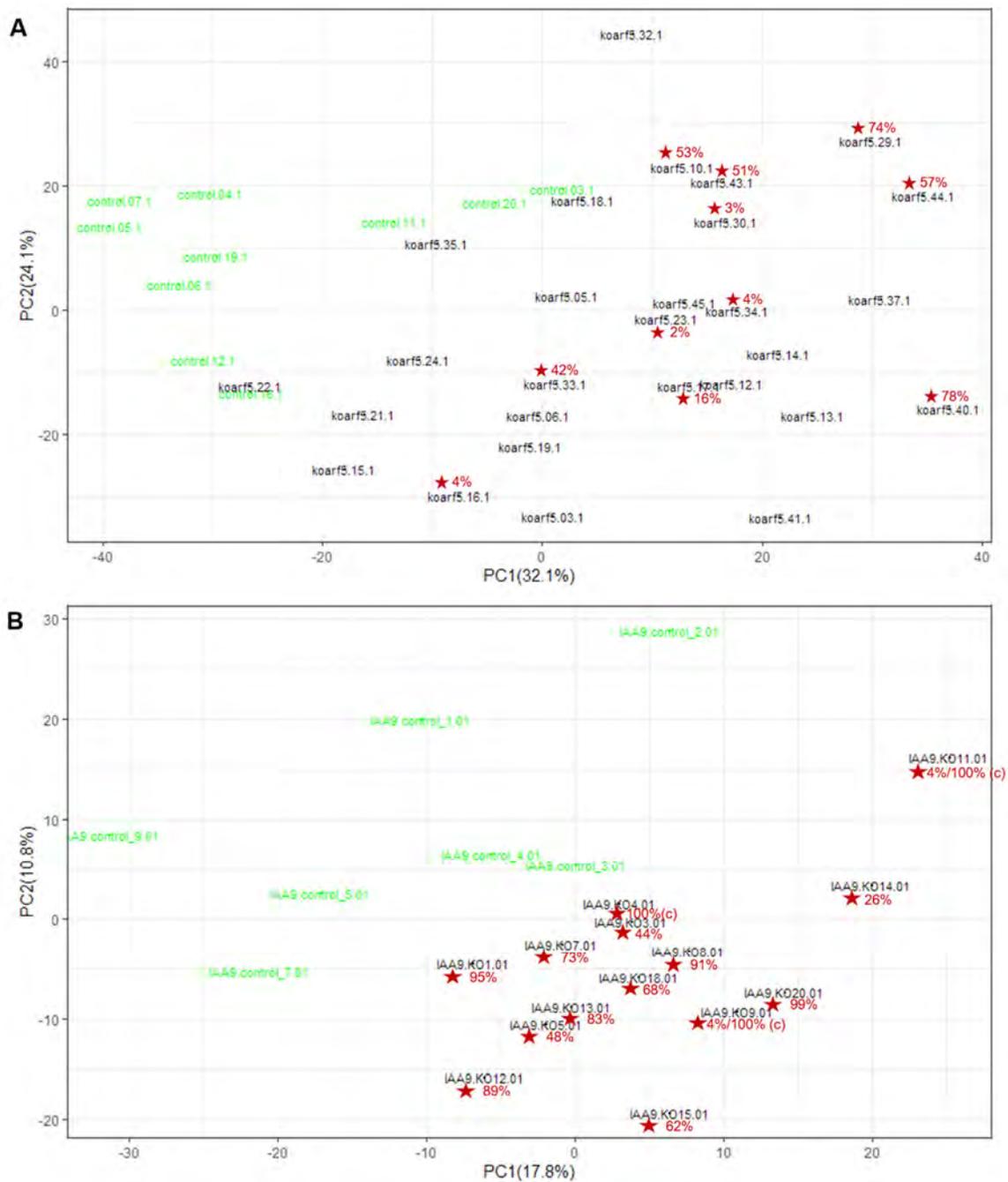


Figure III-9. PLSDA analyses were performed using normalized values of FT-IR absorption spectrum ($800\text{-}4000\text{cm}^{-1}$) obtained from controls and *ARF5_lines* (A); *IAA9A_edited* lines (B) transgenic hairy roots. (A) The first principal component (PC1) separate (A) most controls from *ARF5_lines*, and (B) all *IAA9A_edited* lines from controls. The edited lines noted using red stars were detected by ICE, along with various edited frequencies. The control samples were in green, and transgenic lines were in black. The mutation rate ‘100% (c)’ noted for *IAA9A_4*, *IAA9A_9*, *IAA9A_11* meant that: using sub-cloning method, all 10 subclones of *IAA9A_4*, all 10 subclones of *IAA9A_9* and all 3 subclones of *IAA9A_11* were mutated.

b. Identification of wavelength numbers associated polymers

Absorption spectra comparison between transgenic lines and controls highlighted discriminant wavelength numbers (WN), which are potentially associated with chemical cell wall polymers. Comparisons between *ARF5*_lines/*IAA9A*_edited lines and controls highlighted the most discriminant WNs gathered around 990-1735 cm^{-1} , with many corresponding to the lignin polymer.

Among the twelve most discriminant WNs (Figure III-10 (up)) between *ARF5* and control, eight were related to lignin. The median values of these WNs were higher in controls than in *ARF5*_lines, suggesting that the lignin content was decreased in CRISPR generated transgenic *ARF5*_lines. However, not all WNs related to lignin in controls were higher than in *ARF5*_lines (e.g., WN 2850 cm^{-1} and 3440 cm^{-1}). The WNs associated cellulose were lower in controls than *ARF5*_lines (990 cm^{-1} , 1040 cm^{-1}), in contrast to what was observed in *IAA9A*_edited lines (990 cm^{-1} , 1160 cm^{-1}). Pectin (related WN 1410 cm^{-1}) were decreased in *ARF5*_lines. Among twelve significant WNs in Figure III-10 (down) of *IAA9A* and control, nine WNs positions showed higher absorption spectra values of control than of transgenic *IAA9A*_edited lines, and six WNs were related to lignin. Some WNs were also reported to cellulose and pectin. There still need for more bibliography to mean bounds and related chemical compounds for WNs between 1735 cm^{-1} and 2850 cm^{-1} .

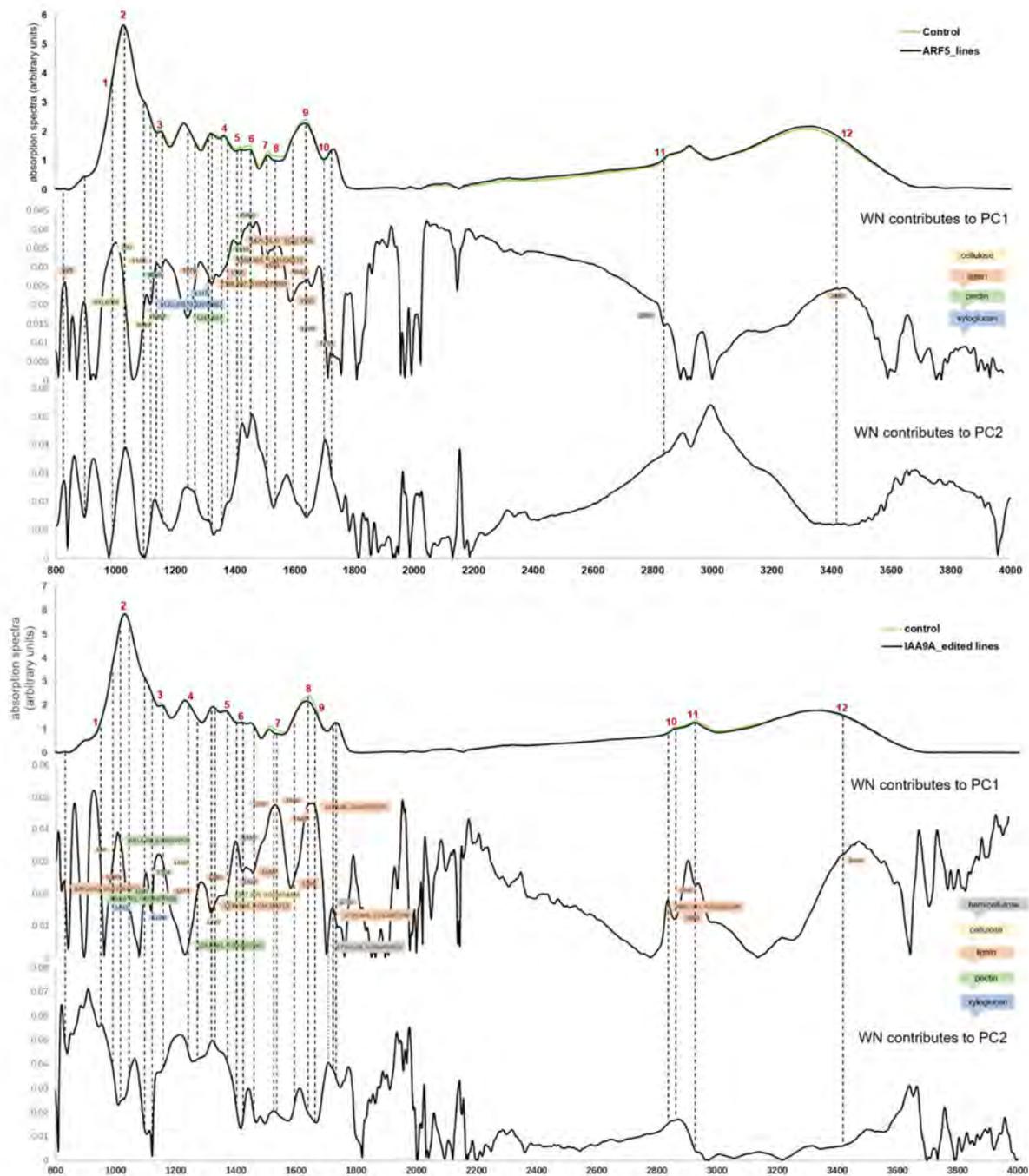


Figure III-10. Wavelength numbers (WN) from FT-IR absorption spectra discriminating between controls and *ARF5_lines* (up)/*IAA9A_edited lines* (down). The first curve in A and B showed FT-IR absorption spectra of controls (in green) and transgenic lines (in black), using the median values, respectively. The second and third curves in A and B presented the contribution of each WN to PC1 and PC2 axis. And the dotted lines along with numbers (1-12 in red) were significant WNs which are associated with polymers of SCW or xyloglucan (primary cell wall) (reference of (Largo-Gosens et al., 2014; Dai et al., 2020)), and/or had most effects in discrimination between transgenic lines and control samples.

3.3 Histology analysis of *IAA9*_edited lines and *ARF5*_edited lines

In order to evaluate the potential impacts of CRISPR-knock out/down of targeting genes on wood formation, we next examined the vascular tissues of CRISPR/Cas9 edited roots (*IAA9A*_lines, *ARF5*_lines) by histological analyses using either phloroglucinol-HCl, which stains lignin polymers in red-purple or the natural auto fluorescence of phenolic compounds (including lignin) under UV-light. As the root anatomy is development (position from the root apex) dependent, we compared the *IAA9A*_edited lines with its corresponding control at three different positions named old and medium and young stage. When compared with control similar position 1 (15-20cm from apex), *IAA9A*_edited lines appeared older than control, displaying larger stele sizes with more xylem cells. In the medium stage (position 2) around 10cm from apex, *IAA9A*_edited lines seemed also more developed than control plants, showing larger stele diameter and more xylem vessel cells (Figure III-11). Only a few young roots were harvested (we lost the young roots samples during the harvesting) for both control and *IAA9A*_edited lines and we have only few sections corresponding this stage. For the few sections obtained at this stage we did not observe obvious difference between the *IAA9A*_edited lines and wild type control. It seemed that at the same position from the root apex, the sections from *IAA9A*_edited lines seemed older than the wild-type, suggesting an accelerated xylem development. There is no much difference in the intensity of phloroglucinol-HCL staining and UV signal between *IAA9A*_lines and control, indicating no evident difference of xylem cell walls lignin content. In the young stage (Position 3), the lignification of central vessel cells was earlier than in control.

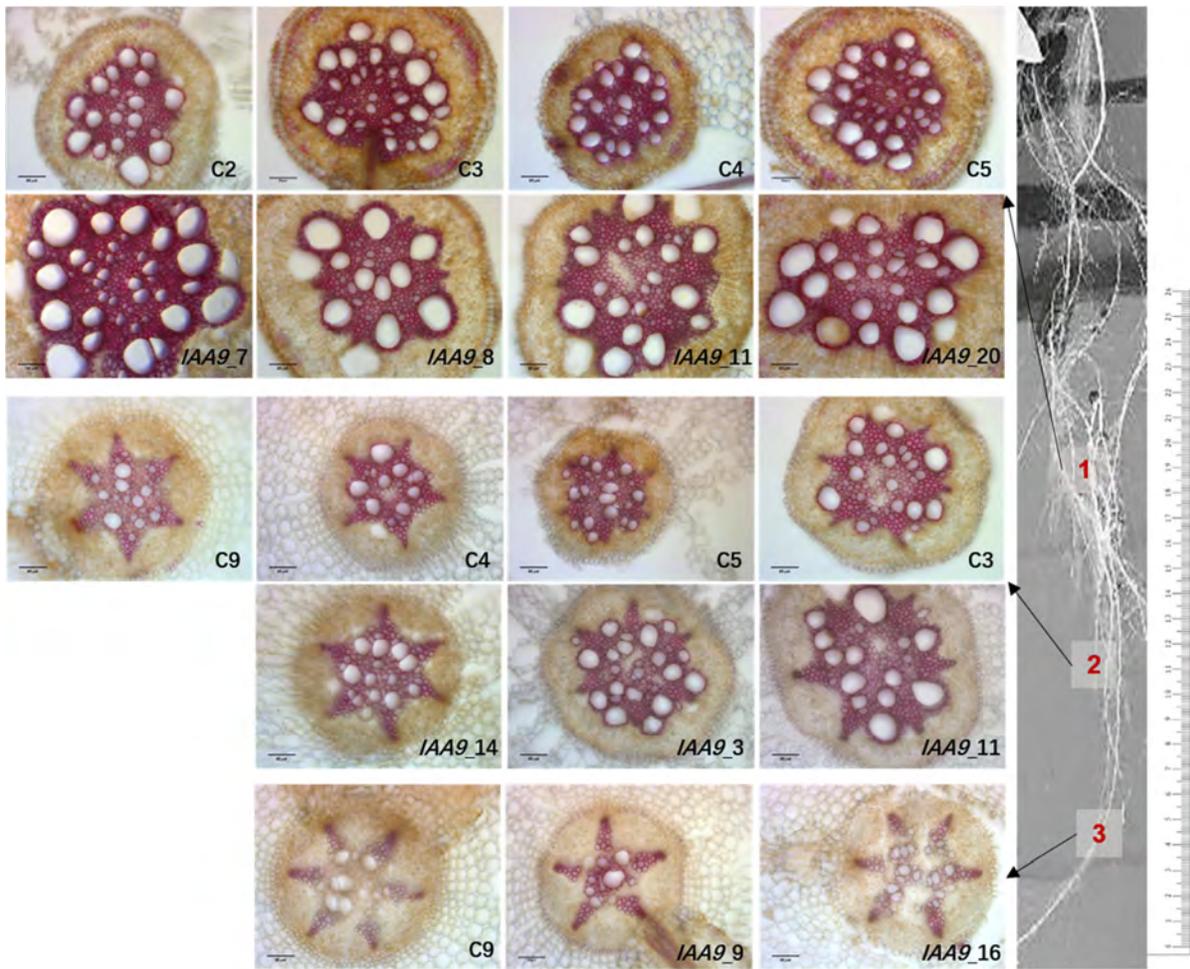


Figure III-11. Comparison of xylem development and lignification of xylem cells especially vessels between transgenic roots of control and *IAA9A*_edited lines. The transversal sections were obtained at three positions/ developmental stages (old (15-20 cm from apex), medium (~10 cm from apex), young (5 cm from apex)) from root apex and were stained by phloroglucinol-HCl. Here the images were observed under 20x magnification. C2, control 2; Scale bar=50 μ m.

We further measured the vessel diameters of *IAA9A*_lines and control (over 5 μ m) using Image J. In the general boxplot, vessel sizes of *IAA9A*_edited lines were larger than controls in both medium and old developing stages (Figure III-12). The median value of xylem vessel cells size in *IAA9A*_lines was 35 μ m, while in control it was 21 μ m (Figure III-S7).

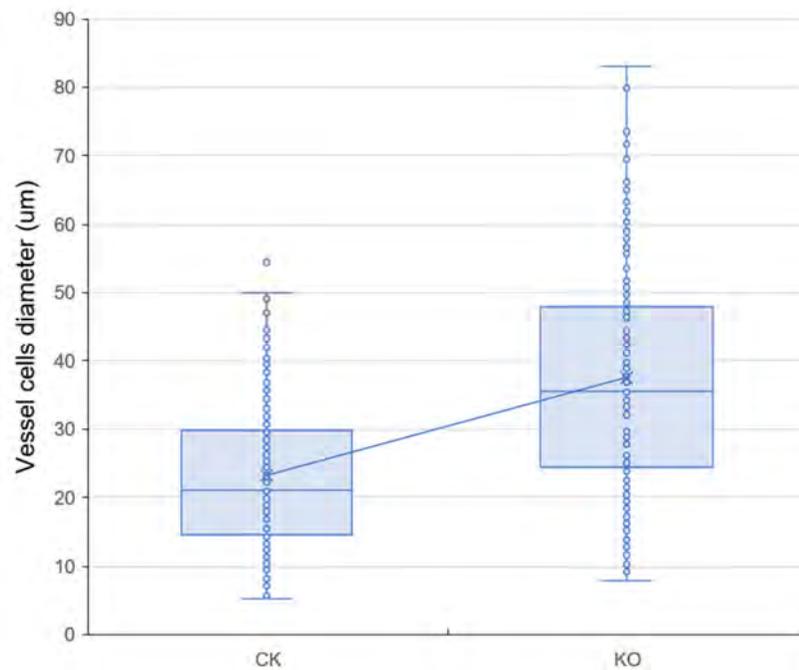


Figure III-12. The median value comparison of xylem vessel cells diameters between all *IAA9A*_edited lines and control. CK, control; KO, *IAA9A*_edited lines.

We also examined wood cells formation in *ARF5*_edited lines and control, however we did not find any obvious differences between control and *ARF5*_edited lines (Figure III-13). There is neither difference in the intensity of phloroglucinol-HCl staining and UV signal between *IAA9A*_lines and control, indicating that at least in our experimental conditions, there is no difference of xylem cell walls lignin content (Figure III-S8).

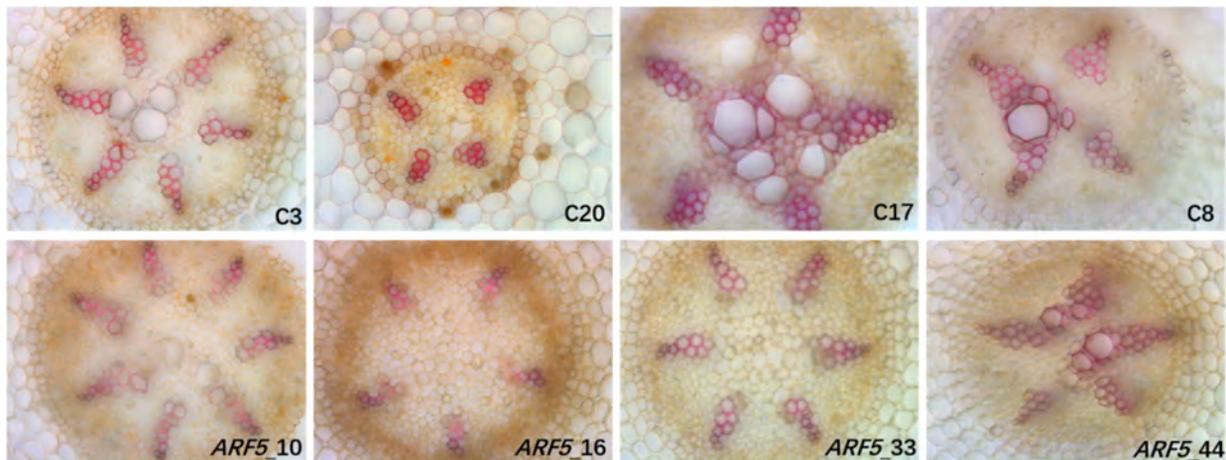


Figure III-13. Comparison of xylem development and lignification of xylem cells, especially vessels between transgenic roots of controls and *ARF5*_edited lines. The transversal sections were stained by phloroglucinol-HCl, and observed under 40x magnification. C3, control_3.

3.4 Y2H results

In order to identify potential protein partners of *EgrIAA9A* and *EgrIAA20* in developing xylem of Eucalyptus, we performed Yeast two Hybrid screening using *EgrIAA9A* and *EgrIAA20* as baits to screen Eucalyptus developing xylem Y2H library.

For *EgrIAA9A*, 37 potential binding partners were identified, including notably *IAA9* itself, *IAA23*, histone linker, *CCoAOMT2*, *TPR2* (TOPLESS-Like protein 2), Enzyme *E1/E2*, *SGT1* and its chaperone *HSP90*.

Remarkably, using *EgrIAA20* as a bait to screen the xylem library, one of the main partners identified was *EgrIAA9* (ten out of 46 positive clones corresponded to *EgrIAA9*) (Table III-1 highlight in yellow). However, it should be noted that when using *EgrIAA9* as bait, we did not obtain *EgrIAA20* as partner. This is possibly due to the relative low expression level of *IAA20* in developing xylem. *IAA20* is more specifically and highly expressed in cambium (Table III-2 highlight in yellow).

Four other clones were also identified as *IAA1*, *IAA11* and *IAA16* (two clones) suggesting that *IAA20* may form dimers with *IAA9* and/or other Aux/IAA as a part of transcription regulator complex. Three clones corresponding to *IRX10* were also found as potential partners. Strikingly, no ARF proteins was pull out from our Y2H screening whereas they are well-known as interactants of IAAs.

We further performed targeted Y2H with Aux/IAA and ARF as bait or pray. The *IAA9-IAA9*, *IAA20-IAA9* protein-protein interactions further confirmed the results from library screening (Figure III-14). Both *IAA9* and *IAA20* interacted with *EgrARF4*, *EgrARF5*, *EgrARF10* and *EgrARF19*, all putative transcription activators selected based on their expression profiles strongly related to the wood formation.

Table III-1. Sequencing results of PCR products from Y2H library (secondary xylem) screening colonies using *EgrIAA20* as a bait

Colony No.	Sequence length	NCBI Blast	Range (total length)	Gap	Identities
1	771	PREDICTED: Eucalyptus grandis probable beta-1,4-xylosyltransferase IRX10 (LOC104453563), mRNA	1131-1609 (1920)	0	467/479(97%)
3	935	PREDICTED: Eucalyptus grandis NHP2-like protein 1 (LOC104436675), mRNA	94-322 (847)	1	226/229(99%)
4	798	PREDICTED: Eucalyptus grandis uncharacterized LOC104456788 (LOC104456788), transcript variant X2, mRNA	390 - 528 (4040)	0	137/139(99%)
5	807	PREDICTED: Eucalyptus grandis uncharacterized LOC104456788 (LOC104456788), transcript variant X2, mRNA	390 - 528 (4040)	0	137/139(99%)
6	1054	PREDICTED: Eucalyptus grandis auxin-responsive protein IAA9 (LOC104449204), transcript variant X1, mRNA	905 - 1396 (1733)	3	473/495(96%)
7	1046	PREDICTED: Eucalyptus grandis auxin-responsive protein IAA9 (LOC104449204), transcript variant X1, mRNA	658 - 1516 (1733)	8	826/866(95%)
8	1043	PREDICTED: Eucalyptus grandis N-acetyltransferase 9-like protein (LOC104424664), mRNA	208- 570 (1460)	0	353/363(97%)
9	1074	PREDICTED: Eucalyptus grandis auxin-responsive protein IAA9 (LOC104449204), transcript variant X1, mRNA	1073 - 1393 (1733)	12	298/333(89%)
10	1040	PREDICTED: Eucalyptus grandis auxin-responsive protein IAA9 (LOC104449204), transcript variant X1, mRNA	958 - 1588 (1733)	2	626/633(99%)
11	1018	PREDICTED: Eucalyptus grandis auxin-responsive protein IAA1 (LOC104423134), mRNA	286 - 827 (1050)	0	540/542(99%)
12	895	PREDICTED: Eucalyptus grandis auxin-responsive protein IAA9 (LOC104414090), transcript variant X2, mRNA	1078 - 1437 (1742)	4	330/364(91%)
13	668	PREDICTED: Eucalyptus grandis uncharacterized LOC104456788 (LOC104456788), transcript variant X2, mRNA	422 - 528 (4040)	0	105/107(98%)
14	768	Yeast two-hybrid vector pGADc, complete sequence	3660 - 3742 (9637)	0	83/83(100%)
15	1039	PREDICTED: Eucalyptus grandis auxin-responsive protein IAA16 (LOC104415430), mRNA	705 - 1374 (1587)	8	653/677(96%)
16	937	PREDICTED: Eucalyptus grandis auxin-responsive protein IAA9 (LOC104449204), transcript variant X1, mRNA	958 - 1588 (1733)	2	627/633(99%)
17	1068	PREDICTED: Eucalyptus grandis auxin-responsive protein IAA9 (LOC104449204), transcript variant X1, mRNA	688 - 1351 (1733)	3	641/667(96%)
19	1038	PREDICTED: Eucalyptus grandis auxin-responsive protein IAA16 (LOC104415430), mRNA	705 - 1388 (1587)	8	667/691(97%)
20	1018	PREDICTED: Eucalyptus grandis auxin-responsive protein IAA11 (LOC104425226), mRNA	516 - 1243 (1375)	1	714/729(98%)
22	678	PREDICTED: Eucalyptus grandis protein SGT1 homolog (LOC104448141), transcript variant X2, mRNA	559 - 807 (1523)	0	249/249(100%)

Chapter III:

Functional Characterization of Wood-Associated *EgrIAA9A*, *EgrIAA20* and *EgrARF5* in Eucalyptus hairy roots

23	1093	PREDICTED: Eucalyptus grandis auxin-responsive protein IAA9 (LOC104449204), transcript variant X1, mRNA	694 - 1454 (1733)	6	736/767(96%)
24	818	PREDICTED: Eucalyptus grandis 14-3-3-like protein (LOC104447984), mRNA	462 - 553 (1292)	1	91/93(98%)
25	913	PREDICTED: Eucalyptus grandis protein C2-DOMAIN ABA-RELATED 7-like (LOC104453695), mRNA	531 - 1060 (1117)	0	522/530(98%)
26	720	PREDICTED: Eucalyptus grandis glyceraldehyde-3-phosphate dehydrogenase, cytosolic (LOC104417117), mRNA	947 - 1238 (1479)	2	285/292(98%)
27	1067	PREDICTED: Eucalyptus grandis auxin-responsive protein IAA9 (LOC104449204), transcript variant X1, mRNA	658 - 1507 (1733)	2	825/851(97%)
28	1067	PREDICTED: Eucalyptus grandis auxin-responsive protein IAA9 (LOC104449204), transcript variant X1, mRNA	658 - 1507 (1733)	2	825/851(97%)
30	817	PREDICTED: Eucalyptus grandis protein SGT1 homolog (LOC104448141), transcript variant X2, mRNA	559 - 807 (1523)	0	249/249(100%)
31	954	PREDICTED: Eucalyptus grandis probable beta-1,4-xylosyltransferase IRX10 (LOC104453563), mRNA	1360 - 1620 (1920)	2	245/263(93%)
33	790	PREDICTED: Eucalyptus grandis circumsporozoite protein (LOC104441474), mRNA	311 - 419 (1136)	0	107/109(98%)
34	840	PREDICTED: Eucalyptus grandis isoflavone reductase-like protein (LOC104448770), mRNA	803 - 1226 (1279)	0	418/424(99%)
35	504	PREDICTED: Eucalyptus grandis probable beta-1,4-xylosyltransferase IRX10 (LOC104453563), mRNA	1364 - 1646 (1920)	0	279/283(99%)

Table III-2. Sequencing results of PCR products from Y2H library (secondary xylem) screening colonies using *EgrIAA9A* as a bait

Colony No.	Sequence length	NCBI Blast	Range (total length)	Gap	Identities
2	331bp	PREDICTED: Eucalyptus grandis heat shock cognate 70 kDa protein 2 (LOC104420897), mRNA	617-855 (2510)	0	209/239(87%)
3	69bp	PREDICTED: Eucalyptus grandis protein AUXIN-REGULATED GENE INVOLVED IN ORGAN SIZE-like (LOC104454499), mRNA	365-414 (813)	0	49/50(98%)
4	110bp	PREDICTED: Eucalyptus grandis 50S ribosomal protein L1, chloroplastic (LOC104454616), mRNA	407-434 (1497)	0	28/28(100%)
5	273bp	PREDICTED: Eucalyptus grandis abscisic acid receptor PYL9 (LOC104443371), mRNA	501-696 (1130)	0	194/196(99%)
6	226bp	PREDICTED: Eucalyptus grandis fructose-bisphosphate aldolase 6, cytosolic (LOC104438376), mRNA	967-1022 (1545)	0	53/56(95%)
7	212bp	PREDICTED: Eucalyptus grandis eukaryotic translation initiation factor 5A-2 (LOC104425989), mRNA	95-215 (869)	4	117/123(95%)
9	188bp	PREDICTED: Eucalyptus grandis dnaJ homolog subfamily B member 1 (LOC104433137), mRNA	487-594 (1566)	0	103/108(95%)
10	198bp	PREDICTED: Eucalyptus grandis putative lactoylglutathione lyase (LOC104441412), transcript variant X2, mRNA	579-688 (1136)	0	109/110(99%)
11	180bp	PREDICTED: Eucalyptus grandis 2-hydroxy-6-oxononadienedioate/2-hydroxy-6-oxononatrienedioate hydrolase (LOC104418300), mRNA	648-741 (1272)	0	93/94(99%)
14	174bp	PREDICTED: Eucalyptus grandis ubiquitin-conjugating enzyme E2 variant 1D (LOC104454089), mRNA	306-391 (879)	0	85/86(99%)
15	224bp	PREDICTED: Eucalyptus grandis protein SPA, chloroplastic (LOC104436690), mRNA	248-378 (745)	0	128/131(98%)
16	400bp	PREDICTED: Eucalyptus grandis clathrin light chain 2 (LOC104445304), mRNA	546-785 (1521)	0	234/240(98%)
17	767bp	PREDICTED: Eucalyptus grandis putative lactoylglutathione lyase (LOC104441412), transcript variant X2, mRNA	579-863 (1136)	0	279/285(98%)
18	770bp	PREDICTED: Eucalyptus grandis protein TPR2 (LOC104424284), transcript variant X3, mRNA	664-853 (4086)	0	181/190(95%)
19	608bp	PREDICTED: Eucalyptus grandis auxin-responsive protein IAA23 (LOC104436630), mRNA	186-672 (1354)	2	463/488(95%)
20	767bp	PREDICTED: Eucalyptus grandis auxin-responsive protein IAA9 (LOC104449204), transcript variant X1, mRNA	1155-1544 (1733)	2	371/391(95%)
22	717bp	PREDICTED: Eucalyptus grandis ribulose-phosphate 3-epimerase, cytoplasmic isoform (LOC104419149), mRNA	224-480 (1050)	0	254/257(99%)
24	558bp	PREDICTED: Eucalyptus grandis signal recognition particle 19 kDa protein (LOC104453607), mRNA	288-603 (851)	10	304/325(94%)

Chapter III:

Functional Characterization of Wood-Associated *EgrIAA9A*, *EgrIAA20* and *EgrARF5* in *Eucalyptus* hairy roots

25	729bp	<i>Eucalyptus grandis</i> beta-tubulin (TUB1) mRNA, complete cds	140-603 (1583)	1	442/465(95%)
27	364bp	PREDICTED: <i>Eucalyptus grandis</i> D-3-phosphoglycerate dehydrogenase 1, chloroplastic (LOC104455297), partial mRNA	809-1065 (1476)	0	256/257(99%)
28	649bp	PREDICTED: <i>Eucalyptus grandis</i> auxin-repressed 12.5 kDa protein (LOC104440998), mRNA	45-589 (688)	1	535/546(98%)
29	772bp	PREDICTED: <i>Eucalyptus grandis</i> cysteine protease RD19A (LOC104448798), mRNA	737-1398 (1404)	2	647/664(97%)
31	649bp	PREDICTED: <i>Eucalyptus grandis</i> xyloglucan endotransglucosylase/hydrolase protein 9 (LOC104424824), mRNA	160-665 (1235)	1	498/507(98%)
32	788bp	<i>Eucalyptus globulus</i> caffeoyl-CoA O-methyltransferase (CCoAOMT2) mRNA, complete cds	405-991 (1013)	0	571/587(97%)
33	770bp	<i>Eucalyptus globulus</i> caffeoyl-CoA O-methyltransferase (CCoAOMT2) mRNA, complete cds	405-991 (1013)	0	571/587(97%)
34	587bp	PREDICTED: <i>Eucalyptus grandis</i> auxin-responsive protein IAA9 (LOC104449204), transcript variant X1, mRNA	1155-1599 (1733)	1	440/446(99%)
35	572bp	PREDICTED: <i>Eucalyptus grandis</i> activator of 90 kDa heat shock protein ATPase homolog 1 (LOC104446028), mRNA	713-1044 (1582)	0	332/332(100%)
36	428bp	PREDICTED: <i>Eucalyptus grandis</i> polyubiquitin-like (LOC104445529), misc_RNA	696-883 (1418)	2	185/190(97%)
37	428bp	PREDICTED: <i>Eucalyptus grandis</i> polyubiquitin-like (LOC104445529), misc_RNA	696-883 (1418)	2	185/190(97%)
38	496bp	PREDICTED: <i>Eucalyptus grandis</i> ubiquitin-activating enzyme E1 1-like (LOC104422383), transcript variant X4, mRNA	3497-3806 (3826)	0	308/310(99%)
39	563bp	PREDICTED: <i>Eucalyptus grandis</i> GDP-L-galactose phosphorylase 2 (LOC104450258), mRNA	1034-1366 (2469)	4	331/336(99%)
40	615bp	PREDICTED: <i>Eucalyptus grandis</i> F-box/LRR-repeat protein 4 (LOC104439928), mRNA	392-892 (1089)	4	490/505(97%)
41	808bp	PREDICTED: <i>Eucalyptus grandis</i> receptor-like kinase TMK3 (LOC104421346), mRNA	2880-3464 (3464)	0	578/585(99%)
42	773bp	PREDICTED: <i>Eucalyptus grandis</i> receptor-like kinase TMK3 (LOC104421346), mRNA	2880-3464 (3464)	0	578/585(99%)
44	686bp	PREDICTED: <i>Eucalyptus grandis</i> histone H1 (LOC104419672), mRNA	444-182 (1051)	5	323/344(94%)
47	474bp	PREDICTED: <i>Eucalyptus grandis</i> protein SGT1 homolog (LOC104448141), transcript variant X2, mRNA	559-807 (1523)	0	249/249(100%)
48	404bp	PREDICTED: <i>Eucalyptus grandis</i> histone H1 (LOC104419672), mRNA	444-725 (1051)	0	275/282(98%)

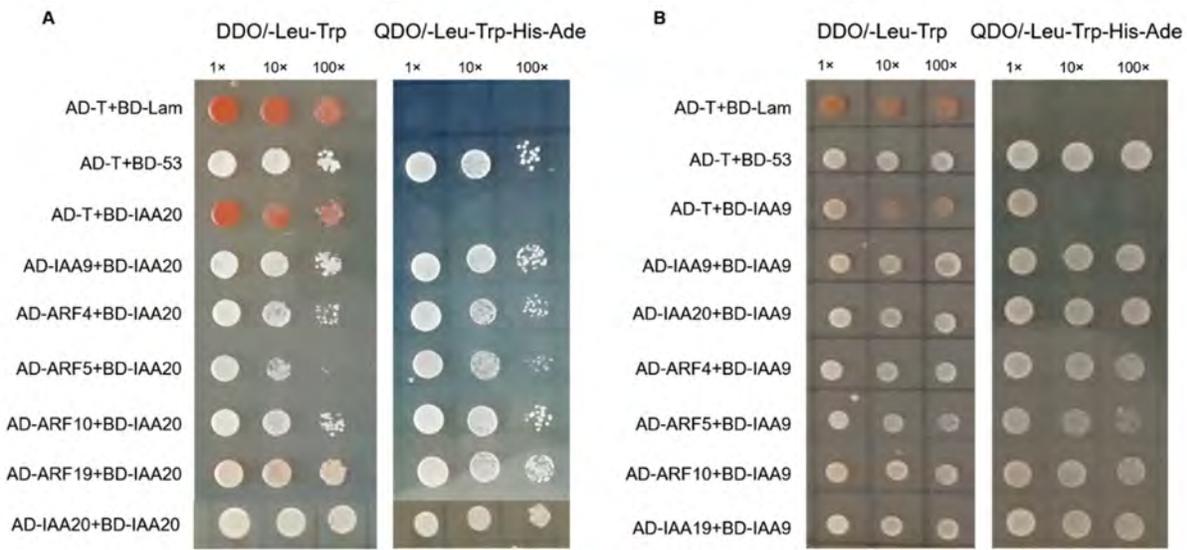


Figure III-14. Protein-protein interactions of *EgrIAA9A* (A) and *EgrIAA20* (B) with potential candidates. Co-transformed yeast of *EgrIAA9*-BD and candidates-AD, *EgrIAA20*-BD and candidates-AD grew on selected DDO medium (lacking Leu and Trp). Yeast with interactions grew on QDO medium (lacking Leu, Trp, His and Ade). AD-T+BD-Lam was the negative control which the yeast grew on DDO but no QDO; AD-T+BD-53 was the positive control; AD-T+BD-*EgrIAA9*/*EgrIAA20* was the negative control to verify that no interactions between bait and AD-T vector. 1x, 10x, 100x represented dilution ratio. Red yeasts were probably due to long time culture and accumulation of purine metabolism.

4. Discussion

The purposes of this chapter were: 1) to explore the regulatory function of *EgrIAA9A*, *EgrIAA20*, *EgrARF5* in vascular cambium and wood formation, through the generation of gain-of-function transgenic plants (35S promoter + coding sequence) and loss-of-function transgenic plants (CRISPR/Cas9 generated as described in Chapter II); 2) to seek for protein partners of *EgrIAA9A* and *EgrIAA20* by Y2H in woody tissues. *EgrIAA9A* belongs to Cluster III (Yu et al., 2015) which contains members relatively highly expressed in xylem tissue. In tomato, loss-of-function mutants of *SlIAA9* showed increased vascular veins and more xylem strands in leaves. Loss-of-function mutants of *AtIAA8* (*iaa8-1*) induced by T-DNA insertion exhibited no obvious phenotype even in *iaa8xiao9* double mutant. This is possibly due to the residual function of the truncated protein of IAA9 since the T-DNA is inserted at the end of coding region and thus the corresponding truncated protein lost only a few amino acids at the C terminal and might stay functional (Arase et al., 2012). In *Eucalyptus* hairy roots, loss-of-function *EgrIAA9A* lines revealed an accelerated xylem development, exhibiting a larger stele and increased diameter of xylem vessel cells when compared with its wild-type control. The cambium layers seemed more actively dividing in some mutants, which is consistent with poplar overexpressing *PtoIAA9m* (a stabilized IAA9 protein) showing repressed cambial proliferation and secondary xylem development (Xu et al., 2019). Altogether, these results suggest that IAA9 is one of the auxin signaling actors which explains the role of auxin in vessel formation. Vessels are derived from tracheids, and the diameter and distribution of vessels in woody plants provide determines xylem area-

specific conductivity, as well as response to environmental cues such as freeze and drought. While we have little knowledge about vessel cell fates and its developmental regulation. Auxin is involved in vessel patterning in wood, and promotes cell expansion. Auxin activates H⁺-ATPases in plasma membrane to form the extracellular acidification, leading to water uptake and cell volume expansion, based on the acid growth model (Spartz et al., 2014). Previous evidences showed polar auxin transport is required in vascular development in primary tissues, especially in leaves and vasculature, auxin action (through PIN-FORMED (PIN) and PAT carriers) forms a feedback loop ‘canalization’ (Scarpella et al., 2010). Recent studies explained that auxin transport is crucial for vessel cells spatial patterning, size and differentiation, which is associated with foliar development and stem hydraulic capacity (Johnson et al., 2018).

EgrIAA20 harbors no domain II, and as such it is more stable than other canonical *EgrIAAs*. Unfortunately, we could not progress on *IAA20* functional analysis since during the first COVID lock-down period, we lost all the overexpressing lines and only 3 *IAA20* CRISPR lines were saved with only one out of 3 clearly harboring a mutation (large deletion). In *Arabidopsis*, overexpression of *AtIAA20* and *AtIAA30* impaired root growth and decreased the cotyledon venation (Sato and Yamamoto, 2008).

Finally, we did not observe clear phenotypes in CRISPR/Cas9 generated *ARF5* lines using histochemical analyses of transgenic hairy roots, this may be due, at least in part, to the fact that we did not obtain complete knock out plants with bi allelic mutations. However, it should be noted that using FTIR screening and multivariate analysis we clearly separated the transgenic lines from the control lines, suggesting that chemical changes occurred in response to *ARF5* down-regulation.

Further investigations are needed to better understand the auxin associated candidate genes roles and regulation mechanisms during vascular development and wood formation.

When using *EgrIAA9A* as a bait for Eucalyptus xylem library screening, we identified several auxin-signaling components that, based on bibliographic data, have good chance to be true functional partners:

- 1) SGT1 (SUPPRESSOR OF G2 ALLELE SKIP1) and its partner HSP90, which targets auxin co-receptor TIR1 (Aux/IAA interacting protein) and also forms a complex with TIR1. Earlier studies revealed that HSP90 co-chaperone SGT1 is required for auxin response in *Arabidopsis*, and recent research showed that HSP90 positively regulates auxin receptor function (Watanabe et al., 2017), and SGT1-HSP90 complex is required for TIR1 stability. Also, HSP90-TIR1 module integrates environmental temperature and auxin signaling for plant development (Wang et al., 2016);
- 2) TPR2 (TOPLESS-Like protein 2) is regarded as an Aux/IAA co-repressor. Interestingly, TPL/TPR protein-protein interactions in *Arabidopsis* plant using yeast two-hybrid identified 17 distinct family partners including a Aux/IAA member *AtIAA8* (Causier et al., 2020);
- 3) Ubiquitination activating Enzyme E1 and conjugating Enzyme E2 are involved in Aux/IAA protein degradation.

Besides, two SCW biosynthetic enzymes were also identified as potential partners: CCoAOMT2 (Caffeoyl coenzyme A O-methyltransferase) and IRX10 playing key roles in lignin biosynthesis pathway and glucuronoxylan synthesis, respectively. The question whether they are real protein partners of *IAA9* and *IAA20*, or artifacts needs further investigation.

Chapter III:

Functional Characterization of Wood-Associated *EgrIAA9A*, *EgrIAA20* and *EgrARF5* in Eucalyptus hairy roots

Surprisingly, we did not identify any ARF members as we expected, probably due to the limited number of positive clones obtained and the library coverage. Indeed, further investigations of targeted protein-protein interactions showed that both *EgrIAA9A* and *EgrIAA20* interacted with our xylem highly expressed Eucalyptus ARF members including *ARF5*.

General Discussion and Perspectives

My PhD project mainly focused on CRISPR/Cas9 implementation in transgenic Eucalyptus hairy roots to generate loss-of-function mutants for functional characterization of genes involved in wood formation, and further use this technology to functionally characterize some auxin-dependent candidates (Aux/IAA and ARF) in the regulation of xylem/wood differentiation. We also identified some potential interacting protein partners of Eucalyptus IAA9A and IAA20 by screening a Y2H Eucalyptus developing xylem library and confirmed targeted protein-protein interactions between wood associated Aux/IAA members and wood associated ARF members.

I will start by some comments on the advantages and limits of the methods used during my PhD. To reveal the CRISPR/Cas9 system genome edition and characterize the mutation types, we extracted genomic DNA from E. grandis transgenic roots. The best method to detect editions (mutations) is high-throughput NGS method, but it is very expensive. Thus, we first used the conventional subcloning method which consists in PCR amplification of the target gene region including predicted cleavage/edition sites, then subcloning the PCR amplicon and sequencing of 6 to 10 subclones. This method is accurate but time-consuming and laborious and also limited by the cloning number and the costs. The prevalent methods now are web-based tools to directly analyze PCR amplicons sequencing results. Among the various analytic on-lines tools, we used the two well-accepted free tools which were widely used for their rapidity (DSDecode ‘Degenerate Sequence Decoding’ (<http://dsdecode.scgene.com>), and ICE ‘Inference of CRISPR Edits’ (<https://ice.synthego.com/#/>)). From the analyses of the sequencing chromatograms, the inferences for editing events and obtained by the two analytic tools were not always same. DSDecode works more efficient for stable mutation such as homozygous or heterozygous mutations but is not suitable to detect chimeric mutations. ICE analysis is more suitable to detect monoallelic and chimera mutations, along with frequencies for each mutation type. Therefore, ICE seems more adapted when working with heterozygous organisms such as trees.

In the CRISPR/Cas9 system we set up for Eucalyptus hairy roots, we made the choice to use two sgRNAs. In theory, this strategy should increase the editing efficiency and lead to large defined deletions. Indeed, large deletions were obtained for CRISPR-generated *ccr1*, *iaa9a*, *iaa20*, *arf5* lines. The risk of off-targets depends on sgRNA specificity, and we chose ‘CRISPOR’ online tool to evaluate the potential genome-scale off-targets risk for each gRNA. We choose the ones with low off-target score gRNA. NGS would be the perfect tool to detect off-target events at the whole genome scale but it’s a very expensive checking. An alternative strategy would be to perform some verifications targeted at locations presenting relative high similarity to the target gene such as the loci of paralogs genes.

FTIR spectrometry lies in the region between wavelength $4000-666\text{cm}^{-1}$ (infrared absorption spectrum) to identify chemical bonds in a molecule (Gorzsás et al., 2011) representing for instance chemical characteristics of cell wall components such as lignin, cellulose, etc (Mohamed et al., 2017). We have shown in my PhD that the combination of FTIR and multivariate analyses such as PCA or PLSDA is a powerful, non-destructive and rapid tool to discriminate between transgenics and control chemotypes prior to other analyses such as histology.

An example is the case of ARF5 that was identified hitherto as a key regulator in vascular tissues specification in embryo, leaf and root (Ruonala and Helariutta, 2017). In our experiment, histological analyses of CRISPR-generated *arf5* mutants did not reveal any obvious phenotypes in xylem cells development. However, the chemotypes obtained by FT-IR spectroscopy analysed by PLS-DA analysis were clearly separated between transgenic lines and controls. Several

discriminating wavelength numbers were identified as related with cell wall components including lignin and cellulose, indicating chemical difference between *ARF5*_edited lines and wildtype control. The profiling result of this new fast pre-screening method revealed differences in the chemical composition between transgenic lines and controls, providing information for further chemical analyses. We have to stress out that because of the lockdown, the harvest of the roots was badly postponed, the competition for growth between *ARF5*-impaired roots and WT roots likely was in favour of the latter and may explain that few *ARF5*_edited lines (11/30) were found with low allele-editing rates. In fact, for one line *ARF5_33*, we detected the edition types at two developmental stage during *in vitro* culture (around 70 days) and hydroponic culture (around 190 days). In the first stage, we detected monoallelic mutations, whereas in old transgenic hairy roots harvested late, we detected chimeric mutations. This is similar to what was observed for CRISPR-generated *ccr1* lines. We hypothesize that edited and non-edited cell lineages co-exist in the same fluorescent root, may result in mixed tissues to form unstable chimera mutation. If the target gene is crucial for roots growth and development, we may have less and less edited transgenic cells production compared with WT non-edited cell lineages in the same root.

To investigate the effects of *ARF5* loss of function at the histological level, it would be thus important to harvest early (3-6 weeks) since chemotypes differences were revealed even in late harvested roots. It would be interesting too, to analyze gene expression of *ARF5*/MP related (regulators, interactors, downstream targets) such as *PIN*, *HB8*, *BDL*...

In our experiment, loss-of-function *iaa9* mutants displayed accelerated xylem development and increased xylem vessel cell diameters, which is partial consistent with previous results: (i) increased vascular development in secondary veins and hypertrophic vascular xylem tissue found in tomato AS-*IAA9* lines (Wang, 2005) (ii) inhibition of cambium periclinal division and secondary xylem development in *PtoIAA9m* overexpressing poplar (Xu et al., 2019) (iii) reduction or complete inhibition of secondary fiber cells, while vessels were promoted in *Arabidopsis* OE-*EgrIAA9Am* (Mingjun Liu, 2015). Altogether these indicate that *IAA9* is a repressor in regulation of vascular development, especially during secondary xylem development.



References

- Abel, S., Nguyen, M.D., and Theologis, A. (1995). ThePS-IAA4/5-like Family of Early Auxin-inducible mRNAs in *Arabidopsis thaliana*. *Journal of Molecular Biology* **251**: 533–549.
- Adamowski, M. and Friml, J. (2015). PIN-Dependent Auxin Transport: Action, Regulation, and Evolution. *Plant Cell* **27**: 20–32.
- Agusti, J., Herold, S., Schwarz, M., Sanchez, P., Ljung, K., Dun, E.A., Brewer, P.B., Beveridge, C.A., Sieberer, T., Sehr, E.M., and Greb, T. (2011). Strigolactone signaling is required for auxin-dependent stimulation of secondary growth in plants. *Proceedings of the National Academy of Sciences* **108**: 20242–20247.
- Ain, Q.U., Chung, J.Y., and Kim, Y.-H. (2015). Current and future delivery systems for engineered nucleases: ZFN, TALEN and RGEN. *Journal of Controlled Release* **205**: 120–127.
- Arabidopsis Interactome Mapping Consortium et al. (2011). Evidence for Network Evolution in an Arabidopsis Interactome Map. *Science* **333**: 601–607.
- Arase, F., Nishitani, H., Egusa, M., Nishimoto, N., Sakurai, S., Sakamoto, N., and Kaminaka, H. (2012). IAA8 Involved in Lateral Root Formation Interacts with the TIR1 Auxin Receptor and ARF Transcription Factors in Arabidopsis. *PLoS ONE* **7**: e43414.
- Bajguz, A. and Piotrowska, A. (2009). Conjugates of auxin and cytokinin. *Phytochemistry* **70**: 957–969.
- Bali, G., Khunsapat, R., Akinosho, H., Payyavula, R.S., Samuel, R., Tuskan, G.A., Kalluri, U.C., and Ragauskas, A.J. (2016). Characterization of cellulose structure of *Populus* plants modified in candidate cellulose biosynthesis genes. *Biomass and Bioenergy* **94**: 146–154.
- Barra-Jiménez, A. and Ragni, L. (2017). Secondary development in the stem: when Arabidopsis and trees are closer than it seems. *Current Opinion in Plant Biology* **35**: 145–151.
- Bennett, T., Hines, G., and Leyser, O. (2014). Canalization: what the flux? *Trends in Genetics* **30**: 41–48.
- Bernard, G., Gagneul, D., Alves Dos Santos, H., Etienne, A., Hilbert, J.-L., and Rambaud, C. (2019). Efficient Genome Editing Using CRISPR/Cas9 Technology in Chicory. *IJMS* **20**: 1155.
- Bewg, W.P., Ci, D., and Tsai, C.-J. (2018). Genome Editing in Trees: From Multiple Repair Pathways to Long-Term Stability. *Front. Plant Sci.* **9**: 1732.
- Bhalerao, R.P. and Fischer, U. (2014). Auxin gradients across wood - instructive or incidental? *Physiologia Plantarum* **151**: 43–51.
- Bi, Z., Tahir, A.T., Huang, H., and Hua, Y. (2019). Cloning and functional analysis of five *TERMINAL FLOWER 1/CENTRORADIALIS*-like genes from *Hevea brasiliensis*. *Physiol Plantarum* **166**: 612–627.
- Bishopp, A., Help, H., El-Showk, S., Weijers, D., Scheres, B., Friml, J., Benková, E., Mähönen, A.P., and Helariutta, Y. (2011). A Mutually Inhibitory Interaction between Auxin and Cytokinin Specifies Vascular Pattern in Roots. *Current Biology* **21**: 917–926.
- Björklund, S., Antti, H., Uddestrand, I., Moritz, T., and Sundberg, B. (2007). Cross-talk between gibberellin and auxin in development of *Populus* wood: Gibberellin stimulates polar auxin transport and has a common transcriptome with auxin. *Plant Journal* **52**: 499–511.
- Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., and Bonas, U. (2009). Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science* **326**: 1509–1512.
- Boer, D.R., Freire-Rios, A., van den Berg, W.A.M., Saaki, T., Manfield, I.W., Kepinski, S., López-Vidriero, I., Franco-Zorrilla, J.M., de Vries, S.C., Solano, R., Weijers, D., and Coll, M. (2014). Structural Basis for DNA Binding Specificity by the Auxin-Dependent ARF Transcription Factors. *Cell* **156**: 577–589.

- Boerjan, W., Ralph, J., and Baucher, M.** (2003). Lignin Biosynthesis. *Annual Review of Plant Biology* **54**: 519–546.
- Bossinger, G. and Spokevicius, A.V.** (2018). Sector analysis reveals patterns of cambium differentiation in poplar stems. *Journal of Experimental Botany* **69**: 4339–4348.
- Brackmann, K. et al.** (2018). Spatial specificity of auxin responses coordinates wood formation. *Nat Commun* **9**: 875.
- Breitler, J.-C., Dechamp, E., Campa, C., Zebal Rodrigues, L.A., Guyot, R., Marraccini, P., and Etienne, H.** (2018). CRISPR/Cas9-mediated efficient targeted mutagenesis has the potential to accelerate the domestication of *Coffea canephora*. *Plant Cell Tiss Organ Cult* **134**: 383–394.
- Brocken, D.J.W., Tark-Dame, M., and Dame, R.T.** (2018). dCas9: A Versatile Tool for Epigenome Editing. *Current Issues in Molecular Biology*: 15–32.
- Bruegmann, T., Deecke, K., and Fladung, M.** (2019). Evaluating the Efficiency of gRNAs in CRISPR/Cas9 Mediated Genome Editing in Poplars. *IJMS* **20**: 3623.
- Bryan, A.C. et al.** (2016). Knockdown of a laccase in *Populus deltoides* confers altered cell wall chemistry and increased sugar release. *Plant Biotechnol J* **14**: 2010–2020.
- Busse-Wicher, M., Grantham, N.J., Lyczakowski, J.J., Nikolovski, N., and Dupree, P.** (2016). Xylan decoration patterns and the plant secondary cell wall molecular architecture. *Biochemical Society Transactions* **44**: 74–78.
- Buttò, V., Deslauriers, A., Rossi, S., Rozenberg, P., Shishov, V., and Morin, H.** (2020). The role of plant hormones in tree-ring formation. *Trees* **34**: 315–335.
- Cai, Y., Chen, L., Liu, X., Sun, S., Wu, C., Jiang, B., Han, T., and Hou, W.** (2015). CRISPR/Cas9-Mediated Genome Editing in Soybean Hairy Roots. *PLoS ONE* **10**: e0136064.
- Calderón Villalobos, L.I.A. et al.** (2012). A combinatorial TIR1/AFB–Aux/IAA co-receptor system for differential sensing of auxin. *Nat Chem Biol* **8**: 477–485.
- Camargo, E.L.O., Ployet, R., Cassan-Wang, H., Mounet, F., and Grima-Pettenati, J.** (2018). Digging in wood: New insights in the regulation of wood formation in tree species. In *Advances in Botanical Research* (Elsevier).
- Cano-Delgado, A., Penfield, S., Smith, C., Catley, M., and Bevan, M.** (2003). Reduced cellulose synthesis invokes lignification and defense responses in *Arabidopsis thaliana*. *Plant J* **34**: 351–362.
- Cao, P.B., Ployet, R., Nguyen, C., Dupas, A., Ladouce, N., Martinez, Y., Grima-Pettenati, J., Marque, C., Mounet, F., and Teulières, C.** (2020). Wood architecture and composition are deeply remodeled in frost sensitive *Eucalyptus* overexpressing CBF/DREB1 transcription factors. *International Journal of Molecular Sciences*.
- Carocha, V., Soler, M., Hefer, C., Cassan-Wang, H., Fevereiro, P., Myburg, A.A., Paiva, J.A.P., and Grima-Pettenati, J.** (2015). Genome-wide analysis of the lignin toolbox of *Eucalyptus grandis*. *New Phytologist* **206**: 1297–1313.
- Carraro, N., Tisdale-Orr, T.E., Clouse, R.M., Knöller, A.S., and Spicer, R.** (2012). Diversification and Expression of the PIN, AUX/LAX, and ABCB Families of Putative Auxin Transporters in *Populus*. *Front. Plant Sci.* **3**.
- Casimiro, I., Beeckman, T., Graham, N., Bhalerao, R., Zhang, H., Casero, P., Sandberg, G., and Bennett, M.J.** (2003). Dissecting *Arabidopsis* lateral root development. *Trends in Plant Science* **8**: 165–171.
- Causier, B., Ashworth, M., Guo, W., and Davies, B.** (2012a). The TOPLESS Interactome: A Framework for Gene Repression in *Arabidopsis*. *Plant Physiol.* **158**: 423–438.
- Causier, B., Lloyd, J., Stevens, L., and Davies, B.** (2012b). TOPLESS co-repressor interactions and their evolutionary conservation in plants. *Plant Signaling & Behavior* **7**: 325–328.

- Cermak, T., Doyle, E.L., Christian, M., Wang, L., Zhang, Y., Schmidt, C., Baller, J.A., Somia, N.V., Bogdanove, A.J., and Voytas, D.F. (2011). Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Research* **39**: e82–e82.
- Cerna, D. and Wilson, D.K. (2005). The Structure of Sif2p, a WD Repeat Protein Functioning in the SET3 Corepressor Complex. *Journal of Molecular Biology* **351**: 923–935.
- Chaffey, N., Cholewa, E., Regan, S., and Sundberg, B. (2002). Secondary xylem development in Arabidopsis: a model for wood formation. *Physiologia Plantarum* **114**: 594–600.
- Charrier, A., Vergne, E., Dousset, N., Richer, A., Petiteau, A., and Chevreau, E. (2019). Efficient Targeted Mutagenesis in Apple and First Time Edition of Pear Using the CRISPR-Cas9 System. *Front. Plant Sci.* **10**: 40.
- Chen, Q., Dai, X., De-Paoli, H., Cheng, Y., Takebayashi, Y., Kasahara, H., Kamiya, Y., and Zhao, Y. (2014). Auxin Overproduction in Shoots Cannot Rescue Auxin Deficiencies in Arabidopsis Roots. *Plant and Cell Physiology* **55**: 1072–1079.
- Chen, Q., Westfall, C.S., Hicks, L.M., Wang, S., and Jez, J.M. (2010). Kinetic Basis for the Conjugation of Auxin by a GH3 Family Indole-acetic Acid-Amido Synthetase. *J. Biol. Chem.* **285**: 29780–29786.
- Chen, Y., Shen, H., Wang, M., Li, Q., and He, Z. (2013a). Salicyloyl-aspartate synthesized by the acetyl-amido synthetase GH3.5 is a potential activator of plant immunity in Arabidopsis. *Acta Biochimica et Biophysica Sinica* **45**: 827–836.
- Chen, Y., Yordanov, Y.S., Ma, C., Strauss, S., and Busov, V.B. (2013b). DR5 as a reporter system to study auxin response in Populus. *Plant Cell Rep* **32**: 453–463.
- Cherenkov, P., Novikova, D., Omelyanchuk, N., Levitsky, V., Grosse, I., Weijers, D., and Mironova, V. (2018). Diversity of cis-regulatory elements associated with auxin response in Arabidopsis thaliana. *Journal of Experimental Botany* **69**: 329–339.
- Chiatante, D., Rost, T., Bryant, J., and Scippa, G.S. (2018). Regulatory networks controlling the development of the root system and the formation of lateral roots: a comparative analysis of the roles of pericycle and vascular cambium. *Annals of Botany*.
- Concordet, J.-P. and Haussler, M. (2018). CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Research* **46**: W242–W245.
- Courtois-Moreau, C.L., Pesquet, E., Sjödin, A., Muñoz, L., Bollhöner, B., Kaneda, M., Samuels, L., Jansson, S., and Tuominen, H. (2009). A unique program for cell death in xylem fibers of *Populus* stem. *The Plant Journal* **58**: 260–274.
- Daboussi, F., Stoddard, T.J., and Zhang, F. (2015). Engineering Meganuclease for Precise Plant Genome Modification. In *Advances in New Technology for Targeted Modification of Plant Genomes*, F. Zhang, H. Puchta, and J.G. Thomson, eds (Springer New York: New York, NY), pp. 21–38.
- Dai, X., Zhang, Y., Zhang, D., Chen, J., Gao, X., Estelle, M., and Zhao, Y. (2015). Embryonic lethality of Arabidopsis *abp1-1* is caused by deletion of the adjacent BSM gene. *Nature Plants* **1**: 15183.
- Dai, Y., Hu, G., Dupas, A., Medina, L., Blandels, N., San Clemente, H., Ladouce, N., Badawi, M., Hernandez-Raquet, G., Mounet, F., Grima-Pettenati, J., and Cassan-Wang, H. (2020). Implementing the CRISPR/Cas9 Technology in Eucalyptus Hairy Roots Using Wood-Related Genes. *IJMS* **21**: 3408.
- Darwin, C. (1897). *The power of movement in plants* (Appleton).
- De Rybel, B. et al. (2014). Integration of growth and patterning during vascular tissue formation in Arabidopsis. *Science* **345**: 1255215–1255215.

- De Smet, I. et al.** (2010). Bimodular auxin response controls organogenesis in *Arabidopsis*. *Proceedings of the National Academy of Sciences* **107**: 2705–2710.
- De Smet, I.** (2010). Multimodular auxin response controls lateral root development in *Arabidopsis*. *Plant Signaling & Behavior* **5**: 580–582.
- DeBolt, S., Cook, D.R., and Ford, C.M.** (2006). L-Tartaric acid synthesis from vitamin C in higher plants. *Proceedings of the National Academy of Sciences* **103**: 5608–5613.
- Deltcheva, E., Chylinski, K., Sharma, C.M., Gonzales, K., Chao, Y., Pirzada, Z.A., Eckert, M.R., Vogel, J., and Charpentier, E.** (2011). CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **471**: 602–607.
- Demura, T. et al.** (2002). Visualization by comprehensive microarray analysis of gene expression programs during transdifferentiation of mesophyll cells into xylem cells. *Proceedings of the National Academy of Sciences* **99**: 15794–15799.
- Derbyshire, P., Ménard, D., Green, P., Saalbach, G., Buschmann, H., Lloyd, C.W., and Pesquet, E.** (2015). Proteomic Analysis of Microtubule Interacting Proteins over the Course of Xylem Tracheary Element Formation in *Arabidopsis*. *Plant Cell*: tpc.15.00314.
- Devillard, C. and Walter, C.** (2014). Formation of plant tracheary elements in vitro – a review. *N.Z. j. of For. Sci.* **44**: 22.
- Ding, L., Chen, Y., Ma, Y., Wang, H., and Wei, J.** (2020). Effective reduction in chimeric mutants of poplar trees produced by CRISPR/Cas9 through a second round of shoot regeneration. *Plant Biotechnol Rep.*
- Doblin, M.S., Kurek, I., Jacob-Wilk, D., and Delmer, D.P.** (2002). Cellulose Biosynthesis in Plants: from Genes to Rosettes. *Plant and Cell Physiology* **43**: 1407–1420.
- Doench, J.G., Hartenian, E., Graham, D.B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B.L., Xavier, R.J., and Root, D.E.** (2014). Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat Biotechnol* **32**: 1262–1267.
- Domingo, C., Andrés, F., Tharreau, D., Iglesias, D.J., and Talón, M.** (2009). Constitutive Expression of *OsGH3.1* Reduces Auxin Content and Enhances Defense Response and Resistance to a Fungal Pathogen in Rice. *MPMI* **22**: 201–210.
- Donaldson, L., Hague, J., and Snell, R.** (2001). Lignin Distribution in Coppice Poplar, Linseed and Wheat Straw. *Holzforschung* **55**: 379–385.
- Du, H., Wu, N., Fu, J., Wang, S., Li, X., Xiao, J., and Xiong, L.** (2012). A GH3 family member, OsGH3-2, modulates auxin and abscisic acid levels and differentially affects drought and cold tolerance in rice. *Journal of experimental botany* **63**: 6467–6480.
- Ebringerová, A. and Heinze, T.** (2000). Xylan and xylan derivatives–biopolymers with valuable properties, 1. Naturally occurring xylans structures, isolation procedures and properties. *Macromolecular rapid communications* **21**: 542–556.
- Elorriaga, E., Klocko, A.L., Ma, C., and Strauss, S.H.** (2018). Variation in Mutation Spectra Among CRISPR/Cas9 Mutagenized Poplars. *Front. Plant Sci.* **9**: 594.
- Engler, C., Kandzia, R., and Marillonnet, S.** (2008). A One Pot, One Step, Precision Cloning Method with High Throughput Capability. *PLoS ONE* **3**: e3647.
- Eriksson, M.E., Israelsson, M., Olsson, O., and Moritz, T.** (2000). Increased gibberellin biosynthesis in transgenic trees promotes growth, biomass production and xylem fiber length. *Nature Biotechnology* **18**: 784–788.

- Espinosa-Ruiz, A., Martínez, C., de Lucas, M., Fàbregas, N., Bosch, N., Caño-Delgado, A.I., and Prat, S.** (2017). TOPLESS mediates brassinosteroid control of shoot boundaries and root meristem development in *Arabidopsis thaliana*. *Development* **144**: 1619–1628.
- Evert, R.F.** (2006). *Esau's plant anatomy: meristems, cells, and tissues of the plant body: their structure, function, and development* (John Wiley & Sons).
- Fàbregas, N., Formosa-Jordan, P., Confraria, A., Siligato, R., Alonso, J.M., Swarup, R., Bennett, M.J., Mähönen, A.P., Caño-Delgado, A.I., and Ibañez, M.** (2015). Auxin Influx Carriers Control Vascular Patterning and Xylem Differentiation in *Arabidopsis thaliana*. *PLoS Genet* **11**: e1005183.
- Fan, D., Liu, T., Li, C., Jiao, B., Li, S., Hou, Y., and Luo, K.** (2015). Efficient CRISPR/Cas9-mediated Targeted Mutagenesis in *Populus* in the First Generation. *Scientific Reports* **5**.
- Fan, Y., Xin, S., Dai, X., Yang, X., Huang, H., and Hua, Y.** (2020). Efficient genome editing of rubber tree (*hevea brasiliensis*) protoplasts using CRISPR/Cas9 ribonucleoproteins. *Industrial Crops and Products* **146**: 112146.
- Farboud, B. and Meyer, B.J.** (2015). Dramatic Enhancement of Genome Editing by CRISPR/Cas9 Through Improved Guide RNA Design. *Genetics* **199**: 959–971.
- Farboud, B., Severson, A.F., and Meyer, B.J.** Strategies for Efficient Genome Editing Using CRISPR-Cas9.: 27.
- Farcot, E., Lavedrine, C., and Vernoux, T.** (2015). A Modular Analysis of the Auxin Signalling Network. *PLoS ONE* **10**: e0122231.
- Fauser, F., Schiml, S., and Puchta, H.** (2014). Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in *Arabidopsis thaliana*. *The Plant Journal* **79**: 348–359.
- Fendrych, M., Akhmanova, M., Merrin, J., Glanc, M., Hagihara, S., Takahashi, K., Uchida, N., Torii, K.U., and Friml, J.** (2018). Rapid and reversible root growth inhibition by TIR1 auxin signalling. *Nature Plants* **4**: 453–459.
- Feng, Z. et al.** (2014). Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in *Arabidopsis*. *Proceedings of the National Academy of Sciences* **111**: 4632–4637.
- Fischer, U., Kucukoglu, M., Helariutta, Y., and Bhalerao, R.P.** (2019). The Dynamics of Cambial Stem Cell Activity. *Annu. Rev. Plant Biol.* **70**: 293–319.
- Fister, A.S., Landherr, L., Maximova, S.N., and Guiltinan, M.J.** (2018). Transient Expression of CRISPR/Cas9 Machinery Targeting TeNPR3 Enhances Defense Response in *Theobroma cacao*. *Frontiers in Plant Science* **9**.
- Friml, J.** (2003). Auxin transport & shaping the plant. *Current Opinion in Plant Biology*: 6.
- Friml, J. and Palme, K.** (2002). Polar auxin transport — old questions and new concepts? In *Auxin Molecular Biology*, C. Perrot-Rechenmann and G. Hagen, eds (Springer Netherlands: Dordrecht), pp. 273–284.
- Fu, X., Shi, Z., Jiang, Y., Jiang, L., Qi, M., Xu, T., and Li, T.** (2019). A family of auxin conjugate hydrolases from *Solanum lycopersicum* and analysis of their roles in flower pedicel abscission. *BMC Plant Biol* **19**: 233.
- Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K., and Sander, J.D.** (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* **31**: 822–826.
- Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M., and Joung, J.K.** (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol* **32**: 279–284.
- 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* **48**: 380–389.

- Fukuda, H. and Komamine, A.** (1980). Establishment of an Experimental System for the Study of Tracheary Element Differentiation from Single Cells Isolated from the Mesophyll of *Zinnia elegans*. *PLANT PHYSIOLOGY* **65**: 57–60.
- Gaj, T., Gersbach, C.A., and Barbas, C.F.** (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends in Biotechnology* **31**: 397–405.
- Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V.** (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences* **109**: E2579–E2586.
- Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., and Liu, D.R.** (2017). Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551**: 464–471.
- Gee, M.A., Hagen, G., and Guilfoyle, T.J.** (1991). Tissue-specific and organ-specific expression of soybean auxin-responsive transcripts GH3 and SAURs. *The Plant Cell* **3**: 419–430.
- Geisler, M., Aryal, B., di Donato, M., and Hao, P.** (2017). A Critical View on ABC Transporters and Their Interacting Partners in Auxin Transport. *Plant and Cell Physiology* **58**: 1601–1614.
- Geisler, M. and Murphy, A.S.** (2006). The ABC of auxin transport: The role of p-glycoproteins in plant development. *FEBS Letters* **580**: 1094–1102.
- Ghannoum, O., Phillips, N.G., Sears, M.A., Logan, B.A., Lewis, J.D., Conroy, J.P., and Tissue, D.T.** (2010). Photosynthetic responses of two eucalypts to industrial-age changes in atmospheric [CO₂] and temperature: Eucalyptus photosynthesis in past and future climates. *Plant, Cell & Environment* **33**: 1671–1681.
- Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C.A., and Smith, H.O.** (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**: 343–345.
- Gil, P. and Green, P.J.** (1997). Regulatory activity exerted by the SAUR-AC1 promoter region in transgenic plants. *Plant molecular biology* **34**: 803–808.
- Girijashankar, V.** (2011). Genetic transformation of eucalyptus. *Physiology and Molecular Biology of Plants* **17**: 9–23.
- Goicoechea, M., Lacombe, E., Legay, S., Mihaljevic, S., Rech, P., Jauneau, A., Lapiere, C., Pollet, B., Verhaegen, D., Chaubet-Gigot, N., and Grima-Pettenati, J.** (2005). EgMYB2, a new transcriptional activator from Eucalyptus xylem, regulates secondary cell wall formation and lignin biosynthesis. *Plant Journal* **43**: 553–567.
- Gorzsás, A., Stenlund, H., Persson, P., Trygg, J., and Sundberg, B.** (2011). Cell-specific chemotyping and multivariate imaging by combined FT-IR microspectroscopy and orthogonal projections to latent structures (OPLS) analysis reveals the chemical landscape of secondary xylem: The chemical landscape of secondary xylem. *The Plant Journal* **66**: 903–914.
- Grattapaglia, D., Mamani, E.M.C., Silva-Junior, O.B., and Faria, D.A.** (2015). A novel genome-wide microsatellite resource for species of *Eucalyptus* with linkage-to-physical correspondence on the reference genome sequence. *Mol Ecol Resour* **15**: 437–448.
- Gray, W.M., Kepinski, S., Rouse, D., Leyser, O., and Estelle, M.** (2001). Auxin regulates SCFTIR1-dependent degradation of AUX/IAA proteins. *Nature* **414**: 271–276.
- Gray, W.M., del Pozo, J.C., Walker, L., Hobbie, L., Risseuw, E., Banks, T., Crosby, W.L., Yang, M., Ma, H., and Estelle, M.** (1999). Identification of an SCF ubiquitin-ligase complex required for auxin response in *Arabidopsis thaliana*. *Genes & Development* **13**: 1678–1691.
- Greaves, B.L., Borralho, N.M.G., and Raymond, C.A.** Breeding Objective for Plantation Eucalypts Grown for Production of Kraft Pulp.: 8.

- Groover, A. and Robischon, M.** (2006). Developmental mechanisms regulating secondary growth in woody plants. *Current Opinion in Plant Biology* **9**: 55–58.
- Groover, A.T., Pattishall, A., and Jones, A.M.** (2003). IAA8 expression during vascular cell differentiation. *Plant molecular biology* **51**: 427–435.
- Guell, M., Yang, L., and Church, G.M.** (2014). Genome editing assessment using CRISPR Genome Analyzer (CRISPR-GA). *Bioinformatics* **30**: 2968–2970.
- Gui, J., Luo, L., Zhong, Y., Sun, J., Umezawa, T., and Li, L.** (2019). Phosphorylation of LTF1, an MYB Transcription Factor in Populus, Acts as a Sensory Switch Regulating Lignin Biosynthesis in Wood Cells. *Molecular Plant* **12**: 1325–1337.
- Guilfoyle, T.J. and Hagen, G.** (2007). Auxin response factors. *Current Opinion in Plant Biology* **10**: 453–460.
- Guilfoyle, T.J. and Hagen, G.** (2012). Getting a grasp on domain III/IV responsible for Auxin Response Factor–IAA protein interactions. *Plant Science* **190**: 82–88.
- Hacke, U.G., Spicer, R., Schreiber, S.G., and Plavcová, L.** (2017). An ecophysiological and developmental perspective on variation in vessel diameter: Variation in xylem vessel diameter. *Plant, Cell & Environment* **40**: 831–845.
- Haeussler, M., Schönig, K., Eckert, H., Eschstruth, A., Mianné, J., Renaud, J.-B., Schneider-Maunoury, S., Shkumatava, A., Teboul, L., Kent, J., Joly, J.-S., and Concordet, J.-P.** (2016). Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biology* **17**.
- Hagen, G. and Guilfoyle, T.** (2002). Auxin-responsive gene expression: genes, promoters and regulatory factors. In *Auxin Molecular Biology*, C. Perrot-Rechenmann and G. Hagen, eds (Springer Netherlands: Dordrecht), pp. 373–385.
- Hamann, T.** (2002). The Arabidopsis BODENLOS gene encodes an auxin response protein inhibiting MONOPTEROS-mediated embryo patterning. *Genes & Development* **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.
- Han, W. and She, Q.** (2017). CRISPR History: Discovery, Characterization, and Prosperity. In *Progress in Molecular Biology and Translational Science* (Elsevier), pp. 1–21.
- Hansen, J., Jørgensen, J.-E., Stougaard, J., and Marcker, K.A.** (1989). Hairy roots – a short cut to transgenic root nodules. *Plant Cell Reports* **8**: 12–15.
- Hardtke, C.S.** (2004). Overlapping and non-redundant functions of the Arabidopsis auxin response factors MONOPTEROS and NONPHOTOTROPIC HYPOCOTYL 4. *Development* **131**: 1089–1100.
- Hess, G.T., Tycko, J., Yao, D., and Bassik, M.C.** (2017). Methods and Applications of CRISPR-Mediated Base Editing in Eukaryotic Genomes. *Molecular Cell* **68**: 26–43.
- Hsu, P.D. et al.** (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* **31**: 827–832.
- Hu, W., Yan, H., Luo, S., Pan, F., Wang, Y., and Xiang, Y.** (2018). Genome-wide analysis of poplar SAUR gene family and expression profiles under cold, polyethylene glycol and indole-3-acetic acid treatments. *Plant Physiology and Biochemistry* **128**: 50–65.
- Humphreys, J.M. and Chapple, C.** (2002). Rewriting the lignin roadmap. *Current Opinion in Plant Biology* **5**: 224–229.
- Hussey, S.G., Mizrahi, E., Creux, N.M., and Myburg, A.A.** (2013). Navigating the transcriptional roadmap regulating plant secondary cell wall deposition. *Front. Plant Sci.* **4**.

- Hussey, S.G., Saïdi, M.N., Hefer, C.A., Myburg, A.A., and Grima-Pettenati, J. (2015). Structural, evolutionary and functional analysis of the NAC domain protein family in *Eucalyptus*. *New Phytologist* **206**: 1337–1350.
- Ibanes, M., Fabregas, N., Chory, J., and Cano-Delgado, A.I. (2009). Brassinosteroid signaling and auxin transport are required to establish the periodic pattern of Arabidopsis shoot vascular bundles. *Proceedings of the National Academy of Sciences* **106**: 13630–13635.
- Igarashi, M., Hatsuyama, Y., Harada, T., and Fukasawa-Akada, T. (2016). Biotechnology and apple breeding in Japan. *Breed. Sci.* **66**: 18–33.
- Immanen, J. et al. (2016). Cytokinin and Auxin Display Distinct but Interconnected Distribution and Signaling Profiles to Stimulate Cambial Activity. *Current Biology* **26**: 1990–1997.
- Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *Journal of Bacteriology* **169**: 5429–5433.
- Ito, J., Fukaki, H., Onoda, M., Li, L., Li, C., Tasaka, M., and Furutani, M. (2016). Auxin-dependent compositional change in Mediator in ARF7- and ARF19-mediated transcription. *Proc Natl Acad Sci USA* **113**: 6562–6567.
- Jacobs, T.B., LaFayette, P.R., Schmitz, R.J., and Parrott, W.A. (2015). Targeted genome modifications in soybean with CRISPR/Cas9. *BMC Biotechnol.* **15**: 16.
- Jansen, Ruud., Embden, Jan.D.A. van, Gaastra, Wim., and Schouls, Leo.M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* **43**: 1565–1575.
- Jansson, S. and Douglas, C.J. (2007). *Populus*: A Model System for Plant Biology. *Annu. Rev. Plant Biol.* **58**: 435–458.
- Jia, H. and Wang, N. (2014). Targeted Genome Editing of Sweet Orange Using Cas9/sgRNA. *PLoS ONE* **9**: e93806.
- Jia, H., Xu, J., Orbović, V., Zhang, Y., and Wang, N. (2017a). Editing Citrus Genome via SaCas9/sgRNA System. *Front. Plant Sci.* **8**: 2135.
- Jia, H., Zhang, Y., and Orbovi, V. (2016). Genome editing of the disease susceptibility gene *CsLOB1* in citrus confers resistance to citrus canker.: 7.
- Jia, H., Zhang, Y., Orbović, V., Xu, J., White, F.F., Jones, J.B., and Wang, N. (2017b). Genome editing of the disease susceptibility gene *CsLOB1* in citrus confers resistance to citrus canker. *Plant Biotechnol J* **15**: 817–823.
- Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B., and Weeks, D.P. (2013). Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. *Nucleic Acids Research* **41**: e188–e188.
- Jiang, Y., Guo, L., Ma, X., Zhao, X., Jiao, B., Li, C., and Luo, K. (2017). The WRKY transcription factors PtrWRKY18 and PtrWRKY35 promote Melampsora resistance in Populus. *Tree Physiology* **37**: 665–675.
- Jin, H., Cominelli, E., Bailey, P., Parr, A., Mehrtens, F., Jones, J., Tonelli, C., Weisshaar, B., and Martin, C. (2000). Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in Arabidopsis. *The EMBO journal* **19**: 6150–6161.
- Jin, Y.-L., Tang, R.-J., Wang, H.-H., Jiang, C.-M., Bao, Y., Yang, Y., Liang, M.-X., Sun, Z.-C., Kong, F.-J., Li, B., and Zhang, H.-X. (2017). Overexpression of *Populus trichocarpa* *CYP85A3* promotes growth and biomass production in transgenic trees. *Plant Biotechnol J* **15**: 1309–1321.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**: 816–821.

- Johnson, D., Eckart, P., Alsamadisi, N., Noble, H., Martin, C., and Spicer, R.** (2018). Polar auxin transport is implicated in vessel differentiation and spatial patterning during secondary growth in *Populus*. *American Journal of Botany* **105**: 186–196.
- Johnson, L.A. and Douglas, C.J.** (2007). *Populus trichocarpa* MONOPTEROS/AUXIN RESPONSE FACTOR5 (ARF5) genes: comparative structure, sub-functionalization, and *Populus* – *Arabidopsis* microsynteny. This article is one of a selection of papers published in the Special Issue on Poplar Research in Canada. *Can. J. Bot.* **85**: 1058–1070.
- Johnsson, C., Jin, X., Xue, W., Dubreuil, C., Lezhneva, L., and Fischer, U.** (2018). The plant hormone auxin directs timing of xylem development by inhibition of secondary cell wall deposition through repression of secondary wall NAC-domain transcription factors. *Physiologia plantarum*: 1–24.
- Jones, A.M. and Herman, E.M.** (1993). KDEL-Containing Auxin-Binding Protein Is Secreted to the Plasma Membrane and Cell Wall. *Plant Physiology* **101**: 595–606.
- Joshi, C.P., Bhandari, S., Ranjan, P., Kalluri, U.C., Liang, X., Fujino, T., and Samuga, A.** (2004). Genomics of cellulose biosynthesis in poplars: Research review. *New Phytologist* **164**: 53–61.
- Junghans, U., Langenfeld-Heyser, R., Polle, A., and Teichmann, T.** (2004). Effect of auxin transport inhibitors and ethylene on the wood anatomy of poplar. *Plant Biology* **6**: 22–29.
- Jurado, S., Abraham, Z., Manzano, C., López-Torrejón, G., Pacios, L.F., and Del Pozo, J.C.** (2010). The *Arabidopsis* Cell Cycle F-Box Protein SKP2A Binds to Auxin. *Plant Cell* **22**: 3891–3904.
- Jurado, S., Diaz-Trivino, S., Abraham, Z., Manzano, C., Gutierrez, C., and del Pozo, J. c.** (2008). SKP2A protein, an F-box that regulates cell division, is degraded via the ubiquitin pathway. *Plant Signaling & Behavior* **3**: 810–812.
- 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.
- Kalluri, U.C., Payyavula, R.S., Labbé, J.L., Engle, N., Bali, G., Jawdy, S.S., Sykes, R.W., Davis, M., Ragauskas, A., Tuskan, G.A., and Tschaplinski, T.J.** (2016). Down-Regulation of KORRIGAN-Like Endo- β -1,4-Glucanase Genes Impacts Carbon Partitioning, Mycorrhizal Colonization and Biomass Production in *Populus*. *Front. Plant Sci.* **7**.
- Kaneda, M., Schuetz, M., Lin, B.S.P., Chanis, C., Hamberger, B., Western, T.L., Ehlting, J., and Samuels, A.L.** (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* **101**: 12381–12386.
- Kepinski, S. and Leyser, O.** (2002). Ubiquitination and Auxin Signaling. *The Plant Cell* **14**: S81–S95.
- Kieffer, M., Stern, Y., Cook, H., Clerici, E., Maulbetsch, C., Laux, T., and Davies, B.** (2006). Analysis of the Transcription Factor WUSCHEL and Its Functional Homologue in *Antirrhinum* Reveals a Potential Mechanism for Their Roles in Meristem Maintenance. *Plant Cell* **18**: 560–573.
- Kim, H., Kim, S.-T., Ryu, J., Kang, B.-C., Kim, J.-S., and Kim, S.-G.** (2017). CRISPR/Cpf1-mediated DNA-free plant genome editing. *Nat Commun* **8**: 14406.
- Kim, H., Shim, D., Moon, S., Lee, J., Bae, W., Choi, H., Kim, K., and Ryu, H.** (2019). Transcriptional network regulation of the brassinosteroid signaling pathway by the BES1–TPL–HDA19 co-repressor complex. *Planta* **250**: 1371–1377.
- Ko, J.-H., Jeon, H.-W., Kim, W.-C., Kim, J.-Y., and Han, K.-H.** (2014). The MYB46/MYB83-mediated transcriptional regulatory programme is a gatekeeper of secondary wall biosynthesis. *Annals of Botany* **114**: 1099–1107.

- Ko, J.-H., Kim, W.-C., Keathley, D.E., and Han, K.-H.** (2016). Chapter 10 - Genetic Engineering for Secondary Xylem Modification: Unraveling the Genetic Regulation of Wood Formation. In *Secondary Xylem Biology*, Y.S. Kim, R. Funada, and A.P. Singh, eds (Academic Press: Boston), pp. 193–211.
- Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R.** (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**: 420–424.
- Konishi, M., Donner, T.J., Scarpella, E., and Yanagisawa, S.** (2015). MONOPTEROS directly activates the auxin-inducible promoter of the Dof5.8 transcription factor gene in *Arabidopsis thaliana* leaf provascular cells. *Journal of Experimental Botany* **66**: 283–291.
- Korasick, D.A., Jez, J.M., and Strader, L.C.** (2015). Refining the nuclear auxin response pathway through structural biology. *Current Opinion in Plant Biology* **27**: 22–28.
- Korasick, D.A., Westfall, C.S., Lee, S.G., Nanao, M.H., Dumas, R., Hagen, G., Guilfoyle, T.J., Jez, J.M., and Strader, L.C.** (2014). Molecular basis for AUXIN RESPONSE FACTOR protein interaction and the control of auxin response repression. *Proceedings of the National Academy of Sciences* **111**: 5427–5432.
- Krogan, N.T., Yin, X., Ckurshumova, W., and Berleth, T.** (2014). Distinct subclades of *Aux/IAA* genes are direct targets of ARF5/MP transcriptional regulation. *New Phytol* **204**: 474–483.
- Kubo, M.** (2005). Transcription switches for protoxylem and metaxylem vessel formation. *Genes & Development* **19**: 1855–1860.
- Kubo, M., Udagawa, M., Nishikubo, N., Horiguchi, G., Yamaguchi, M., Ito, J., Mimura, T., Fukuda, H., and Demura, T.** (2005). Transcription switches for protoxylem and metaxylem vessel formation. *Genes & Development* **19**: 1855–1860.
- Kumar, M., Mishra, L., Carr, P., Pilling, M., Gardner, P., Mansfield, S.D., and Turner, S.** (2018). Exploiting CELLULOSE SYNTHASE (CESA) Class Specificity to Probe Cellulose Microfibril Biosynthesis. *Plant Physiol.* **177**: 151–167.
- Lacombe, E., Hawkins, S., Van Doorselaere, J., Piquemal, J., Goffner, D., Poeydomenge, O., Boudet, A.-M., and Grima-Pettenati, J.** (1997). Cinnamoyl CoA reductase, the first committed enzyme of the lignin branch biosynthetic pathway: cloning, expression and phylogenetic relationships. *The Plant Journal* **11**: 429–441.
- Largo-Gosens, A., Hernández-Altamirano, M., García-Calvo, L., Alonso-Simón, A., Álvarez, J., and Acebes, J.L.** (2014). Fourier transform mid infrared spectroscopy applications for monitoring the structural plasticity of plant cell walls. *Front. Plant Sci.* **5**.
- Larson, P.R.** (2012). *The vascular cambium: development and structure* (Springer Science & Business Media).
- Lavy, M. and Estelle, M.** (2016). Mechanisms of auxin signaling. *Development* **143**: 3226–3229.
- Lawrenson, T., Shorinola, O., Stacey, N., Li, C., Østergaard, L., Patron, N., Uauy, C., and Harwood, W.** (2015). Induction of targeted, heritable mutations in barley and *Brassica oleracea* using RNA-guided Cas9 nuclease. *Genome Biol* **16**: 258.
- Ledford, H.** (2015). Alternative CRISPR system could improve genome editing. *Nature* **526**: 17–17.
- Lee, C., Teng, Q., Zhong, R., and Ye, Z.-H.** (2011). Molecular Dissection of Xylan Biosynthesis during Wood Formation in Poplar. *Molecular Plant* **4**: 730–747.
- Lee, Hyoshin, 문흥규, and Park, So-Young** (2014). Agrobacterium-mediated Transformation via Somatic Embryogenesis System in Korean fir (*Abies koreana* Wil.), A Korean Native Conifer. *한국자원식물학회지* **27**: 242–248.

- Legay, S. et al.** (2010). EgMYB1, an R2R3 MYB transcription factor from eucalyptus negatively regulates secondary cell wall formation in Arabidopsis and poplar. *The New phytologist* **188**: 774–786.
- Legay, S., Lacombe, E., Goicoechea, M., Brière, C., Séguin, A., Mackay, J., and Grima-Pettenati, J.** (2007). Molecular characterization of EgMYB1, a putative transcriptional repressor of the lignin biosynthetic pathway. *Plant Science* **173**: 542–549.
- Lei, Y., Lu, L., Liu, H.-Y., Li, S., Xing, F., and Chen, L.-L.** (2014). CRISPR-P: A Web Tool for Synthetic Single-Guide RNA Design of CRISPR-System in Plants. *Molecular Plant* **7**: 1494–1496.
- Lekha, P., Bush, T., Pammenter, N., Sitholè, B., and Berjak, P.** (2017). Quantitative assessment of xylan distribution across the secondary cell wall layers of Eucalyptus dissolving pulp fibres. *Holzforschung* **72**: 1–8.
- Leysner, O.** (2018). Auxin Signaling. *Plant Physiology* **176**: 465–479.
- Li, C., Wang, X., Ran, L., Tian, Q., Fan, D., and Luo, K.** (2015). PtoMYB92 is a Transcriptional Activator of the Lignin Biosynthetic Pathway During Secondary Cell Wall Formation in *Populus tomentosa*. *Plant and Cell Physiology* **56**: 2436–2446.
- Li, E., Bhargava, A., Qiang, W., Friedmann, M.C., Forneris, N., Savidge, R. a, Johnson, L. a, Mansfield, S.D., Ellis, B.E., and Douglas, C.J.** (2012). The Class II KNOX gene KNAT7 negatively regulates secondary wall formation in Arabidopsis and is functionally conserved in *Populus*. *The New phytologist* **7**: 102–115.
- Li, J.-F.** (2013). Multiplex and homologous recombination-mediated genome editing in Arabidopsis and *Nicotiana benthamiana* using guide RNA and Cas9. *Current Biology* **23**: 42–48.
- Li, J.-F., Bush, J., Xiong, Y., Li, L., and McCormack, M.** (2011). Large-Scale Protein-Protein Interaction Analysis in Arabidopsis Mesophyll Protoplasts by Split Firefly Luciferase Complementation. *PLoS ONE* **6**: e27364.
- Liang, G., Zhang, H., Lou, D., and Yu, D.** (2016). Selection of highly efficient sgRNAs for CRISPR/Cas9-based plant genome editing. *Sci Rep* **6**: 21451.
- Liang, Z., Chen, K., Zhang, Y., Liu, J., Yin, K., Qiu, J.-L., and Gao, C.** (2018). Genome editing of bread wheat using biolistic delivery of CRISPR/Cas9 in vitro transcripts or ribonucleoproteins. *Plant Biotechnology Journal* **16**: 183–192.
- Liao, C.-Y., Smet, W., Brunoud, G., Yoshida, S., Vernoux, T., and Weijers, D.** (2015). Reporters for sensitive and quantitative measurement of auxin response. *Nature Methods* **12**: 207–210.
- Lin, Y.-C. et al.** (2014). A simple improved-throughput xylem protoplast system for studying wood formation. *Nat Protoc* **9**: 2194–2205.
- Liu, K., Yuan, C., Feng, S., Zhong, S., Li, H., Zhong, J., Shen, C., and Liu, J.** (2017). Genome-wide analysis and characterization of Aux/IAA family genes related to fruit ripening in papaya (*Carica papaya* L.). *BMC Genomics* **18**.
- 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: microRNA in Arabidopsis seed germination. *The Plant Journal* **52**: 133–146.
- Liu, Q., Wang, C., Jiao, X., Zhang, H., Song, L., Li, Y., Gao, C., and Wang, K.** (2019). Hi-TOM: a platform for high-throughput tracking of mutations induced by CRISPR/Cas systems. *Sci. China Life Sci.* **62**: 1–7.
- Liu, W., Xie, X., Ma, X., Li, J., Chen, J., and Liu, Y.-G.** (2015). DSDecode: A Web-Based Tool for Decoding of Sequencing Chromatograms for Genotyping of Targeted Mutations. *Molecular Plant* **8**: 1431–1433.
- Liu, Z. and Karmarkar, V.** (2008). Groucho/Tup1 family co-repressors in plant development. *Trends in Plant Science* **13**: 137–144.

- Ljung, K., Bhalerao, R.P., and Sandberg, G.** (2002a). Sites and homeostatic control of auxin biosynthesis in *Arabidopsis* during vegetative growth: Auxin biosynthesis and distribution in *Arabidopsis*. *The Plant Journal* **28**: 465–474.
- Ljung, K., Hull, A.K., Celenza, J., Yamada, M., Estelle, M., Normanly, J., and Sandberg, G.** (2005). Sites and Regulation of Auxin Biosynthesis in *Arabidopsis* Roots. *Plant Cell* **17**: 1090–1104.
- Ljung, K., Hull, A.K., Kowalczyk, M., Marchant, A., Celenza, J., Cohen, J.D., and Sandberg, G.** (2002b). Biosynthesis, conjugation, catabolism and homeostasis of indole-3-acetic acid in *Arabidopsis thaliana*. In *Auxin Molecular Biology*, C. Perrot-Rechenmann and G. Hagen, eds (Springer Netherlands: Dordrecht), pp. 249–272.
- Lourenço, A., Rencoret, J., Chemetova, C., Gominho, J., Gutiérrez, A., del Río, J.C., and Pereira, H.** (2016). Lignin Composition and Structure Differs between Xylem, Phloem and Pith in *Quercus suber* L. *Front. Plant Sci.* **7**.
- 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*. *Proceedings of the National Academy of Sciences* **106**: 5984–5989.
- Lu, H., Klocko, A.L., Dow, M., Ma, C., Amarasinghe, V., and Strauss, S.H.** (2016). Low frequency of zinc-finger nuclease-induced mutagenesis in *Populus*. *Mol Breeding* **36**: 121.
- Ludwig-Müller, J.** (2011). Auxin conjugates: their role for plant development and in the evolution of land plants. *Journal of Experimental Botany* **62**: 1757–1773.
- Luo, J., Zhou, J.-J., and Zhang, J.-Z.** (2018). Aux/IAA Gene Family in Plants: Molecular Structure, Regulation, and Function. *IJMS* **19**: 259.
- Ma, X. et al.** (2015). A Robust CRISPR/Cas9 System for Convenient, High-Efficiency Multiplex Genome Editing in Monocot and Dicot Plants. *Molecular Plant* **8**: 1274–1284.
- Ma, X., Zhu, Q., Chen, Y., and Liu, Y.-G.** (2016). CRISPR/Cas9 Platforms for Genome Editing in Plants: Developments and Applications. *Molecular Plant* **9**: 961–974.
- Mach, J.** (2010). Auxin Binding by SKP2A Activates Proteolysis of Downstream Cell Cycle Regulators and Promotes Cell Division. *Plant Cell* **22**: 3877–3877.
- Majda, M. and Robert, S.** (2018). The Role of Auxin in Cell Wall Expansion. *IJMS* **19**: 951.
- Maleki, S.S., Mohammadi, K., Movahedi, A., Wu, F., and Ji, K.S.** (2020). Increase in Cell Wall Thickening and Biomass Production by Overexpression of PmCesA2 in Poplar. *Front. Plant Sci.* **11**: 110.
- Malnoy, M., Viola, R., Jung, M.-H., Koo, O.-J., Kim, S., Kim, J.-S., Velasco, R., and Nagamangala Kanchiswamy, C.** (2016). DNA-Free Genetically Edited Grapevine and Apple Protoplast Using CRISPR/Cas9 Ribonucleoproteins. *Front. Plant Sci.* **7**.
- Mano, Y. and Nemoto, K.** (2012). The pathway of auxin biosynthesis in plants. *Journal of Experimental Botany* **63**: 2853–2872.
- Mao, Y., Zhang, H., Xu, N., Zhang, B., Gou, F., and Zhu, J.-K.** (2013). Application of the CRISPR–Cas System for Efficient Genome Engineering in Plants. *Molecular Plant* **6**: 2008–2011.
- Martin-Arevalillo, R., Nanao, M.H., Larriec, A., Vinos-Poyo, T., Mast, D., Galvan-Ampudia, C., Brunoud, G., Vernoux, T., Dumas, R., and Parcy, F.** (2017). Structure of the *Arabidopsis* TOPLESS corepressor provides insight into the evolution of transcriptional repression. *Proc Natl Acad Sci USA* **114**: 8107–8112.
- Mashiguchi, K. et al.** (2011). The main auxin biosynthesis pathway in *Arabidopsis*. *Proceedings of the National Academy of Sciences* **108**: 18512–18517.

Matchmaker® Gold Yeast Two-Hybrid System User Manual: 41.

- Mattsson, J., Ckurshumova, W., and Berleth, T.** (2003). Auxin Signaling in Arabidopsis Leaf Vascular Development. *Plant Physiol.* **131**: 1327–1339.
- Mauriat, M. et al.** (2014). Wood Formation in Trees. In *Tree biotechnology*, K.G. Ramawat, J.M. Merillon, and M.R. Ahuja, eds (CRC Press), p. 656.
- 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.
- Mauseth, J.D.** (1988). *Plant anatomy*.
- McCarthy, R.L., Zhong, R., Fowler, S., Lyskowski, D., Piyasena, H., Carleton, K., Spicer, C., and Ye, Z.H.** (2010). The poplar MYB transcription factors, PtrMYB3 and PtrMYB20, are involved in the regulation of secondary wall biosynthesis. *Plant and Cell Physiology* **51**: 1084–1090.
- McClure, B.A. and Guilfoyle, T.** (1987). Characterization of a class of small auxin-inducible soybean polyadenylated RNAs. *Plant Mol Biol* **9**: 611–623.
- Mellerowicz, E. and Sundberg, B.** (2008). Wood cell walls: biosynthesis, developmental dynamics and their implications for wood properties. *Current Opinion in Plant Biology* **11**: 293–300.
- Michalko, J., Dravecká, M., Bollenbach, T., and Friml, J.** (2015). Embryo-lethal phenotypes in early *abp1* mutants are due to disruption of the neighboring BSM gene. *F1000Res* **4**: 1104.
- Milhinhos, A. and Miguel, C.M.** (2013). Hormone interactions in xylem development: A matter of signals. *Plant Cell Reports* **32**: 867–883.
- Mironova, V.V., Omelyanchuk, N.A., Wiebe, D.S., and Levitsky, V.G.** (2014). Computational analysis of auxin responsive elements in the Arabidopsis thaliana L. genome. *BMC Genomics* **15**: S4.
- Mitsuda, N., Iwase, A., Yamamoto, H., Yoshida, M., Seki, M., Shinozaki, K., and Ohme-Takagi, M.** (2007). NAC transcription factors, NST1 and NST3, are key regulators of the formation of secondary walls in woody tissues of Arabidopsis. *The Plant Cell* **19**: 270–280.
- Mitsuda, N., Seki, M., Shinozaki, K., and Ohme-Takagi, M.** (2005). The NAC transcription factors NST1 and NST2 of Arabidopsis regulate secondary wall thickenings and are required for anther dehiscence. *The Plant cell* **17**: 2993–3006.
- Mohamed, M.A., Jaafar, J., Ismail, A.F., Othman, M.H.D., and Rahman, M.A.** (2017). Fourier Transform Infrared (FTIR) Spectroscopy. In *Membrane Characterization* (Elsevier), pp. 3–29.
- Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J., and Almendros, C.** (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155**: 733–740.
- Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J., and Soria, E.** (2005). Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. *J Mol Evol* **60**: 174–182.
- Moreno-Piovano, G.S., Moreno, J.E., Cabello, J.V., Arce, A.L., Otegui, M.E., and Chan, R.L.** (2017). A role for LAX2 in regulating xylem development and lateral-vein symmetry in the leaf. *Annals of Botany* **120**: 577–590.
- Moscou, M.J. and Bogdanove, A.J.** (2009). A Simple Cipher Governs DNA Recognition by TAL Effectors. *Science* **326**: 1501–1501.
- Moyle, R., Schrader, J., Stenberg, A., Olsson, O., Saxena, S., Sandberg, G., and Bhalerao, R.P.** (2002). Environmental and auxin regulation of wood formation involves members of the *Aux/IAA* gene family in hybrid aspen. *The Plant*

Journal 31: 675–685.

- Muhr, M., Paulat, M., Awwanah, M., Brinkkötter, M., and Teichmann, T.** (2018). CRISPR/Cas9-mediated knockout of *Populus* BRANCHED1 and BRANCHED2 orthologs reveals a major function in bud outgrowth control. *Tree Physiology* **38**: 1588–1597.
- Müller, C.J., Valdés, A.E., Wang, G., Ramachandran, P., Beste, L., Uddenberg, D., and Carlsbecker, A.** (2016). PHABULOSA Mediates an Auxin Signaling Loop to Regulate Vascular Patterning in Arabidopsis. *Plant Physiol.* **170**: 956–970.
- Müller, M., Burghammer, M., and Sugiyama, J.** (2006). Direct investigation of the structural properties of tension wood cellulose microfibrils using microbeam X-ray fibre diffraction. *Holzforschung* **60**: 474–479.
- Mushtaq, M., Bhat, J.A., Mir, Z.A., Sakina, A., Ali, S., Singh, A.K., Tyagi, A., Salgotra, R.K., Dar, A.A., and Bhat, R.** (2018). CRISPR/Cas approach: A new way of looking at plant-abiotic interactions. *Journal of Plant Physiology* **224–225**: 156–162.
- Myburg, A.A. et al.** (2014). The genome of *Eucalyptus grandis*. *Nature* **510**: 356–362.
- Myburg, A.A., Potts, B.M., Marques, C.M., Kirst, M., Gion, J.-M., Grattapaglia, D., and Grima-Pettenatti, J.** (2007). *Eucalypts*. In *Forest Trees* (Springer), pp. 115–160.
- Nakano, Y., Yamaguchi, M., Endo, H., Rejab, N.A., and Ohtani, M.** (2015). NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants. *Frontiers in plant science* **6**: 288.
- Nanao, M.H. et al.** (2014). Structural basis for oligomerization of auxin transcriptional regulators. *Nat Commun* **5**: 3617.
- Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J.D.G., and Kamoun, S.** (2013). Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nature Biotechnology* **31**: 691–693.
- Nguyen, N.H. and Cheong, J.-J.** (2018). AtMYB44 interacts with TOPLESS-RELATED corepressors to suppress protein phosphatase 2C gene transcription. *Biochemical and Biophysical Research Communications* **507**: 437–442.
- Nicol, F.** (1998). A plasma membrane-bound putative endo-1,4-beta -D-glucanase is required for normal wall assembly and cell elongation in Arabidopsis. *The EMBO Journal* **17**: 5563–5576.
- Nieminen, K. et al.** (2008). Cytokinin signaling regulates cambial development in poplar. *Proceedings of the National Academy of Sciences* **105**: 20032–20037.
- Nieminen, K., Blomster, T., Helariutta, Y., and Mähönen, A.P.** (2015). Vascular Cambium Development. *The Arabidopsis Book* **13**: e0177.
- Nieminen, K., Robischon, M., Immanen, J., and Helariutta, Y.** (2012). Towards optimizing wood development in bioenergy trees. *New Phytologist* **194**: 46–53.
- Nilsson, J., Karlberg, A., Antti, H., Lopez-Vernaza, M., Mellerowicz, E., Perrot-Rechenmann, C., Sandberg, G., and Bhalerao, R.P.** (2008). Dissecting the Molecular Basis of the Regulation of Wood Formation by Auxin in Hybrid Aspen. *THE PLANT CELL ONLINE* **20**: 843–855.
- Nishitani, C., Hirai, N., Komori, S., Wada, M., Okada, K., Osakabe, K., Yamamoto, T., and Osakabe, Y.** (2016). Efficient Genome Editing in Apple Using a CRISPR/Cas9 system. *Sci Rep* **6**: 31481.
- Nixon, B.T. et al.** (2016). Comparative Structural and Computational Analysis Supports Eighteen Cellulose Synthases in the Plant Cellulose Synthesis Complex. *Sci Rep* **6**: 28696.
- Normanly, J., Cohen, J.D., and Fink, G.R.** (1993). Arabidopsis thaliana auxotrophs reveal a tryptophan-independent biosynthetic pathway for indole-3-acetic acid. *Proceedings of the National Academy of Sciences* **90**: 10355–10359.

- Normanly, J., Slovin, J.P., and Cohen, J.D.** (2010). Auxin biosynthesis and metabolism. In *Plant Hormones* (Springer), pp. 36–62.
- Oeller, P.W., Keller, J.A., Parks, J.E., Silbert, J.E., and Theologis, A.** (1993). Structural characterization of the early indoleacetic acid-inducible genes, PS-IAA4/5 and PS-IAA6, of pea (*Pisum sativum* L.). *Journal of molecular biology* **233**: 789–798.
- Oh, S., Park, S., and Han, K.-H.** (2003). Transcriptional regulation of secondary growth in *Arabidopsis thaliana*. *Journal of Experimental Botany* **54**: 2709–2722.
- Ohashi-Ito, K., Oda, Y., and Fukuda, H.** (2010). *Arabidopsis* VASCULAR-RELATED NAC-DOMAIN6 Directly Regulates the Genes That Govern Programmed Cell Death and Secondary Wall Formation during Xylem Differentiation. *Plant Cell* **22**: 3461–3473.
- Öhman, D., Demedts, B., Kumar, M., Gerber, L., Gorzsás, A., Goeminne, G., Hedenström, M., Ellis, B., Boerjan, W., and Sundberg, B.** (2013). MYB103 is required for *FERULATE-5-HYDROXYLASE* expression and syringyl lignin biosynthesis in *Arabidopsis* stems. *Plant J* **73**: 63–76.
- Ohtani, M., Nishikubo, N., Xu, B., Yamaguchi, M., Mitsuda, N., Goué, N., Shi, F., Ohme-Takagi, M., and Demura, T.** (2011). A NAC domain protein family contributing to the regulation of wood formation in poplar. *The Plant Journal* **67**: 499–512.
- O'Malley, R.C., Huang, S.C., Song, L., Lewsey, M.G., Bartlett, A., Nery, J.R., Galli, M., Gallavotti, A., and Ecker, J.R.** (2016). Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell* **165**: 1280–1292.
- Paponov, I.A., Teale, W., Lang, D., Paponov, M., Reski, R., Rensing, S.A., and Palme, K.** (2009). The evolution of nuclear auxin signalling. *BMC Evol Biol* **9**: 126.
- Park, J., Lim, K., Kim, J.-S., and Bae, S.** (2017). Cas-analyzer: an online tool for assessing genome editing results using NGS data. *Bioinformatics* **33**: 286–288.
- Park, S.-H., Chung, P.J., Juntawong, P., Bailey-Serres, J., Kim, Y.S., Jung, H., Bang, S.W., Kim, Y.-K., Do Choi, Y., and Kim, J.-K.** (2012). Posttranscriptional Control of Photosynthetic mRNA Decay under Stress Conditions Requires 3' and 5' Untranslated Regions and Correlates with Differential Polysome Association in Rice. *Plant Physiol.* **159**: 1111–1124.
- Parry, G., Calderon-Villalobos, L.I., Prigge, M., Peret, B., Dharmasiri, S., Itoh, H., Lechner, E., Gray, W.M., Bennett, M., and Estelle, M.** (2009). Complex regulation of the TIR1/AFB family of auxin receptors. *PNAS* **106**: 22540–22545.
- Pauly, M., Gille, S., Liu, L., Mansoori, N., de Souza, A., Schultink, A., and Xiong, G.** (2013). Hemicellulose biosynthesis. *Planta* **238**: 627–642.
- Pauwels, L. et al.** (2010). NINJA connects the co-repressor TOPLESS to jasmonate signalling. *Nature* **464**: 788–791.
- Paux, E., Carocha, V., Marques, C., Mendes de Sousa, A., Borralho, N., Sivadon, P., and Grima-Pettenati, J.** (2005). Transcript profiling of *Eucalyptus* xylem genes during tension wood formation. *New Phytologist* **167**: 89–100.
- 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.
- Peer, R., Rivlin, G., Golobovitch, S., Lapidot, M., Gal-On, A., Vainstein, A., Tzfira, T., and Flaishman, M.A.** (2015). Targeted mutagenesis using zinc-finger nucleases in perennial fruit trees. *Planta* **241**: 941–951.
- Pěňčík, A. et al.** (2013). Regulation of Auxin Homeostasis and Gradients in *Arabidopsis* Roots through the Formation of the Indole-3-Acetic Acid Catabolite 2-Oxindole-3-Acetic Acid. *Plant Cell* **25**: 3858–3870.

- Peng, A., Chen, S., Lei, T., Xu, L., He, Y., Wu, L., Yao, L., and Zou, X. (2017). Engineering canker-resistant plants through CRISPR/Cas9-targeted editing of the susceptibility gene *CsLOB1* promoter in citrus. *Plant Biotechnol J* **15**: 1509–1519.
- Pierre-Jerome, E., Moss, B.L., Lanctot, A., Hageman, A., and Nemhauser, J.L. (2016). Functional analysis of molecular interactions in synthetic auxin response circuits. *Proc Natl Acad Sci USA* **113**: 11354–11359.
- Pinello, L., Canver, M.C., Hoban, M.D., Orkin, S.H., Kohn, D.B., Bauer, D.E., and Yuan, G.-C. (2016). Analyzing CRISPR genome-editing experiments with CRISPResso. *Nat Biotechnol* **34**: 695–697.
- Plasencia, A., Soler, M., Dupas, A., Ladouce, N., Silva-Martins, G., Martinez, Y., Lapierre, C., Franche, C., Truchet, I., and Grima-Pettenati, J. (2016). *Eucalyptus* hairy roots, a fast, efficient and versatile tool to explore function and expression of genes involved in wood formation. *Plant Biotechnology Journal* **14**: 1381–1393.
- Porco, S. et al. (2016). Dioxygenase-encoding *AtDAO1* gene controls IAA oxidation and homeostasis in *Arabidopsis*. *Proc Natl Acad Sci USA* **113**: 11016–11021.
- del Pozo, J.C., Diaz-Trivino, S., Cisneros, N., and Gutierrez, C. (2006). The Balance between Cell Division and Endoreplication Depends on E2FC-DPB, Transcription Factors Regulated by the Ubiquitin-SCF^{SKP2A} Pathway in *Arabidopsis*. *Plant Cell* **18**: 2224–2235.
- Przemeck, Gerhard K.H., Mattsson, J., Hardtke, Christian 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**.
- Puchta, H. and Fauser, F. (2014). Synthetic nucleases for genome engineering in plants: prospects for a bright future. *Plant J* **78**: 727–741.
- Qiu, T., Qi, M., Ding, X., Zheng, Y., Zhou, T., Chen, Y., Han, N., Zhu, M., Bian, H., and Wang, J. (2020). The SAUR41 subfamily of SMALL AUXIN UP RNA genes is abscisic acid inducible to modulate cell expansion and salt tolerance in *Arabidopsis thaliana* seedlings. *Annals of Botany* **125**: 805–819.
- Ralph, J., Lapierre, C., and Boerjan, W. (2019). Lignin structure and its engineering. *Current Opinion in Biotechnology* **56**: 240–249.
- Ralph, J., Lundquist, K., Brunow, G., Lu, F., Kim, H., Schatz, P.F., Marita, J.M., Hatfield, R.D., Ralph, S.A., Christensen, J.H., and Boerjan, W. (2004). Lignins: Natural polymers from oxidative coupling of 4-hydroxyphenyl-propanoids. *Phytochemistry Reviews* **3**: 29–60.
- Ramachandran, P., Carlsbecker, A., and EtcHELLS, J.P. (2017). Class III HD-ZIPs govern vascular cell fate: an HD view on patterning and differentiation. *EXBOTJ* **68**: 55–69.
- Ramos, J.A., Zenser, N., Leyser, O., and Callis, J. Rapid Degradation of Auxin/Indoleacetic Acid Proteins Requires Conserved Amino Acids of Domain II and Is Proteasome Dependent.: 13.
- Ran, F.A., Hsu, P.D., Lin, C.-Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., and Zhang, F. (2013). Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. *Cell* **154**: 1380–1389.
- Ranocha, P. et al. (2013). *Arabidopsis* WAT1 is a vacuolar auxin transport facilitator required for auxin homeostasis. *Nat Commun* **4**: 2625.
- Ranocha, P., Denancé, N., Vanholme, R., Freyrier, A., Martinez, Y., Hoffmann, L., Köhler, L., Pouzet, C., Renou, J.-P., Sundberg, B., Boerjan, W., and Goffner, D. (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: Tonoplastic WAT1 and secondary wall formation. *The Plant Journal* **63**: 469–483.
- Ratke, C. et al. (2018). Downregulating aspen xylan biosynthetic GT43 genes in developing wood stimulates growth via

- reprogramming of the transcriptome. *New Phytologist* **219**: 230–245.
- Raymond, C.A.** (2000). Tree breeding issues for solid wood production. In *The Future of Eucalypts for Wood Products. Proceedings of an IUFRO Conference, Launceston, Tasmania*, pp. 19–24.
- Ren, C., Liu, X., Zhang, Z., Wang, Y., Duan, W., Li, S., and Liang, Z.** (2016). CRISPR/Cas9-mediated efficient targeted mutagenesis in Chardonnay (*Vitis vinifera* L.). *Scientific reports* **6**: 32289.
- Ren, F., Ren, C., Zhang, Z., Duan, W., Lecourieux, D., Li, S., and Liang, Z.** (2019). Efficiency Optimization of CRISPR/Cas9-Mediated Targeted Mutagenesis in Grape. *Front. Plant Sci.* **10**: 612.
- Renault, H., Werck-Reichhart, D., and Weng, J.-K.** (2019). Harnessing lignin evolution for biotechnological applications. *Current Opinion in Biotechnology* **56**: 105–111.
- Rencoret, J., Gutiérrez, A., Nieto, L., Jiménez-Barbero, J., Faulds, C.B., Kim, H., Ralph, J., Martínez, Á.T., and del Río, J.C.** (2011). Lignin Composition and Structure in Young versus Adult *Eucalyptus globulus* Plants. *Plant Physiol.* **155**: 667–682.
- Richard Wenck, A., Quinn, M., Whetten, R.W., Pullman, G., and Sederoff, R.** (1999). High-efficiency *Agrobacterium*-mediated transformation of Norway spruce (*Picea abies*) and loblolly pine (*Pinus taeda*). *Plant Molecular Biology* **39**: 407–416.
- Robert, S. et al.** (2010). ABP1 Mediates Auxin Inhibition of Clathrin-Dependent Endocytosis in *Arabidopsis*. *Cell* **143**: 111–121.
- Roginsky, J.** (2018). Analyzing CRISPR Editing Results: Synthego Developed a Tool Called ICE to Be More Efficient Than Other Methods. *Genetic Engineering & Biotechnology News* **38**: S24–S26.
- Ron, M. et al.** (2014). Hairy Root Transformation Using *Agrobacterium rhizogenes* as a Tool for Exploring Cell Type-Specific Gene Expression and Function Using Tomato as a Model. *PLANT PHYSIOLOGY* **166**: 455–469.
- Ruonala, R., Ko, D., and Helariutta, Y.** (2017). Genetic Networks in Plant Vascular Development. *Annual Review of Genetics* **51**: 335–359.
- Rushton, P.J., Somssich, I.E., Ringler, P., and Shen, Q.J.** (2010). WRKY transcription factors. *Trends in Plant Science* **15**: 247–258.
- Ryder, P., McHale, M., Fort, A., and Spillane, C.** (2017). Generation of stable nulliplex autopolyploid lines of *Arabidopsis thaliana* using CRISPR/Cas9 genome editing. *Plant Cell Rep* **36**: 1005–1008.
- Ryu, H., Cho, H., Bae, W., and Hwang, I.** (2014). Control of early seedling development by BES1/TPL/HDA19-mediated epigenetic regulation of ABI3. *Nat Commun* **5**: 4138.
- Salehin, M., Bagchi, R., and Estelle, M.** (2015). SCF^{TIR1/AFB}-Based Auxin Perception: Mechanism and Role in Plant Growth and Development. *The Plant Cell Online* **27**: 9–19.
- Sander, J.D., Maeder, M.L., Reyon, D., Voytas, D.F., Joung, J.K., and Dobbs, D.** (2010). ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. *Nucleic Acids Research* **38**: W462–W468.
- Sankar, M., Nieminen, K., Ragni, L., Xenarios, I., and Hardtke, C.S.** (2014). Automated quantitative histology reveals vascular morphodynamics during *Arabidopsis* hypocotyl secondary growth. *eLife* **3**: e01567.
- dos Santos Maraschin, F., Memelink, J., and Offringa, R.** (2009). Auxin-induced, SCF^{TIR1}-mediated polyubiquitination marks AUX/IAA proteins for degradation. *The Plant Journal* **59**: 100–109.
- 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.

- Sato, S. et al.** (2001). Role of the Putative Membrane-Bound Endo-1,4- β -Glucanase KORRIGAN in Cell Elongation and Cellulose Synthesis in *Arabidopsis thaliana*. *Plant and Cell Physiology* **42**: 251–263.
- Sauer, M., Balla, J., Luschnig, C., Wisniewska, J., Reinohl, V., Friml, J., and Benkova, E.** (2006). Canalization of auxin flow by Aux/IAA-ARF-dependent feedback regulation of PIN polarity. *Genes & Development* **20**: 2902–2911.
- Savidge, R.A.** (1983). The role of plant hormones in higher plant cellular differentiation. II. Experiments with the vascular cambium, and sclereid and tracheid differentiation in the pine, *Pinus contorta*. *Histochem J* **15**: 447–466.
- Scarpella, E., Barkoulas, M., and Tsiantis, M.** (2010). Control of Leaf and Vein Development by Auxin. *Cold Spring Harbor Perspectives in Biology* **2**: a001511–a001511.
- Scheller, H.V. and Ulvskov, P.** (2010). Hemicelluloses. *Annu. Rev. Plant Biol.* **61**: 263–289.
- Scheres, B., Wolkenfelt, H., Willemsen, V., Terlouw, M., Lawson, E., Dean, C., and Weisbeek, P.** (1994). Embryonic origin of the *Arabidopsis* primary root and root meristem initials. *Development* **120**: 2475–2487.
- Schuetz, M., Smith, R., and Ellis, B.** (2013). Xylem tissue specification, patterning, and differentiation mechanisms. *Journal of experimental botany* **64**: 11–31.
- Seyfferth, C., Wessels, B., Jokipii-Lukkari, S., Sundberg, B., Delhomme, N., Felten, J., and Tuominen, H.** (2018). Ethylene-Related Gene Expression Networks in Wood Formation. *Frontiers in Plant Science* **9**.
- Sghaier, N., Ben Ayed, R., Ben Marzoug, R., and Rebai, A.** (2018). Dempster-Shafer Theory for the Prediction of Auxin-Response Elements (AuxREs) in Plant Genomes. *BioMed Research International* **2018**: 1–13.
- Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., Zhang, K., Liu, J., Xi, J.J., Qiu, J.-L., and Gao, C.** (2013). Targeted genome modification of crop plants using a CRISPR-Cas system. *Nat Biotechnol* **31**: 686–688.
- Shen, Y., Li, Y., Xu, D., Yang, C., Li, C., and Luo, K.** (2018). Molecular cloning and characterization of a brassinosteroid biosynthesis-related gene *PtoDWF4* from *Populus tomentosa*. *Tree Physiology* **38**: 1424–1436.
- Shepherd, M., Bartle, J., Lee, D.J., Brawner, J., Bush, D., Turnbull, P., Macdonel, P., Brown, T.R., Simmons, B., and Henry, R.** (2011). Eucalypts as a biofuel feedstock. *Biofuels* **2**: 639–657.
- Shi, D., Lebovka, I., López-Salmerón, V., Sanchez, P., and Greb, T.** (2019). Bifacial cambium stem cells generate xylem and phloem during radial plant growth. *Development* **146**: dev171355.
- Shinya, T., Iwata, E., Nakahama, K., Fukuda, Y., Hayashi, K., Nanto, K., Rosa, A.C., and Kawaoka, A.** (2016). Transcriptional Profiles of Hybrid Eucalyptus Genotypes with Contrasting Lignin Content Reveal That Monolignol Biosynthesis-related Genes Regulate Wood Composition. *Front. Plant Sci.* **7**.
- Shu, W., Liu, Y., Guo, Y., Zhou, H., Zhang, J., Zhao, S., and Lu, M.** (2015). A *Populus* TIR1 gene family survey reveals differential expression patterns and responses to 1-naphthaleneacetic acid and stress treatments. *Front. Plant Sci.* **6**: 719.
- Shu, W., Zhou, H., Jiang, C., Zhao, S., Wang, L., Li, Q., Yang, Z., Groover, A., and Lu, M.-Z.** (2019). The auxin receptor TIR1 homolog (*PagFBL 1*) regulates adventitious rooting through interactions with *Aux/IAA28* in *Populus*. *Plant Biotechnol J* **17**: 338–349.
- Si, J., Sun, Y., Wang, L., Qin, Y., Wang, C., and Wang, X.** (2016). Functional analyses of *Populus euphratica* brassinosteroid biosynthesis enzyme genes *DWF4* (*PeDWF4*) and *CPD* (*PeCPD*) in the regulation of growth and development of *Arabidopsis thaliana*. *J Biosci* **41**: 727–742.
- Simon, S. and Petrášek, J.** (2011). Why plants need more than one type of auxin. *Plant Science* **180**: 454–460.

- Smet, W. and De Rybel, B.** (2016). Genetic and hormonal control of vascular tissue proliferation. *Current Opinion in Plant Biology* **29**: 50–56.
- Smetana, O. et al.** (2019). High levels of auxin signalling define the stem-cell organizer of the vascular cambium. *Nature* **565**: 485–489.
- Smith, R.A., Schuetz, M., Karlen, S.D., Bird, D., Tokunaga, N., Sato, Y., Mansfield, S.D., Ralph, J., and Samuels, A.L.** (2017). Defining the Diverse Cell Populations Contributing to Lignification in Arabidopsis Stems. *Plant Physiol.* **174**: 1028–1036.
- Soler, M., Camargo, E.L.O., Carocha, V., Cassan-Wang, H., San Clemente, H., Savelli, B., Hefer, C.A., Paiva, J.A.P., Myburg, A.A., and Grima-Pettenati, J.** (2015). The *Eucalyptus grandis* R2R3-MYB transcription factor family: evidence for woody growth-related evolution and function. *New Phytologist* **206**: 1364–1377.
- Soler, M., Plasencia, A., Larbat, R., Pouzet, C., Jauneau, A., Rivas, S., Pesquet, E., Lapierre, C., Truchet, I., and Grima-Pettenati, J.** (2017). The *Eucalyptus* linker histone variant EgH1.3 cooperates with the transcription factor EgMYB1 to control lignin biosynthesis during wood formation. *New Phytologist* **213**: 287–299.
- Song, C. et al.** (2020). Genome-wide identification and expression profiling of the YUCCA gene family in *Malus domestica*. *Sci Rep* **10**: 10866.
- Song, C., Zhang, D., Zhang, J., Zheng, L., Zhao, C., Ma, J., An, N., and Han, M.** (2016). Expression analysis of key auxin synthesis, transport, and metabolism genes in different young dwarfing apple trees. *Acta Physiol Plant* **38**: 43.
- 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.
- 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.
- Soyano, T., Thitamadee, S., Machida, Y., and Chua, N.-H.** (2008). *ASYMMETRIC LEAVES2-LIKE19/LATERAL ORGAN BOUNDARIES DOMAIN30* and *ASL20/LBD18* Regulate Tracheary Element Differentiation in *Arabidopsis*. *Plant Cell* **20**: 3359–3373.
- Spartz, A.K., Ren, H., Park, M.Y., Grandt, K.N., Lee, S.H., Murphy, A.S., Sussman, M.R., Overvoorde, P.J., and Gray, W.M.** SAUR Inhibition of PP2C-D Phosphatases Activates Plasma Membrane H⁺-ATPases to Promote Cell Expansion in ArabidopsisC W.: 15.
- Staswick, P.E.** (2009). The Tryptophan Conjugates of Jasmonic and Indole-3-Acetic Acids Are Endogenous Auxin Inhibitors. *Plant Physiol.* **150**: 1310–1321.
- Staswick, P.E., Serban, B., Rowe, M., Tiryaki, I., Maldonado, M.T., Maldonado, M.C., and Suza, W.** (2005). Characterization of an Arabidopsis Enzyme Family That Conjugates Amino Acids to Indole-3-Acetic Acid. *Plant Cell* **17**: 616–627.
- Staswick, P.E., Tiryaki, I., and Rowe, M.L.** (2002). Jasmonate Response Locus *JAR1* and Several Related Arabidopsis Genes Encode Enzymes of the Firefly Luciferase Superfamily That Show Activity on Jasmonic, Salicylic, and Indole-3-Acetic Acids in an Assay for Adenylation. *Plant Cell* **14**: 1405–1415.
- Stepanova, A.N., Robertson-Hoyt, J., Yun, J., Benavente, L.M., Xie, D.-Y., Doležal, K., Schlereth, A., Jürgens, G., and Alonso, J.M.** (2008). TAA1-Mediated Auxin Biosynthesis Is Essential for Hormone Crosstalk and Plant Development. *Cell* **133**: 177–191.
- Subburaj, S., Chung, S.J., Lee, C., Ryu, S.-M., Kim, D.H., Kim, J.-S., Bae, S., and Lee, G.-J.** (2016). Site-directed mutagenesis in *Petunia × hybrida* protoplast system using direct delivery of purified recombinant Cas9 ribonucleoproteins. *Plant Cell Rep* **35**: 1535–1544.

- Sun, X., Hu, Z., Chen, R., Jiang, Q., Song, G., Zhang, H., and Xi, Y.** (2015). Targeted mutagenesis in soybean using the CRISPR-Cas9 system. *Sci Rep* **5**: 10342.
- Sundell, D. et al.** (2017). AspWood: High-Spatial-Resolution Transcriptome Profiles Reveal Uncharacterized Modularity of Wood Formation in *Populus tremula*. *The Plant Cell* **29**: 1585–1604.
- Suzuki, S., Li, L., Sun, Y.-H., and Chiang, V.L.** (2006). The Cellulose Synthase Gene Superfamily and Biochemical Functions of Xylem-Specific Cellulose Synthase-Like Genes in *Populus trichocarpa*. *Plant Physiol.* **142**: 1233–1245.
- Svitashev, S., Schwartz, C., Lenderts, B., Young, J.K., and Mark Cigan, A.** (2016). Genome editing in maize directed by CRISPR–Cas9 ribonucleoprotein complexes. *Nat Commun* **7**: 13274.
- Swarup, R. and Bhosale, R.** (2019). Developmental Roles of AUX1/LAX Auxin Influx Carriers in Plants. *Front. Plant Sci.* **10**: 1306.
- Swarup, R. and Péret, B.** (2012). AUX/LAX family of auxin influx carriers—an overview. *Front. Plant Sci.* **3**.
- 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.J., 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. *Plant J* **37**: 730–740.
- Takase, T., Nakazawa, M., Ishikawa, A., Manabe, K., and Matsui, M.** (2003). DFL2, a New Member of the Arabidopsis GH3 Gene Family, is Involved in Red Light-Specific Hypocotyl Elongation. *Plant and Cell Physiology* **44**: 1071–1080.
- Takata, N., Awano, T., Nakata, M.T., Sano, Y., Sakamoto, S., Mitsuda, N., and Taniguchi, T.** (2019). Populus NST/SND orthologs are key regulators of secondary cell wall formation in wood fibers, phloem fibers and xylem ray parenchyma cells. *Tree Physiology* **39**: 514–525.
- Tan, X., Calderon-Villalobos, L.I.A., Sharon, M., Zheng, C., Robinson, C.V., Estelle, M., and Zheng, N.** (2007). Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature* **446**: 640–645.
- Tang, X., Zhuang, Y., Qi, G., Wang, D., Liu, H., Wang, K., Chai, G., and Zhou, G.** (2015). Poplar PdMYB221 is involved in the direct and indirect regulation of secondary wall biosynthesis during wood formation. *Scientific reports* **5**: 12240.
- Tao, Y. et al.** (2008). Rapid Synthesis of Auxin via a New Tryptophan-Dependent Pathway Is Required for Shade Avoidance in Plants. *Cell* **133**: 164–176.
- 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* **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**: 2529–2539.
- Taylor, N.G., Scheible, W.-R., Cutler, S., Somerville, C.R., and Turner, S.R.** (1999). The irregular xylem3 Locus of Arabidopsis Encodes a Cellulose Synthase Required for Secondary Cell Wall Synthesis. *The Plant Cell* **11**: 769–779.
- Taylor-Teeples, M. et al.** (2015). An Arabidopsis gene regulatory network for secondary cell wall synthesis. *Nature* **517**: 571–575.
- Timell, T.E.** (1967). Recent progress in the chemistry of wood hemicelluloses. *Wood Sci. Technol.* **1**: 45–70.

- Tiwari, S.B., Hagen, G., and Guilfoyle, T.** (2003). The Roles of Auxin Response Factor Domains in Auxin-Responsive Transcription. *Plant Cell* **15**: 533–543.
- Tiwari, S.B., Hagen, G., and Guilfoyle, T.J.** (2004). Aux/IAA Proteins Contain a Potent Transcriptional Repression Domain. *Plant Cell* **16**: 533–543.
- Tonn, N. and Greb, T.** (2017). Radial plant growth. *Current Biology* **27**: R878–R882.
- de la Torre, F., Rodríguez, R., Jorge, G., Villar, B., Álvarez-Otero, R., Grima-Pettenati, J., and Gallego, P.P.** (2014). Genetic transformation of *Eucalyptus globulus* using the vascular-specific EgCCR as an alternative to the constitutive CaMV35S promoter. *Plant Cell Tiss Organ Cult* **117**: 77–84.
- Tournier, V., Grat, S., Marque, C., El Kayal, W., Penchel, R., de Andrade, G., Boudet, A.-M., and Teulière, C.** (2003). An Efficient Procedure to Stably Introduce Genes into an Economically Important Pulp Tree (*Eucalyptus grandis* × *Eucalyptus urophylla*). *Transgenic Research* **12**: 403–411.
- Trenner, J., Poeschl, Y., Grau, J., Gogol-Döring, A., Quint, M., and Delker, C.** (2016). Auxin-induced expression divergence between *Arabidopsis* species may originate within the TIR1/AFB–AUX/IAA–ARF module. *EXBOTJ*: erw457.
- Tronchet, M., Balaguá, C., Kroj, T., Jouanin, L., and Roby, D.** (2010). Cinnamyl alcohol dehydrogenases-C and D, key enzymes in lignin biosynthesis, play an essential role in disease resistance in *Arabidopsis*. *Molecular Plant Pathology* **11**: 83–92.
- Tsai, C.-J. and Xue, L.-J.** (2015). CRISPRing into the woods. *GM Crops & Food* **6**: 206–215.
- Turner, S., Gallois, P., and Brown, D.** (2007). Tracheary Element Differentiation. *Annu. Rev. Plant Biol.* **58**: 407–433.
- Turner, S.R. and Somerville, C.R.** (1997). Collapsed xylem phenotype of *Arabidopsis* identifies mutants deficient in cellulose deposition in the secondary cell wall. *The Plant Cell* **9**: 689–701.
- Twumasi, P., Schel, J.H.N., van Ieperen, W., Woltering, E., Van Kooten, O., and Emons, A.M.C.** (2009). Establishing in vitro *Zinnia elegans* cell suspension culture with high tracheary element differentiation. *Cell Biology International* **33**: 524–533.
- Ulmasov, T., Hagen, G., and Guilfoyle, T.J.** (1999). Dimerization and DNA binding of auxin response factors. *Plant J* **19**: 309–319.
- 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. *The Plant Cell* **9**: 1963–1971.
- Vanholme, R. et al.** (2013). Caffeoyl Shikimate Esterase (CSE) Is an Enzyme in the Lignin Biosynthetic Pathway in *Arabidopsis*. *Science* **341**: 1103–1106.
- Vanholme, R., De Meester, B., Ralph, J., and Boerjan, W.** (2019). Lignin biosynthesis and its integration into metabolism. *Current Opinion in Biotechnology* **56**: 230–239.
- Vanholme, R., Demedts, B., Morreel, K., Ralph, J., and Boerjan, W.** (2010). Lignin Biosynthesis and Structure. *Plant Physiol.* **153**: 895–905.
- Vanneste, S. and Friml, J.** (2009). Auxin: A Trigger for Change in Plant Development. *Cell* **136**: 1005–1016.
- Verma, P., Sharma, M.P., and Dwivedi, G.** (2016). Potential use of eucalyptus biodiesel in compressed ignition engine. *Egyptian Journal of Petroleum* **25**: 91–95.
- Vernoux, T. et al.** (2011). The auxin signalling network translates dynamic input into robust patterning at the shoot apex. *Mol Syst Biol* **7**: 508.

- Voytas, D.F. (2013). Plant Genome Engineering with Sequence-Specific Nucleases. *Annual Review of Plant Biology* **64**: 327–350.
- Wan, S., Li, C., Ma, X., and Luo, K. (2017). PtrMYB57 contributes to the negative regulation of anthocyanin and proanthocyanidin biosynthesis in poplar. *Plant Cell Rep* **36**: 1263–1276.
- Wang, B., Chu, J., Yu, T., Xu, Q., Sun, X., Yuan, J., Xiong, G., Wang, G., Wang, Y., and Li, J. (2015). Tryptophan-independent auxin biosynthesis contributes to early embryogenesis in *Arabidopsis*. *Proc Natl Acad Sci USA* **112**: 4821–4826.
- Wang, H. (2020). Regulation of vascular cambium activity. *Plant Science* **291**: 110322.
- Wang, H. (2005). The Tomato Aux/IAA Transcription Factor IAA9 Is Involved in Fruit Development and Leaf Morphogenesis. *THE PLANT CELL ONLINE* **17**: 2676–2692.
- Wang, H., Tang, R., Wang, C., Qi, Q., Gai, Y., Jiang, X., and Zhang, H. (2014). Functional repression of PtSND2 represses growth and development by disturbing auxin biosynthesis, transport and signaling in transgenic poplar. *Tree Physiology* **35**: 95–105.
- Wang, H., Zhao, Q., Chen, F., Wang, M., and Dixon, R.A. (2011). NAC domain function and transcriptional control of a secondary cell wall master switch: NST1 function and regulation. *The Plant Journal* **68**: 1104–1114.
- Wang, L., Chen, S., Peng, A., Xie, Z., He, Y., and Zou, X. (2019). CRISPR/Cas9-mediated editing of CsWRKY22 reduces susceptibility to *Xanthomonas citri* subsp. *citri* in Wanjincheng orange (*Citrus sinensis* (L.) Osbeck). *Plant Biotechnol Rep* **13**: 501–510.
- Wang, P., Lu, S., Xie, M., Wu, M., Ding, S., Khaliq, A., Ma, Z., Mao, J., and Chen, B. (2020). Identification and expression analysis of the small auxin-up RNA (SAUR) gene family in apple by inducing of auxin. *Gene* **750**: 144725.
- Wang, R. and Estelle, M. (2014). Diversity and specificity: auxin perception and signaling through the TIR1/AFB pathway. *Current Opinion in Plant Biology* **21**: 51–58.
- Wang, X., Guo, R., Tu, M., Wang, D., Guo, C., Wan, R., Li, Z., and Wang, X. (2017). Ectopic Expression of the Wild Grape WRKY Transcription Factor VqWRKY52 in *Arabidopsis thaliana* Enhances Resistance to the Biotrophic Pathogen Powdery Mildew But Not to the Necrotrophic Pathogen *Botrytis cinerea*. *Front. Plant Sci.* **8**.
- Wang, Y., Wang, N., Xu, H., Jiang, S., Fang, H., Su, M., Zhang, Z., Zhang, T., and Chen, X. (2018a). Auxin regulates anthocyanin biosynthesis through the Aux/IAA–ARF signaling pathway in apple. *Hortic Res* **5**: 59.
- Wang, Z., Wang, S., Li, D., Zhang, Q., Li, L., Zhong, C., Liu, Y., and Huang, H. (2018b). Optimized paired-sgRNA/Cas9 cloning and expression cassette triggers high-efficiency multiplex genome editing in kiwifruit. *Plant Biotechnol J* **16**: 1424–1433.
- Wei, L., Yang, B., Jian, H., Zhang, A., Liu, R., Zhu, Y., Ma, J., Shi, X., Wang, R., Li, J., and Xu, X. (2019). Genome-wide identification and characterization of *Gretchen Hagen3* (GH3) family genes in *Brassica napus*. *Genome* **62**: 597–608.
- Weijers, D., Benkova, E., Jäger, K.E., Schlereth, A., Hamann, T., Kientz, M., Wilmoth, J.C., Reed, J.W., and Jürgens, G. (2005). Developmental specificity of auxin response by pairs of ARF and Aux/IAA transcriptional regulators. *EMBO J* **24**: 1874–1885.
- Weng, J.-K. and Chapple, C. (2010). The origin and evolution of lignin biosynthesis: Tansley review. *New Phytologist* **187**: 273–285.
- Went, F. (1926). On growth-accelerating substances in the coleoptile of *Avena sativa*. In *Proc Kon Akad Wetensch Amsterdam*, pp. 10–19.

- Wenzel, C.L., Schuetz, M., Yu, Q., and Mattsson, J. (2007). Dynamics of MONOPTEROS and PIN-FORMED1 expression during leaf vein pattern formation in *Arabidopsis thaliana*: MP and PIN1 expression in *Arabidopsis* leaves. *The Plant Journal* **49**: 387–398.
- Wilmoth, J.C., Wang, S., Tiwari, S.B., Joshi, A.D., Hagen, G., Guilfoyle, T.J., Alonso, J.M., Ecker, J.R., and Reed, J.W. (2005). NPH4/ARF7 and ARF19 promote leaf expansion and auxin-induced lateral root formation: ARF proteins regulate auxin-induced root formation. *The Plant Journal* **43**: 118–130.
- Wiltshire, R. (2004). *Eucalypts*. Encyclopedia of Forest Sciences 4 Volume Set **1**: 2000.
- Won, C., Shen, X., Mashiguchi, K., Zheng, Z., Dai, X., Cheng, Y., Kasahara, H., Kamiya, Y., Chory, J., and Zhao, Y. (2011). Conversion of tryptophan to indole-3-acetic acid by TRYPTOPHAN AMINOTRANSFERASES OF ARABIDOPSIS and YUCCAs in *Arabidopsis*. *Proceedings of the National Academy of Sciences* **108**: 18518–18523.
- Woo, E.-J. (2002). Crystal structure of auxin-binding protein 1 in complex with auxin. *The EMBO Journal* **21**: 2877–2885.
- Woodward, A.W. (2005). Auxin: Regulation, Action, and Interaction. *Annals of Botany* **95**: 707–735.
- Wu, W., Liu, Y., Wang, Y., Li, H., Liu, J., Tan, J., He, J., Bai, J., and Ma, H. (2017). Evolution Analysis of the Aux/IAA Gene Family in Plants Shows Dual Origins and Variable Nuclear Localization Signals. *IJMS* **18**: 2107.
- Wu, X., Ma, W., Mei, C., Chen, X., Yao, Y., Liu, Y., Qin, X., and Yuan, Y. (2020). Description of CRISPR/Cas9 development and its prospect in hepatocellular carcinoma treatment. *J Exp Clin Cancer Res* **39**: 97.
- Xi, W., Song, D., Sun, J., Shen, J., and Li, L. (2017). Formation of wood secondary cell wall may involve two type cellulose synthase complexes in *Populus*. *Plant Mol Biol* **93**: 419–429.
- Xie, K., Minkenberg, B., and Yang, Y. (2015). Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proc Natl Acad Sci USA* **112**: 3570–3575.
- Xie, K. and Yang, Y. (2013). RNA-Guided Genome Editing in Plants Using a CRISPR–Cas System. *Molecular Plant* **6**: 1975–1983.
- Xing, H.-L., Dong, L., Wang, Z.-P., Zhang, H.-Y., Han, C.-Y., Liu, B., Wang, X.-C., and Chen, Q.-J. (2014). A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC plant biology* **14**: 327.
- Xu, C., Shen, Y., He, F., Fu, X., Yu, H., Lu, W., Li, Y., Li, C., Fan, D., Wang, H.C., and Luo, K. (2019). Auxin-mediated Aux/IAA-ARF-HB signaling cascade regulates secondary xylem development in *Populus*. *New Phytologist*.
- Xu, T. et al. (2014). Cell Surface ABP1-TMK Auxin-Sensing Complex Activates ROP GTPase Signaling. *Science* **343**: 1025–1028.
- Xue, L.-J. and Tsai, C.-J. (2015). AGESeq: Analysis of Genome Editing by Sequencing. *Molecular Plant* **8**: 1428–1430.
- Yamaguchi, M., Goué, N., Igarashi, H., Ohtani, M., Nakano, Y., Mortimer, J.C., Nishikubo, N., Kubo, M., Katayama, Y., Kakegawa, K., Dupree, P., and Demura, T. (2010). VASCULAR-RELATED NAC-DOMAIN6 and VASCULAR-RELATED NAC-DOMAIN7 Effectively Induce Transdifferentiation into Xylem Vessel Elements under Control of an Induction System. *Plant Physiol.* **153**: 906–914.
- Yamaguchi, M., Kubo, M., Fukuda, H., and Demura, T. (2008). Vascular-related NAC-DOMAIN7 is involved in the differentiation of all types of xylem vessels in *Arabidopsis* roots and shoots. *The Plant journal: for cell and molecular biology* **55**: 652–664.
- Yamamoto, R., Fujioka, S., Iwamoto, K., Demura, T., Takatsuto, S., Yoshida, S., and Fukuda, H. (2006). Co-Regulation of Brassinosteroid Biosynthesis-Related Genes During Xylem Cell Differentiation. *Plant and Cell Physiology* **48**: 74–83.

- Yamamoto, T. ed** (2015). Targeted genome editing using site-specific nucleases: ZFNs, TALENs, and the CRISPR/Cas9 system (Springer: Tokyo ; New York).
- Yang, L., Zhao, X., Ran, L., Li, C., Fan, D., and Luo, K.** (2017). PtoMYB156 is involved in negative regulation of phenylpropanoid metabolism and secondary cell wall biosynthesis during wood formation in poplar. *Scientific Reports* **7**: 1–14.
- Ye, X., Kang, B., Osburn, L.D., Li, Y., and Zong-Ming (Max), C.** (2009). Identification of the flavin-dependent monooxygenase-encoding YUCCA gene family in *Populus trichocarpa* and their expression in vegetative tissues and in response to hormone and environmental stresses. *Plant Cell Tiss Organ Cult* **97**: 271–283.
- Ye, Z.-H. and Zhong, R.** (2015). Molecular control of wood formation in trees. *Journal of Experimental Botany* **66**: 4119–4131.
- Yi Chou, E., Schuetz, M., Hoffmann, N., Watanabe, Y., Sibout, R., and Samuels, A.L.** (2018). Distribution, mobility, and anchoring of lignin-related oxidative enzymes in *Arabidopsis* secondary cell walls. *Journal of Experimental Botany* **69**: 1849–1859.
- Yin, K., Han, X., Xu, Z., and Xue, H.** (2009). *Arabidopsis* GLP4 is localized to the Golgi and binds auxin in vitro. *Acta Biochimica et Biophysica Sinica* **41**: 478–487.
- Yu, H.** (2014). Caractérisation des familles de facteurs de transcription ARF et Aux/IAA chez l'eucalyptus, rôles dans la formation du bois. PhD Thesis of Université de Toulouse, Université Toulouse III-Paul Sabatier.
- Yu, H., Soler, M., Clemente, H.S., Mila, I., Paiva, J.A.P., Myburg, A.A., Bouzayen, M., Grima-Pettenati, J., and Cassan-Wang, H.** (2015). Comprehensive genome-wide analysis of the Aux/IAA gene family in eucalyptus: Evidence for the role of EgrIAA4 in wood formation. *Plant and Cell Physiology* **56**: 700–714.
- Yu, H., Soler, M., Mila, I., San Clemente, H., Savelli, B., Dunand, C., Paiva, J.A.P., Myburg, A.A., Bouzayen, M., Grima-Pettenati, J., and Cassan-Wang, H.** (2014). Genome-Wide Characterization and Expression Profiling of the AUXIN RESPONSE FACTOR (ARF) Gene Family in *Eucalyptus grandis*. *PLoS ONE* **9**: e108906.
- 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–a001552.
- van Zeijl, A., Wardhani, T.A.K., Seifi Kalhor, M., Rutten, L., Bu, F., Hartog, M., Linders, S., Fedorova, E.E., Bisseling, T., Kohlen, W., and Geurts, R.** (2018). CRISPR/Cas9-Mediated Mutagenesis of Four Putative Symbiosis Genes of the Tropical Tree *Parasponia andersonii* Reveals Novel Phenotypes. *Front. Plant Sci.* **9**: 284.
- Zetsche, B., Gootenberg, J.S., Abudayyeh, O.O., Slaymaker, I.M., Makarova, K.S., Essletzbichler, P., Volz, S.E., Joung, J., van der Oost, J., Regev, A., Koonin, E.V., and Zhang, F.** (2015). Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell* **163**: 759–771.
- Zhang, F., LeBlanc, C., Irish, V.F., and Jacob, Y.** (2017). Rapid and efficient CRISPR/Cas9 gene editing in Citrus using the YAO promoter. *Plant Cell Rep* **36**: 1883–1887.
- Zhang, H. et al.** (2020). Efficient Generation of CRISPR/Cas9-Mediated Homozygous/Biallelic *Medicago truncatula* Mutants Using a Hairy Root System. *Front. Plant Sci.* **11**: 294.
- Zhang, J., Lin, J.E., Harris, C., Campos Mastrotti Pereira, F., Wu, F., Blakeslee, J.J., and Peer, W.A.** (2016a). DAO1 catalyzes temporal and tissue-specific oxidative inactivation of auxin in *Arabidopsis thaliana*. *Proc Natl Acad Sci USA* **113**: 11010–11015.
- Zhang, J. and Peer, W.A.** (2017). Auxin homeostasis: the DAO of catabolism. *Journal of Experimental Botany* **68**: 3145–3154.
- Zhang, J., Xie, M., Tuskan, G.A., Muchero, W., and Chen, J.-G.** (2018a). Recent Advances in the Transcriptional

- Regulation of Secondary Cell Wall Biosynthesis in the Woody Plants. *Front. Plant Sci.* **9**: 1535.
- Zhang, T., Li, R., Xing, J., Yan, L., Wang, R., and Zhao, Y.** (2018b). The YUCCA-Auxin-WOX11 Module Controls Crown Root Development in Rice. *Front. Plant Sci.* **9**: 523.
- Zhang, Y., Ge, X., Yang, F., Zhang, L., Zheng, J., Tan, X., Jin, Z.-B., Qu, J., and Gu, F.** (2015). Comparison of non-canonical PAMs for CRISPR/Cas9-mediated DNA cleavage in human cells. *Sci Rep* **4**: 5405.
- Zhang, Z., Mao, Y., Ha, S., Liu, W., Botella, J.R., and Zhu, J.-K.** (2016b). A multiplex CRISPR/Cas9 platform for fast and efficient editing of multiple genes in Arabidopsis. *Plant Cell Rep* **35**: 1519–1533.
- Zhao, Y.** (2012). Auxin Biosynthesis: A Simple Two-Step Pathway Converts Tryptophan to Indole-3-Acetic Acid in Plants. *Molecular Plant* **5**: 334–338.
- Zhao, Y.** (2010). Auxin Biosynthesis and Its Role in Plant Development. *Annu. Rev. Plant Biol.* **61**: 49–64.
- Zhao, Y., Man, Y., Wen, J., Guo, Y., and Lin, J.** (2019). Advances in Imaging Plant Cell Walls. *Trends in Plant Science* **24**: 867–878.
- Zhao, Z. et al.** (2013). A Role for a Dioxygenase in Auxin Metabolism and Reproductive Development in Rice. *Developmental Cell* **27**: 113–122.
- Zhong, R., Demura, T., and Ye, Z.-H.** (2006). SND1, a NAC Domain Transcription Factor, Is a Key Regulator of Secondary Wall Synthesis in Fibers of Arabidopsis[W]. *The Plant Cell* **18**: 3158–3170.
- Zhong, R., Lee, C., McCarthy, R.L., Reeves, C.K., Jones, E.G., and Ye, Z.-H.** (2011a). Transcriptional activation of secondary wall biosynthesis by rice and maize NAC and MYB transcription factors. *Plant & cell physiology* **52**: 1856–1871.
- Zhong, R., Lee, C., and Ye, Z.-H.** (2010a). Evolutionary conservation of the transcriptional network regulating secondary cell wall biosynthesis. *Trends in Plant Science* **15**: 625–632.
- Zhong, R., Lee, C., and Ye, Z.-H.** (2010b). Functional Characterization of Poplar Wood-Associated NAC Domain Transcription Factors. *Plant Physiology* **152**: 1044–1055.
- Zhong, R., Lee, C., and Ye, Z.-H.** (2010d). Global Analysis of Direct Targets of Secondary Wall NAC Master Switches in Arabidopsis. *Molecular Plant* **3**: 1087–1103.
- Zhong, R., Lee, C., Zhou, J., McCarthy, R.L., and Ye, Z.-H.** (2008). A Battery of Transcription Factors Involved in the Regulation of Secondary Cell Wall Biosynthesis in Arabidopsis. *The Plant Cell* **20**: 2763–2782.
- Zhong, R., McCarthy, R.L., Haghghat, M., and Ye, Z.H.** (2013). The Poplar MYB Master Switches Bind to the SMRE Site and Activate the Secondary Wall Biosynthetic Program during Wood Formation. *PLoS ONE* **8**: 18–21.
- Zhong, R., McCarthy, R.L., Lee, C., and Ye, Z.-H.** (2011b). Dissection of the Transcriptional Program Regulating Secondary Wall Biosynthesis during Wood Formation in Poplar. *Plant Physiol.* **157**: 1452–1468.
- Zhong, R., Morrison, W.H., Freshour, G.D., Hahn, M.G., and Ye, Z.-H.** (2003). Expression of a Mutant Form of Cellulose Synthase AtCesA7 Causes Dominant Negative Effect on Cellulose Biosynthesis. *Plant Physiol.* **132**: 786–795.
- Zhong, R. and Ye, Z.-H.** (2001). Alteration of Auxin Polar Transport in the Arabidopsis *ifl1* Mutants. *Plant Physiol.* **126**: 549–563.
- Zhong, R. and Ye, Z.-H.** (2014). Complexity of the transcriptional network controlling secondary wall biosynthesis. *Plant Science* **229**: 193–207.

- Zhong, R. and Ye, Z.-H.** (2010). The poplar PtrWNDs are transcriptional activators of secondary cell wall biosynthesis. *Plant signaling & behavior* **5**: 469–472.
- Zhou, G.-K., Zhong, R., Richardson, E.A., Morrison, W.H., Nairn, C.J., Wood-Jones, A., and Ye, Z.-H.** (2006). The Poplar Glycosyltransferase GT47C is Functionally Conserved with Arabidopsis Fragile Fiber8. *Plant and Cell Physiology* **47**: 1229–1240.
- Zhou, J., Lee, C., Zhong, R., and Ye, Z.-H.** (2009). MYB58 and MYB63 Are Transcriptional Activators of the Lignin Biosynthetic Pathway during Secondary Cell Wall Formation in *Arabidopsis*. *Plant Cell* **21**: 248–266.
- Zhou, J.-J. and Luo, J.** (2018). The PIN-FORMED Auxin Efflux Carriers in Plants. *IJMS* **19**: 2759.
- Zhou, X., Broadbelt, L.J., and Vinu, R.** (2016). Mechanistic Understanding of Thermochemical Conversion of Polymers and Lignocellulosic Biomass. In *Advances in Chemical Engineering* (Elsevier), pp. 95–198.
- Zhou, X., Jacobs, T.B., Xue, L.-J., Harding, S.A., and Tsai, C.-J.** (2015). Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial *Populus* reveals 4-coumarate: CoA ligase specificity and redundancy. *New Phytologist* **208**: 298–301.
- Zhou, X., Li, W., Mabon, R., and Broadbelt, L.J.** (2017). A Critical Review on Hemicellulose Pyrolysis. *Energy Technol.* **5**: 52–79.
- Zhu, Y., Song, D., Sun, J., Wang, X., and Li, L.** (2013). PtrHB7, a class III HD-Zip Gene, Plays a Critical Role in Regulation of Vascular Cambium Differentiation in *Populus*. *Molecular Plant* **6**: 1331–1343.
- Zuo, J., Niu, Q.-W., Nishizawa, N., Wu, Y., Kost, B., and Chua, N.-H.** (2000). KORRIGAN, an Arabidopsis endo-1, 4- β -glucanase, localizes to the cell plate by polarized targeting and is essential for cytokinesis. *The Plant Cell* **12**: 1137–1152.
- Zwiewka, M., Bilanovičová, V., Seifu, Y.W., and Nodzyński, T.** (2019). The Nuts and Bolts of PIN Auxin Efflux Carriers. *Front. Plant Sci.* **10**: 985.



Supplemental Data

Supplemental Figures and Tables in Chapter I

Table I-S1. ZFNs, TALENs, CRISPR-Cas9 pros and cons comparison

Traits	ZFN	TALEN	CRISPR/Cas9
Backbone origin	Highly prevalent in eukaryotes	Bacteria (<i>Xanthomonas</i> spp.)	Bacteria (<i>Streptococcus pyogenes</i>)
Type of recognition	DNA/protein	DNA/protein	DNA/RNA
DNA binding determinant	Zinc-finger protein (FokI fused)	Transcription-activator like effectors (FokI fused)	sgRNA complementary to the target sequence with Cas9
Nuclease	FokI	FokI	Cas9
Recognition site size	18-36bp	30-40bp	20bp+PAM
Binding specificity	3 nucleotides	1 nucleotide	1:1 nucleotide pairing
Target sequence	guanine-rich	start w/thymine, end w/adenine	end w/PAM sequence (NGG/NAG)
Off-target effects	specific	more specific than ZFNs	Variable
Multiple gene mutations	rarely used	Moderate	Efficient
Cytotoxicity	Variable to high	Low	Low
Time and cost	High	High	Low
Methylation Sensitivity	Unknown	Sensitive to CpG methylation	No
Design availability	More complex	Complex	Simple
Targeting Efficiency	Low	High	High
Ease of delivery	difficult, need to link zinc finger modules together	difficult due to TALE repeat sequences	Easy, design sgRNA and use standard cloning technic

Table I-S2. CRISPR/Cas9 applications in trees using different targets, associated phenotypes and mutated efficiency

tree species	target gene(s)	gene function	knockout traits	T0 Efficiency	reference
citrus	<i>CsPDS</i>	Phytoene Desaturase	albino leaves, chimerism of variegated (white, pale green, and green) leaves	3.2-79.7%	Jia and Wang, 2014, Zhang et al., 2017, Jia et al., 2017a
	<i>CsLOB1</i>	Lateral Organ Boundaries 1, citrus disease susceptibility gene for citrus canker	reduced or no canker symptoms observed in leaves, reduced pustules	23.8-89.4%	Jia et al., 2017b
	<i>CsLOB1 promoter</i>	<i>CsLOB1 cis-element</i>	alleviated canker symptoms	11.5-64.7%	Jia et al., 2016, Peng et al., 2017
	<i>CsWRKY22</i>	transcription factor, negatively correlated with citrus resistance to citrus canker	reduced diseased areas, significantly decreased susceptibility to citrus canker	68.2-85.7%	Wang et al., 2019
apple	<i>MdPDS</i>	Phytoene Desaturase	albino leaves, chimerism of variegated (white, pale green, and green) leaves	13.6-85.1%	Nishitani et al., 2016, Charrier et al., 2019
	<i>MdTFL1.1</i>	Terminal Flower, floral repressor	accelerate flowering, early flowering	100.0%	Charrier et al., 2019
	<i>DIPM1, DIPM2, DIPM4</i>	DspE-interacting proteins from Malus	predicted: resistance to fire blight disease	0.5-6.9%	Malnoy et al., 2016
grape	<i>VvPDS</i>	Phytoene Desaturase	pale or white leaves, chlorophyll deficiency	1.7-86.6%	Nakajima et al., 2017, Ren et al., 2019
	<i>MLO-7</i>	Mildew Locus O	predicted: reduce susceptibility gene to powdery mildew	0.1%	Malnoy et al., 2016
	<i>VqWRKY52</i>	Transcription factor, biotic stress responses	enhance resistance to <i>Botrytis cinerea</i>	31.0%	Wang et al., 2018a
	<i>IdnDH</i>	L-idonate dehydrogenase, tartaric acid (TA) biosynthesis	decrease of TA content	100.0%	Ren et al., 2016
kiwifruit	<i>AcPDS</i>	Phytoene Desaturase	albino phenotype	0-91.7%	Wang et al., 2018
pear	<i>PcTFL1.1(MdTFL1.1)</i>	Terminal Flower, floral repressor	accelerate flowering, early flowering	9.0%	Charrier et al., 2019
cassava	<i>MePDS</i>	Phytoene Desaturase	albino phenotype	90-100%	Odipio et al., 2017
coffee	<i>CcPDS</i>	Phytoene Desaturase	albino phenotype	30.4(28/92)%	Breitler et al., 2018
cacao	<i>TcNPR3</i>	Non-Expressor of Pathogenesis-Related 3, suppressor of defense response	reduced lesion sizes in leaf, significant reduction ($p < 0.05$) in pathogen DNA	27.0%	Fister et al., 2018
rubber tree	<i>HbFT1, HbFT2</i>	Flowering Locus T, florigen signal for flowering initiation	delayed-flowering	2.7-5.6%	Fan et al., 2020

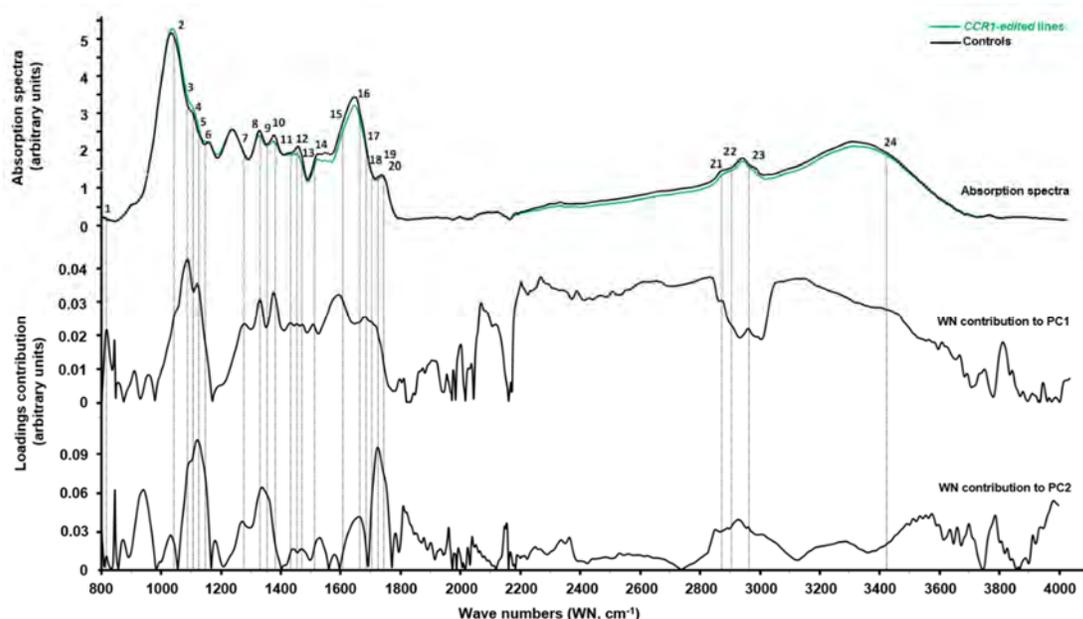
Supplemental Data

	<i>HbTFL1-1, HbTFL1-2, HbTFL1-3</i>	Terminal Flower, floral repressor	accelerate flowering, early flowering	1.3-20.1%	Fan et al., 2020
<i>Parasponia andersonii</i>	<i>PanNSP1, PanNSP2</i>	GRAS-type transcriptional regulators NODULATION SIGNALLING PATHWAY	unable to form root nodules	48.3-55.2%	van Zeijl et al., 2018
	<i>PanHK4</i>	<i>HISTIDINE KINASE 4, cytokinin receptor</i>	procambium activity is reduced	88.9%	van Zeijl et al., 2018
	<i>PanEIN2</i>	<i>ETHYLENE INSENSITIVE 2, a central component of the ethylene signaling pathway</i>	ethylene insensitive, bisexual flowers forming, impaired nodule development	50.0%	van Zeijl et al., 2018
poplar	<i>PtPDS1, PtPDS2, PtoPDS</i>	Phytoene Desaturase	albino shoots and leaves	50-86.4%	Liu et al., 2015, Fan et al., 2015
	<i>4CL1</i>	4-coumarate:CoA ligase, lignin biosynthesis	reddish brown wood color, lignin content and syringyl-to-guaiacyl (S : G) monolignol ratio were reduced drastically (52–92%)	100.0%	Zhou et al., 2015, Tsai and Xue, 2015,
	<i>4CL2</i>	4-coumarate:CoA ligase, flavonoid biosynthesis	reduced Condensed tannins levels in roots	100.0%	Zhou et al., 2015, Tsai and Xue, 2015,
	<i>PtoDWF4</i>	cytochrome P450, brassinosteroid biosynthesis, a positive regulator for improving xylem development in poplar	reduced shoot dry weight, significantly smaller the size of stem diameter, reduced xylem development		Shen et al., 2018
	<i>PtoMYB156</i>	a R2R3-MYB transcription factor	ectopic deposition of lignin, xylan and cellulose during secondary cell wall formation	48.0%	Yang et al., 2017
	<i>PtMYB57</i>	a R2R3-MYB transcription factor predominantly expressed in mature leaves	more proanthocyanidins (PA, also called condensed tannins) compounds were accumulated in the leaves, increased expression of flavonoid biosynthetic activators		Wan et al., 2017
	<i>PtMYB115</i>	a R2R3-MYB transcription factor, specifically regulate proanthocyanidins (PA) metabolism	significantly reduced expression of PA structural genes and a decrease in PA content	45.4%	Wang et al., 2017a
	<i>PtMYB170</i>	a R2R3-MYB pleiotropic regulator, preferentially expressed in young leaves and xylem tissues	weakened lignin deposition, a more flexible and collapsed xylem phenotype		Xu et al., 2017
	<i>LTF (MYB TF)</i>	a lignin biosynthesis associated transcription factor, binds the promoter of a key lignin biosynthetic gene encoding 4-coumarate-CoA ligase (4CL)	Increased Lignin Deposition in Developing Xylem of Populus		Gui et al., 2019

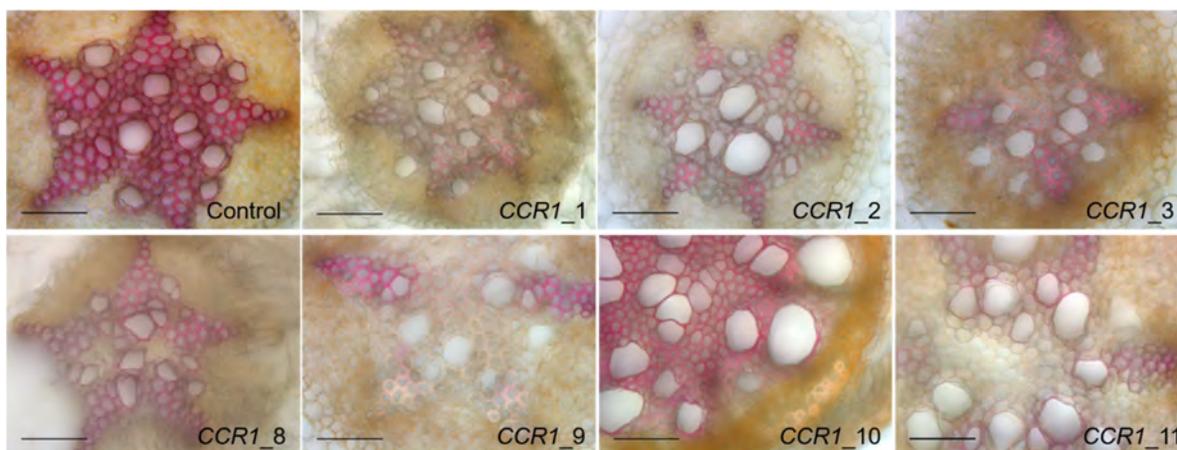
Supplemental Data

	<i>VNS09/SND1</i> , <i>VNS10/SND1</i> , <i>VNS11/SND1</i> , <i>VNS12/SND1</i>	NAC SECONDARY WALL THICKENING PROMOTING FACTOR3/SECONDARY WALL-ASSOCIATED NAC DOMAIN PROTEIN 1, control SCW formation in wood fibers and phloem fibers	SCW was severely suppressed in wood cells, some fiber cells formed thick cell walls.	12-87%	Takata et al., 2019
	<i>PLFY</i>	floral meristem identity gene, LEAFY (LFY), strong expression in developing inflorescences	predicted: female trees with completely sterile flowers and apparently normal growth	62.3-77%	Elorriaga et al., 2018
	<i>PAGs</i>	floral organ identity gene AGAMOUS (AG)	predicted: lose determinacy of the floral meristem in Arabidopsis; healthy trees with completely sterile flowers	81.30%	Elorriaga et al., 2018
	<i>SOCI</i>	flowering time genes SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1 (CO), floral integrators	/	56.4%	Bruegmann et al., 2019
	<i>AG8.1</i> , <i>AG8.2</i>	two paralogs of FUL, which play an additional role in biomass formation	/	73.4%	Bruegmann et al., 2019
	<i>NFP-like</i>	associated to mycorrhization	/	30.0%	Bruegmann et al., 2019
	<i>TOZ19</i>	TORMOZEMBRYO DEFECTIVE 19 (TOZ19) putatively correlated with sex determination in poplar	/	61.5%	Bruegmann et al., 2019
	<i>BRC1</i> , <i>BRC2</i>	The TCP-type transcription factors BRANCHED1 and BRANCHED2, 2 shape plant architecture by suppressing bud outgrowth	BRANCHED1-1 mutants exhibited strongly enhanced bud outgrowth. BRANCHED2-1 mutants had an extreme bud outgrowth phenotype and possessed two ectopic leaves at each node.	69.6-89.7%	Muhr et al., 2018
Eucalyptus	<i>EgrCCR1</i>	<i>Cinnamoyl-CoA Reductase1 (CCR1)</i> , a key lignin biosynthetic gene	decreased lignification, collapsed xylem vessels, irregular shapes for both xylem vessels and fibers	100.0%	Dai et al., 2020
	<i>EgrIAA9A</i>	an auxin dependent transcription factor of Aux/IAA family	/	92.3%	Dai et al., 2020

Supplemental Figures and Tables in Chapter II



Supplementary Figure S1: Wave numbers from FT-IR absorption spectra discriminating between controls and *CCR1*-edited lines. Upper graph represents FT-IR absorption spectra of controls (black) and *CCR1*-edited lines (green), respectively. The curves were drawn using the median of controls and *CCR1*-edited lines absorption values, respectively. Lower graphs represent the contribution of each wave numbers to PC1 and PC2 axis of Figure 2. Dotted lines and numbers 1 to 24 represent the most significant wave numbers related polymers of the secondary cell walls, the most involved in the discrimination between controls and *CCR1*-edited lines (see Figure 2 and Supplementary Table S6).



Supplementary Figure S2. Comparison of xylem development and lignification of xylem cells between control and *CCR1*-edited lines' roots. Transversal root sections made at around 10 cm from the root apex of control and edited lines. Lignified cell walls are visualized in red/purple by phloroglucinol-HCl. Scale bar = 50 μ m.

The complete version of Table S1 and Table S4 were found in Polint link:

\\polint.lrsv.ups-tlse.fr\Users\ying.dai\Supplemental data

Table S1. Details of mutations detected in 3 batches of *EgrCCR1* transgenic roots and control

Table S4. Details of mutations detected in *EgrIAA9A* transgenic roots and control

(‘Polint’ link: \\polint.lrsv.ups-tlse.fr\Users\ying.dai\Supplemental data)

Table S2. The edition rate decrease along the transgenic *ccr1* plants age

Transformation batch¹	Transgenic roots age (days)	Total subclone (N)	Edited subclone (N)	Edition rate
3rd	54	26	21	80.8%
2nd	73	39	18	46.2%
1st	153	213	50	23.5%

¹The batch 1 had 15 transgenic plants, batch two had 5 transgenic plants and batch 3 had 4 transgenic plants.

Table S3. Mortality of *ccr1* mutants during the transfert from in vitro culture to hydroponic culture

transformants	In vitro culutre (N)	hydroponic cultrue survival (N)	dead after transfert (N)	Mortality
CCR1	49	20	29	59.20%
Control	11	8	3	27.30%

Table S5. The edition type comparison of KOEgrIAA9A transgenic plants between DSDecode and sub-cloning methods

Edition types	DSDecode Genotyping	Allele	Edited Position	KO_EgrIAA9A	Edited Position	Sequenced clones	Sbuclones (x/y)	Sub-cloning Genotyping	Edition types
73bp del	ATGCTCCACCACCTCTCGG---73bp---CTTGGC monozygous mutation	Allele 1	gR1&2	IAA9A_1	gR1&2 gR2	9	8/9 1/9	ATGCTCCACCACCTCTCGG-----CTTGGC ATGCTCCACCACCTCTCGGAGGAAGGGGGTGGCAATACCTCTACAGTAGCCACCTCACCTCCATAGATGGCGCCTCTCAT-----TTGG	73bp del 6bp del
1bp del	ATGCTCCACCACCTCTCGG-TGTCAGGAAGGGG monozygous mutation	Allele 1	gR1	IAA9A_3	gR1&2 gR1&2	10	9/10 1/10	ATGCTCTTCAAGCTCTATGCTCTCTCATTTGCTGAGGAAGGGGGTGGCAATACCTCTACAGTAGCCACCTCACCTCCATAGATGGCGCCTCTCATGACTCTTGG ATGCTCCACCAC-----AATACCTCTACAGTAGCCACCTCACCTCCATAGATGGCGCCTCTCATGACTG-----	7bp sub, 16bp ins, 1bp del 33bp del
WT	ATGCTCCACCACCTCTCGGTTGGAGGAAGGGG monozygous mutation	Allele 1	no	IAA9A_5	WT gR1	3	1/3 2/3	ATGCTCCACCACCTCTCGGTTGGAGGAAGGGGGTGGCAATACCTCTACAGTAGCCACCTCACCTCCATAGATGGCGCCTCTCATGACTCTTGG ATGCTCCACCACCTCTCGG-TGTCAGGAAGGGGGTGGCAATACCTCTACAGTAGCCACCTCACCTCCATAGATGGCGCCTCTCATGACTCTTGG	WT 1bp del
1bp del	ATGCTCCACCACCTCTCGG-TGTCAGGAAGGGG	Allele 1	gR1	IAA9A_8	gR1&2	10	5/10	ATGCTCCACCACCTCTCGG-TGTCAGGAAGGGGGTGGCAATACCTCTACAGTAGCCACCTCACCTCCATAGATGGCGCCTCTCATGACTCTTGG	3bp del
2bp del	ATGCTCCACCACCTCTCGG--TGTCAGGAAGGGG	Allele 2	gR1		gR2		4/10 1/10	ATGCTCCACCACCTCTCGG-TGTCAGGAAGGGGGTGGCAATACCTCTACAGTAGCCACCTCACCTCCATAGATGGCGCCTCTCATG-----CTTGG ATGCTCCACCACCTCTCGGTTGGAGGAAGGGGGTGGCAATACCTCTACAGTAGCCACCTCACCTCCATAGATGGCGCCTCTCATGACTCTTGG	1bp del 4bp del
	No mutation, a wild-type sequence			IAA9A_10	WT	6	6/6	ATGCTCCACCACCTCTCGGTTGGAGGAAGGGGGTGGCAATACCTCTACAGTAGCCACCTCACCTCCATAGATGGCGCCTCTCATGACTCTTGG	WT
WT	ATGCTCCACCACCTCTCGGTTGGAGGAAGGGG	Allele 1	no	IAA9A_11	gR1&2	3	2/3	ATGCTCCACCACCTCTCG-----TTGGAGGAAGGGGGTGGCAATACCTCTACAGTAGCCACCTCACCTCCATAGATGGCGCCTCTCATGACT-----TTGG	7bp del
1bp sub, 1bp	ATGCTCCACCACCTCTCGGTTGGAGG-AGGGG	Allele 2	gR1		gR1&2		1/3	-----GTTGGAGGAAGGGGGTGGCAATACCTCTACAGTAGCCACCTCACCTCCATAGATGGCGCCTCTCATGATTGCTTGG	24bp del, 1bp sub
WT	ATGCTCCACCACCTCTCGGTTGGAGGAAGGGG	Allele 1	no	IAA9A_12	gR1&2	9	5/9	ATGCTCCACCACCTCTCGG-TGTCAGGAAGGGGGTGGCAATACCTCTACAGTAGCCACCTCACCTCCATAGATGGCGCCTCTCATGACT-----TTGG	9bp del
1bp del	ATGCTCCACCACCTCTCGG-TGTCAGGAAGGGG	Allele 2	gR1		gR1		2/9 1/9	ATGCTCCACCACCTCTCGG-TGTCAGGAAGGGGGTGGCAATACCTCTACAGTAGCCACCTCACCTCCATAGATGGCGCCTCTCATGACT-----TTGG ATGCTCCACCACCTCTCGGTTGGAGGAAGGGGGTGGCAATACCTCTACAGTAGCCACCTCACCTCCATAGATGGCGCCTCTCATGACTCTTGG	4bp del 1bp del
					gR1		1/9	ATGCTCCACCACCTCTCGG-TGTCAGGAAGGGGGTGGCAATACCTCTACAGTAGCCACCTCACCTCCATAGATGGCGCCTCTCATGACTCTTGG	1bp del
1bp del	ATGCTCCACCACCTCTCGG-TGTCAGGAAGGGG	Allele 1	gR1	IAA9A_15	gR1&2	4	2/4	ATGCTCCACCACCTCTCG-----CTTGG	73bp del
73bp del	ATGCTCCACCACCTCTCGG---73bp---CTTGGC	Allele 2	gR1&2		gR1&2		2/4	ATGCTCCACCACCTCTCGG-TGTCAGGAAGGGGGTGGCAATACCTCTACAGTAGCCACCTCACCTCCATAGATGGCGCCTCTCATGACTCTTGG	1bp del, 1bp ins
1bp del	ATGCTCCACCACCTCTCGG-TGTCAGGAAGGGG monozygous mutation	Allele 1	gR1	IAA9A_20	gR1&2 gR1	6	4/6 1/6	ATGCTCCACCACCTCTCGG-TGTCAGGAAGGGGGTGGCAATACCTCTACAGTAGCCACCTCACCTCCATAGATGGCGCCTCTCATGACTCTTGG ATGCTCCACCACCTCTCGGTTGGAGGAAGGGGGTGGCAATACCTCTACAGTAGCCACCTCACCTCCATAGATGGCGCCTCTCATGATTGCTTGG	2bp del 1bp ins
					gR2		1/6	ATGCTCCACCACCTCTCGGTTGGAGGAAGGGGGTGGCAATACCTCTACAGTAGCCACCTCACCTCCATAGATGGCGCCTCTCATGATTGCTTGG	1bp sub

Among 13 PCR amplicons direct sequencing, nine PCR amplicons were successfully treated by DsDecode. Lines highly lighted in green indicate the same results obtained from two methods; Lines highly lighted in yellow indicate different edited types were obtained from two methods; Lines no highlighting with color indicate the partial same edited type obtained from two methods. Abbreviation 'sub' means bases substitution, 'del' means deletion, and 'ins' means insertion. The underlined and bold characters are sgRNA sequences, and PAM sequences are in blue.

Table S6: Wave numbers from FT-IR absorption spectra discriminating between controls and *CCRI*-edited lines and their related compounds.

n°	WN (cm-1)	Meaning Bound	Related compound	References
1	815	C-H bending of S units	Lignin	[33]
	870	?	?	
	950	?	?	
2	1030	C-O primary alcohol	Lignin	[33]
3	1086	C-O deformation, secondary alcohol and aliphatic ether	Lignin	[33]
4	1100	C-O stretch, C-H stretch	Pectin	[34]
5	1115	Aromatic C-H deformation in S ring	Lignin	[35]
6	1140	C-H stretch in G ring	Lignin	[36,37]
7	1275	Aryl ring C=O Stretch, Guaiacyl units	Lignin	[38,39]
8	1320	CH in-plane bending / Cellulose	Cellulose	[40]
9	1330	Aryl ring breathing with C-O stretch, Syringyl units	Lignin	[38]
10	1360-1380	Phenolic OH and aliphatic C-H in methyl groups	Lignin	[35,41]
11	1410	COO- symmetric stretch / Pectin ester group	Pectin	[34]
12	1425	O-CH ₃ , C-H deformation, phenolic compounds	Lignin	[33]
13	1460	C-H deformation, asymmetric	Lignin	[33]
	1490	?	?	
14	1510	Aryl ring stretch, phenolic compounds	Lignin	[33]
15	1595	Aryl ring stretch, phenolic compounds	Lignin	[38]
16	1640	C=O stretch Conjugated carbonyl	Lignin	[42]
		Ring-conjugated C=O stretch of		
17	1670	coniferaldehyde/sinapaldehyde	Lignin	[41]
18	1708	Non-conjugated carbonyl	Lignin	[41]
19	1710-1730	C=O stretch Non-conjugated carbonyl	Hemicellulose	[33,43]
		C=O stretch Non-conjugated ketone, carboxyl and ester		
20	1735	groups	Lignin	[41]
	1770	?	?	
	2040	?	?	
	2060	?	?	
21	2850	C-H stretch O-CH ₃ group	Liginin	[41]
22	2880	C-H stretch methyl and methylene groups	Liginin	[41]
23	2945	C-H stretch methyl and methylene groups	Liginin	[41]
24	3440	O-H stretch, H-bonded	Lignin	[41]

We identified wave numbers involved in differences between controls and *CCRI*-edited lines absorption spectra, based on PLS-DA loadings contribution to PC1 and PC2 axis (see Figure 2 and Supplementary Figure S1). The wave numbers listed in this table are the most discriminant (top 50% among the highest contribution). Numbers 1 to 24 represents chemical bounds related to secondary cell wall structure and composition already reported in literature. Five wave numbers were not attributed to any reported compounds absorption.

Table S7. The selected sgRNAs sequences designed using CRISPOR (<http://crispor.tefor.net/>)

sgRNA of target gene	Position Strand	Guide Sequence + PAM+ Restriction Enzymes	Specificity Score ¹	Predicted Efficiency ²	Out-of-Frame score ³	Off-targets for 0-1-2-3-4 mismatches+ next to PAM ⁴	Genome Browser links to matches sorted by CFD off-target score ⁵
<i>EgrCCR1 (Eucgr.J03114)</i>							
CCR1_gRNA1	582 / rev	GCGGTCCAAGCACGAGCACA GGG	85	56	44	0-0-0-2-43	4:scaffold 11 2.01 Mbp
		Enzymes: Alw21I, Tru1I				0-0-0-0-2	4:scaffold 8 68.16 Mbp
						3:scaffold 6 3.22 Mbp	
						45 off-targets	
CCR1_gRNA2	666 / rev	GACCGAGTTGGCGTAGGTCT TGG	97	38	46	0-0-0-2-26	4:scaffold 8 10.92 Mbp
		Enzymes: BshFI, NmeAIII,				0-0-0-0-0	3:scaffold 3 71.35 Mbp
						4:scaffold 2 61.86 Mbp	
						28 off-targets	
<i>EgrIAA9A (Eucgr.H02407)</i>							
IAA9A_gRNA1	23 / fw	GTCTCCACCACTTCTGGGTG TGG	98	45	63	0-0-0-1-12	4:scaffold 540 9.82 Kbp
		Enzymes: BslI				0-0-0-0-0	4:scaffold 1 11.89 Mbp
						4:scaffold 7 10.33 Mbp	
						13 off-targets	show all...
IAA9A_gRNA2	96 / fw	GGCGCCTCTCATGACTGCTT GGG	99	38	69	0-0-1-0-2	2:scaffold 4 14.05 Mbp
						0-0-1-0-0	4:scaffold 8 71.33 Mbp
						4:scaffold 8 9.98 Mbp	
						3 off-targets	

¹. Specificity score: The specificity score ranges from 0-100 and measures the uniqueness of a guide in the genome [66]. The higher the specificity score, the lower are off-target effects in the genome (we recommend values >50, where possible). ². Predicted efficiency score: It is a prediction of how well this target may be cut by its RNA guide sequence. It ranges from 0-100 with 100 being the best. ³. Out-of-frame score: this score (0-100) is a prediction how likely a guide is to lead to out-of-frame deletions, especially for doing gene knockout with single guide RNA. ⁴. Off-targets for 0-1-2-3-4 mismatches: the number of possible off-targets in the genome, for each number of mismatches. This is a summary of the whole-genome search for sequences similar to the guide target sequence. It is best explained by an example: a description “0 - 1 - 2 - 9 - 28” means that the target matches 0 locations in the genome with no mismatch, 1 location in the genome with 1 mismatch, 2 locations with 2 mismatches, 9 with 3 and 28 locations with 4 mismatches. ⁵. CRISPOR lists the locations of all possible off-targets with up to four mismatches. CFD, Cutting frequency determination.

Table S8. The primers and sgRNA scaffold used for generating sgRNA intermediary vectors

Primers to generate sgRNA intermediary vectors ¹	sequence 5' - 3'
CCR1_gRNA1_F	tgtggtctca ATTGCGGTCCAAGCACGAGCACAG gttttagagctagaaatagcaag
CCR1_gRNA2_F	tgtggtctca ATTGACCGAGTTGGCGTAGGTCT gttttagagctagaaatagcaag
IAA9A_gRNA1_F	tgtggtctca ATTGTCTCCACCACCTTCTGGGTG gttttagagctagaaatagcaag
IAA9A_gRNA2_F	tgtggtctca ATTGGCGCCTCTCATGACTGCTT gttttagagctagaaatagcaag
Rreverse primer	tgtggtctca AGCG TAATGCCAACTTTGTAC
sgRNA scaffold ²	tgtggtctcaATTGNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGTTAAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTG GCACCGAGTCGGTGCTTTTTTTCTAGACCCAGCTTTCTT GTACAAAGTTGGCATTACGCTt gagaccaca

¹: Design the forward sgRNA primer as following: tgtggtctca ATTG NNNN NNNNN NNNNN NNNNN gtttagagctagaaatagcaag (BsaI site is in blue, the 20 bp guide sequence is in red).

²: Amplify an sgRNA using a pair of primers and pICH86966::AtU6p::sgRNA_PDS construct (Addgene plasmid 46966) as a template. The resulting PCR product will be as following (sequence in bold belongs to primers).

Table S9. The primers used in CRISPR/Cas9 system, the primer names, sequences and the specific uses

Primer name	Primer sequence (5' to 3')	Use(s)
gRNA_R	tgtggtctcaAGCGTAATGCCAACTTTGTAC	Clone PCR product within sgRNA
Cas9_F	ACAAGAAGTACTCCATTGGGCTCGA	Verify Cas9 sequence
Cas9_R	GCTTGACAAGCTCACACCTTCCTCTTCT	Verify Cas9 sequence
DsRed-seq-F	CGGTTCTTTCATACTGCTCAACGAT	Verify DsRed sequence
CCR1_gRNA1_F	tgtggtctcaATTGCGGTCCAAGCACGAGCACAgtttttagagctagaatagcaag	EgrCCR1 sgRNA1
CCR1_gRNA2_F	tgtggtctcaATTGACCGAGTTGGCGTAGGTCTgtttttagagctagaatagcaag	EgrCCR1 sgRNA2
IAA9A_gRNA1_F	tgtggtctcaATTGTCTCCACCACTTCTGGGTGgttttagagctagaatagcaag	EgrIAA9A sgRNA1
IAA9A_gRNA2_F	tgtggtctcaATTGGCGCCTCTCATGACTGCTTgtttttagagctagaatagcaag	EgrIAA9A sgRNA2
L2_gRNA_seq_R	TCGGTCACATGTGCATCCTCTC	Amplify and sequence sgRNA in level 2 acceptor
L2_Terminator_F	GCATGACGTTATTTATGAGATGGGT	Amplify and sequence sgRNA in level 2 acceptor
pICSL4723_seq_F	GTGGTGTAACAATAATTGACGC	Primer on LB, Verify acceptor sequence
pICSL4723_seq_R	GGATAAACCTTTTCACGCC	Primer on RB, Verify acceptor sequence
Level 1_seq_F	CTGGTGGCAGGATATATTGTGGTG	Amplify and sequence AtU6 promoter and sgRNA in level 1
Level 1_seq_R	GAACCTGTGGTTGGCATGCACATAC	Amplify and sequence AtU6 promoter and sgRNA in level 1
CCR1_edit check_NCBI_F	TTTGGTCCCGAGAAGTGGT	Amplify and sequence clones; Editing event genotyping
CCR1_edit check_NCBI_R	CGCTTGACCACAGCTTGAGTA	Amplify and sequence clones; Editing event genotyping
nested primer CCR1_R	GCGGGTCATAGCGAAAGAGT	Sequence PCR products; Editing event genotyping
IAA9A_edit check_NCBI F	AGATGCAGTGTGCTTTGCG	Amplify and sequence clones; Editing event genotyping
IAA9A_edit check_NCBI R	TGAGTTGACGCTTGGTCTCC	Amplify and sequence clones; Editing event genotyping
nested primer IAA9A_F	TGTTGCCAGTTGTAGCGTTG	Sequence PCR products; Editing event genotyping

Supplemental Figures and Tables in Chapter III

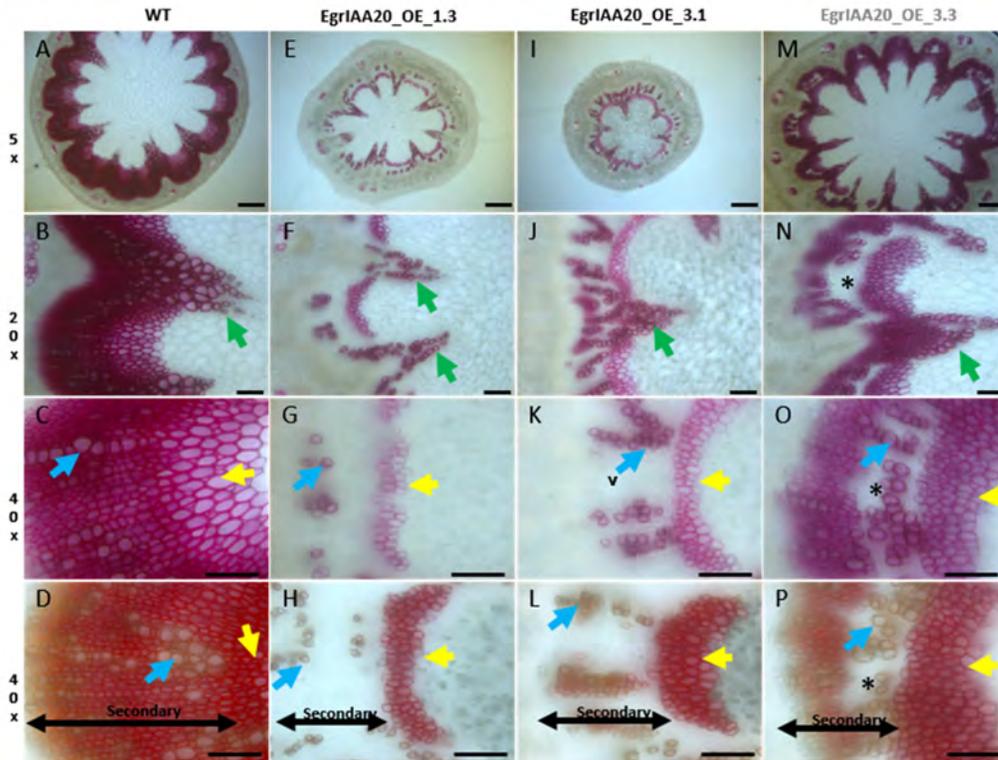


Figure III-S1. Cross sections of inflorescence stems at the basal part showing dramatically inhibited fiber cells of secondary growth in *EgrIAA20* overexpressing lines. Yellow arrows indicated primary infascicular fiber cells, green arrows showed vessels originated from primary growth and blue arrows indicate vessels of secondary growth. For 5x pictures scale bar = 200 μm , and for 20x and 40x pictures scale bar = 50 μm .

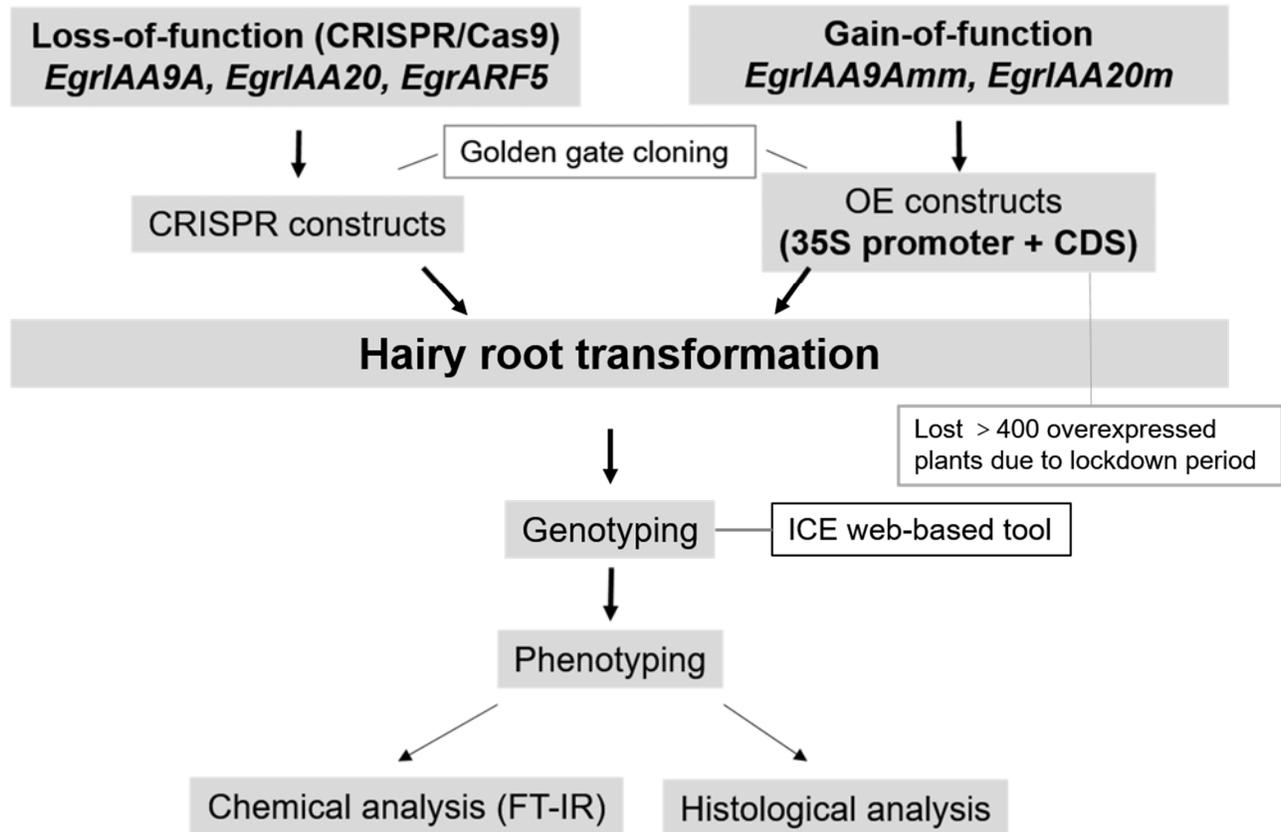


Figure III-S2. The procedure used in Chapter III for loss-of-function and gain-of-function mutants' functional characterization



Figure III-S3. Editing events of *ARF5_17*, *ARF5_29*, *ARF5_33*, *ARF5_43*, *ARF5_44* by ICE analysis. As for all *ARF5_* lines and *IAA9A_* edited lines analyzed by ICE, find the results in Polint Link ([\polint.lrsv.ups-tlse.fr\Users\ying.dai\Supplemental data](http://\polint.lrsv.ups-tlse.fr/Users\ying.dai\Supplemental data)).

ARF5_43 **gRNA1**
Allele 1: GATGTATTCCAGTACCTGACTT**TGG**ACTAAGGCCA (WT)
Allele 2: GATGTATTCCAGTACCTGA-**TTTGG**ACTAAGGCCA (deletion)

ARF5_44 **gRNA1** **gRNA2**
Allele 1: GATGTATTCCAGTACCTGACTT**TGG**NNN₄₉**GCGAGTGATACAAGTACACA**TGG**** (WT)
Allele 2: GATGTATTCCAG-----**CACA**TGG**** (78bp del)

Figure III-S4. Allele mutations of *ARF5_43* and *ARF5_44* detected by DSDcode. Red characters were gRNAs, bold characters were PAM sequences.

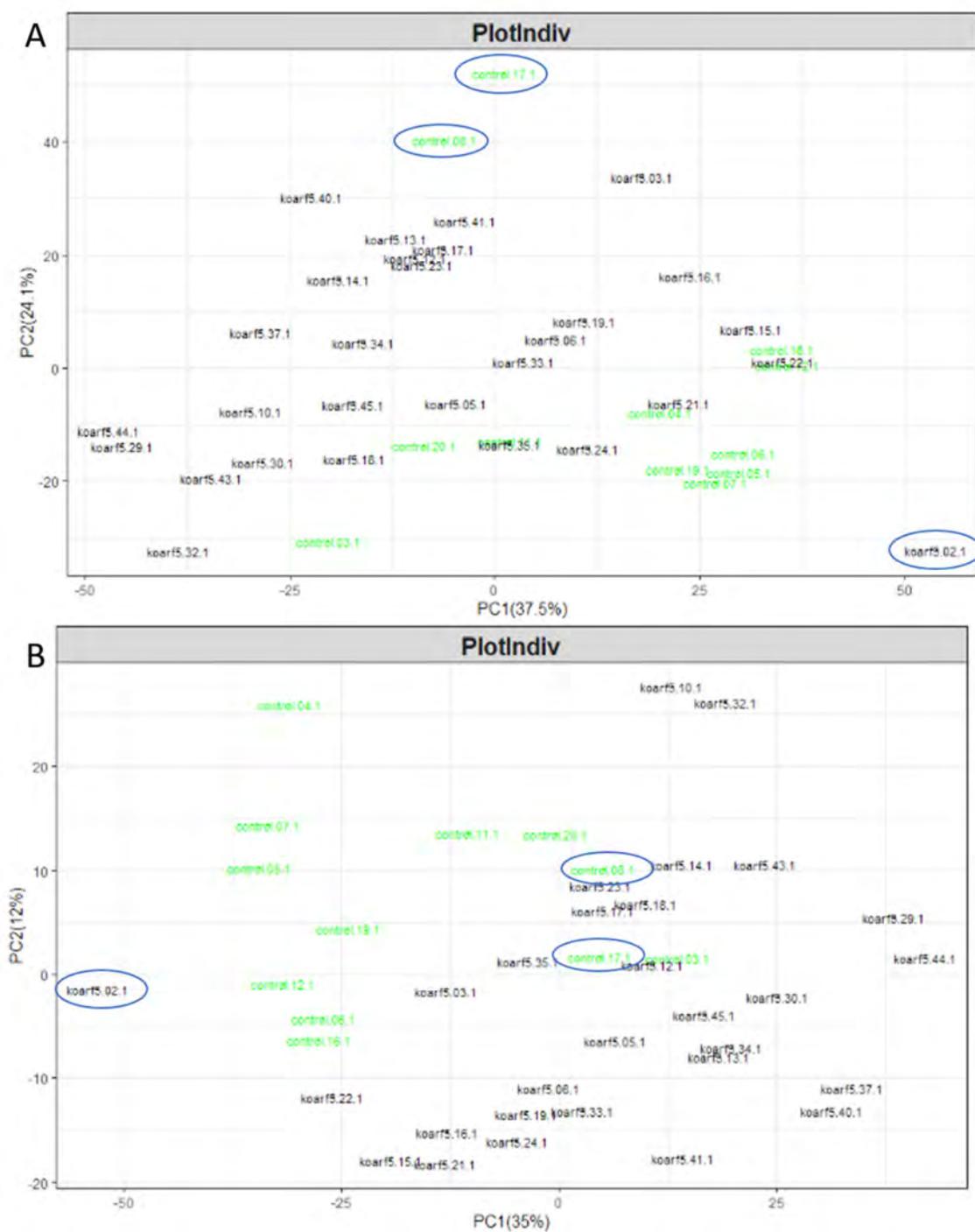


Figure III-S5. Combination of FT-IR and PCA (A) /PLS-DA (B) analysis for all *ARF5* lines. The line *ARF5_2* and control samples (control_8, control_17) in blue circles were outliers, which were discarded in further FT-IR_PLS-DA analysis.

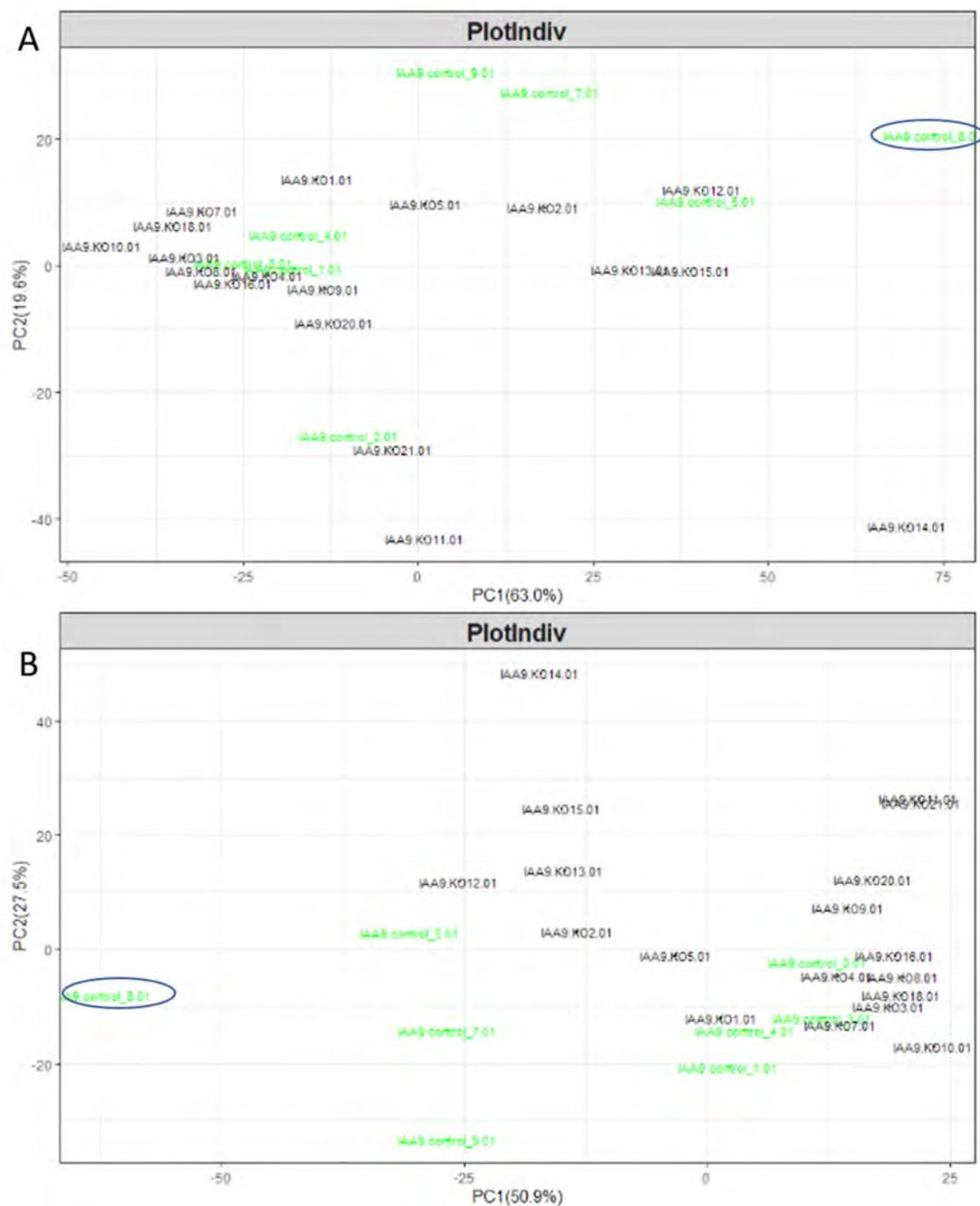


Figure III-S6. Combination of FT-IR and PCA (A) /PLS-DA (B) analysis for all *IAA9A* lines. The control sample (control_8) in blue circles was outlier, which was discarded in further FT-IR_PLS-DA analysis.

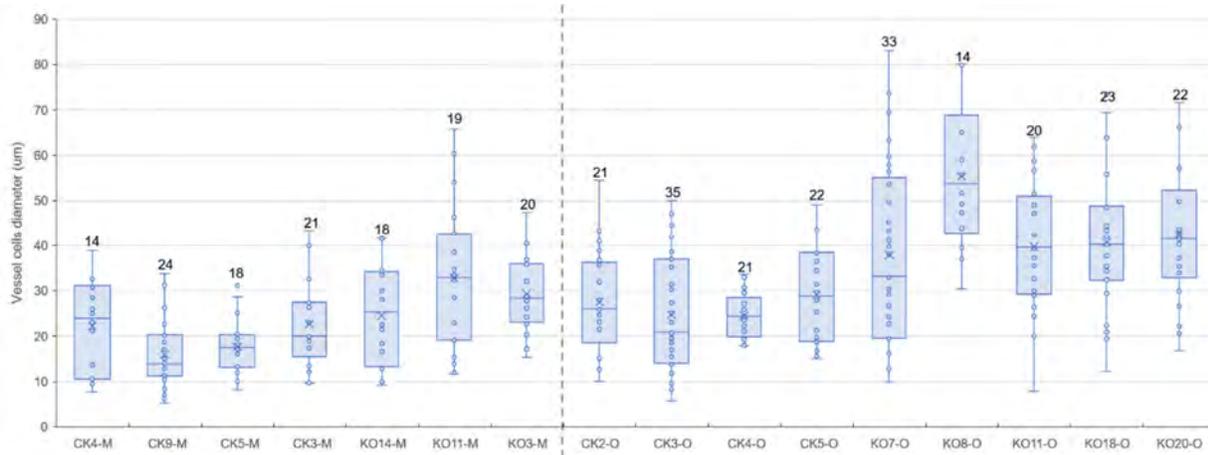


Figure III-S7. The comparison of xylem vessel cells diameters between *IAA9A* edited lines and control in medium and old developing stages. The samples of medium stages are left the gray lines, old stages are right. CK4-M, control 4 of medium stage; KO14-M, *IAA9A*_14 of medium stage; CK2-O, control 2 of old stage; KO8-O, *IAA9A*_8 old stage. The number on each box means the xylem vessel cell numbers in each section. The blue lines linked the median value of each sample, and the point in each box represented the diameter size of every vessel cell.

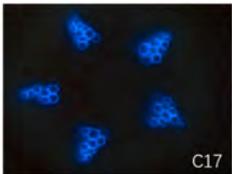
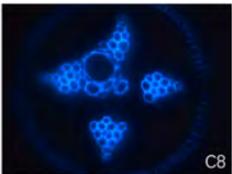
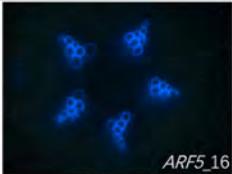
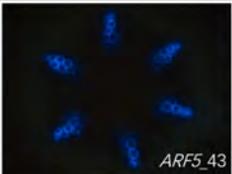
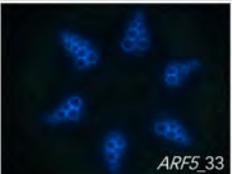
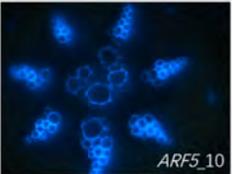
				
proto/metaxylem numbers	44	54	45	55
Vessel cells formation	x	x	√	√
				
proto/metaxylem numbers	43	50	46	70
Vessel cells formation	x	x	x	√

Figure III-S8. Comparison of *ARF5* lines and control under 40x magnification using UV light.

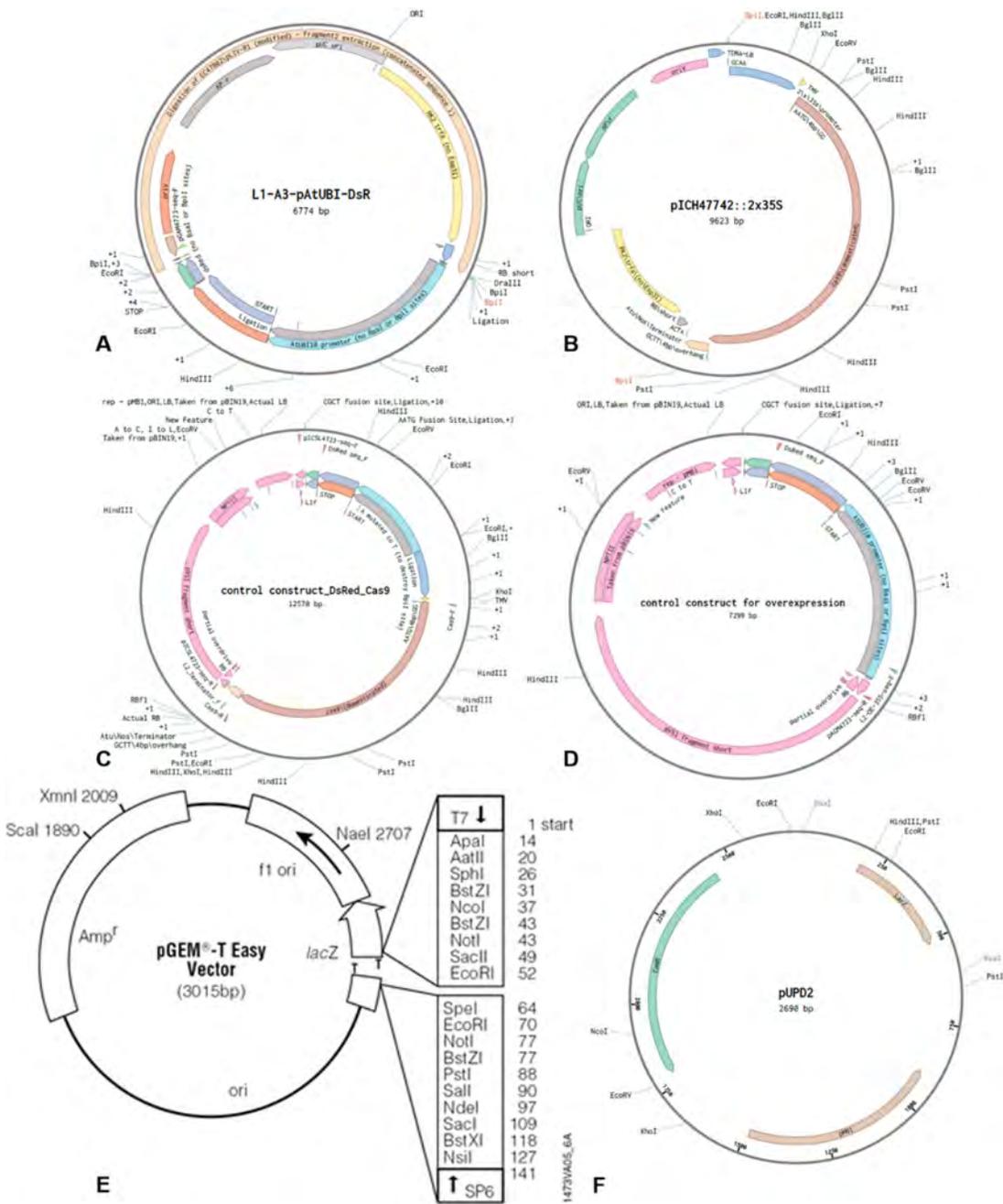


Figure III-S9. Vectors maps used in Chapter III. A, DsRED plasmid; B, Cas9 which is driven by 2x35S promoter; C, control construct of CRISPR/Cas9; D, control construct of overexpression (harboring DsRed); E, pGEM-T Easy Vector for subcloning method (clone PCR amplicon into this vector); F, pUPD2 vector which was used for constructing *EgrIAA9Amm* overexpression in Level 1.

Table III-S1A. Information of CRISPR/Cas9 mediated alived *ARF5* (30 fluorescenced roots / 50 in total), *IAA20* transgenic roots (three alive), and controls

No.	DNA	PCR	Edition by ICE (edit%, mutation types)	Edition by DSDdecode	FTIR powder quantity (1'30" grinding)	Microscope	plant size	
							total	DsRed part
<i>ARF5_2</i>	√	√	WT	WT	normal	/	e	e
<i>ARF5_5</i>	√	√	WT	WT	normal	/	c	c
<i>ARF5_6</i>	√	√	WT	WT	normal	/	d	d
<i>ARF5_10</i>	√	√	53%, chimeric	succeed but wrong cut site	normal	√	d	d
<i>ARF5_12</i>	√	√	WT	WT	normal	/	e	e
<i>ARF5_13</i>	√	√	WT	WT	normal	√	d	c
<i>ARF5_14</i>	√	√	WT	WT	much	/	e	e
<i>ARF5_15</i>	√	√	WT	WT	normal	/	e	e
<i>ARF5_16</i>	√	√	4%, monoallelic (1bp del at sR2), 93% WT	WT	normal	√	d	c
<i>ARF5_17</i>	√	√	16%, chimeric, 82% WT	WT	normal	/	b	b
<i>ARF5_18</i>	√	√	WT	x	normal	/	e	d
<i>ARF5_19</i>	√	√	WT	WT	normal	√	d	d
<i>ARF5_21</i>	√	√	WT	WT	normal	/	c	c
<i>ARF5_22</i>	√	√	WT	WT	normal	√	b	b
<i>ARF5_23</i>	√	√	2%, monoallelic, 96% WT	WT	normal	/	e	c
<i>ARF5_24</i>	√	√	WT	WT	normal	/	d	d
<i>ARF5_29</i>	√	√	74%, chimeric, 22% LDT (71bp and 72bp del), 15% WT	x	normal	√	e	c
<i>ARF5_30</i>	√	√	3%, monoallelic, 95% WT	x	normal	/	e	e
<i>ARF5_31</i>	√	√	WT	WT	normal	/	e	e
<i>ARF5_32</i>	√	√	x	x	normal	√	c	b
<i>ARF5_34</i>	√	√	4%, monoallelic, 93% WT	WT	normal	/	e	b
<i>ARF5_35</i>	√	√	WT	WT	normal	/	e	d
<i>ARF5_37</i>	√	√	WT	WT	normal	/	b	c

Supplemental Data

<i>ARF5_38</i>	√	√	WT	x	/ (no materials for FTIR)	/	e	a
<i>ARF5_40</i>	√	√	78%, chimeric, 7% LDT (71bp del), 4% WT	x	normal	√	e	c
<i>ARF5_41</i>	√	√	x	x	normal	/	e	c
<i>ARF5_43</i>	√	√	51%, chimeric, 45% WT	monoallelic (1bp del at sR1)	less	√	b	b
<i>ARF5_44</i>	√	two bands	57%, monoallelic, 57% LDT (78bp del), 38% WT	monoallelic (LDT, 78bp del)	normal	√	e	d
<i>ARF5_45</i>	√	√	92% WT	x	normal	√	d	d
1st <i>ARF5_33</i> (<i>in vitro</i>)	√	√	56%, monoallelic, 56% LDT (78bp del), 39% WT	monoallelic (LDT, 78bp del)	/		b	b
2nd <i>ARF5_33</i>	√	√	42%, chimeric, 1% large deletion (23bp del at sR1)	succeed but wrong cut site	normal	√	c	c
<i>IAA20_2</i>	√	√	x	x	normal	√	e	e
<i>IAA20_3</i>	√	x	x	x	normal	√	e	d
<i>IAA20_4</i>	√	small size band	67%, chimeric, 55% LDT (65bp del), various large deletions	x	normal	√	e	e
control 3 (for <i>IAA20</i>)	√	√	WT	WT	normal	√	c	c
control 4 (for <i>ARF5</i>)	√	√	WT	x	normal	√	c	c
control 5 (for <i>ARF5</i>)	√	√	WT	x	normal	√	d	d
control 6 (for <i>ARF5</i>)	√	√	/	/	less	√	d	b
control 7 (for <i>ARF5</i>)	√	√	WT	x	normal	/	c	b
control 8 (for <i>ARF5</i>)	√	√	WT	x	normal	√	d	d
control 11 (for <i>IAA20</i>)	√	√	x	x	normal	√	e	c
control 12 (for <i>IAA20</i>)	√	√	x	x	normal	/	e	c
control 16 (for <i>IAA20</i>)	√	√	x	x	normal	/	e	d
control 17	√	√	x	x	normal	√	e	d
control 19	√	/	/	/	less	√	c	c
control 20	√	/	/	/	normal	√	e	e

Notes: LDT, large deletion between two sgRNAs; WT, wide-type; sR1, sgRNA1; sR2, sgRNA2; x, failed; √, succeed/done; /, not done. The plant sizes are classified into five stages: a (very small), b (small), c (normal), d (large) and e (very large).

Table III-S1B. Information of CRISPR/Cas9 mediated alived *IAA9A*, and controls

No.	DNA	PCR	Edition by ICE (edit%, mutation types)	Edition by DSDecode	Edition by subcloning (edit%, mutation types)	FTIR powder quantity	Microscope	plant size
								DsRed part
<i>IAA9A_1</i>	√	low band	95%, 95% LDT (73bp del)	homozygous (73bp del)	100% (9/9), heterozygous, LDT	normal	/	e
<i>IAA9A_2</i>	x	/	/	/	/	less	/	a
<i>IAA9A_3</i>	√	√	44%, chimeric, 28% WT	homozygous (1bp del)	100% (10/10), heterozygous, LDT	less	√	a
<i>IAA9A_4</i>	√	two bands	x	x	100% (10/10), heterozygous, LDT	normal	√	b
<i>IAA9A_5</i>	√	√	48%, chimeric (2 types small deletions), 48% WT	monoallelic (1bp del, WT)	66.7% (2/3), monoallelic	normal	√	a
<i>IAA9A_7</i>	√	weak	73%, chimeric, 67% LDT (73bp del)	complicated variant	100% (8/8), chimeric, LDT	normal	√	b
<i>IAA9A_8</i>	√	two bands	91%, chimeric, 6% WT	heterozygous (1bp del, 2bp del)	100% (10/10), chimeric	much	√	e
<i>IAA9A_9</i>	√	low band	4%, 4% LDT (78bp del)	x	100% (10/10), heterozygous, LDT, large deletion (ATG lost)	normal	√	c
<i>IAA9A_10</i>	√	√	100% WT	WT	WT	less	√	a
<i>IAA9A_11</i>	√	√	4%, chimeric, 93% WT	monoallelic (1bp sub and 1bp del, WT)	100% (3/3), heterozygous, large deletion (ATG lost)	much	√	e
<i>IAA9A_12</i>	√	√	89%, chimeric, 7% WT	monoallelic (1bp del, WT)	100% (9/9), chimeric	less	/	a
<i>IAA9A_13</i>	√	√	83%, chimeric, 68% LDT (73bp del)	complicated, succeed but wrong cut site	100% (7/7), heterozygous, LDT	less	/	a
<i>IAA9A_14</i>	√	√	26%, monoallelic, 72% WT	monoallelic (1bp sub and 1bp del, WT)	/	normal	√	b
<i>IAA9A_15</i>	√	two bands	62%, chimeric, 38% LDT (7bp del), 11% WT	heterozygous (1bp del, 73bp del)	100% (4/4), heterozygous, LDT	less	√	a
<i>IAA9A_16</i>	x	/	/	/	/	less	√	a
<i>IAA9A_18</i>	√	√	68%, chimeric, 12% large deletion at sR2, 30% WT	x	/	normal	√	c
<i>IAA9A_20</i>	√	√	99%, biallelic	homozygous (1bp del)	100% (6/6), chimeric	normal	√	b
<i>IAA9A_21</i>	√	√	100% WT	WT	/	less	/	a
control 1	x	x	/	/	/	normal	√	c
control 2	√	x	/	/	/	less	√	a
control 3	√	√	WT	WT	WT (9/9)	much	√	d

Supplemental Data

control 4	x	x	/	/	/	normal	√	c
control 5	x	x	/	/	/	less	√	a
control 7	√	x	/	/	/	normal	√	b
control 8	x	x	/	/	/	less	√	a
control 9	√	√	WT	x	/	normal	√	c

Notes: LDT, large deletion between two sgRNAs; WT, wide-type; sR1, sgRNA1; sR2, sgRNA2; x, failed; √, succeed/done; /, not done. The plant sizes are classified into five stages: a (very small), b (small), c (normal), d (large) and e (very large).

Table III-S2. The selected sgRNAs sequences of *EgrARF5* and *EgrIAA20* designed using CRISPOR (<http://crispor.tefor.net/>)

sgRNA of target gene	Position Strand	Guide Sequence + PAM+ Restriction Enzymes	Specificity Score ¹	Predicted Efficiency ²	Out-of-Frame score ³	Off-targets for 0-1-2-3-4 mismatches+ next to PAM ⁴	Genome Browser links to matches sorted by CFD off-target score ⁵
<i>EgrARF5 (Eucgr.F02090)</i>							
ARF5_gRNA1	117 / fw	GTATTTCCAGTACCTGACTT TGG	96	45	26	0-0-0-2-29	4:scaffold_6 39.29 Mbp
		Enzymes: BshFI, DseDI				0-0-0-0-2	4:scaffold_6 16.31 Mbp
						4:scaffold_1 39.43 Mbp	
						31 off-targets	
ARF5_gRNA2	189 / fw	GCGAGTGATACAAGTACACA TGG	96	66	32	0-0-0-3-18	4:scaffold_3 62.46 Mbp
		Enzymes: MspI, NciI, BseDI, StyD4I				0-0-0-0-1	4:scaffold_6 35.10 Mbp
						4:scaffold_1 19.91 Mbp	
						21 off-targets	
<i>EgrIAA20(Eucgr.K00561)</i>							
IAA20_gRNA1	58/ rev	GAGACGCCGAGACATGGACG AGG	94	62	50	0-0-1-3-28	3:scaffold_6 12.13 Mbp
		Enzymes: BslI				0-0-0-0-2	4:scaffold_4 6.13 Mbp
						4:scaffold_4 6.95 Mbp	
						32 off-targets	
IAA20_gRNA2	124 / rev	GCCATCTTGGTGAGAAGCCG GGG	98	58	37	0-0-0-2-18	4:scaffold_7 644.29 Kbp
		Enzymes: TatI, Hpy166II, NlaIII, DraIII				0-0-0-0-1	4:scaffold_7 606.77 Kbp
						4:scaffold_7 616.14 Kbp	
						20 off-targets	

1. Specificity score: The specificity score ranges from 0-100 and measures the uniqueness of a guide in the genome. The higher the specificity score, the lower are off-target effects in the genome (we recommend values >50, where possible). 2. Predicted efficiency score: It is a prediction of how well this target may be cut by its RNA guide sequence. It ranges from 0-100 with 100 being the best. 3. Out-of-frame score: this score (0-100) is a prediction how likely a guide is to lead to out-of-frame deletions, especially for doing gene knockout with single guide RNA. 4. Off-targets for 0-1-2-3-4 mismatches: the number of possible off-targets in the genome, for each number of mismatches. This is a summary of the whole-genome search for sequences similar to the guide target sequence. It is best explained by an example: a description “0 - 1 - 2 - 9 - 28” means that the target matches 0 locations in the genome with no mismatch, 1 location in the genome with 1 mismatch, 2 locations with 2 mismatches, 9 with 3 and 28 locations with 4 mismatches. 5. CRISPOR lists the locations of all possible off-targets with up to four mismatches. CFD, Cutting frequency determination.

Table III-S3A. The primers used in CRISPR/Cas9 system, the primer names, sequences and the specific uses

Primer name	Primer sequence (5' to 3')	Use(s)
gRNA_R	tgtggtctcaAGCGTAATGCCAACTTTGTAC	Clone PCR product within sgRNA
Cas9_F	ACAAGAAGTACTCCATTGGGCTCGA	Verify Cas9 sequence
Cas9_R	GCTTGACAAGCTCACACCTTCCTCTTCT	Verify Cas9 sequence
DsRed-seq-F	CGGTTCTTTTCATACTGCTCAACGAT	Verify DsRed sequence
ARF5_gRNA1_F	tgtggtctcaATTGTATTTCCAGTACTGACTTgtttttagagctagaaatagcaag	EgrARF5 sgRNA1
ARF5_gRNA2_F	tgtggtctcaATTGCGAGTGATACAAGTACACAgtttttagagctagaaatagcaag	EgrARF5 sgRNA2
IAA20_gRNA1_F	tgtggtctcaATTGAGACGCCGAGACATGGACGgttttagagctagaaatagcaag	EgrIAA20 sgRNA1
IAA20_gRNA2_F	tgtggtctcaATTGCCATCTTGGTGAGAAGCCGgttttagagctagaaatagcaag	EgrIAA20 sgRNA2
L2_gRNA_seq_R	TCGGTCACATGTGCATCCTCTC	Amplify and sequence sgRNA in level 2 acceptor
L2_Terminator_F	GCATGACGTTATTTATGAGATGGGT	Amplify and sequence sgRNA in level 2 acceptor
pICSL4723_seq_F	GTGGTGTAACAAATTGACGC	Primer on LB, Verify acceptor sequence
pICSL4723_seq_R	GGATAAACCTTTTCACGCC	Primer on RB, Verify acceptor sequence
Level 1_seq_F	CTGGTGGCAGGATATATTGTGGTG	Amplify and sequence AtU6 promoter and sgRNA in level 1
Level 1_seq_R	GAACCTGTGGTTGGCATGCACATAC	Amplify and sequence AtU6 promoter and sgRNA in level 1
ARF5_edit check_NCBI_F	GCCAAGTTCACAACGTGACTC	Amplify and sequence clones; Editing event genotyping
ARF5_edit check_NCBI_R	GCACAATTAACGACTCTCACCC	Amplify and sequence clones; Editing event genotyping
nested primer ARF5_F	TAGGCCGACAGAGACAGAT	Sequence PCR products; Editing event genotyping
IAA20_edit check_NCBI F	ACGGAGAAGATGGGCAAAGG	Amplify and sequence clones; Editing event genotyping
IAA20_edit check_NCBI R	GCTCCAGTCAAACCCCTTCA	Amplify and sequence clones; Editing event genotyping
nested primer IAA20_R	TTCACGTAGAAGTAGCCGC	Sequence PCR products; Editing event genotyping

Table III-S3B. The primers used in overexpressed constructions, the primer names, sequences and the specific uses

Primer name	Primer sequence (5' to 3')	Use(s)
dome-EgrIAA9A-PCR1-F	GCGCCGTCTCGCTCGAATGTCTCCACCCTTCTGGG	Amplify partial EgrIAA9A CDS sequence for domestication
dome-EgrIAA9A-PCR1-R	GCGCCGTCTCGGCAGACCAAGCGTCAACTCA	Amplify partial EgrIAA9A CDS sequence for domestication
dome-EgrIAA9A-PCR2-F	GCGCCGTCTCGCTGCCTGGATCTCAATCTCC	Amplify partial EgrIAA9A CDS sequence for domestication
dome-EgrIAA9A-PCR2-R	GCGCCGTCTCGCTCAAAGCCTATTGTCTCACTAGTCCAATTGC	Amplify partial EgrIAA9A CDS sequence for domestication
dome-EgrIAA20-PCR1-F	TGTGGTCTCGAATGATGGCAAAGGAAGCA	Amplify partial EgrIAA20 CDS sequence for domestication
dome-EgrIAA20-PCR1-R	CCTTATCTTCATATAACAACACATGACACCC	Amplify partial EgrIAA20 CDS sequence for domestication
dome-EgrIAA20-PCR2-F	ATGAAGATAAGGAAGGGGACTGGTTGATGG	Amplify partial EgrIAA20 CDS sequence for domestication
dome-EgrIAA20-PCR2-R	ACAGGTCTCGAAGCTCACTCTATTCTTGTAAATC	Amplify partial EgrIAA20 CDS sequence for domestication
pUPD2_Seq_F	GCTTTCGCTAAGGATGATTTCTGG	primer on pUPD2, verify overexpressed Level 0 sequence
pUPD2_Seq_R	GGCAAGGTGTCACCACCCTG	primer on pUPD2, verify overexpressed Level 0 sequence
Level 1_35S_seq_F	GGTGATTTGTGCCGAGCTG	Verify 35S plasmid sequence
CDS_seq_L1&2_F	ATTTGGAGAGGACACGCTCG	Verify CDS sequence in Level 1 and Level 2 acceptor
CDS_seq_L1&2_R	CAAGACCGCAACAGGATTC	Primer on terminator, verify CDS sequence in Level 1 and Level 2 acceptor
DsRed-seq-F	CGGTTCTTTCATACTGCTCAACGAT	Verify DsRed sequence
L2_Terminator_F	GCATGACGTTATTTATGAGATGGGT	Amplify and sequence CDS in overexpressed level 2 acceptor

Table III-S4. Mutation types, occurrence, total Indel percentage and presumed knockout score of 11 *ARF5*_edited lines and one *IAA20*_edited line

Transformants	mutation	Indel%	Knockout score	Major editions (bp)			Model Fit (R ²)
				gRNA1	gRNA2	gRNA1&2	
<i>ARF5_40</i>	chimeric	78	67	-2, -13		-71	0.82
<i>ARF5_29</i>	chimeric	74	74	-1		-71, -72	0.96
<i>ARF5_44</i>	monoallelic	57	57			-78	0.95
<i>ARF5_10</i>	chimeric	53	46	-2	-2	-1 &-1	0.53
<i>ARF5_43</i>	chimeric	51	51	-1	-1		0.96
<i>ARF5_33</i>	chimeric	42	36	-11	-1	-2 &-3	0.42
<i>ARF5_17</i>	chimeric	16	16		+1, -5		0.98
<i>ARF5_16</i>	monoallelic	4	4		-1		0.97
<i>ARF5_34</i>	monoallelic	4	4		-1		0.97
<i>ARF5_30</i>	monoallelic	3	3		-4		0.98
<i>ARF5_23</i>	monoallelic	2	2		-1		0.98
<i>IAA20_4</i>	chimeric	67	64	-18	-19	-65, -5 &-7	0.67

Examples of plant sizes (from a-e):



