



**HAL**  
open science

# Étude de la régulation transcriptionnelle de la formation du bois (Xylème secondaire)

Hua Cassan-Wang

► **To cite this version:**

Hua Cassan-Wang. Étude de la régulation transcriptionnelle de la formation du bois (Xylème secondaire). Botanique. Université Toulouse 3 - Paul Sabatier, 2023. tel-04670685

**HAL Id: tel-04670685**

**<https://hal.science/tel-04670685v1>**

Submitted on 12 Aug 2024

**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.



**Université Toulouse III - Paul Sabatier**

**École Doctorale SEVAB**

**Discipline** : Sciences de la vie ; **Spécialité** : Physiologie Végétale

**ÉTUDE DE LA RÉGULATION TRANSCRIPTIONNELLE  
DE LA FORMATION DU BOIS (XYLÈME SECONDAIRE)**

***Mémoire en vue de soutenir l'habilitation à diriger des recherches***

***(HDR)***

**Hua Cassan-Wang**

Soutenue le 20 décembre 2023 devant le Jury composé de :

Mme. <b>Marie BAUCHER</b> , Professeur à l'Université libre de Bruxelles	Rapporteur
M. <b>Christian DUBOS</b> , DR INRAe Montpellier	Rapporteur
M. <b>Richard SIBOUT</b> , DR INRAe Nantes	Rapporteur
Mme. <b>Marie-Chantal TEULIERES</b> , Professeur à l'Université Toulouse III	Examinatrice
M. <b>Mondher BOUZAYEN</b> , Professeur à l'ENSAT	Examineur



**Laboratoire de rattachement :**

UMR 5546 Laboratoire de Recherche en Sciences Végétales (LRSV), UPS-CNRS-INP

24, Chemin de Borde Rouge, 31320 Auzeville Tolosane - FRANCE





## Remerciements

Je tiens d'abord à exprimer toute ma gratitude à mes superviseurs et/ou mentors, aux équipes d'accueil et à leurs personnels. Leur confiance partagée tout au long des projets de recherche m'a permis de me former et de progresser dans ma carrière professionnelle.

Aussi je souhaiterais remercier dans l'ordre chronologique :

\* **Professeur Weixin CHEN**, Laboratoire « Après-récolte et Transformation des Fruits », Université Agronomie de Chine du Sud, Guangzhou, Chine, qui m'a fait découvrir le potentiel futur de la biotechnologie des plantes à travers la Recherche scientifique.

\* **Professeur Mondher BOUZAYEN**, Directeur du laboratoire Génomique et Biotechnologie des fruits, ENSAT-INP Toulouse France, qui m'a ouvert le monde de la biologie moléculaire, de la signalisation des hormones, et qui m'a accompagnée et toujours soutenue jusqu'à présent dans mes travaux de recherche sur les Aux/IAAs, les ARF et la signalisation de l'auxine.

\* **Dr. Jacqueline GRIMA-PETTENATI**, Directrice de recherche CNRS, Chef équipe Génomique Fonctionnelle de l'*Eucalyptus*, Laboratoire de recherche en Sciences Végétales, Toulouse France, avec qui j'ai pu développer des axes de recherche sur la régulation de l'auxine dans la formation du bois. Ce voyage scientifique a changé mon regard sur le monde végétal à jamais. Sa passion ainsi que celle de son équipe pour la régulation transcriptionnelle dans la formation du bois ont été un soutien scientifique et humain précieux dans mes travaux de recherche.

Je tiens particulièrement à remercier l'ensemble des membres des équipes d'accueil, qui m'a accompagnée ces quinze dernières années au laboratoire LRSV et pendant huit ans au laboratoire GBF, pour leurs collaborations, leurs aides, leurs soutenance et leur bienveillance.

Merci Fabien et Chantal pour les travaux de recherche et d'enseignement en équipe. J'en profite aussi pour remercier particulièrement Fabien, qui m'a précédée récemment dans l'exercice de la rédaction et de la soutenance de l'HDR.

Merci Annabelle et Nathalie pour votre bonne humeur quotidienne, vos sourires, vos accompagnements, vos partages, et vos engagements qui éclairent le labo et pour votre ingéniosité qui amènent toujours les solutions originales et efficaces ! Vos expertises par exemple le Fluidigm, RNAseq cellule unique, clonage, transformation hairy root et la transformation permanente, opération FT-IR, la culture *in vitro*, la culture hydroponique... Votre patience et vos accompagnements des stagiaires de notre équipe, nous apportent vraiment beaucoup. Merci beaucoup pour vos aides et vos soutiens à notre équipe et à nos travaux de recherche.

Grands mercis aussi à tous les anciens membres de l'équipe :

Merci Guojian Hu, pour ton aide à la mise en place des technique CRISPR-Cas9 à l'équipe, les expériences de Chromatin immunoprecipitation (ChIP), des trouble-shooting et des discussions stimulantes.

Merci à Marçal et Anna, pour les aventures de doubles hybrides, de MYB, des discussions stimulantes et l'exploitation des familles de facteurs de transcription du génome de l'*Eucalyptus* alors tout juste séquencé dans son intégralité.

Merci à Philippe Ranocha, tu m'as faite entrer dans le monde des parois, appris à travailler avec la tenace *Arabidopsis* pour utiliser ses performantes ressources génomiques.

Merci à Isabelle Mila pour tous les bons moments qu'on partage autour du même bureau pour tes collaborations de localisation subcellulaire et d'essais de transactivation.

Merci à Yves Barriere pour tous vos encouragements et pour les discussions sur le facteur de transcription de lignification chez les arbres, chez l'*Arabidopsis* et sur le maïs.

Merci à Lydie, Pierre, Dominique, Isabelle T., Christiane, Déborah, Laurent, Alain, Jean-Claude, Corinne, Inès, Raphaël, Audrey, Nadia, Edouardo, Victor, Sylvain, Nicolas... C'est avec vous que nous avons pu réaliser ensemble tous ces travaux merveilleux.

Des remerciements tous particuliers à l'attention des jeunes docteurs Hong Yu et Ying Dai que j'ai eu la chance d'accompagner dans l'aventure de leur thèse. Merci à vous toutes les deux, vous avez été des moteurs pour moi. Sur ce sujet, j'en profite pour remercier aussi les membres des comités de thèse de Hong : Françoise Monéger, Didier Aldon, Mondher Bouzayen ; et les membres des comités de thèse de Ying : Christian Dubos, Andreas Niebel, et Nicolas Frei dit Frey, pour leurs précieux conseils et aides qui ont fait avancer nos travaux des thèses, ainsi les jurys de soutenances de thèse : Marie Baucher, Catherine Bellini, Zhengguo Li, Vincent Burlat, Christian Dubos, Virginie Lauvergeat, Richard Sibout pour évaluer nos travaux de recherche, et nous éclairer pour la suite de ces recherche.

Aussi je remercie tous les étudiants, stagiaires que j'ai pu (co-)encadrer : Zhangsheng Zhu, Ipeleng Randome, Yuan Yuan, Yingwe Yang, Lanlan Shan, Andrea Balogh, Hugo Génie, Yawang Wang, Zhengrong Yao, Alexandre Lage, Mingjun Liu, Mengqiang Zhang, Ruta Kardinskaite, Matthieu Bensussan, Chunhong Lin, Tom Lalum, Cristina Barsan, Thierry Nohasiarivelo, Tiffany Yan, Adeline Cros, Alexandre Pillon, Françoise Mesquida, et Marie Alix Marmois. Il serait difficile de défendre une HDR sans leurs contributions.

Je tiens à remercier tout particulièrement mes collègues du LRSV, qui ont participé, aidé de près ou de loin à ces réalisations scientifiques. Un grand merci à Yves Martinez, qui m'a appris les analyses d'histochimie et de microscopie avec une confiance et un soutien constant jusqu'à présent, quelle chance nous avons de pouvoir toujours compter sur lui. Merci à Hélène San-Clemente, Marielle Aguilar et Christophe Dunand pour vos expertises bio-informatiques, notamment vous nous avez beaucoup aidé à réaliser des analyses de facteurs de transcription du génome de l'*Eucalyptus*. Merci à Pierre-Marc Delaux et Malick Mbengue pour leurs aides au cours des expériences de Golden gate cloning et leurs partages des vecteurs. Merci à Jean-Philippe Combier pour son aide aux analyses de microRNAs ciblant les ARF chez l'*Eucalyptus* et sa confiance pour mener le projet commun de miPep-Bois. Merci aussi à Laurent Deslandes, actuellement collègue du LIPMe mais peut être bientôt collègue du notre grand laboratoire commun TPS (Toulouse Plant Science), pour nous aider aux expériences de double hybride.

Je remercie mes collègues enseignant(e)s en équipe pédagogique de Physiologie végétale à l'Université Toulouse III - Paul Sabatier.

J'en profite également pour remercier toutes les personnes bienveillantes qui ont passé du temps à m'aider sur les préparations et réalisations de projets, de soutenances et de concours. Ces vingt dernières années de recherche en sciences végétales m'ont permis de rencontrer des

collaborateurs/trices formidables : chercheurs/euses, des technicien(ne)s, des ingénieurs et des personnels d'appui à la recherche, confirmés ou débutants, dans de nombreuses disciplines scientifiques. Je voudrais les remercier chaleureusement pour m'avoir aidée et accompagnée durant ces années.

Je remercie chaleureusement les membres du jury experts avertis des domaines de recherche comme la régulation transcriptionnelle, la signalisation auxine, la formation des parois, la formation du bois et la régulation de la tolérance au froid chez *Eucalyptus*, qui ont accepté de juger ce travail. Je suis très reconnaissante envers Dr. Marie BAUCHER, Dr. Christian DUBOS, et Dr. Richard SIBOUT pour avoir endossé les rôles de rapporteurs de mon dossier HDR. Je suis très honorée de la venue du Professeur Chantal TEULIÈRES et du Professeur Mondher BOUZAYEN qui apportent une expertise précieuse en tant qu'examineurs. Enfin, je tiens à exprimer ma profonde gratitude au Professeur Chantal TEULIÈRES d'avoir accepté d'être ma marraine du programme du diplôme HDR.

Je remercie ma famille, Cédric, Thomas et Julia, qui m'encouragent toujours et me font progresser aussi sur la communication et la gestion du temps ce qui est très utile pour avancer dans la recherche. Mes remerciements se dirigent aussi pour mes amis, pour leur bienveillance et leurs amitiés qui m'ont soutenue pour ces réalisations scientifiques ; mes parents loin en Chine : Zhengjiao WU et Dingbao WANG, qui s'intéressent toujours à mes travaux et qui me demandent souvent leurs objectifs ; pour mes grands-parents qui ont toujours cru en moi ; et pour mes beaux-parents Gérard et Pierre qui m'ont beaucoup soutenu humainement.

## Liste des annexes

1. Article de thèse : IAA9 tomate
2. Article de premier Post-doc : Fructification transcriptome métabolome
3. Article de second Post-doc : Facteurs de transcription & Formation parois secondaires
4. Un exemple d'articles d'intégration de l'équipe d'accueil : Gènes références *Eucalyptus*
5. Article *ARF* genome wide *Eucalyptus*
6. Article *Aux/IAA Eucalyptus EgrIAA4*
7. Article CRISPR/Cas9 Chez l'*Eucalyptus*
8. Article *EgrIAA20* hypo-lignification fibre secondaire
9. Article IAA9 chez le peuplier
10. Projets financés en tant que co-porteurs du projet
11. Citations des différents travaux de recherche

## Sommaire

<b>REMERCIEMENTS.....</b>	<b>3</b>
<b>LISTE DES ANNEXES.....</b>	<b>6</b>
<b>SOMMAIRE.....</b>	<b>7</b>
<b>CHAPITRE 1. SYNTHÈSE DE MA CARRIÈRE ET DE MES ACTIVITÉS.....</b>	<b>11</b>
I. CURRICULUM VITAE .....	13
A. <i>État Civil.....</i>	13
B. <i>Cursus Professionnel.....</i>	13
C. <i>Formations.....</i>	14
D. <i>Bilan Publications Scientifiques.....</i>	15
<b>1. Dans des journaux internationaux à comité de lecture :</b> .....	15
<b>2. Chapitre de livre : 4.....</b>	15
<b>3. Statistiques des Citations par Google scholar (19.10.2023).....</b>	15
II. LISTE DE PUBLICATIONS .....	16
A. <i>Revue internationale indexées.....</i>	16
B. <i>Contributions à des ouvrages ou des chapitres.....</i>	16
C. <i>Actes de colloques et congrès nationaux et internationaux.....</i>	18
III. ENCADREMENTS .....	21
A. <i>Co-Encadrement de doctorats.....</i>	21
B. <i>Encadrement de Masters.....</i>	22
C. <i>Autres Encadrements.....</i>	23
IV. ACTIVITÉS ET RESPONSABILITÉS D'ENSEIGNEMENT.....	24
A. <i>Activités d'enseignement.....</i>	24
B. <i>Responsabilités pédagogiques.....</i>	25
V. RESPONSABILITÉS COLLECTIVES ET ACTIVITÉ D'EXPERTISE .....	25
A. <i>Participation à des jurys et des comités de thèse.....</i>	25
B. <i>Responsabilités scientifiques.....</i>	26
<b>B1. Obtention de contrats de recherche sur appels d'offre.....</b>	26
<b>B2. Participation aux contrats de recherche.....</b>	26
C. <i>Participation à des conseils d'établissement, à des comités scientifiques.....</i>	28
D. <i>Expertises d'articles ou de projets.....</i>	28
<b>D1. Expertises d'articles scientifiques.....</b>	28
<b>D2. Invitation dans des universités étrangères.....</b>	28

<b>D3. Participation à des projets Européens Erasmus+</b> .....	29
<b>D4. Participation aux activités de valorisation et de vulgarisation scientifiques</b> .....	29
<i>E. Récompenses</i> .....	29
Prix « Best poster » .....	29
Prix conference internationale.....	30
<b>CHAPITRE 2 - PRÉSENTATION ET ANALYSE DE MES ACTIVITÉS DE RECHERCHE AVANT MON RECRUTEMENT À L'UNIVERSITÉ TOULOUSE III PAUL SABATIER</b> .....	<b>32</b>
I. TRAVAUX DE DOCTORAT : CARACTÉRISATION FONCTIONNELLE DU GÈNE <i>IAA9</i> , FACTEUR TRANSCRIPTIONNEL DE LA SIGNALISATION DE L'AUXINE CHEZ LA TOMATE.....	34
<i>Résumé de la thèse</i> .....	34
<i>Contexte scientifique</i> .....	35
<i>Travaux de thèse</i> .....	36
1. Clonage et analyse phylogénétique du gène DR4 chez la tomate .....	37
2. Études d'expression et localisation subcellulaire de la protéine DR4.....	37
3. Génération et caractérisation des plantes transgéniques sous-exprimant le gène DR4 .....	37
<i>Annexe 1 Article de thèse</i> .....	39
II. PREMIER POST-DOCTORAT (2004-2008) : ÉTUDE DE GÉNOMIQUE FONCTIONNELLE DU DÉVELOPPEMENT DU FRUIT DE TOMATE PAR DES APPROCHES COMBINÉES DE TRANSCRIPTOMIQUE ET DE MÉTABOLOMIQUE.....	41
<i>Résumé de Postdoc I</i> .....	41
<i>Travaux de Postdoc I</i> .....	41
1. Analyse cytologique de l'expression de <i>IAA9</i> au cours de la fructification .....	41
2. Etude des facteurs hormonaux impliqués dans le développement du fruit de tomate et leur incidence sur la qualité du fruit .....	41
3. Identification des facteurs clés de la fructification indépendante de la pollinisation chez les plantes AS- <i>IAA9</i> .....	42
4. La régulation transcriptionnelle et post-traductionnelle de l'expression du gène <i>IAA9</i> .....	42
<i>Annexe 2 Article de Postdoc I</i> .....	43
III. DEUXIÈME POST-DOCTORAT (10.2008-08.2009): IDENTIFICATION DE GÈNES RÉGULATEURS IMPLIQUÉS DANS LA FORMATION DE LA PAROI CHEZ <i>ARABIDOPSIS</i> PAR DES APPROCHES POST-GÉNOMIQUES. ....	45
<i>Résumé de Post-doctorat II</i> .....	45
<i>Annexe 3 Article de Post-doctorat II</i> .....	46
<b>CHAPITRE 3 - TRAVAUX DE RECHERCHE EN TANT QUE MAÎTRE DE CONFÉRENCES</b> .....	<b>47</b>
A. PARTICIPATION AUX PROJETS DE L'ÉQUIPE D'ACCUEIL .....	49
<i>Les résumés des publications réalisées dans cet axe</i> :.....	49

1. Cassan-Wang, H., et al. (2012) .....	49
2. Grima-Pettenati, J., et al. (2012) .....	50
3. Courtial, A., et al. (2013) .....	50
4. Myburg, A. A., et al. (2014) .....	51
5. Yu, H., et al. (2014) .....	52
6. Carocha, V., et al. (2015) .....	52
7. Li, Q., et al. (2015) .....	53
8. Soler, M., et al. (2015) .....	53
9. Yu, H., et al. (2015) .....	53
10. Camargo, E., et al. (2019) .....	54
11. Dai, Y., et al (2020) .....	54
12. Hadj Bachir, I., et al (2022) .....	55
<i>Annexe 4. Un exemple des Articles suite à mon intégration dans l'équipe d'accueil.....</i>	<i>57</i>
<b>B. DÉVELOPPEMENT D'UN NOUVEL AXE DE RECHERCHE VISANT À IDENTIFIER DES RÉGULATEURS DE LA VOIE DE L'AUXINE IMPLIQUÉS DANS LA FORMATION DU BOIS D'EUCALYPTUS.....</b>	<b>59</b>
<i>Annexe 5. Article ARF Eucalyptus à l'échelle du génome .....</i>	<i>61</i>
<i>Annexe 6. Article Aux/IAA Eucalyptus à l'échelle du génome et études fonctionnelles du gène EgrIAA4.....</i>	<i>61</i>
<i>Annexe 7. CRISPR/Cas9 Eucalyptus .....</i>	<i>61</i>
<i>Annexe 8. Article EgrIAA20.....</i>	<i>61</i>
<b>C. COLLABORATIONS NATIONALES ET INTERNATIONALES .....</b>	<b>63</b>
<i>Résumés des publications issues des collaborations nationales et internationales.....</i>	<i>63</i>
1. Karannagoda, N., et al. (2022) .....	63
2. Lin, Z., et al. (2019) .....	63
3. Xu, C., et al. (2019) .....	64
4. Zouine, M., et al. (2014) .....	64
<i>Annexe 9. Article IAA9 chez Peuplier .....</i>	<i>67</i>
<b>CHAPITRE 4 - PROJETS DE RECHERCHE EN COURS ET À VENIR .....</b>	<b>69</b>
INTRODUCTION.....	71
<b>A : LA DIFFÉRENTIATION CELLULAIRE AU COURS DE LA FORMATION DU BOIS ET LORS D'UNE RÉPONSE AUX STRESS ABIOTIQUES. ....</b>	<b>72</b>
<i>A1 : la poursuite des travaux sur la caractérisation de la régulation auxine dépendante de la formation du bois. ....</i>	<i>74</i>
1. La surexpression de EgrIAA9A régule sélectivement la maturation et la lignification des vaisseaux et des fibres secondaires dans le xylème secondaire .....	74



2. mi-PEP, la nouvelle clé pour découvrir un nouveau mécanisme de la régulation auxine-dépendant post-transcriptionnelle du développement des xylèmes secondaires ?.....	80
3. Vers de nouveaux modes de régulation de la signalisation de l'auxine par les membres non-classiques des familles Aux/IAA et ARF, possible par trans-localisation subcellulaire de ces facteurs de transcription ?.....	83
<i>A2. Trade-off entre croissance et adaptation aux stress dans le bois chez les arbres : via la signalisation de l'auxine ?.....</i>	<i>84</i>
<i>A3 : Régulation et dynamique de la différenciation cellulaire au cours de la formation du bois par approche transcriptomique type séquençage de l'ARN unicellulaire (scRNA-seq) .....</i>	<i>87</i>
<b>B : PARTICIPATION AUX PROJETS DE NOTRE ÉQUIPE.....</b>	<b>91</b>
<i>B1. Régulation post transcriptionnelle de la formation du bois : Rôle des protéines Musashi</i>	<i>91</i>
<i>B2 : Projet « TYPEX » (2023-2028) : Vers une édition spécifique et précise du génome végétal. ....</i>	<i>92</i>
<b>C : CARACTÉRISATION DES MODIFICATIONS ET ADAPTATIONS DES PAROIS AU COURS DU DÉVELOPPEMENT ET EN RÉPONSE AUX STRESS BIOTIQUES ET ABIOTIQUES.....</b>	<b>94</b>
<i>C1. Caractérisation parois en réponse aux stress biotiques .....</i>	<i>94</i>
<i>C2. Compréhension des modifications de la paroi associées au développement du gel loculaire : Formation et régulation des tissus gel loculaire, la partie manquante du développement du fruit de la tomate.....</i>	<i>98</i>
<b>RÉFÉRENCES .....</b>	<b>101</b>

## CHAPITRE 1. Synthèse de ma carrière et de mes activités



## I. Curriculum Vitae

### Hua Cassan-Wang

#### A. État Civil

Née le 03.05.1976

à Jiangxi, Chine

Mariée - 2 enfants

5 Allée de Nègue Saume

31320 Auzeville Tolosane

Tél. : 06 84 04 01 66

Email : [huawang76@yahoo.com](mailto:huawang76@yahoo.com)

#### **Coordonnées professionnelles :**

UMR5546 UPS-CNRS-INP - Laboratoire de  
Recherche en Sciences Végétales

24, chemin de Borde-Rouge.

31320, Auzeville-Tolosane.

Tél. : 05 34 32 38 51

Email : [hua.cassan@univ-tlse3.fr](mailto:hua.cassan@univ-tlse3.fr)

ORCID <https://orcid.org/0000-0002-7395-9285>

Page Google scholar: <https://scholar.google.com/citations?user=4hsptFOAAAJ&hl=fr>

#### B. Cursus Professionnel

##### Depuis le 1<sup>er</sup> Septembre 2009

Maître de Conférence, Université Toulouse III.

UMR 5546 Laboratoire de Recherche en Sciences Végétales, Équipe « Régulation et Dynamique de la formation du bois chez l'*Eucalyptus* ».

Sujet « Régulation transcriptionnelle de la formation du bois »

##### Octobre 2008 - Aout 2009

Stage Postdoctoral au laboratoire UMR 5546 Surfaces Cellulaires et Signalisation chez les Végétaux, CNRS/ Université P. Sabatier, Toulouse, France :

Sujet Identification des gènes régulateurs impliqués dans la formation de la paroi cellulaire secondaire, dans le cadre du projet Européen FP7 : « Improving plant cell walls for use as a renewable industrial feedstock »

##### Avril 2004 – Septembre 2008

Stage Postdoctoral dans le laboratoire Génomique et Biotechnologie des Fruits INRA/INP-ENSAT, Toulouse, France :

Chargée d'analyse transcriptomique par puces à ADN et d'analyse fonctionnelle des gènes par transgénèse dans le cadre du projet Trilatéral sur le sujet : « L'exploration de la variabilité naturelle du développement et de la qualité du fruit de tomate par des approches de transcriptomique et de métabolomique », puis du projet européen EU-SOL sur le sujet « mécanisme moléculaire du développement du fruit de tomate ».

## C. Formations

Sep 2000 - Mars 2004 **Doctorat** en Biologie cellulaire et moléculaire végétale (Bourses : Egide et AFCRST, Mention Très Honorable)

UMR990 Génomique et Biotechnologie des Fruit, ENSAT INP Toulouse,  
Directeur de thèse : Prof Bouzayen MONDHER.

Sujet : La Régulation Transcriptionnelle Dépendante de L'auxine Lors du Développement du Fruit : Caractérisation Fonctionnelle du Gène *DR4*, un Homologue d'*Aux/IAA* Chez la Tomate (*Lycopersicon esculentum*, Mill).

Sep 1997 - Juillet 2000 **Master** en Science Alimentaire (Bac + 7, Mention TB, major de promo)

Laboratory « Fruits Storage and Transformation », South China Agricultural University, Guangzhou, China. Directeur du stage de recherche (2 ans): Prof Weixin CHEN

Sujet : Studies on Postharvest pollution-free treatments for Banana disease (*Brazilian, Musa spp. AAB Group*) and its mechanisms.

Sep 1993 - Juillet 1997 **Bachelor** en "Horticulture and Gardening" (Bac +4, major de promo) South China Agricultural University, Guangzhou, China

## D. Bilan Publications Scientifiques

### 1. Dans des journaux internationaux à comité de lecture :

Articles en tant que 1<sup>er</sup> Auteur : **5** (Plant Cell, PCP, Front Plant Sci)

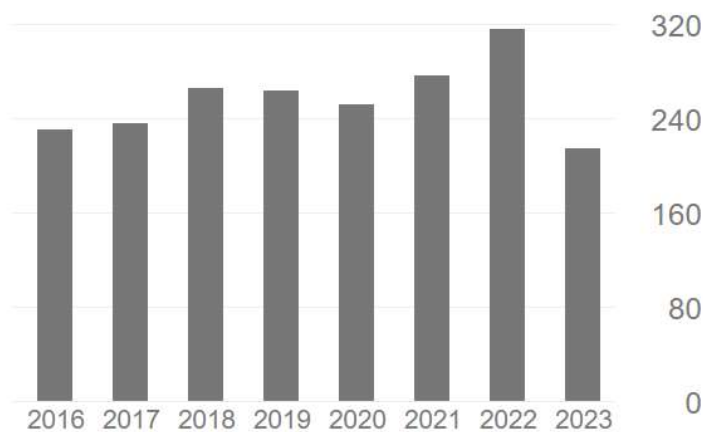
Articles en tant que dernier auteur/corresponding author : **6** (New Phytol. IJMS. PCP, PLoS One)

Articles en tant que co-auteur : **22** (Nature, Plant Mol Bio, J Exp Bot)

### 2. Chapitre de livre : 4

### 3. Statistiques des Citations par Google scholar (19.10.2023)

	Toutes	Depuis 2018
Citations	2857	1592
indice h	17	15
indice i10	19	19



## II. Liste de publications

### A. Revue internationale indexée

1. Yu, H., Liu, M., Zhu, Z., Wu, A., Mounet, F., Pesquet, E., Grima-Pettenati, J., & **Cassan-Wang, H\***. (2022). Overexpression of EgrIAA20 from *Eucalyptus grandis*, a Non-Canonical Aux/IAA Gene, Specifically Decouples Lignification of the Different Cell-Types in Arabidopsis Secondary Xylem. *Int J Mol Sci*, 23(9). <https://doi.org/10.3390/ijms23095068> (IF=6.2)
2. Karannagoda, N., Spokevicius, A., Hussey, S., **Cassan-Wang, H.**, Grima-Pettenati, J., & Bossinger, G. (2022). *Eucalyptus grandis* AUX/INDOLE-3-ACETIC ACID 13 (EgrIAA13) is a novel transcriptional regulator of xylogenesis. *Plant Mol Biol*, 109(1-2), 51-65. <https://doi.org/10.1007/s11103-022-01255-y> (IF=4.1)
3. Dai, Y., Hu, G., Dupas, A., Medina, L., Blandels, N., Clemente, H. S., Ladouce, N., Badawi, M., Hernandez-Raquet, G., Mounet, F., Grima-Pettenati, J., & **Cassan-Wang, H.\*** (2020). Implementing the CRISPR/Cas9 Technology in *Eucalyptus* Hairy Roots Using Wood-Related Genes. *Int J Mol Sci*, 21(10). <https://doi.org/10.3390/ijms21103408> (IF=6.2)
4. Xu, C., Shen, Y., He, F., Fu, X., Yu, H., Lu, W., Li, Y., Li, C., Fan, D., **Cassan-Wang, H.\*** & Luo, K. (2019). Auxin-mediated Aux/IAA-ARF-HB signaling cascade regulates secondary xylem development in *Populus*. *New Phytol*, 222(2), 752-767. <https://doi.org/10.1111/nph.15658> (IF=7.4)
5. Lin, Z., Long, J. M., Yin, Q., Wang, B., Li, H. L., Luo, J. Z., **Cassan-Wang, H.**, & Wu, A. M. (2019). Identification of novel lncRNAs in *Eucalyptus grandis*. *Industrial Crops and Products*, 129, 309-317. <https://doi.org/10.1016/j.indcrop.2018.12.016> (IF=3.9)
6. Yu, H., Soler, M., San Clemente, H., Mila, I., Paiva, J. A., Myburg, A. A., Bouzayen, M., Grima-Pettenati, J., & **Cassan-Wang, H.\*** (2015). Comprehensive genome-wide analysis of the Aux/IAA gene family in *Eucalyptus*: evidence for the role of EgrIAA4 in wood formation. *Plant Cell Physiol*, 56(4), 700-714. <https://doi.org/10.1093/pcp/pcu215> (IF=5.0)
7. Myburg, A. A., Grattapaglia, D., Tuskan, G. A., Hellsten, U., Hayes, R. D., Grimwood, J., Jenkins, J., Lindquist, E., Tice, H., Bauer, D., Goodstein, D. M., Dubchak, I., Poliakov, A., Mizrachi, E., Kullán, A. R., Hussey, S. G., Pinard, D., van der Merwe, K., Singh, P... **Cassan-Wang, H.** ... Schmutz, J. (2014). The genome of *Eucalyptus grandis*. *Nature*, 510(7505), 356-362. <https://doi.org/10.1038/nature13308> (IF=42)
8. Soler, M., Camargo, E. L., Carocha, V., **Cassan-Wang, H.**, San Clemente, H., Savelli, B., Hefer, C. A., Paiva, J. A., Myburg, A. A., & Grima-Pettenati, J. (2015). The *Eucalyptus grandis* R2R3-MYB transcription factor family: evidence for woody growth-related evolution and function. *New Phytol*, 206(4), 1364-1377. <https://doi.org/10.1111/nph.13039> (IF=6.6)
9. Carocha, V., Soler, M., Hefer, C., **Cassan-Wang, H.**, Fevereiro, P., Myburg, A. A., Paiva, J. A., & Grima-Pettenati, J. (2015). Genome-wide analysis of the lignin toolbox of *Eucalyptus grandis*. *New Phytol*, 206(4), 1297-1313. <https://doi.org/10.1111/nph.13313> (IF=6.6)

10. Li, Q., Yu, H., Cao, P. B., Fawal, N., Mathe, C., Azar, S., **Cassan-Wang, H.**, Myburg, A. A., Grima-Pettenati, J., Marque, C., Teulieres, C., & Dunand, C. (2015). Explosive tandem and segmental duplications of multigenic families in *Eucalyptus grandis*. *Genome Biol Evol*, 7(4), 1068-1081. <https://doi.org/10.1093/gbe/evv048> (IF=4.5)
11. Yu, H., Soler, M., Mila, I., San Clemente, H., Savelli, B., Dunand, C., Paiva, J. A., Myburg, A. A., Bouzayen, M., Grima-Pettenati, J., & **Cassan-Wang, H.\*** (2014). Genome-wide characterization and expression profiling of the AUXIN RESPONSE FACTOR (ARF) gene family in *Eucalyptus grandis*. *PLoS One*, 9(9), e108906. <https://doi.org/10.1371/journal.pone.0108906> (IF=4.2)
12. Zouine, M., Fu, Y., Chateigner-Boutin, A. L., Mila, I., Frasse, P., **Cassan-Wang, H.**, Audran, C., Roustan, J. P., & Bouzayen, M. (2014). Characterization of the tomato ARF gene family uncovers a multi-levels post-transcriptional regulation including alternative splicing. *PLoS One*, 9(1), e84203. <https://doi.org/10.1371/journal.pone.0084203> (IF=4.2)
13. **Cassan-Wang, H.**, Goue, N., Saidi, M. N., Legay, S., Sivadon, P., Goffner, D., & Grima-Pettenati, J. (2013). Identification of novel transcription factors regulating secondary cell wall formation in *Arabidopsis*. *Front Plant Sci*, 4, 189. <https://doi.org/10.3389/fpls.2013.00189> (IF=4.5)
14. Courtial, A., Soler, M., Chateigner-Boutin, A. L., Reymond, M., Mechin, V., **Wang, H.**, Grima-Pettenati, J., & Barriere, Y. (2013). Breeding grasses for capacity to biofuel production or silage feeding value: an updated list of genes involved in maize secondary cell wall biosynthesis and assembly. *Maydica*, 58(1-4), 67-102 <https://hal.inrae.fr/hal-02651049> (IF=0.9)
15. **Cassan-Wang, H.**, Soler, M., Yu, H., Camargo, E. L., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A., Leple, J. C., & Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of *Eucalyptus* grown in various environmental conditions. *Plant Cell Physiol*, 53(12), 2101-2116. <https://doi.org/10.1093/pcp/pcs152> (IF=5.0)
16. **Wang, H.**, Soler, M., Yu, H., Camargo, E. L. O., Clemente, H. S., Savelli, B., Ladouce, N., Paiva, J., & Grima-Pettenati, J. (2011). Master regulators of wood formation in *Eucalyptus*. *BMC Proceedings*, 5(S7), P110. <https://doi.org/10.1186/1753-6561-5-s7-p110>
17. Barriere, Y., Mechin, V., Lafarguette, F., Manicacci, D., Guillon, F., **Wang, H.**, Lauressergues, D., Pichon, M., Bosio, M., & Tatout, C. (2009). Toward the discovery of maize cell wall genes involved in silage quality and capacity to biofuel production. *Maydica*, 54(2-3), 161-198. <Go to ISI>://000274052700004 (IF=0.9)
18. Chaabouni, S., Jones, B., Delalande, C., **Wang, H.**, Li, Z., Mila, I., Frasse, P., Latche, A., Pech, J. C., & Bouzayen, M. (2009). SI-IAA3, a tomato Aux/IAA at the crossroads of auxin and ethylene signalling involved in differential growth. *J Exp Bot*, 60(4), 1349-1362. <https://doi.org/10.1093/jxb/erp009> (IF=6.0)
19. **Wang, H.**, Schauer, N., Usadel, B., Frasse, P., Zouine, M., Hernould, M., Latche, A., Pech, J. C., Fernie, A. R., & Bouzayen, M. (2009). Regulatory features underlying pollination-dependent and -independent tomato fruit set revealed by transcript and primary metabolite profiling. *Plant Cell*, 21(5), 1428-1452. <https://doi.org/10.1105/tpc.108.060830> (IF=9.3)



20. **Wang, H.**, Jones, B., Li, Z., Frasse, P., Delalande, C., Regad, F., Chaabouni, S., Latche, A., Pech, J. C., & Bouzayen, M. (2005). The tomato Aux/IAA transcription factor IAA9 is involved in fruit development and leaf morphogenesis. *Plant Cell*, 17(10), 2676-2692. <https://doi.org/10.1105/tpc.105.033415> (IF=9.3)

## B. Contributions à des ouvrages ou des chapitres

1. Bachir, I. H., Ployet, R., Teulieres, C., **Cassan-Wang, H.**, Mounet, F., & Grima-Pettenati, J. (2022). Regulation of secondary cell wall lignification by abiotic and biotic constraints. In R. Sibout (Ed.), *LIGNIN AND HYDROXYCINNAMIC ACIDS: Biosynthesis and the Buildup of the Cell Wall* (Vol. 104, pp. 363-392). Academic Press Ltd-Elsevier Science Ltd. <https://doi.org/10.1016/bs.abr.2022.03.008>
2. Camargo, E. L. O., Ployet, R., **Cassan-Wang, H.**, Mounet, F., & Grima-Pettenati, J. (2019). Digging in wood: new insights in the regulation of wood formation in tree species. *Molecular Physiology and Biotechnology of Trees*, 89, 201-233. <https://doi.org/10.1016/bs.abr.2018.11.007>
3. Grima-Pettenati, J., Soler, M., Camargo, E. L. O., & **Wang, H.** (2012). Transcriptional Regulation of the Lignin Biosynthetic Pathway Revisited: New Players and Insights. *Lignins: Biosynthesis, Biodegradation and Bioengineering*, 61, 173-218. <https://doi.org/10.1016/B978-0-12-416023-1.00006-9>
4. **Wang, H.**, Li, Z., Frasse, P., Audran, C., Jones, B., Jauneau, A., & Bouzayen, M. (2003). Characterisation of Aux/IAA-like genes expressed in tomato fruit. *Biology and Biotechnology of the Plant Hormone Ethylene*, 349, 309-310. <Go to ISI>://WOS:000226793100082

## C. Actes de colloques et congrès nationaux et internationaux

1. 2023 **Cassan-Wang H\*** (conférence invitée, présentation orale) Auxin-dependant Regulatory networks underlying wood formation during development and in response to abiotic stress in *Eucalyptus*. In Journées Xylogénèse 12-13 Jun 2023, Nancy, France
2. 2023 H Yu, M Li, Z Zhu, A Wu, F Mounet, E Pesquet, J Grima-Pettenati, **H Cassan-Wang\***, 2022 Overexpression of EgrIAA20 from *Eucalyptus grandis*, a Non-Canonical Aux/IAA Gene, Specifically Decouples Lignification of the Different Cell-Types in *Arabidopsis* Secondary Xylem (**Conférence invitée**). SouthWest University, Chongqing
3. 2022 H Yu, M Li, Z Zhu, A Wu, F Mounet, E Pesquet, J Grima-Pettenati, **H Cassan-Wang\***, 2022 Overexpression of EgrIAA20 from *Eucalyptus grandis*, a Non-Canonical Aux/IAA Gene, Specifically Decouples Lignification of the Different Cell-Types in *Arabidopsis* Secondary Xylem (Poster). The 20th IUFRO Tree Biotech and the 2nd FTMB (Forest Tree Molecular Biology and Biotechnology) Conference; Harbin, China; 6-9 July 2022.
4. 2019 Dai Y, Dupas A, N. L, Hu G, San Clemente H, Grima-Pettenati J, **Cassan-Wang H\*** (Poster) Knock-out of Cinnamoyl CoA Reductase (CCR) using CRISPR/Cas9 gene editing technology in *Eucalyptus grandis* hairy roots In 2019 IUFRO Tree Biotechnology Conference, 23-28 June 2019. Raleigh, NC, USA

5. 2019 Dai Y, Dupas A, N. L, San Clemente H, Grima-Pettenati J, **Cassan-Wang H\* (présentation orale sélectionnée)** Knock out of Cinnamoyl CoA reductase, a key lignin biosynthetic gene, using CRISPR/Cas9 gene editing technology in *Eucalyptus* hairy roots In 12e journées du Réseau Français des Parois, 14-16 mai 2016, Roscoff, France
6. 2018 **Cassan-Wang H\* (conférence invitée, présentation orale)** Regulatory networks underlying wood formation during development and in response to abiotic stress in *Eucalyptus*. In Journées Xylogénèse 12-13 Jun 2018, Paris
7. 2018 **Cassan-Wang H\***, Grima-Pettenati J (**conférence invitée, présentation orale**) Regulatory networks underlying wood formation during development and in response to abiotic stress in *Eucalyptus*. In International Symposium on Forest Tree Molecular Biology and Biotechnology, 24-26 July 2018, Harbin, China
8. 2018 Yu H, Liu M, Bouzayen M, Grima-Pettenati J, **Cassan-Wang H\*** (Poster) New Auxin Transcription Factors Regulating Lignified Wood Cell Formation. In International Symposium on Forest Tree Molecular Biology and Biotechnology, 24-26 July 2018, Harbin, China
9. 2017 **Cassan-Wang H**, Yu H, Liu M, Yao Z, Lemonnier L, Saint-Martin D, Pirrello J, Bouzayen M, Grima-Pettenati J (**présentation orale sélectionnée**) New Auxin Transcription factors regulating lignified secondary cell wall formation In 11es Journées du Réseau Français des Parois, 27-29 June 2017, Orléans, France
10. 2017 **Cassan-Wang H**, Yu H, Liu M, Grima-Pettenati J, (**Conférence invitée**) New Auxin Transcription factors regulating lignified secondary cell wall formation In Invited seminar, South China Agriculture University, Guangzhou, China
11. Oct 2016 **Cassan-Wang H (conférence invitée)** « Functional characterization of ARF and *Aux/IAA* gene families in *Eucalyptus* and their roles in wood formation » à l'Académie de foresterie de Guangdong, Canton, Chine.
12. Oct 2016 **Cassan-Wang H (conférence invitée)** « Higher education in France and general introduction of France agriculture and foresterie » au Guangdong Ecological Engineering College, Canton, Chine.
13. Oct 2016 **Cassan-Wang H (conférence invitée)** « L'exploitation du bois en France » à l'Institut des arbres à Canton, Chine.
14. 2014 Carocha, V., M. Soler, C. Hefer, H. **Cassan-Wang**, P. Faveiro, A. A. Myburg, J. Paiva and J. Grima-Pettenati (**présentation orale**). Gene diversity and regulation underlying phenylpropanoid metabolism and lignin biosynthesis in *Eucalyptus grandis*. Lignin 2014: Biosynthesis and utilization. 24-28 August 2014, Umea, Sweden.
15. 2013 **Hua Cassan-Wang**, Marçal Soler, Hong Yu, Anna Plasencia, Hélène San Clemente, Nathalie Ladouce, Charles Hefer, Zander Myburg, Jorge Pinto Paiva and Jacqueline Grima-Pettenati (**conférence invitée, présentation orale**) New transcription factors regulating secondary xylem differentiation in *Eucalyptus*. International Plant Vascular Biology conferences 2013, 25-30 July, Helsinki, Finland
16. 2013 **Hua Cassan-Wang**, Marçal Soler, Hong Yu, Anna Plasencia, Hélène San Clemente, Nathalie Ladouce, Charles Hefer, Zander Myburg, Jorge Pinto Paiva and Jacqueline Grima-Pettenati (Poster) New transcription factors regulating lignified secondary cell wall formation in *Eucalyptus*. The XIII Cell Wall Meeting, 7-12 July 2013, Nantes, France

17. 2013 **Hua Cassan-Wang**, Hong Yu, Marçal Soler, Jacqueline Grima-pettenati (**conférence invitée, présentation orale**) Construction and functional characterization of *Eucalyptus* ARF and *Aux/IAA* genes during wood formation. Tree For Joules 4th Meeting, 22 - 24 Avril 2013, Hamburg, Germany
18. 2012 **Hua Wang**, Hong Yu, Marçal Soler, Jacqueline Grima-pettenati (Présentation orale) Expression profiling and functional characterization of Auxin Response Factor (ARF) gene family in *Eucalyptus*. Tree For Joules 3rd Meeting, 11 - 13 Avril 2012, Malaga, Spain
19. 2012 **Hua Wang**, Marçal Soler, Najib Saïdi, Steven Hussey, Hong Yu, Victor Carocha, Eduardo Camargo, Nathalie Ladouce, Hélène San clemente, Charles Hefer, Bruno Savelli, Zander Myburg, Jorge Pinto Paiva, Jacqueline Grima-pettenati (**Présentation orale**) Master Regulators of Wood Formation in *Eucalyptus*. International Plant and Animal genome XX, Jan. 2012, San Diego, USA
20. 2011 **Hua Wang**, Hong Yu, Marçal Soler, Hélène San clemente, Eduardo Camargo, Najib Saïdi, Christophe Dunand, Nathalie Ladouce, Jorge Pinto Paiva, Jacqueline Grima-pettenati (**Présentation orale**) Auxin Signaling Transcription Factors in *Eucalyptus* Wood Formation. 9ème Colloque National de la SFBV, 12 - 14 Déc. 2011, Clermont Ferrand, France
21. 2011 **Hua Wang**, Hong Yu, Marçal Soler, Najib Saïdi, Nathalie Ladouce, Jean-Charles Leplé, Jacqueline Grima-pettenati (**Présentation orale**) Identification and evaluation of new reference genes in *Eucalyptus* for accurate normalization of quantitative RT-qPCR data. Tree For Joules 2nd Meeting, 11 - 13 Oct. 2011, Lisbonne, Portugal
22. 2011 Hong Yu, Hélène San Clemente, Christophe Dunand, Marçal Soler, Eduardo Camargo, Jorge Pinto Paiva, Jacqueline Grima-Pettenati, **Hua Wang** (Poster) Studies of auxin responses mediators in wood formation in *Eucalyptus*. 26th International New Phytologist Symposium "Bioenergy Trees", 17-19 Mai 2011, Nancy France
23. 2011 **Hua Wang**, Nathalie Ladouce, Jacqueline Grima-pettenati (Présentation orale) Fluidigm: a powerful tool for high throughput quantitative RT-qPCR analysis. Tree For Joules 1st Meeting, 11 - 12 April. 2011, Toulouse, France
24. 2010 **Hua Wang**, Sylvain Legay, Nadia Goué, Philippe Ranocha, Jacqueline Grima Pettenati, Deborah Goffner (Poster) Identification of new transcription factors involved in secondary cell wall formation in *Arabidopsis* through post-genomic approaches The XII International Cell Wall Meeting, 25-30 Jul. 2010, Porto, Portugal
25. 2010 **Hua Wang**, Sylvain Legay, Philippe Ranocha, Deborah Goffner, Jacqueline Grima Pettenati (Présentation orale) To Identify novel transcription factors involved in secondary cell wall formation using postgenomic approaches. RENEWALL Meeting (Eu projet meeting), Jan 2010, Toulouse, France
26. 2009 **Hua Wang**, Philippe Ranocha, Jacqueline Grima Pettenati, Deborah Goffner (Présentation orale) Looking for key regulatory genes involved in secondary cell wall formation. 1st Consortium of RENEWALL (International), Jan 2009, York, Royaume-Uni
27. 2007 **Wang H**, Schauer N, Frasse P, Zouine M, Latché A, Pech J, Fernie A and Bouzayen M (2007) Combined transcriptomic & metabolomic analyses define specific signaling & metabolic pathways associated with fruit set in tomato. Conférence invitée, présentation orale. The 4th Solanacea Genome Workshop, Jeju Island, Korea.
28. 2006 Audran C, Chaabouni S, **Wang H**, Mila I, Regad F, Bouzayen M (2006) Ethylene and auxin interplay throughout fruit development and Ripening. Conférence invitée, Présentation orale. 7th

International Symposium on the Plant Hormone Ethylene. Pisa, Italy

29. 2005 **Wang H**, Jones B, Li ZG, Frasse P, Delalande C, Regad F, Latché A, Pech JC and Bouzayen M. LeIAA9, a tomato Aux/IAA gene involved in Fruit Development and leaf Morphogenesis. Poster. Xth France-Japan Workshop on Plant Sciences. Toulouse, France.
30. 2004 **Wang H**, Jones B, Li Z, Frasse P, Delalande C, Laché A, Pech J, Bouzayen M. *DR4*, a tomato Aux/IAA involved in a specific set of auxin-dependent developmental processes. Poster. 2nd International Conference of Auxin. Crète, Grèce.
31. 2003 **Wang H**, Jones B, Frasse P, Bouzayen M. Characterization of DR4 a tomato Aux/IAA like gene. Poster. 7th international congress of plant molecular biology. Barcelona (Espagne).
32. 2002 **Wang H**, Jones B, Bouzayen M. Characterization of two Aux/IAA genes *DR3* and *DR4* Functions in Tomato. Presentation orale. Second international consortium of tomato. Nottingham University, Royaume-Uni.

### III. Encadrements

#### A. Co-Encadrement de doctorats

- **Sep. 2017 – dec. 2020 Ying DAI** « Mise au point de la technologie CRISPR/Cas9 dans des racines « hairy root » d'*Eucalyptus grandis* et caractérisation fonctionnelle de facteurs de transcription de la signalisation auxinique potentiellement impliqués dans la formation du bois ». Thèse co-encadrée avec Dr. Jacqueline Grima-Pettenati (50%). Suite à un appel d'offre gouvernemental Chinois (China Scholarship Council, CSC), nous avons obtenu l'allocation de thèse pour Mlle Ying DAI. Malgré des difficultés liées à la crise sanitaire de COVID-19 (nous avons perdu de nombreux matériels biologiques spécifiques (>400 plantes transgéniques) pendant les périodes de confinement), la thèse a été soutenue comme prévu en décembre 2020. **Nous avons publié un article en premier auteur (2020 IF=4.5) (Dai et al., 2020).**
- **Oct. 2010 – oct. 2014 Hong YU** « Régulation transcriptionnelle auxine dépendante lors de la formation du bois chez l'*Eucalyptus* ». Thèse co-encadrée avec Jacqueline Grima-Pettenati (50%). Suite à un appel d'offre gouvernemental Chinois (China Scholarship Council, CSC), j'ai obtenu l'allocation de thèse (4 ans) pour Hong Yu, doctorante chinoise. Malgré des difficultés liées à des changements culturels et de sujet d'étude, la thèse se déroule bien et la thèse est achevée comme prévu en septembre 2014. **Nous avons publié 7 articles dont 5 en premiers auteurs (2014, 2015a, 2015b, 2019, 2022)** deux autres articles en 2011 et 2012 (Cassan-Wang et al., 2012; Li et al., 2015; Xu et al., 2019; Yu et al., 2022; Yu et al., 2014; Yu et al., 2015). Après la thèse elle a décroché un poste permanent de maître de conférences à l'Université médical de Sichuan en Chine.
- **Aout 2018 – 2024 Ipeleng RANDEME** "Evaluation of novel ARF and MYB *Eucalyptus* wood formation transcription factor candidates using hairy root and poplar reverse genetics systems" Co-supervisor (25%) avec Dr. Steven Hussey (main supervisor) et Prof. Zander Myburg, Department of Biochemistry, Genetics & Microbiology, University of Pretoria, South Africa.

- **Nov. 2017 – oct. 2018 Zhangsheng ZHU** « Mise en place de la technique DAPseq au labo et criblage des partenaires protéiques des gènes de réponse à l'auxine par approche de double hybride chez la levure ». Co-encadrée avec Jacqueline Grima-Pettenati et Fabien Mounet. Stage de doctorat en visite de 12 mois avec financement chinois. Une partie de son travail a été publiée en 2020 (Yu et al., 2022).
- **Mars – oct. 2018 Yuan YUAN** « Caractérisation des peupliers transgéniques sur-exprimant le gène IAA13 de *Eucalyptus* ». Stage de doctorante en visite de 6 mois avec financement chinois.
- **2007-2008 Yingwu YANG** Functional characterisation of MicroRNA167 and its regulation in fruit development in tomato. Co-encadrement pour un stage de doctorant en visite de 12 mois avec financement chinois.
- **2007-2008 Lanlan SHAN** Functional characterisation of MicroRNA160 and its regulation in fruit development in tomato. Co-encadrement pour un stage de doctorante en visite de 6 mois avec financement chinois.
- **2004 Andrea BALOGH** (Doctorante Hongroise dans le cadre d'un projet Marie Curie) Microarray identification of key factors in transgenic plants AS-IAA9. Co-encadrement pour un stage de 6 mois en anglais, rédaction d'un rapport et soutenance orale

## B. Encadrement de Masters

- **Jan. – juin 2023 Hugo GÉNIE** (Master II) Auxin regulation of wood cell differentiation: functional characterization of *EgrIAA9A* and *EgrIAA20* during wood formation in *Eucalyptus*. Co-encadrement avec Dr. Fabien Mounet.
- **Jan. – juin 2018 Yawang WANG** (master I international) « Cribler les lignes transgéniques d'*Eucalyptus* par la technique FT-IR couplée avec l'analyse PCA pour valider préliminaire les fonctions des gènes des facteurs de transcription de la famille MYB ». Stage de master international de recherche de 6 mois dans le cadre de la collaboration avec l'Université de Chongqing (programme échange international). Après ce stage, il a poursuivi avec une thèse aux États-Unis.
- **Jan. – juin 2017 Zhengrong YAO** (Master I international) « A New Plant Model to Decipher the Auxin-Dependent Molecular Mechanisms of Secondary Xylem Differentiation ». Stage de recherche de 6 mois du Master international dans le cadre de la collaboration avec l'Université de Chongqing (programme échange international). Sur un axe de recherche innovant que j'ai élaboré pour répondre à l'appel d'offre des projets IFR2016, qui propose d'utiliser la plante de tomate comme plante modèle pour l'étude du développement du bois au lieu de *Arabidopsis*, et de cribler une collection locale unique au monde des plantes de tomate affectant des facteurs de transcriptions impliqués dans la signalisation de l'auxine. L'avantage est que cette plante de tomate possède une croissance secondaire véritable que n'a pas *Arabidopsis* tout en conservant le bénéfice d'une plante modèle pratique à utiliser en laboratoire pour l'analyse du développement des arbres.

- **Mars - mai 2017 Alexandre LAGE** (Master I) « Tomato as an alternative model to study auxin-dependent mechanisms involved in wood formation » Stage M1 de l'Université Toulouse III (3 mois).
- **2015 Mingjun LIU** (Master I international) Caractérisation fonctionnelle des gènes *EgrIAA20* et *EgrIAA9A* dans la formation du bois. Stage de recherche de 6 mois du Master international dans le cadre de la collaboration avec l'Université de Chongqing (programme échange international). Le stage s'est bien déroulé et il a produit des très bons résultats qui ont fait l'objet de l'article en deuxième auteur (Yu et al., 2022). Après ce stage, il a poursuivi avec une thèse à Northwest University, Washington, United States.
- **2014 MengQiang ZHANG** (Master I international) « Étude d'expression des gènes de biosynthèse des parois secondaires chez les *Arabidopsis* OE *EgrIAA4m*, OE *EgrIAA20* et OE *EgrIAA9A* ». Stage de recherche de 6 mois dans le cadre de la collaboration avec l'Université de Chongqing. Le stage s'est bien déroulé et les résultats produits font partie du projet thèse de Mlle Hong YU.
- **2011-2012 Ruta KARDINSKAITE** (Master 2) international. « Regulation involved in secondary wall formation in *Arabidopsis* and study of impact on saccharification » Master international (AgroFood Chain, UPS), formation en anglais.
- **2010 Matthieu BENSUSSAN** (Master 2) Recherche. « Identification des partenaires de master régulateurs *EgMYB1* et *EgMYB2* lors de la formation du bois chez l'*Eucalyptus* ». Devenir de l'étudiant : Doctorant CIFRE-INRA Versailles.
- **2009 Chunhong LIN** (Master 1) « Identification chez l'*Arabidopsis* des facteurs de transcription impliqués dans la formation de la paroi secondaire ». Master international, formation délivrée en anglais. Devenir de l'étudiant : CDI en entreprise en Chine
- **2008 Tom LALOUM** (Master1) Caractérisation de la séquence du gène *LeIAA9* chez le mutant de tomate *entire*. Rédaction d'un rapport et soutenance orale.
- **2005-2006 Cristina BARSAN** (Master2) Role of the *SlIAA9* transcription factor on fruit quality using reverse genetics in tomato. Co-encadrement en anglais, rédaction d'un rapport et soutenance orale.
- **2005-2006 Thierry NOHASIARIVELLO** (Master2) Étude d'expression de gènes en réponse à un traitement aux cytokinines chez la tomate. Rédaction d'un rapport et soutenance orale.

### C. Autres Encadrements

- **2023 Tiffany YAN**, Licence 2 SdV BBE, Université Toulouse III - Paul Sabatier - FSI – Department Biologie et Geosciences- Auxin regulation of wood cell differentiation: functional characterization of *EgrIAA4* and *EgrIAA13* during wood formation in Poplar, Stage facultatif (semestre 2).
- **2010 Adeline CROS**, Licence 3 professionnelle « Recherche de partenaires des facteurs de transcriptions impliqués dans la formation du bois chez l'*Eucalyptus* ». Rédaction d'un rapport et soutenance orale.
- **2006 Tom LALOUM** (Licence) Rôle du gène *LeIAA9* chez la tomate. Rédaction d'un rapport et soutenance orale.

- **2006 Alexandre PILLON** (CCP3) Influence de l'auxine et des gibbérellines dans le processus de fructification chez la tomate. Rédaction d'un rapport et soutenance orale.
- **2005 Françoise MESQUIDA** (IUT) Étude des facteurs hormonaux impliqués dans le processus de nouaison de la tomate. Rédaction d'un rapport et soutenance orale.
- **2003 Marie Alix MARMOIS** (Maîtrise). Analyse de l'expression du gène DR4 chez la tomate.

## IV. Activités et responsabilités d'enseignement

### A. Activités d'enseignement

Depuis mon recrutement en tant que MCF, j'interviens au sein d'une équipe pédagogique de 20 MCF et PR, pour enseigner la physiologie végétale :

- **En L1** Licence 1<sup>ère</sup> année, je réalise des Travaux Dirigés (TD) et des Travaux pratiques (TP) dans l'Unité d'Enseignement (UE) de "Physiologie Végétale" dispensée à plus de 900 étudiants chaque année (responsables : F. Mounet, N. Frei-Dit-Frey). Cette UE représente >600h annuelles réalisées par une équipe d'environ 15 enseignants, doctorants et vacataires. Nous y abordons des notions fondamentales de la biologie des plantes : le fonctionnement des méristèmes, les grandes étapes du développement des organismes végétaux, les voies de signalisations hormonales, la réponse des plantes aux contraintes de l'environnement et la germination des graines.

- **En L2** Licence 2<sup>ème</sup> année, je réalise des cours magistraux (CM), des Travaux Dirigés (TD) et des Travaux pratiques (TP) dans l'UE de « Physiologie Végétale II : Photosynthèse et nutrition en eau et azote » (responsable : H. Cassan-Wang). Dans cette UE, j'ai créé une pièce de théâtre pour illustrer les mécanismes de la photosynthèse, comment les plantes transforment l'énergie solaire en l'énergie interne chimique (ATP) et l'énergie redox (NADPH) grâce aux différents acteurs. J'ai créé et j'assure des CM, TD et TP dans l'UE de « Biologie du Développement » (co-responsables : D. Aldon, H. Cassan-Wang). Nous y abordons les mécanismes impliqués dans le fonctionnement des méristèmes, la perception du signal lumière, la fécondation et la remobilisation des réserves lors de la germination.

- **En L3** Licence 3<sup>ème</sup> année, j'ai réalisé des CM dans l'UE de "Valorisation du Végétal" (responsable : F. Mounet). Dans cette UE, nous essayons de sensibiliser les étudiants aux procédés de valorisation agro-industrielles des ressources végétales telles que les biomasses et le bois.

- **En Master 1**, j'ai créé et j'ai réalisé des CM et TP sur les signalisations des hormones (Auxine et Éthylène) et travaux pratiques « maturation des fruits » dans l'UE « Biologie du Développement Végétal » (responsable D. Aldon ; 2009-2022) ; j'ai créé et je réalise actuellement des CM et TD sur les croissances secondaires et la formation du bois dans l'UE Morphologie, Anatomie & Morphogenèse Végétales (responsable H. Gryta). J'ai encadré des stages M1 classique et 4 stages Master 1 International

- **En Master 2**, j'ai créé des CM sur la régulation transcriptionnelle de la formation des parois secondaires ; j'ai encadré des stages (3 M2) et j'ai participé au jury des ateliers et fin d'année régulièrement.

- **En Master International** (cours dispensés en Anglais). J'ai réalisé des CM et TD sur les bases moléculaires pour l'amélioration génétique des plantes (2009-2016). J'ai participé au montage d'un



module « Guided tours » pour la formation master international EUR (L'École Universitaire de Recherche) : « International Master of Functional Biology and Ecology, The "TULIP-Graduate School" of Toulouse and Perpignan » dans le cadre du Programme Investissements d'Avenir 3 (PIA3).

- En **Doctorat**, j'ai participé au jury de concours de l'école doctorale SEVAB (Sciences Ecologiques Vétérinaires Agronomiques & Bioingénieries).

Enfin, j'ai organisé et/ou j'ai participé aux événements pédagogiques organisés par l'université ou laboratoire

- Soirées thématiques pour les activités recherches en biologies du développement végétales  
- Visite du laboratoire LRSV, la Fédération de Recherche IFR3450 et INRAE pour les étudiants en L1 et L2 afin de mieux guider leur orientation ;

- De plus, j'ai participé en 2023 à la formation de la délégation française pour les Olympiades Internationales de Biologie (Lycéens). J'ai leur conçu et encadrer des long travaux pratiques pour leur aider à approfondir leur connaissance en physiologie et biochimie végétales.

## B. Responsabilités pédagogiques

2012 - aujourd'hui, Responsable du UE L2 « Physiologie Végétale » parcours Sciences de la Vie et de la Terre – Enseignement (SVT-E) qui prépare aux concours des métiers de l'enseignement secondaire des SVT (CAPES et agrégation).

2009 - 2016, Responsable du module M2S13 « Molecular basis for plant and animal improvement » dans le cadre de la formation **Master international AgroFood Chain**, enseignement en anglais.

2022- aujourd'hui, co-responsable (avec D. Aldon) de l'UE « Développement végétal » parcours Biodiversité et Biologie Environnementale de la Licence mention Sciences de la Vie ( $\pm 200$  étudiants).

2019 - aujourd'hui, Correspondante de l'enseignement au laboratoire LRSV pour la formation **Master international EUR** (L'École Universitaire de Recherche), module « Guided tours » : « International Master of Functional Biology and Ecology, The "TULIP-Graduate School" of Toulouse and Perpignan » dans le cadre du Programme Investissements d'Avenir 3 (PIA3).

## V. Responsabilités collectives et activité d'expertise

### A. Participation à des jurys et des comités de thèse

- 2023 Jury de recrutement AI (Assistant Ingénieur) Single-cell RNAseq INRAE, au laboratoire LIPME.
- 2019- aujourd'hui Membre de jury pour attribuer des bourses chaque année aux 7 programmes d'appels d'offre: PhD mobility Packages, PhD Research Allocations, Formation Packages, Initiation to Research, PhD Tutorial Assignment, PhD Position Grants, Master Internship Grants, Master Mobility Packages, dans le cadre du projet "TULIP-Graduate School" of Toulouse and Perpignan » Programme Investissements d'Avenir 3 (PIA3).



- **2020-2023** Comité de thèse de Madame GAO Yushuo « Régulation transcriptionnelle du développement du fruit de la tomate » UMR5546 LRSV - Laboratoire de Recherche en Sciences Végétales, UPS/CNRS/INP Toulouse. Ecole doctorale : EDSEVAB - SEVAB - Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingenieries n°458.
- **2023** Jury de thèse de Mme Shiyu YING « Regulation and role of class E Ethylene Response Factors (ERF.E) in climacteric fruit ripening » UMR5546 LRSV - Laboratoire de Recherche en Sciences Végétales, UPS/CNRS/INP Toulouse. Ecole doctorale : EDSEVAB - SEVAB - Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingenieries n°458.
- **2022** Jury de thèse de Mme Ximena Chirinos Herrera « Molecular mechanisms underlying the transition to ripening in tomato fruit » UMR5546 LRSV - Laboratoire de Recherche en Sciences Végétales, UPS/CNRS/INP Toulouse. Ecole doctorale : EDSEVAB - SEVAB - Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingenieries n°458.
- **2022, 2018, 2010** Jury de soutenance Master II Biologie Végétale
- **2020** Jury de concours école doctorale SEVAB (Sciences Ecologiques Vétérinaires Agronomiques & Bioingénieries).
- **2010-2013** Comité de thèse de M. Xinyu Wang “Characterization of protein-protein interactions involved in auxin mediation.” UMR INRA/INP 990 Génomique et Biotechnologie des Fruits, INP Toulouse. Ecole doctorale : EDSEVAB - SEVAB - Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingenieries n°458.

## **B. Responsabilités scientifiques**

### ***B1. Obtention de contrats de recherche sur appels d'offre***

- **2021-2022** Co-coordinatrice avec Dr. Marta MARCHETTI/Richard BERTHOME du projet FR AIB (The Agrobiosciences, Interactions and Biodiversity Research Federation) “Analysis of primary cell WALL in REsistance under combined stress” .
- **2015-2016** Co-coordinatrice avec Dr. Julien PIRRELLO du projet FR AIB (The Agrobiosciences, Interactions and Biodiversity Research Federation). “A new plant model to decipher the auxin-dependent molecular mechanisms of xylem differentiation”.
- **2010-2011** Responsable du projet BQR Université Toulouse III- Paul Sabatier, « Cibles et partenaires de nouveaux régulateurs de la formation du bois chez *Eucalyptus* ». Ce projet a financé les expériences réalisées dans le cadre de la formation M2R (Matthieu Bensussan, 2010) dont J'étais l'encadrante principale.

### ***B2. Participation aux contrats de recherche***

- **2023-2028** Participation au **projet ANR PEPR** (Programmes et équipements prioritaires de recherche) : **TYPEX**, Toward highly Predictable Editing of the plant genome leXicon (Vers une édition spécifique et précise du génome végétal). Responsable du projet : Pierre-Marc Delaux, DR CNRS, Fabien Nogué, DR INRAE. L'objectif du projet sera de développer des outils et des règles générales pour un Prime Editing efficace chez les plantes. Plus particulièrement je me charge à 1) Régénérer des Hairy Root fluorescentes transgéniques avec des constructions d'édition Prime ciblant les gènes EgrCCR1 et EgrIAA9A ; 2) à sélectionner de Hairy Root transgéniques éditées par KO-CCR et gain de fonction et perte de fonction

EgrIAA9A ; 3) à identifier des lignées éditées présentant des tissus vasculaires hypo lignifiés dans les fibres et/ou dans les vaisseaux.

- **2023-2026** Participation au projet **ANR MusaWall** Mechanisms and contributions of Musashi-mediated control in cell wall polymers synthesis in plants (Régulation post transcriptionnelle de la formation du bois : Rôle des protéines Musashi) Coordinateur du projet : Thierry Lagrange, DR CNRS Laboratoire Génome et Développement des Plantes Perpignan ; Partner 2 : Biopolymères Interactions Assemblages (BIA, INRAE, Nantes), PI : Richard Sibout ; Partner 3 : LRSV, UT3/CNRS/INP ENSAT, PI : Fabien Mounet. Plus particulièrement, pour les premières étapes du projet, je me charge de la construction des vecteurs sur-expresseurs sous contrôle de promoteur 35S et knock-out par la technique de CRISPR Cas9 avec une Ingénieure dédiée à ce projet.
- **2019-2028 Labex "TULIP-Graduate School"** of Toulouse and Perpignan, **Programme Investissements d'Avenir 3 (PIA3)**. TULIP-GS a deux objectifs majeurs : promouvoir la recherche et former des étudiants à l'interface entre écologie, évolution et biologie fonctionnelle. Pour l'aspect Recherche, Je participe au jury de nombreux appels d'offres à l'attention des doctorants et des chercheurs et Enseignants-chercheurs des laboratoires TULIP : PhD mobility Packages, PhD Research Allocations, Formation Packages. Pour l'aspect enseignement, TULIP-GS propose un Master ouvert à l'international (voir la rubrique Le Master TULIP-GS) . Ce Master s'appuie sur des méthodes et des concepts permettant de comprendre et de prédire comment les organismes vivants sont affectés par les modifications de leur milieu (biotique ou abiotique) et s'adaptent aux changements de leur environnement, comme le changement climatique. Je suis correspondante de l'enseignement pour laboratoire LRSV. Je participe aussi au jury pour attribuer des bourses chaque année aux programmes d'appels d'offre de ce projet:; Initiation to Research, PhD Tutorial Assignment, PhD Position Grants, Master Internship Grants, Master Mobility Packages.
- **2012-2025** Participation au projet **Labex TULIP**. Laboratoire d'Excellence qui privilégie une approche interdisciplinaire, alliant biologie et écologie, centrée sur les interactions entre organismes ou communautés, dans des milieux naturels ou modifiés par l'humain.
- **2011-2015** Participation au **projet Européen KBBE "Tree for Joules"**: improving *Eucalyptus* and poplar wood properties for bioenergy (13 partenaires 4 pays). Coordinatrice principale Mme Jacqueline Grima-Pettenati, DR CNRS Toulouse. J'étais Responsable du Task 2 dans WP1 : Transcript profiling of selected candidate genes. J'ai aussi participé au montage de ce projet tel que la relecture, la mise en forme et la validation des drafts venant des différents partenaires, et à l'établissement des tableaux des budgets. Nous avons publié 8 articles, dont 3 articles en dernier auteur, et un article en premier auteur (Carocha et al., 2015; Cassan-Wang et al., 2012; Li et al., 2015; Myburg et al., 2014; Soler et al., 2015; Yu et al., 2022; Yu et al., 2014; Yu et al., 2015)
- **2008-2010** Participation au **projet Européen FP7 « RENEWALL »**: Improving plant cell walls for use as a renewable industrial feedstock. (23 partenaires, 6 pays). Coordinateur principal : Simon McQueen-Mason, York University, UK. J'ai encadré une étudiante en Master1 dans ce projet. Ces expériences en cours m'ont permis d'apprendre à bénéficier des ressources bioinformatiques en génomique telles que l'analyse d'expression génique in silico, prédiction fonction des gènes par le system co-expression, et l'utilisation des banques de mutants de type insertion d'ADN-T ou RNAi. J'ai publié un article en premier auteur sur ce projet (Cassan-Wang, Goue, et al., 2013).
- **2004-2008** Participation au **projet Européen EU-SOL (FP6)** sur le sujet « mécanisme moléculaire du développement du fruit de tomate » et au **projet Européen Trilatéral** sur le

sujet : « L'exploration de la variabilité naturelle du développement et de la qualité du fruit de tomate par des approches de transcriptomique et de métabolomique ». J'ai réalisé l'analyse transcriptomique de processus i) de la transformation de fleur en fruit, ii) de régulation hormonale de la fructification, iii) de la maturation du fruit de tomate. En collaboration avec le laboratoire du Dr. Alisdair Fernie (Max Planck Institute, Allemagne) une analyse métabolomique a été réalisée qui vise à révéler le profil des métabolites associés. J'ai publié un article en premier auteur sur ce projet (Wang et al., 2009).

### **C. Participation à des conseils d'établissement, à des comités scientifiques**

- **2019 - aujourd'hui** Membre du Conseil Pédagogique de projet « TULIP-Graduate School of Research » (TULIP-GSR) de Toulouse et de Perpignan. Participation chaque année à 7 programmes d'appels d'offre pour attribution de bourses : PhD mobility Packages, PhD Research Allocations, Formation Packages, Initiation to Research, PhD Tutorial Assignment, PhD Position Grants, Master Internship Grants, Master Mobility Packages.
- **2009 - aujourd'hui** Responsable Axe Asie au service international de l'Université Toulouse III. Participation à des réunions de travail, à l'accueil de à l'accueil des plusieurs délégations universitaires Asiatiques en visite à l'université de Toulouse, et gestion des Appels d'offre pour attribution de bourses.
- **2012 - 2015** Membre élu du conseil administratif national de la SFBV (Société Française de Biologie Végétale)

### **D. Expertises d'articles ou de projets**

#### ***D1. Expertises d'articles scientifiques***

- Éditeur board chez Frontier in Plant Science
- Examinatrice pour plusieurs journaux scientifiques internationaux :  
New phytologist, Frontiers in Plant Science,  
the Journal of Biotechnologies, Plant Cell report, Molecular Biology Report,  
PLoS One, Genetics and Molecular Biology

#### ***D2. Invitation dans des universités étrangères***

- **2023-2024** invited professor à SouthWest University of China (Chine) (60 000 étudiants). Participation au « 2023 South West University Global Immersion Program », Invitation 2 semaines à South West University (Chongqing, Chine). J'ai dispensé 24 heures de cours (1 UE complète) sur « Plant genetic improvement and Biotechnologies » aux étudiants en Bachelor et en Master.
- **2019-2020** invited professor à SouthWest University of China (Chine) (60 000 étudiants). 2019, Invitation de 10 jours sur place et j'ai dispensé 12h cours en anglais sur « Plant genetic improvement and Biotechnologies » aux étudiants en master et en doctorat.

- **2020 - 2021** Invitation 2 semaines à South West University (Chine) (60 000 étudiants) en qualité de « invited Professor ». Mission annulée à cause de crise sanitaire COVID-19. J'ai dispensé 24 heures de cours en ligne sur « Plant genetic improvement and Biotechnologies » aux étudiants en Bachelor et en Master.
- **2012-2020** à South China Agricultural University (> 41 000 étudiants) en qualité de « Ding-Ying International Chair Professor ». J'ai dispensé plusieurs séminaires : « Functional characterisation of *Eucalyptus* auxin dependent transcription factors in wood formation », « How to use known resources in model system of *Arabidopsis* to explore other species », et « Golden gate cloning system » etc. J'ai aussi dispensé des cours de « biologie végétale », « Le développement des fruits » aux étudiants en Licence, en Master et en Doctorat.

### **D3. Participation à des projets Européens Erasmus+**

- **2017** - Organiser et dispenser des cours (30 heures, Intensive advanced training courses: Biotechnology for Plant genetic improvement) aux enseignants des Universités du Cambodge (University of Battambang, Mean Chey University, Svay Rieng University, Chea Sim University of Kamchaymear), à University of Battambang, Cambodge, dans le cadre du projet Erasmus+ Program UNICAM: « Implementing quality of education & training of the Young UNiversities in rural area of CAMbodia ».
- **2018** - Dispenser des cours et organiser des visites de laboratoires et entreprise semencière aux enseignants des Universités de l'Asie sud-est (Vietnam, Cambodge, Myanmar) dans le cadre du projet Européen Erasmus+ NutriSEA "Network of Universities and Enterprises for Food Training in Southeast Asia", qui œuvre pour le développement et la mise en place de formation universitaires au niveau Bachelor et Master au sein des six universités partenaires du Sud-Est de l'Asie, dans le secteur des ressources naturelles et de l'agroalimentaire.

### **D4. Participation aux activités de valorisation et de vulgarisation scientifiques**

- **2014- aujourd'hui** Organisation et animation de l'atelier « Parois cellulaires végétales » au sein du programme « Fête de la Science » tous les 2 ans environ
- **Mars 2015** Participation à la 7ème session (parmi 13 trinômes) du programme « Jumelages Parlementaires-Membres de l'Académie des sciences-Jeunes chercheurs » au Sénat et à l'Assemblée Nationale à Paris, créés par l'Office parlementaire d'évaluation des choix scientifiques et technologiques (OPECST) et l'Académie des sciences, dont l'objectif est d'aider les politiques (parlementaires, sénateurs) à mieux connaître le milieu de la recherche et réciproquement.
- **Mai 2015** Organisation et animation de la visite de Monsieur le Sénateur de Haute Garonne Pierre Médevielle dans notre Laboratoire de Recherche en Sciences Végétales. Participation à la « Table Ronde » sur le sujet « nos activités et nos métiers chercheur / enseignant-chercheur ».
- **Septembre 2015** Participation au débat et à la rencontre des représentants d'agriculteurs dans les circonscriptions de Monsieur le Sénateur de Haute Garonne Pierre Médevielle à la Marie de Boulogne-sur-Gesse.

## **E. Récompenses**

- *Prime d'encadrement doctoral 2017-2020*

- *Prime d'excellence scientifique 2013-2016*

- *Prix « Best poster »* lors du 26th International New *Phytologist* Symposium “Bioenergy Trees”, 17-19 Mai 2011, Nancy France (<http://www.newphytologist.org/bioenergy/default.htm>)

*Prix conference internationale “Earl J. Scherago Travel Grant Award”* lors de l’ « International Plant and Animal genome XX », Jan. 2012, San Diego, CA, USA



CHAPITRE 2 - Présentation et analyse de mes activités de  
recherche avant mon recrutement à l'Université Toulouse III  
Paul Sabatier





## I. Travaux de doctorat : Caractérisation fonctionnelle du gène *IAA9*, facteur transcriptionnel de la signalisation de l'auxine chez la tomate.

Je suis née et j'ai grandi en Chine. Sélectionnée par un concours national j'ai fait mes études Bachelor (BAC +4) à l'Université SCAU (South China Agricultural University). Étant la major de la promo, j'ai été présélectionnée à poursuivre mes études en Master en Science Alimentaire (Bac +7). Au cours de mes 3 ans d'études en Master, j'étais convaincue que la biotechnologie était très importante pour l'avenir et j'ai souhaité poursuivre mes études en doctorat dans ce domaine à l'étranger. J'ai eu la chance d'obtenir une bourse de thèse qui m'a permis de venir en France pour faire mes études en thèse au laboratoire « Génomique et Biotechnologie des Fruit » (UMR990 INP-ENSAT/INRA).

Mes travaux de recherche, réalisés au cours de la thèse sous la direction du Professeur Mondher Bouzayen, au sein du laboratoire Génomique et Biotechnologie des Fruits, s'inscrivent dans la thématique globale de l'étude des bases moléculaires et génétiques dans la construction de la qualité des fruits charnus.

À mon arrivée en France depuis la Chine pour ma thèse en septembre 2000, les activités de recherche de mon laboratoire d'accueil étaient centrées sur la maturation des fruits et l'éthylène. Lors de mes travaux de thèse sur la « Régulation transcriptionnelle auxine dépendante lors du développement du fruit », nous avons **développé un nouvel axe de recherche** visant à étudier la **régulation auxine-dépendante** de la maturation du fruit. Les plantes que j'ai générées ont montré des phénotypes intéressants particulièrement sur leur capacité à développer des fruits sans graines en absence de pollinisation et altération de la morphogénèse des feuilles et tissu vasculaire. Les résultats de ma thèse ont fait l'objet d'une **publication dans la revue « The Plant Cell » (H. Wang et al., 2005)** (annexe 1).

### Résumé de la thèse

Le développement des fruits de l'anthèse à la maturation est un processus génétiquement programmé dont le déroulement se trouve sous un contrôle multi hormonal. Alors que l'éthylène joue un rôle déterminant au cours de la maturation des fruits climactériques comme la tomate, les modalités d'intervention des autres hormones ne sont toujours pas connues. Une approche de criblage différentiel a permis d'isoler au laboratoire plusieurs gènes codant pour des régulateurs transcriptionnels impliqués dans la réponse à l'auxine et montrant un profil d'expression qui varie au cours de la maturation et en réponse à l'éthylène. Le travail présenté dans cette thèse est focalisé sur la caractérisation fonctionnelle du gène *DR4* (appelé *IAA9* plus tard) apparenté à la famille des *Aux/IAA* de la tomate dont le rôle en tant que médiateur de la réponse des plantes à l'auxine est bien connu. Alors que de nombreux mutants « gain de fonction » touchant les gènes *Aux/IAA* ont été décrits chez *Arabidopsis*, de façon remarquable, aucun mutant « perte de fonction » n'a été caractérisé à ce jour. Le travail présenté ici montre que la sous-expression du gène *DR4* chez la tomate affecte plusieurs processus de développement dépendant de l'auxine. Ainsi, les lignées transgéniques exprimant une construction antisens du gène *DR4* (*DR4AS*) montrent un phénotype pléiotrope qui touche essentiellement la morphologie des feuilles et le développement du fruit. Les lignées *DR4AS* possèdent des feuilles simples et non pas composées comme chez le sauvage et le développement des fruits survient en l'absence de toute pollinisation et fécondation ce qui conduit à la production de fruits parthénocarpiques. Les plantes *DR4AS* montrent également plusieurs phénotypes associés à une réponse exacerbée à l'auxine tel qu'une forte élévation de l'hypocotyle, de la tige et de la racine, une

réduction de la dominance apicale, des fusions d'organes et une transition d'un développement basipétal des tiges axillaires à un mode acropétal. Le croisement génétique des plantes *DR4AS* avec des lignées exprimant le gène *GUS* sous contrôle d'un promoteur répondant à l'auxine engendre une augmentation substantielle de l'expression du gène rapporteur dans les lignées hybrides indiquant que la protéine DR4 agit comme régulateur négatif des gènes de réponse à l'auxine. Les données présentées dans cette thèse constituent la première description de phénotypes associés à une sous-expression d'un gène *Aux/IAA* et révèlent pour la première fois le rôle essentiel joué par un membre de la famille des *Aux/IAA* dans la morphogenèse des feuilles et le développement du fruit.

**Mots clé :** Auxine, *Aux/IAA*, tomate, fruit, fructification, parthénocarpié, morphogenèse

## Contexte scientifique

Les fruits sont des composants essentiels dans l'alimentation humaine car ils sont une source importante de vitamines, d'éléments minéraux, de fibres et d'antioxydants. Pour la plupart des fruits c'est au cours de la phase de maturation que s'élaborent les qualités organoleptiques et nutritionnelles. Les changements s'opérant au cours du processus de maturation, en termes de saveur, d'arôme, de couleur et de texture, sont relativement similaires dans tous les fruits, même s'il existe des différences fondamentales aux niveaux des mécanismes de régulation. Par ailleurs, le processus de maturation s'accompagne d'une fragilisation des tissus du fruit ce qui expose ce dernier à une détérioration plus rapide réduisant sa durée de vie post-récolte. De ce point de vue, la maîtrise des processus de développement et de maturation des fruits constitue un enjeu économique de première importance notamment en matière de maintien de la qualité lors du transport et de la distribution. La finalité de la recherche de la qualité des fruits charnus est d'identifier et contrôler les facteurs biotiques ou abiotiques qui conduisent à la mise en place d'une bonne qualité du fruit.

Les fruits peuvent être classés en deux catégories au vu de leur maturation : climactériques et non climactériques. Les fruits dits climactériques sont caractérisés par une synthèse auto catalytique d'éthylène accompagnée d'une crise respiratoire survenant au début de la maturation. Pour cette raison, l'éthylène a été la première cible de nombreux travaux dont les objectifs appliqués visaient le contrôle génétique et technologique de la physiologie pré- et post-récolte des fruits. Plusieurs composants de la voie de signalisation de l'éthylène ont été caractérisés, en particulier chez l'espèce modèle *A. thaliana* et chez la tomate avec l'isolement du mutant de récepteur à l'éthylène *NEVER RIPE (Nr)*, puis la manipulation de leur expression par la voie de la transgénèse chez la tomate a permis de contrôler efficacement le processus de maturation.

Des études plus récentes ont permis de montrer l'existence des voies métaboliques dépendantes de l'éthylène mais également des voies indépendantes de l'éthylène au sein du processus de la maturation des fruits. Par exemple, l'isolement récent par clonage positionnel des gènes responsables de mutations, *ripening inhibitor (rin)* et *non-ripening (nor)*, à l'origine du blocage du mûrissement chez la tomate indique que les produits de ces gènes participent aux mécanismes précoces de régulation de la maturation et que leur expression est indépendante de l'éthylène. Il reste donc à élucider les mécanismes de l'acquisition de la compétence à mûrir qui, à partir d'un stade déterminé du développement, confère au fruit la capacité, à répondre de façon spécifique à l'éthylène pour des fruits climactériques, de déclencher le programme de maturation et d'activer tous les gènes associés aux voies métaboliques. Pour cela, les physiologistes ont avancé depuis longtemps la notion de changement de la balance hormonale. Cette hypothèse est appuyée par le fait que le niveau de

plusieurs phytohormones dans le fruit subit des variations importantes tout au long du développement de cet organe, de l'anthèse à la maturation. Ces données suggèrent l'existence d'un contrôle multi hormonal présidant à ce processus de développement. L'hypothèse des travaux de ma thèse et de mon premier stage post doctoral postule que le changement du statut hormonal du fruit, soit par les traitements des phytohormones, soit par transgénèse en modifiant la sensibilité aux hormones ou leur teneur, serait responsable de la réponse différenciée, qui déclenche un développement ou une maturation et conduit à la modification de la qualité des fruits.

Parmi les espèces utilisées pour l'étude de la maturation des fruits, la tomate (*Solanum lycopersicum*) est le modèle reconnu par l'ensemble de la communauté scientifique mondiale et à ce titre la tomate fait l'objet de programmes d'envergure de génération d'outils et de ressources biologiques (EST, DNA chips, séquençage du génome, cartes génétiques...). Ce modèle doit permettre des transpositions à d'autres fruits d'intérêt agronomique, climactériques ou non telles que la pomme, la pêche, ou le raisin. Dans le domaine du métabolisme secondaire (arômes, colorants...) toutefois, chaque fruit peut présenter des particularités qui requièrent des études spécifiques.

## Travaux de thèse

### **La régulation transcriptionnelle dépendante de l'auxine lors du développement du fruit : caractérisation fonctionnelle du gène *DR4*, un homologue d'*Aux/IAA* chez la tomate (*Lycopersicon esculentum*, Mill)**

Le rôle déterminant de l'éthylène au cours de la maturation des fruits de type climactérique comme la tomate, est maintenant bien démontré. Cependant on connaît peu de choses sur le rôle des autres hormones. On soupçonnait que l'auxine joue un rôle très important tout au long de ce processus, cependant les données les plus précises sur le rôle de l'auxine concernaient essentiellement le développement du fruit lors de la transformation de la fleur en fruit via la pollinisation et la fécondation. Si on applique expérimentalement de l'auxine sur le stigmate de la fleur, le processus de fructification est déclenché malgré l'absence de fécondation véritable. Le traitement de l'auxine est une pratique courante en arboriculture. Elle permet d'intervenir sur une homogénéisation du calibre des fruits, une augmentation de la taille des baies de raisins, et une suppression des pépins (comme pour les clémentines et raisin) etc. Mais jusqu'à présent, au niveau moléculaire, les modalités d'intervention auxinique dans le développement du fruit, restent encore mal connues.

L'auxine naturelle mise en évidence dans les tissus des plantes est l'acide indole acétique, l'AIA, qui est un acide faible. Il s'agit d'une hormone très importante chez les plantes, qui est impliquée dans le contrôle de la croissance végétale, de la division, l'expansion et la différenciation cellulaire, la dominance apicale, le tropisme, la rhizogenèse, la formation des tissus vasculaires et la fructification et le développement du fruit.

Les facteurs transcriptionnels apparaissent comme les médiateurs principaux de la réponse hormonale (Vogler & Kuhlemeier, 2003). Deux types de régulateurs transcriptionnels sont connus comme médiateurs de la réponse à l'auxine : les *Aux/IAA* et les ARF. Les *Aux/IAA* sont des gènes de réponse précoce à l'auxine, qui codent pour des protéines ne se liant pas à l'ADN et qui fonctionnant comme des répresseurs transcriptionnels (Tiwari et al., 2003). Les ARF (Auxin Response Factor) sont des protéines qui se lient directement à l'ADN au niveau de séquences spécifiques situées dans les promoteurs des gènes de réponse à l'auxine (Ulmasov et al., 1999). C'est à travers leurs interactions les AFR que les *Aux/IAA* régulent l'expression génique dépendant de l'auxine.

Des travaux antérieurs au début de ma thèse ayant mis en œuvre des approches de criblage différentiel ont permis d'isoler plusieurs gènes codant pour des régulateurs transcriptionnels impliqués dans la réponse à l'auxine et présentent un profil d'expression qui varie au cours de la maturation et en réponse à l'éthylène (Jones et al., 2002). Parmi les clones isolés, les études d'homologie de séquences ont permis d'identifier des gènes *DR1*, *DR3*, *DR4*, *DR8* (Developmental Regulated) appartenant à la famille d'Aux/IAA, connue de facteurs de transcription de l'auxine. L'exploration fonctionnelle de ces gènes a été mise en œuvre par l'approche génétique inverse consistant à créer des lignées transgéniques altérées dans l'expression de ces gènes. Les premières caractérisations des lignées transgéniques ont orienté les études de façon prioritaire vers le gène *DR4*, qui montre des phénotypes associés au développement du fruit.

Mon travail de thèse a donc consisté à la caractérisation fonctionnelle du gène *DR4* dans le développement des plantes et en particulier dans le fruit de la tomate. Ceci permettra d'explorer le rôle de *DR4* dans la réponse à l'auxine et d'évaluer la régulation auxinique dans le développement du fruit.

### 1. Clonage et analyse phylogénétique du gène *DR4* chez la tomate

L'isolement d'un clone de cDNA pleine taille *DR4* m'a permis de montrer que la protéine codée comporte bien les quatre domaines très conservés, caractéristiques de la famille des Aux/IAA. L'arbre phylogénétique par PAUP et ProDom suggère que les protéines Aux/IAA se répartissent en quatre sous-familles, *DR4* appartenant à la sous-famille IV dont les membres se distinguent par leur plus grande taille due en particulier à une extension de la région N-terminale située avant le domaine I.

### 2. Études d'expression et localisation subcellulaire de la protéine *DR4*

J'ai étudié les profils d'expression spatio-temporelle du gène *DR4* chez la tomate par les techniques de RT-PCR semi quantitative et de Northern blot. Le messenger *DR4* a été détecté dans tous les organes testés suggérant une expression ubiquitaire et constitutive. Ce profil d'expression diffère cependant de celui décrit pour les autres Aux/IAA chez *A. thaliana*, où ces gènes montrent une expression préférentielle dans différents tissus et organes. De plus, contrairement aux autres membres de la famille des Aux/IAA qui montrent des inductions rapides et fortes par l'auxine, le gène *DR4* montre une faible induction par cette hormone.

J'ai réalisé des expériences d'expression transitoire dans le système protoplaste où la protéine GFP (green fluorescent protein) a été fusionnée avec la protéine *DR4* montre que la totalité du signal se trouve exclusivement dans le noyau. Cette localisation nucléaire est conforme avec la fonction supposée de régulateur transcriptionnel de *DR4*.

### 3. Génération et caractérisation des plantes transgéniques sous-exprimant le gène *DR4*

Une approche de génétique inverse a été adoptée qui vise à modifier l'expression endogène du gène *DR4* dans les tomates. Plusieurs lignées transformées ont été obtenues présentant différents niveaux d'inhibition du gène *DR4* endogène, dont deux lignées AS1 et AS2 sont fortement inhibées. Il n'a pas été possible d'obtenir des lignées complètement inhibées pour ce gène ce qui est conforme à l'absence de lignées "perte-de-fonction" (knock-out) issues des nombreux criblages à haut débit des mutants *A. thaliana*. En effet, tous les mutants des *aux/iaas* décrits jusqu'à présent sont des mutants "gain-de-fonction" et les lignées décrites ici constituent le premier exemple de phénotype résultant de la sous-expression d'un gène d'Aux/IAA. Les lignées *DR4AS* présentent un faisceau de phénotypes dont certains touchent le développement du fruit.

### i) Caractérisation des fruits

Les lignées *DR4AS* présentent un développement précoce et accéléré des fruits survenant en l'absence de toute pollinisation et conduisant à la production de fruits parthénocarpiques. Les fruits *DR4AS* présentent les mêmes caractéristiques que les fruits sauvages en ce qui concerne le poids, la taille, la texture, la couleur, et l'émission d'éthylène mais accumulent plus de sucres et d'acides. Il a été vérifié que les plantes *DR4AS* n'ont pas de problème de fertilité, mais plutôt une désynchronisation du développement des appareils mâles et femelles conduisant à une perturbation du processus de fructification.

### ii) Caractérisation au niveau végétatif

Les plantes transgéniques *DR4AS* portent des feuilles simples au lieu de feuilles composées aussi bien dans le fond génétique « microTom » que « Ailsa Craig » mettant en évidence une implication de DR4 dans le processus de morphogenèse des feuilles. Les plantes *DR4AS* montrent de nombreux autres phénotypes qui peuvent être attribués à une réponse exacerbée à l'auxine telle que : forte élongation de l'hypocotyle, de la tige et de la racine, réduction de la dominance apicale, fusions d'organes et développement acropétal (et non basipétal) des tiges axillaires.

Les analyses anatomiques montrent un profil du tissu vasculaire anormal dispersé dans la stèle souvent associé à des triples cotylédons et à une hypertrophie des vaisseaux de la nervure des feuilles.

### iii) Caractérisation de la réponse à l'auxine

Les tests physiologiques basés sur la mesure de l'élongation de l'hypocotyle indiquent que les lignées antisens sont plus sensibles à l'auxine. Une approche génétique, consistant à comparer l'activité d'un promoteur répondant à l'auxine dans le fond génétique sauvage et *DR4AS*, montre que l'activité de ce promoteur est fortement augmentée dans les plantes transgéniques démontrant ainsi que DR4 est un régulateur négatif de la réponse à l'auxine.

Les travaux réalisés au cours ma thèse apportent des preuves solides sur l'implication de l'auxine dans le développement (fructification, parthénocarpie) et la maturation (richesse en sucres et en acides) des fruits et révèlent le rôle de cette hormone dans la mise en place de la qualité des fruits. Ils amènent des éléments nouveaux sur le mode d'action de l'auxine et sur le rôle particulier du gène DR4 dans cette voie de signalisation hormonale. Ces données ont fait l'objet d'une publication dans *Plant Cell* (Wang et al., 2005).

Si on replace la caractérisation du gène DR4 dans le contexte de la maturation des fruits, il est important de rappeler que les effets antagonistes de l'auxine et de l'éthylène ont été souvent évoqués dans la littérature. Toutefois, on constate dans le cadre de notre étude que le phénotype d'hypersensibilité à l'auxine n'engendre pas de perturbation notable dans la chronologie du déroulement de la maturation alors qu'il entraîne une accumulation plus importante de sucre dans les fruits *DR4AS*. Ces résultats n'excluent pas définitivement l'éventualité d'un rôle de l'auxine dans le processus de maturation qui passerait par une voie indépendante de DR4. Il est également possible que l'absence de phénotype visible sur la maturation soit le résultat d'une redondance fonctionnelle au sein de la famille des Aux/IAA. La création de doubles et triples mutants par croisements génétiques entre lignées transgéniques affectées dans différents membres de la famille des Aux/IAA est maintenant envisagée pour essayer d'éviter une éventuelle redondance fonctionnelle.

Annexe 1 Article de thèse



# The Tomato *Aux/IAA* Transcription Factor *IAA9* Is Involved in Fruit Development and Leaf Morphogenesis <sup>W</sup>

Hua Wang,<sup>a</sup> Brian Jones,<sup>a</sup> Zhengguo Li,<sup>b</sup> Pierre Frasse,<sup>a</sup> Corinne Delalande,<sup>a</sup> Farid Regad,<sup>a</sup> Salma Chaabouni,<sup>a</sup> Alain Latché,<sup>a</sup> Jean-Claude Pech,<sup>a</sup> and Mondher Bouzayen<sup>a,1</sup>

<sup>a</sup>Unité Mixte de Recherche 990, Institut National de la Recherche Agronomique/Institut National Polytechnique–Ecole Nationale Supérieure Agronomique Toulouse, Génomique et Biotechnologie des Fruits Pôle de Biotechnologie Végétale, 31326 Castanet-Tolosan Cedex, France

<sup>b</sup>Genetic Engineering Research Center, Chongqing University, 40030 Chongqing, China

**Auxin/indole-3-acetic acid (*Aux/IAA*) proteins are transcriptional regulators that mediate many aspects of plant responses to auxin. While functions of most *Aux/IAAs* have been defined mainly by gain-of-function mutant alleles in *Arabidopsis thaliana*, phenotypes associated with loss-of-function mutations have been scarce and subtle. We report here that the downregulation of *IAA9*, a tomato (*Solanum lycopersicum*) gene from a distinct subfamily of *Aux/IAA* genes, results in a pleiotropic phenotype, consistent with its ubiquitous expression pattern. *IAA9*-inhibited lines have simple leaves instead of wild-type compound leaves, and fruit development is triggered before fertilization, giving rise to parthenocarpy. This indicates that *IAA9* is a key mediator of leaf morphogenesis and fruit set. In addition, antisense plants displayed auxin-related growth alterations, including enhanced hypocotyl/stem elongation, increased leaf vascularization, and reduced apical dominance. Auxin dose–response assays revealed that *IAA9* downregulated lines were hypersensitive to auxin, although the only early auxin-responsive gene that was found to be upregulated in the antisense lines was *IAA3*. The activity of the *IAA3* promoter was stimulated in the *IAA9* antisense genetic background, indicating that *IAA9* acts in planta as a transcriptional repressor of auxin signaling. While no mutation in any member of subfamily IV has been reported to date, the phenotypes associated with the downregulation of *IAA9* reveal distinct and novel roles for members of the *Aux/IAA* gene family.**

## INTRODUCTION

The phytohormone auxin is central to a myriad of aspects of plant growth and developmental processes. At the cellular level, auxin controls cell division, extension, and differentiation. On a whole-plant level, auxin plays an essential role in processes such as apical dominance, lateral/adventitious root formation, tropisms, fruit set and development, vascular differentiation, and embryogenesis (Friml, 2003). While it is clear that auxin plays a pivotal role in plant growth and development, the molecular effectors by which this hormone exerts its effect are still relatively unknown. For example, in the process of fruit set, the onset of ovary development into fruit and its subsequent development are naturally triggered by signals from successful fertilization. These processes can be initiated in the absence of pollination and fertilization by exogenous auxin or auxin transport inhibitors (Gustafson, 1937; Beyer and Quebedeaux, 1974) or by the ovary-specific expression of *Agrobacterium tumefaciens indoleacetamide monooxygenase (iaaM)* or *root loci B (rolB)* genes, which confer higher auxin production or increased auxin sensitivity, respectively (Ficcadenti et al., 1999; Carmi et al., 2003). The mo-

lecular mediators by which auxin impacts this process are still unknown.

The recent discovery that the F-box protein Transport Inhibitor Response1 functions as an auxin receptor represents a major breakthrough (Dharmasiri et al., 2005; Kepinski and Leyser, 2005) in understanding how auxin mediates cellular responses. It has been known for a decade that auxin signaling operates by recruiting specific transcription factors, leading to the expression of downstream genes that perform the required responses (Vogler and Kuhlemeier, 2003). The auxin/indole-3-acetic acid (*Aux/IAA*) and auxin response factor (ARF) families of transcription factors have been shown to be instrumental in auxin-dependent transcriptional regulation. *Aux/IAA* proteins are short-lived transcription factors that share four highly conserved domains. Domain I contains a functionally characterized transcriptional repressor motif (Tiwari et al., 2004), while domain II interacts with a component of the ubiquitin-proteasome protein degradation pathway shown to be essential for auxin signaling. Domains III and IV act as C-terminal dimerization domains, mediating homodimerization and heterodimerization among *Aux/IAA* family members and dimerization with similar domains found in ARFs (Kim et al., 1997; Ulmasov et al., 1997; Ouellet et al., 2001). ARFs are transcriptional activators and repressors that bind with specificity to TGTCTC auxin-responsive *cis*-acting elements found in promoters of primary/early auxin-responsive genes, including members of the *Aux/IAA* family (Ulmasov et al., 1997).

*Aux/IAA* genes are expressed in distinct spatial and temporal patterns, contributing to the diversity of auxin responses in different plant tissues, organs, and developmental stages (Abel et al.,

<sup>1</sup>To whom correspondence should be addressed. E-mail bouzayen@ensat.fr; fax 0033-5-62193573.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Mondher Bouzayen (bouzayen@ensat.fr).

<sup>W</sup>Online version contains Web-only data.

Article, publication date, and citation information can be found at www.plantcell.org/cgi/doi/10.1105/tpc.105.033415.



1995). Screens for *Arabidopsis thaliana* mutants with altered auxin response or morphology have identified mutations in 10 different *Aux/IAA* genes: *IAA1/AXR5* (Park et al., 2002; Yang et al., 2004), *IAA3/SHY2* (Tian and Reed, 1999), *IAA6/SHY1* (Kim et al., 1996), *IAA7/AXR2* (Nagpal et al., 2000), *IAA12/BDL* (Hamann et al., 2002), *IAA14/SLR* (Fukaki et al., 2002), *IAA17/AXR3* (Rouse et al., 1998), *IAA18* (Reed, 2001), *IAA19/MSG2* (Tatematsu et al., 2004), and *IAA28* (Rogg et al., 2001). All of these primary mutations in *Aux/IAA* genes were found in the highly conserved domain II, which is responsible for protein degradation. The mutations stabilize the proteins, resulting in gain-of-function phenotypes. Several revertants have been shown to have mutations in other domains, affecting the capacity of the hyperstable proteins to function (Rouse et al., 1998). The *aux/iaa* mutants exhibit a variety of auxin-related developmental phenotypes, including altered phototropism/gravitropism, root formation, apical dominance, stem/hypocotyl elongation, leaf expansion, and leaf formation in the dark. However, because the stabilization caused by these mutations may not mimic regulatory events actually occurring in wild-type plants, accurate determination of the physiological significance of *Aux/IAA* proteins would benefit from the study of loss-of-function mutants. Unfortunately, the few *Aux/IAA* null mutations that have been characterized display subtle or nonscorable phenotypes, probably due to functional redundancy or to feedback regulatory loops that enable the mutant plants to compensate for absence of a particular *Aux/IAA* protein.

We report here on the functional characterization of *IAA9*, a member of subfamily IV of the tomato (*Solanum lycopersicum*) *Aux/IAA* gene family for which, to our knowledge, no mutation has been described so far. Phenotypes and molecular analyses of the downregulation of *IAA9* in transgenic tomatoes indicated that the *IAA9* protein is a pivotal mediator of auxin in the process of fruit set and leaf morphogenesis. The evidence provided by this work supports the hypothesis that the *IAA9* protein acts to repress the auxin response pathway.

## RESULTS

### *IAA9* Belongs to a Distinct Subfamily of the *Aux/IAA* Gene Family

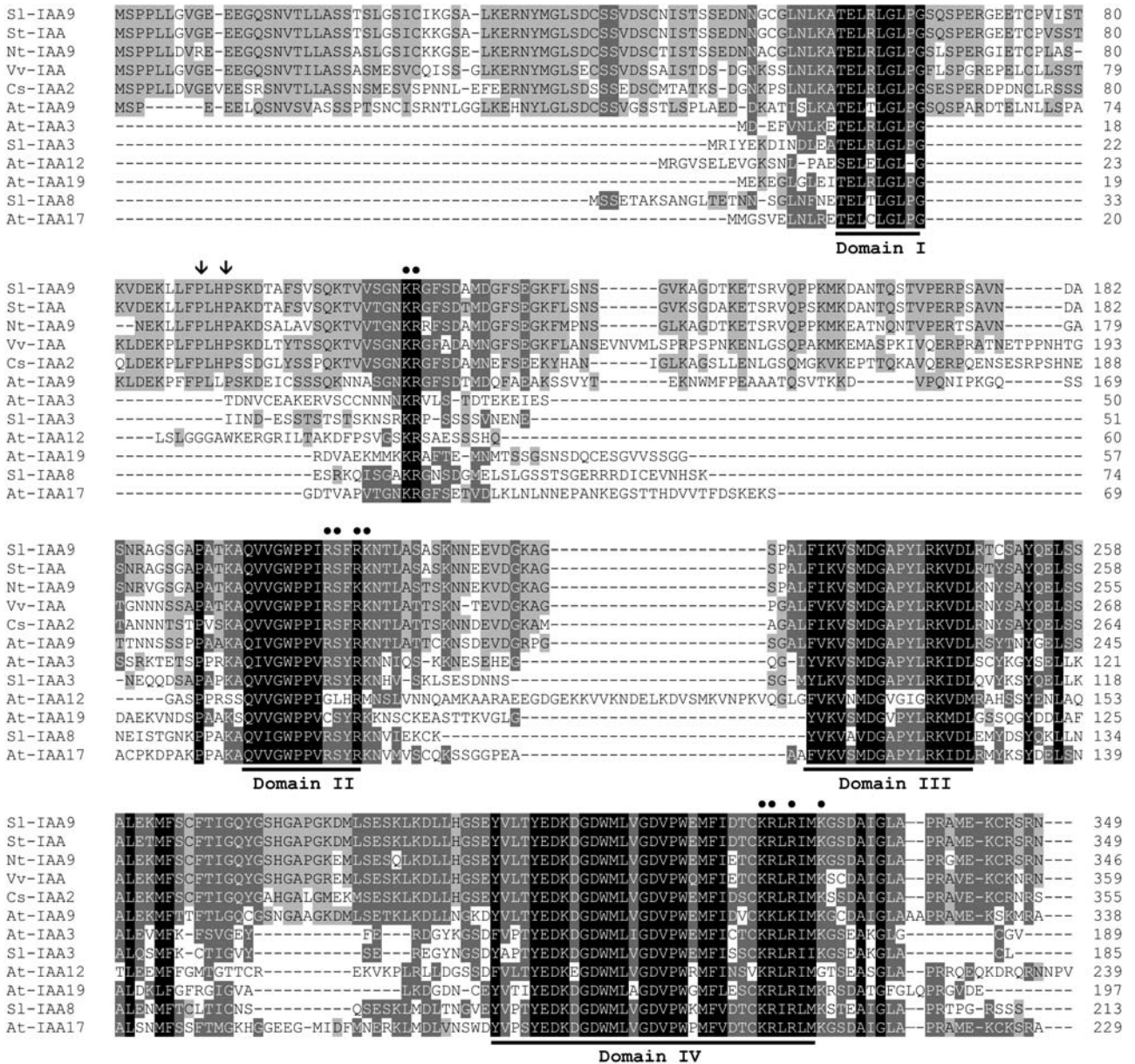
*IAA9* and other partial tomato *Aux/IAA* clones were isolated from tomato fruit using gene family-specific degenerate primers designed to conserve sequences in *Aux/IAAs* from different plant species (Jones et al., 2002). The *IAA9* cDNA clone was found to correspond to a previously isolated 301-bp partial tomato clone (Nebenfuhr et al., 2000), initially named *IAA4* (GenBank accession number AF022015). The complete coding sequence (1047 bases), encoding a putative protein of 349 amino acids, was then isolated and renamed *IAA9* to comply with the numbering retained for the *Arabidopsis Aux/IAA* genes. The derived protein comprises the four conserved domains characteristic of the *Aux/IAA* gene family, domains I to IV, and several nuclear localization signal (NLS) sequences (Figure 1). Phylogenetic analysis was conducted to assess the relationships between tomato *IAA9* and *Aux/IAA* genes from *Arabidopsis* and other plant species. The phylogenetic tree shown in Figure 2

suggests that *Aux/IAA* proteins can be grouped into four subfamilies. Subfamilies II, III, and IV are well defined, and their existence was confirmed by both PAUP and ProDom analyses (Servant et al., 2002). Subfamily I includes up to six subgroups, each of them bearing at least one specific domain. Tomato *IAA9* falls into subfamily IV along with sequences from grape (*Vitis vinifera*), cucumber (*Cucumis sativus*), *Zinnia elegans*, and *Arabidopsis*. Three other tomato *Aux/IAA* proteins, highly similar to their putative *Arabidopsis* homologs, fall into three of the four main subfamilies. Subfamily IV is clearly distinguishable from other *Aux/IAA* subfamilies, as its members are longer (typically >300 amino acids) than members of all other subfamilies (~200 amino acids) and typically contain 40 to 60 additional amino acids N-terminal to domain I and >50 amino acids between domains I and II (Figure 1). This N-terminal region unique to subfamily IV is well conserved among different plant species (shaded in gray in Figure 1), which suggests conserved function. The predicted tomato *IAA9* protein shares 97% identity with potato (*Solanum tuberosum*) *Aux/IAA* and 92% similarity with tobacco (*Nicotiana tabacum*) *IAA9*. It also shares 80, 78, 74, and 70% similarity with grape *Aux/IAA*, cucumber *IAA2*, *Z. elegans* auxin-responsive protein (ARP), and *Arabidopsis IAA9*, respectively. Tomato *IAA9* contains several highly conserved amino acid sequence motifs (Figure 1) unique to subfamily IV such as the short (Pxx)<sub>2</sub> sequences located between domains I and II, which are indicative of polyproline II conformations.

The predicted bipartite NLS comprised an invariant basic doublet KR in interdomain I/II and basic amino acids in domain II. A second SV40-type NLS is located in domain IV (Figure 1). The nuclear localization of the *IAA9* protein was investigated by fusion to the green fluorescent protein (GFP) under the control of the 35S promoter of *Cauliflower mosaic virus* and transient expression in tobacco protoplasts. Fluorescence microscopy analysis associated with image overlay techniques demonstrated that control cells transformed with GFP alone displayed fluorescence throughout the cell (Figure 3), whereas the *IAA9*-GFP fusion protein was exclusively localized to the nucleus, indicating that *IAA9* was able to fully direct the fusion protein to the nucleus (Figure 3). This nuclear targeting of the *IAA9* protein is consistent with a putative transcriptional regulatory function.

### *IAA9* Expression Is Ubiquitous and Constitutive

RT-PCR analyses showed that basal expression of the *IAA9* gene was high in roots, stems, leaves, flowers, and fruit (Figure 4A). Expression data gained from The Institute for Genomic Research tomato EST database (<http://www.tigr.org/tdb/tgi/lgi>) and the Solanaceae Genome Network tomato expression database (<http://www.sgn.cornell.edu>) indicated that *IAA9* is also highly expressed in seeds and in vitro-cultured callus tissues. The low number of cycles (<22) required for successful RT-PCR amplification in all tomato tissues examined suggested that the basal transcript levels for *IAA9* were higher than for other tomato *Aux/IAAs* (H. Wang and M. Bouzayen, unpublished data). *IAA9* expression also showed constitutive expression throughout leaf and fruit ontogeny (Figures 4B and 4C). In contrast with the majority of *Aux/IAA* gene family members, including *IAA2*, *IAA3*, and *IAA8*, that show rapid and strong induction by auxin,



**Figure 1.** Sequence Analysis of IAA9.

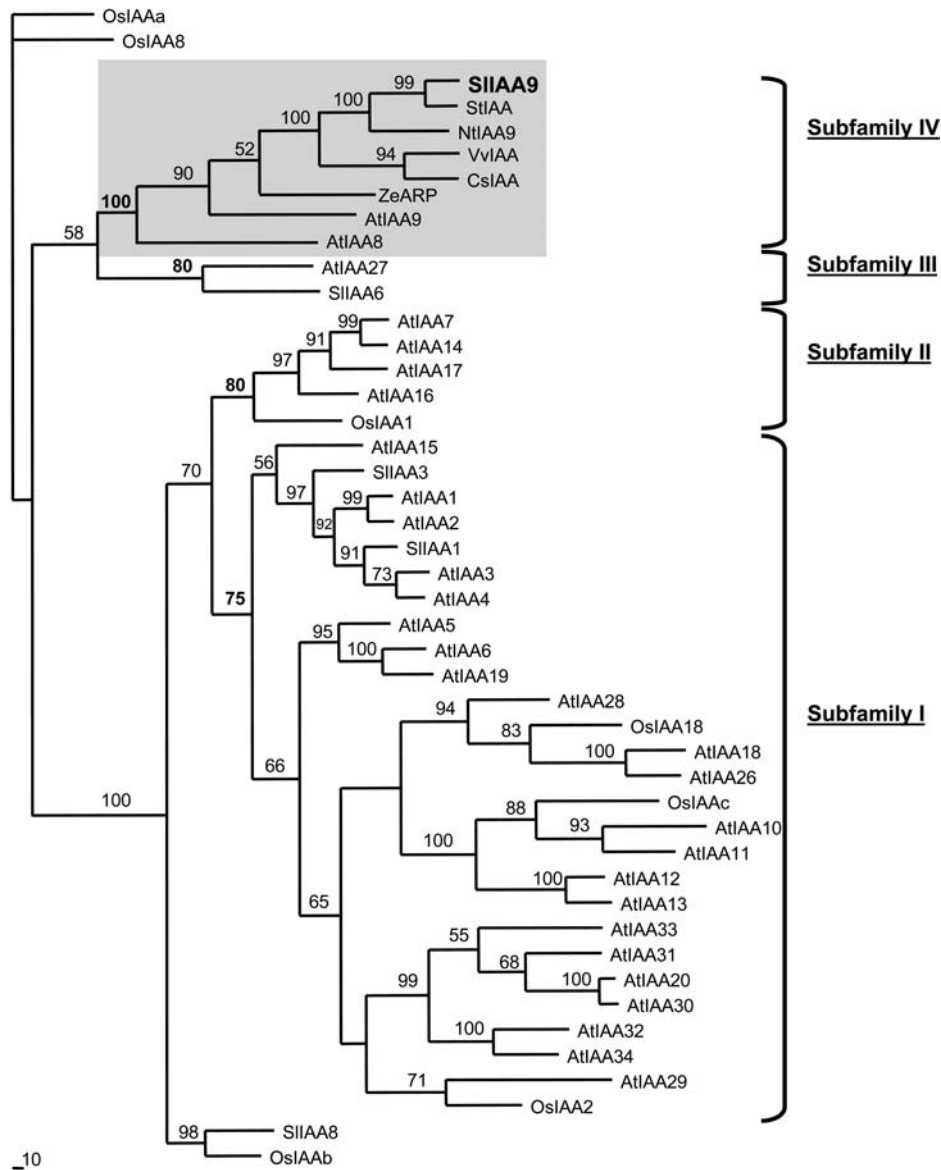
Sequence comparison of IAA9, putative orthologs from other plants, and other members of the Aux/IAA protein family. Conserved residues are shaded in black, dark gray shading indicates similar residues in at least 7 out of 12 of the sequences, and light gray shading indicates similar residues in 4 to 6 out of 12 of the sequences. The four conserved domains I, II, III, and IV are underlined. Conserved basic residues that putatively function as NLS are indicated by closed circles on the top of the alignment. Putative polyproline II (Pxx)<sub>2</sub> conformation residues are indicated by arrows.

IAA9 mRNA levels did not alter markedly after 30 min of auxin treatment, but increased after 3 h (Figure 4D).

**IAA9 Antisense Plants Exhibited Altered Leaf Morphology and Multiple Organ Fusion**

No mutants altered in the expression of subfamily IV members of the Aux/IAA gene family have been described to date. Therefore, an antisense construct of the IAA9 gene was stably transformed

into tomato plants to address the physiological significance of the encoded protein. Several transgenic antisense lines (AS-IAA9) corresponding to independent transformation events were selected and analyzed. AS-IAA9 lines exhibited a wide range of phenotypic effects, suggesting an important role for IAA9 in a number of developmental processes. The most readily visible phenotype was related to leaf morphology. Wild-type tomato leaves are unipinnately compound with a terminal leaflet and three pairs of lobed major lateral leaflets with pinnate venation.

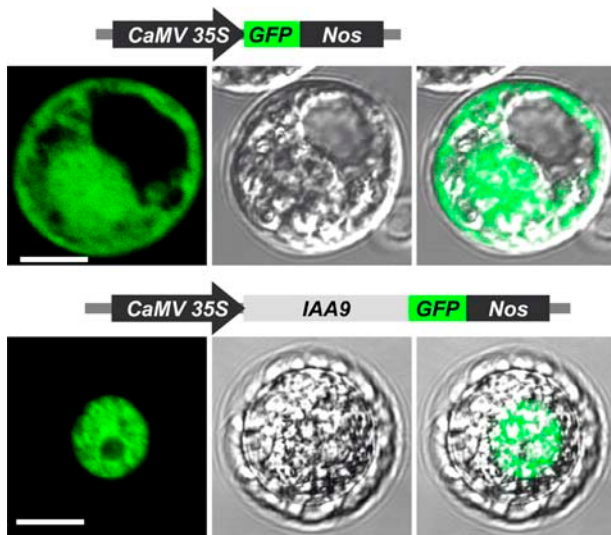


**Figure 2.** *IAA9* and Its Homologs Belong to a Distinct Subfamily of the *Aux/IAA* Gene Family.

The resultant tree was obtained by a character state cladistic approach (Parsimony) using *OslAAa* and *OslAA8* as the outgroups (chosen because of their relative isolation on preliminary calculations). Values above branches are bootstrap percentages (5000 replicates). The score of the best tree found was 4442 with consistency index of 0.6376 and homoplasy index of 0.3624.

Smaller leaflets are often seen between the major leaflets (Figure 5A). By contrast, the leaves of *AS-IAA9* plants were characterized by minimally lobed simple leaves varying from perfect entire-margined simple leaves to compound leaves (Figure 5B) depending on the level of downregulation displayed by the different transgenic lines (Figures 5B and 5C). In the lobed simple leaves, one or more pair of lateral leaflets merges with the terminal leaflet. In these leaves, the borders between the leaflets often remain defined, and the lamina was frequently asymmetrical and wrinkled. The consistency of the phenotypes was supported by reproducing them in two different tomato genetic backgrounds: *MicroTom* and *Ailsa Craig* (Figures 5A and 5B).

The levels of *IAA9* transcripts were significantly reduced in five antisense transgenic lines compared with that of wild-type control plants (Figure 5C), further supporting that downregulation of *IAA9* accounts for the phenotypes displayed by the transgenic lines. *AS111* and *AS213* showed the strongest downregulation of *IAA9* and correspondingly were found to have the most severe phenotypes producing almost complete simple leaf morphology (Figures 5B and 5C). Lines *AS250* and *AS58* showed a less dramatic decrease in *IAA9* transcript accumulation and exhibited only moderate phenotypes with both simple and lobed simple leaves, whereas line *AS70* showed a slight decrease in *IAA9* transcript accumulation and a concomitant weak phenotype with



**Figure 3.** Nuclear Localization of the IAA9 Protein Fused to a GFP Tag.

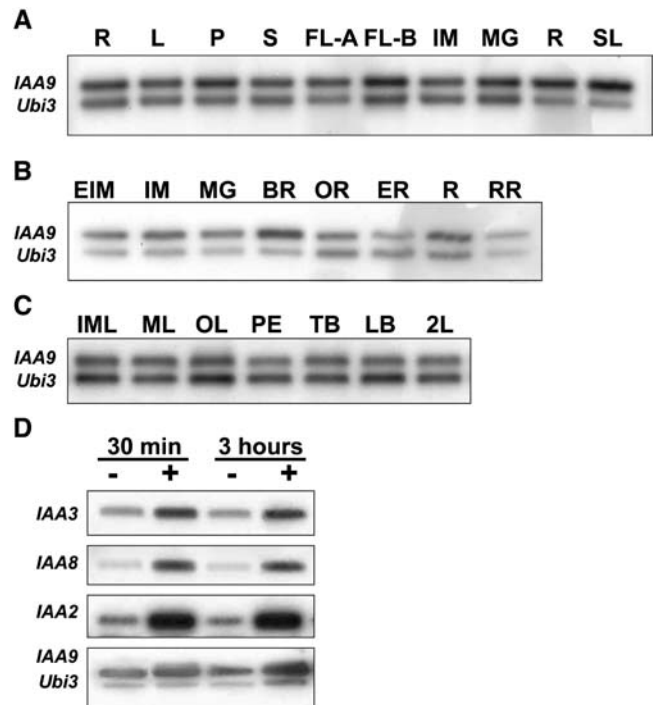
Tobacco protoplasts were transiently transformed with either *Pro*<sub>35S</sub>:*GFP* or *Pro*<sub>35S</sub>:*IAA9-GFP* constructs, and the subcellular localizations of the IAA9-GFP fusion protein or the GFP were analyzed by confocal laser scanning microscopy (left panel). Light micrographs (middle panel) and fluorescence (left panel) images are merged (right panel) to illustrate the different location of the two proteins. Bars = 10  $\mu$ m. CaMV, *Cauliflower mosaic virus*.

only occasional lobed simple leaves. Finally, line AS252 showed a similar level of *IAA9* transcript accumulation to the wild type and was morphologically indistinguishable from wild-type plants. Interestingly, the phenotypes of *AS-IAA9* lines were reminiscent of simple leaves in the recessive spontaneous tomato mutant *entire*, which was also found to accumulate significantly lower levels of *IAA9* mRNA than wild-type plants (Figures 5A and 5C). By contrast, the expression of *IAA9* was not altered in *Lanceolate* and *Petroselinum*, two other leaf-shape mutants (see Supplemental Figure 1 online). Transcript accumulation of *KNOX* genes (*TKN1* and *LeT6/TKN2*) and *PHANTASTICA*, the ortholog of *Arabidopsis ASYMMETRIC LEAVES1*, all known to be involved in leaf morphogenesis (Kessler and Sinha, 2004), was not affected in whole expanding leaves, fruit, and seedlings of *AS-IAA9* lines (see Supplemental Figure 2 online).

*AS-IAA9* lines also displayed multiple organ fusions. The majority of the fused petioles/leaves (Figure 5D) were position dependent and usually found between the first and the third true leaves that correspond to the transition leaves. Such a phenotype is reminiscent of the *Arabidopsis pin1* mutant affected in auxin transport (Galweiler et al., 1998). Fusions of flowers, fruit, cotyledons, and stems were also observed (see Supplemental Figure 3 online). In addition, bifurcated, triple, or quadruple cotyledons were also frequently observed in *AS-IAA9* seedlings (Figure 5E). While wild-type tomato flowers and fruit normally carry five to six symmetrically arranged sepals, *AS-IAA9* tomato invariably had partially fused and asymmetrically arranged sepals (Figure 5F). The tomato *entire* mutant also carries, though at a lower frequency, asymmetrical sepals and a multifusion phenotype affecting leaves, flowers, and fruits.

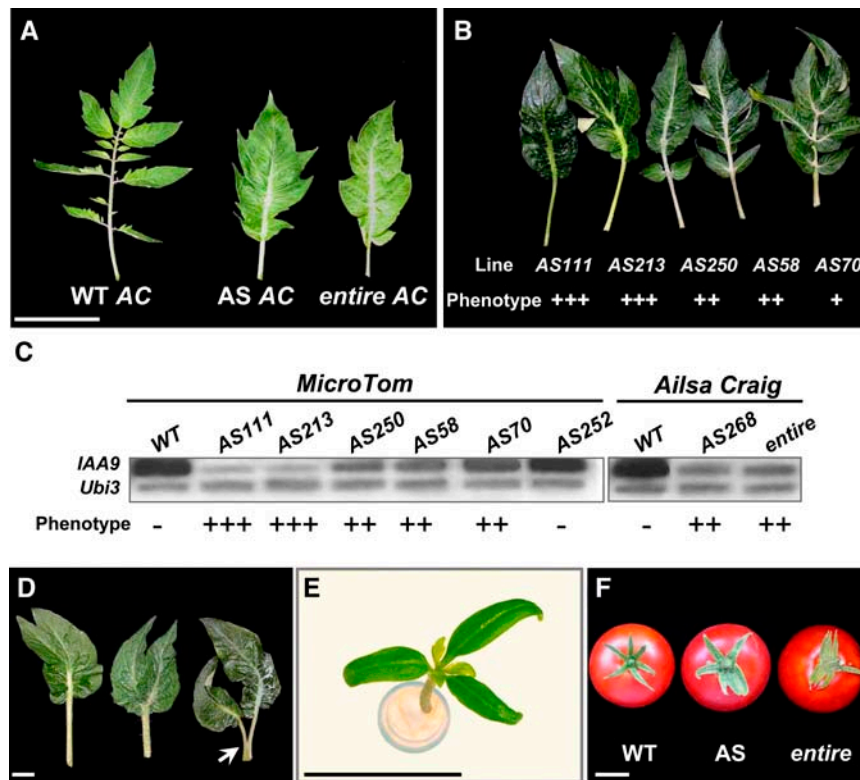
### Fruit Set Prior to Pollination Results in Parthenocarpy in *AS-IAA9* Plants

Downregulation of *IAA9* resulted in a dramatic alteration of early fruit development, with all *AS-IAA9* lines exhibiting precocious fruit set prior to anthesis, resulting in parallel fruit and flower development (Figure 6B). Depending on how early the precocious enlargement of the fruit occurred, the development of the stamens was either completely impaired (Figure 6A), or in less severe cases, the fused stamen cones were torn open by the developing fruit (Figure 6B). The accelerated enlargement of the fruit positioned the stigma out of reach of the stamens, impairing self-pollination and contributing to the development of seedless fruit (Figure 6C). Despite their parthenocarpic character, these fruit were similar in appearance to wild-type tomato in terms of size, skin color, and flesh consistency (Figure 6C) as well as ripening-associated ethylene production (see Supplemental Figure 4 online). In the strongly inhibited line *AS111*, the proportion of parthenocarpic fruit reached 82% (Figure 6D). Once again, the



**Figure 4.** RT-PCR Analysis of *IAA9* Expression Patterns.

**(A)** *IAA9* expression in different organs: roots (R), leaves (L), petiole (P), stem (S), flowers at anthesis (FL-A), flower buds (FL-B), immature green fruit (IM), mature green fruit (MG), red fruit (R), and seedlings (SL). **(B)** *IAA9* expression at different stages of fruit development: early immature green (EIM), immature green (IM), mature green (MG), breaker (BR), orange (OR), early red (ER), red (R), and red ripe (RR). **(C)** *IAA9* expression in expanding immature leaves (IML), fully expanded mature leaves (ML), aged, turning yellow leaves (OL), petioles (PE), terminal-leaflet blades (TB), lateral-leaflet blades (LB), and the second leaf counted from the cotyledon node (2L). **(D)** Transcript accumulation of *IAA3*, *IAA8*, and *IAA9* tomato *Aux/IAA* genes in response to auxin treatment (20  $\mu$ M IAA) of tomato seedlings.



**Figure 5.** Vegetative Growth Phenotypes of *IAA9* Downregulated Plants.

**(A)** Tomato leaf morphology in wild-type *Ailsa Craig* (WT AC), *IAA9* antisense (AS AC), and monogenic spontaneous mutant *entire* (*entire* AC) lines. Bar = 100 mm.

**(B)** Positive correlation between the severity of the simple-leaf phenotype and the level of downregulation of the *IAA9* gene in *MicroTom* genotype. *AS111* and *AS213* are strongly inhibited lines, and *AS250*, *AS58*, and *AS70* are weakly inhibited lines.

**(C)** RT-PCR analysis of *IAA9* transcript accumulation in wild-type and antisense lines either in *MicroTom* (*AS111*, *AS213*, *AS250*, *AS58*, *AS70*, and *AS252*) or *Ailsa Craig* (*AS268*) genetic backgrounds and in *entire* mutant plants. Symbols indicate the presence (+) or absence (-) of leaf and parthenocarpity phenotypes. +++ designates lines with only simple leaves and a high percentage of parthenocarpity (60 to 100%); ++ designates lines showing both lobed and entire margin simple leaves and moderate percentage of parthenocarpity (30 to 40%); + designates lines displaying only lobed simple leaves and occasional parthenocarpity.

**(D)** Leaf fusion in *IAA9* antisense lines consists of either fused twin leaves appearing as a single leaf with two terminal apices (left), one petiole bearing two lamina (medium), or two leaves partially fused at the end of the petiole and forming a *pin-like* structure (right) as indicated by the arrowhead.

**(E)** Three-cotyledon structure often occurring in seedlings of strongly inhibited lines.

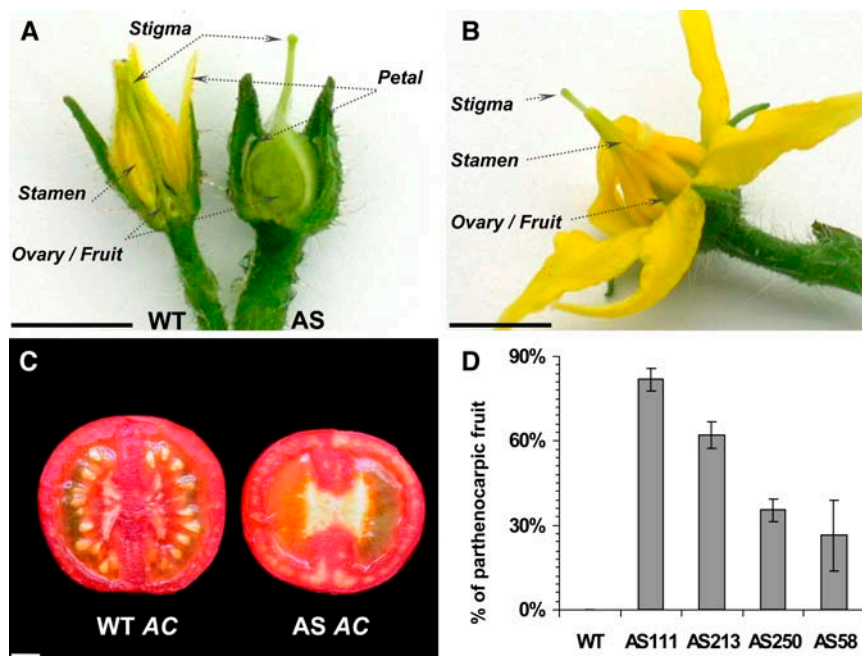
**(F)** Fused sepals in antisense lines. Wild-type fruit bear symmetrically arranged calyx, and *AS-IAA9* (AS) and *entire* plants exhibit asymmetrical and partly fused calyx. Bars = 20 mm in **(D)** to **(F)**.

level of downregulation of the *IAA9* gene strongly correlated with the severity of the phenotypes. The most strongly downregulated lines exhibited the highest percentage of parthenocarpic fruit (Figure 6D).

Emasculations assays were performed to determine whether fruit set was independent of pollination and fertilization. Without pollination, fruit set never occurred in emasculated wild-type flowers, while *AS-IAA9* plants developed fruits from the emasculated, unfertilized flowers in all four independent lines tested (Table 1). These fruit proved to be seedless in all cases. The percentage of fruit set from *AS-IAA9* emasculated flowers ranged from 48 to 68% (Table 1) among the four independent transgenic lines. These data indicated that fruit set in *AS-IAA9* lines can occur independently of pollination and fertilization.

Cross-pollination assays were performed to examine whether pollen or ovule sterility contributed to the parthenocarpity of *AS-IAA9* lines. Flowers were emasculated before dehiscence of the anthers (pre-anthesis) and manually pollinated. A high percentage of successful fruit set was achieved (79%) using wild-type flowers as the female recipient and *AS-IAA9* plants as the pollen donor. Moreover, all the developed fruits were seeded, and when germinated, all these seeds were kanamycin resistant (Table 2), indicating that *AS-IAA9* pollen were viable. When the reciprocal cross was made, 54% of the pollinated *AS-IAA9* flowers set fruit, and among these, 35% were seeded (Table 2), indicating that the ovules of *AS-IAA9* plants were also fertile. These data clearly indicate that parthenocarpity in *AS-IAA9* lines was mainly due to an alteration of the normal fruit set program.





**Figure 6.** Fruit Set and Parthenocarpy in AS-IAA9 Lines.

(A) Flower buds at 1 d before anthesis in wild-type and AS-IAA9 lines (AS), showing dramatically enlarged ovary and underdeveloped stamen in AS-IAA9 lines.

(B) Precocious fruit development in AS-IAA9.

(C) Wild-type *Ailsa Craig* seeded fruit (WT AC) and AS-IAA9 parthenocarpic fruit (AS AC). Bars = 10 mm in (A) to (C).

(D) The level of parthenocarpy in AS-IAA9 tomato lines positively correlates with the level of downregulation of the *IAA9* gene in independent AS-IAA9 lines. Error bars indicate mean  $\pm$  SE of at least five independent trials,  $n \geq 200$  fruits.

### AS-IAA9 Plants Showed Increased Stem Elongation and Reduced Apical Dominance

AS-IAA9 plants were usually taller than wild-type plants (Figure 7A) and exhibited enhanced hypocotyl elongation and longer internodes (Figure 7B). AS-IAA9 plants showed a strongly reduced apical dominance and a dramatically altered pattern of axillary shoot development (Figure 7C). In wild-type tomato plants, lateral shoots appeared only after floral transition and grew out in an apical-basal sequence along the primary shoot axis. In this basipetal growth pattern, the first lateral shoot arises from the last leaf node just below the first inflorescence (Figure 7C). By contrast, AS-IAA9 plants showed an acropetal growth pattern where the lateral shoots grew out in a basal-apical sequence, the first lateral shoot arising from the first leaf node just above the cotyledon (Figure 7C). As a consequence, in all transgenic lines, lateral shoots were always present in the first three internodes, while in wild-type plants, axillary shoots never arose from the first three internodes (Figure 7D). The level of reduction of apical dominance in the transgenic lines clearly correlated with the level of downregulation of the endogenous *IAA9* gene; that is, the strongly inhibited line AS213 showed higher apical dominance than the weakly inhibited line AS250 (Figure 7E). Downregulation of *IAA9* also promoted lateral shoot elongation, resulting in axillary shoots fourfold to fivefold longer than those in control wild-type plants (Figure 7F).

### Downregulation of *IAA9* Resulted in Enhanced Auxin Sensitivity

A number of phenotypes associated with the downregulation of *IAA9* suggested enhanced sensitivity to auxin. To further investigate this, we examined auxin dose response on both root formation and hypocotyl segment elongation. Figure 8A shows that the promotion of root organ regeneration from cotyledon explants was auxin dose dependent in both wild-type and AS-IAA9 plants. However, in wild-type seedlings, the synthetic auxin,  $\alpha$ -naphthalene acetic acid (NAA), promoted root regeneration at concentrations above 0.1  $\mu$ M, while in AS-IAA9, the

**Table 1.** Emasculation Assay to Assess the Ability to Set Fruit in the Absence of Pollination

Lines	Fruit Set (%)	Fruit Set/Emasculated Flower	Seed Number/Fruit
Wild type	0	0/100	0
AS111	63	28/44	0
AS213	68	39/56	0
AS250	49	9/17	0
AS58	48	7/14	0

Wild-type and AS-IAA9 flowers were emasculated 1 to 4 d before anthesis. The results represent two independent trials.

**Table 2.** Cross-Fertilization Assay

Female Recipient	Pollen Donor	Fruit Set/Crossed Flowers	Fruit Set (%)	Seeded Fruit (%)	F1 Kanamycin Resistance (%)
Wild type	AS- <i>IAA9</i>	65/82	79	100	100
AS- <i>IAA9</i>	Wild type	20/37	54	35	100
Wild type	Wild type	7/7	100	100	0

Tomato pollen from wild-type flowers was used to fertilize emasculated AS-*IAA9* flowers and then the number of fruit assessed at the ripe stage. Conversely, emasculated wild-type flowers were fertilized with AS-*IAA9* pollen. In the control assay, wild-type emasculated flowers were fertilized with wild-type pollen. For each cross-fertilization assay, the proportion of seeded fruit was determined and the capacity of the F1 seeds to grow on kanamycin-containing ( $70 \text{ mg L}^{-1}$ ) medium was assessed. The results are representative of data obtained with three independent AS-*IAA9* lines (AS213, AS111, and AS250).

critical concentration was 10-fold lower ( $0.01 \mu\text{M}$ ). At concentrations higher than optimal, auxin still promoted root formation but inhibited root outgrowth, resulting in short and swollen roots. At  $40 \mu\text{M}$ , root regeneration was almost completely suppressed in AS-*IAA9*, while it was still active in the wild type, indicating that AS-*IAA9* lines were more sensitive to the inhibitory effect of high auxin concentrations. The higher sensitivity of AS-*IAA9* plants to auxin was further explored by determining the auxin dose response on hypocotyl segments. Figure 8B shows that after 23 h of auxin treatment the maximum hypocotyl elongation was obtained with  $10^{-4} \text{ M}$  for the wild type, while it was reached with a 10-fold lower concentration ( $10^{-5} \text{ M}$ ) in the AS-*IAA9* lines.

#### **N-1-Naphthylphthalamic Acid–Treated Wild-type Plants Mimicked the Simple-Leaf Phenotype of AS-*IAA9* Plants**

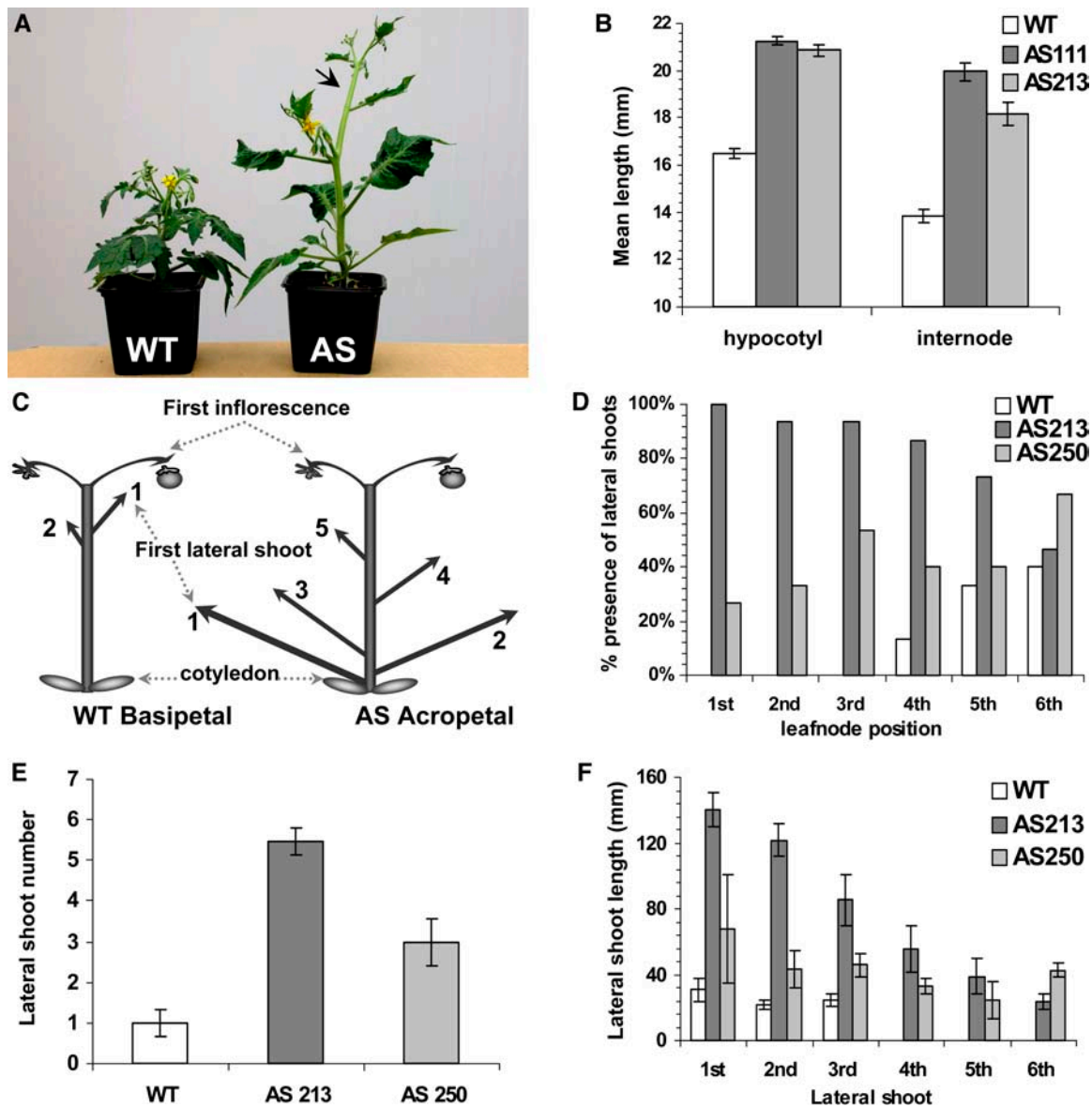
Under normal growth conditions, 19-d-old wild-type seedlings had one pair of either compound or deeply lobed leaves and a well-developed primary root bearing one or two lateral roots (Figures 9A and 9C). AS-*IAA9* seedlings grown under the same condition exhibited simple leaves, enhanced primary root elongation (147% longer than the wild type), and an increased number of lateral roots (Figures 9A to 9C). N-1-naphthylphthalamic acid (NPA), an auxin transport inhibitor, is known to alter the endogenous auxin gradients, resulting in overaccumulation of auxin in root and shoot apices. In wild-type plants, NPA led to the swelling and dramatic reduction of primary root elongation as well as the suppression of lateral root formation. By contrast, primary root growth continued in NPA-treated AS-*IAA9* plants, although lateral root formation was completely inhibited (Figures 9A to 9C). Leaves of wild-type seedlings treated with  $1 \mu\text{M}$  NPA shifted from compound to simple, with an entire leaf margin mimicking the phenotype of *IAA9*-inhibited plants. In AS-*IAA9* lines, NPA treatment enhanced the antisense-associated phenotype by increasing the frequency of leaf fusion and exaggerating the entire leaf margin. It has been postulated that leaf shape is coupled with vascular patterning. We therefore investigated leaf venation in wild-type and antisense plants. Figure 9D shows vein patterning in wild-type, AS-*IAA9*, and NPA-treated leaves. Wild-type leaves had compound unipinnate venation showing discrete vein size orders; the large primary vein or midvein was continuous with the stem vascular bundles and extended along the length of the leaf. Secondary veins branched from the primary vein and extended to the tips of the separate leaflets, which

reflected the compound lamina morphogenesis. Tertiary and quaternary veins interconnected veins of higher orders or formed open ends (Figure 9D, left panels). AS-*IAA9* leaves had unipinnate venation but no separate leaflets. The number of secondary veins was increased from six to nine per terminal lamina half in the wild type to 10 to 13 per lamina half in AS-*IAA9* (Figure 9D, middle panels). Moreover, AS-*IAA9* leaves had more tertiary and quaternary veins connecting higher order veins than the wild type, and the number of parallel vascular strands in the midvein and petiole was increased. NPA-treated wild-type plants also displayed a shift from compound to simple unipinnate venation with a dramatic increase of vein branching (Figure 9D, right panels). Compared with AS-*IAA9* lines, the vascular differentiation of NPA-treated wild-type leaves was more dramatic with the leaf tip displaying the most extensive density of disorganized vein strands, resulting in a complete loss of the central vein order. The central regions and petioles of NPA-treated leaves were broadened, and the lamina surface of these hypervascularized leaves was extensively wrinkled (shown in inset of Figure 9A).

#### **IAA9 Selectively Controls the Expression of *IAA3***

Given the strong auxin-associated phenotypes in the AS-*IAA9* lines, we investigated the effect of downregulation of the *IAA9* gene on the expression of a number of early/primary auxin-responsive genes. Figure 10A indicates that basal accumulation of *IAA3* transcripts in roots was higher in AS-*IAA9* than in the wild type. Nevertheless, *IAA3* retained the capacity to respond to exogenous auxin treatment in the AS-*IAA9* genetic background (Figure 10A). By contrast, downregulation of *IAA9* did not alter the basal expression of other auxin-inducible genes, such as *IAA2*, *IAA6*, *SAUR*, and *GH3*, whose transcript accumulation remained identical to that in the wild type and displayed similar auxin inducibility. Since *IAA9* transcript accumulation is reduced in the *entire* mutant, we investigated whether the expression of some other *Aux/IAA* genes was also affected in this mutant. Semiquantitative analyses by RT-PCR indicated that, with the exception of *IAA9*, none of the other eight *Aux/IAA* genes tested showed significantly altered expression in *entire* leaves (see Supplemental Figure 5 online).

To further investigate the role of *IAA9* in controlling the expression of *IAA3*, we examined the expression of a  $\beta$ -glucuronidase (*GUS*) reporter gene driven by the *IAA3* promoter (*Pro<sub>IAA3</sub>*) in both wild-type and AS-*IAA9* genetic backgrounds. *IAA3* is a tomato *Aux/IAA* gene that showed rapid induction by auxin (Figure 10A).



**Figure 7.** Increased Stem Elongation and Decreased Apical Dominance in AS-*IAA9* Lines.

(A) AS-*IAA9* plants (AS) are taller than the wild type and have longer internodes (indicated by arrow).

(B) Increased hypocotyl and stem elongation in *IAA9* downregulated lines. The data show the hypocotyl mean length of 21-d-old seedlings and internode mean length of 50-d-old plants in wild-type and two independent AS-*IAA9* lines in *MicroTom* genetic background (AS213 and AS111). The data are the mean  $\pm$  SE of at least 68 seedlings or plants and are representative of three independent experiments.

(C) Diagram depicting the inverted pattern of axillary shoot development and reduced apical dominance in AS-*IAA9* (AS) compared with wild-type plants. Numbers indicate the emergence order of lateral shoots.

(D) Percentage of plants displaying a lateral shoot in each of the first six leaf nodes in wild-type and two antisense lines.

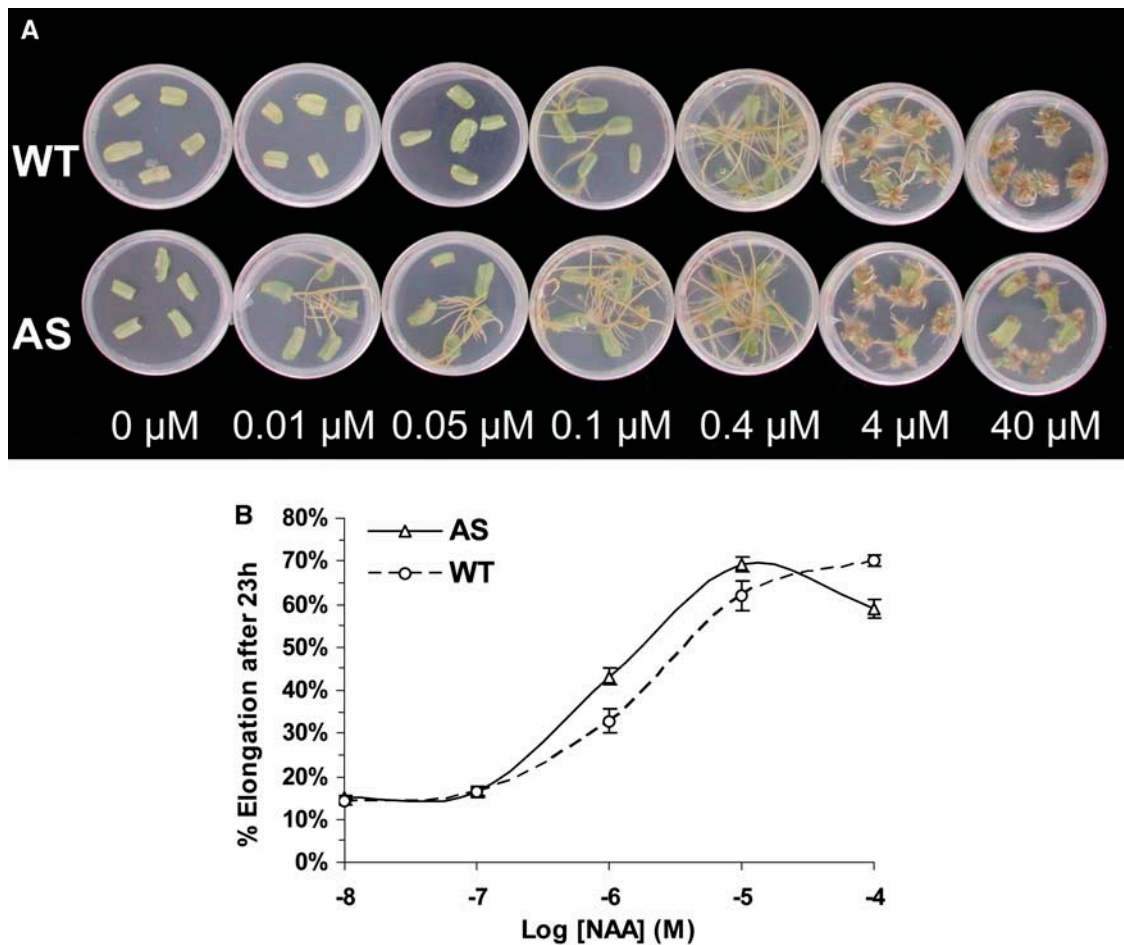
(E) Increased total number of lateral shoots in AS-*IAA9* lines.

(F) AS-*IAA9* lines bear longer lateral shoots. Error bars represent mean  $\pm$  SE of at least 15 plants in (E) and (F).

In the wild-type background, *Pro<sub>IAA3</sub>*-driven *GUS* expression was mainly found in the aerial part of the plantlets, particularly in the shoot apex and in the vascular system (Figure 10B), but was barely detectable in the hypocotyl and in the root system, where *GUS* staining was restricted to small spots corresponding to root

tips and lateral root initiation sites (inset of Figure 10B). Upon exogenous auxin treatment, *GUS* expression in the wild-type background was dramatically increased and spread to all parts of the plantlet (Figure 10C), clearly demonstrating the responsiveness of the *IAA3* promoter to auxin treatment. In AS-*IAA9*





**Figure 8.** Enhanced Auxin Sensitivity in *IAA9*-Inhibited Lines.

**(A)** Auxin dose–response assay of cotyledon explants. Root formation is induced by lower auxin (NAA) concentration in *AS-IAA9* (AS) plants than in the wild type. Root regeneration is promoted at 0.1  $\mu$ M NAA in the wild type and at 10 times lower concentration (0.01  $\mu$ M) in *AS-IAA9* lines.

**(B)** Auxin dose response in hypocotyl segments. Elongation is given as percentage increase in final length over the initial length after 23 h incubation in a solution containing the indicated NAA concentration. The results are representative of data obtained with three independent *AS-IAA9* lines: *AS213*, *AS111*, and *AS250*. Error bars represent mean  $\pm$  SE,  $n \geq 20$ .

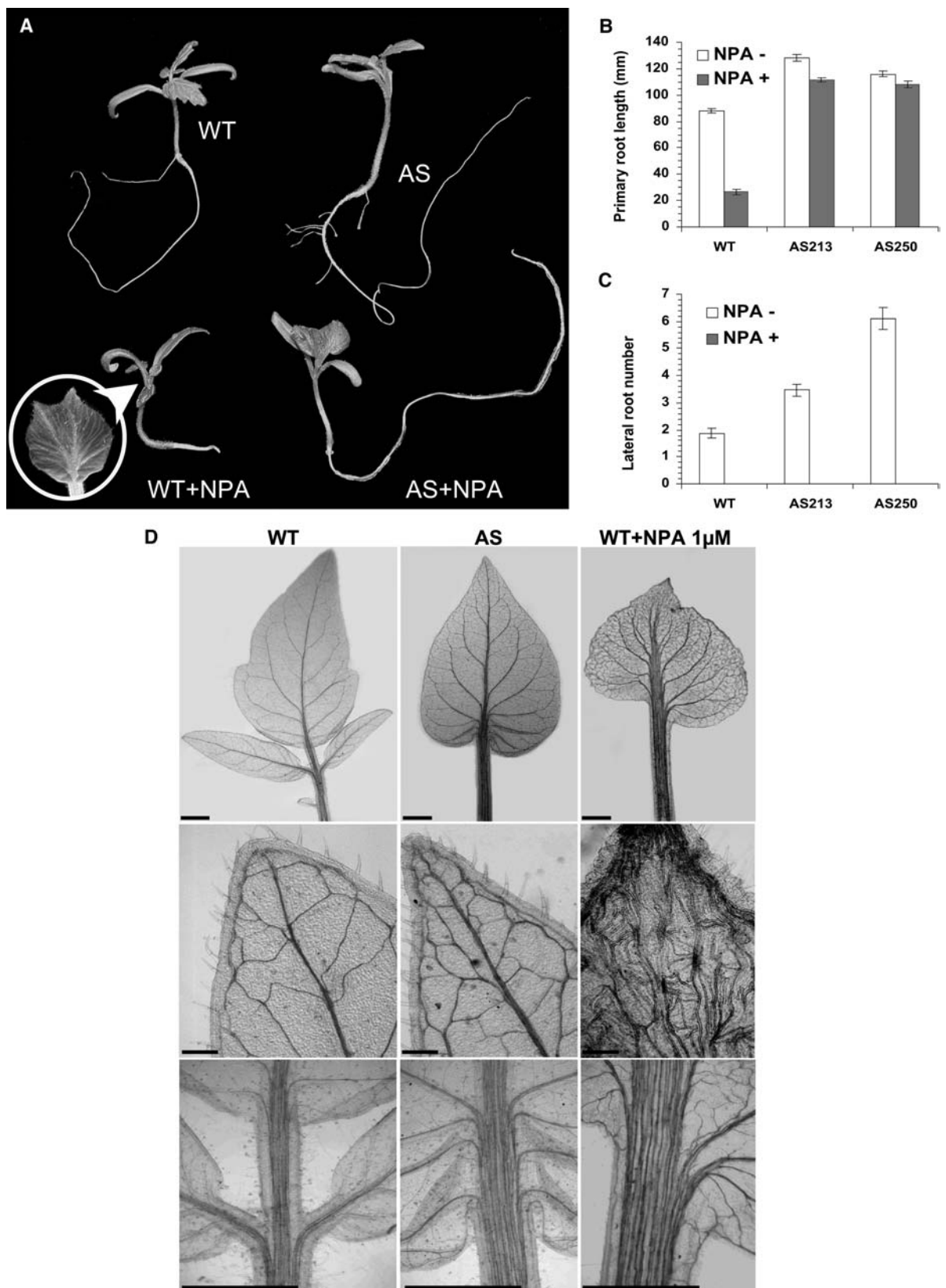
background, *Pro<sub>IAA3</sub>* activity was dramatically enhanced in the absence of auxin treatment and displayed a pattern of expression mimicking that of auxin-treated wild-type lines. In particular, the GUS staining was detected over the entire root system, indicating that downregulation of *IAA9* expression induces a constitutive activation of the auxin-responsive promoter and that *IAA9* acts as a transcriptional repressor of auxin-induced gene expression.

## DISCUSSION

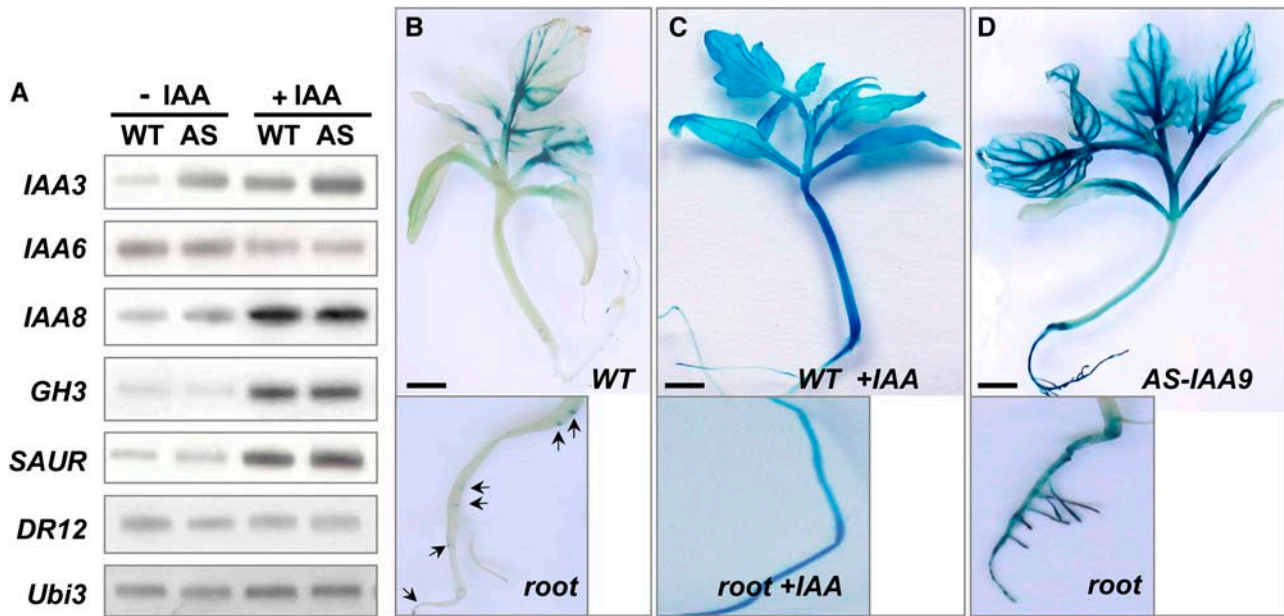
Our understanding of the functional role of Aux/IAA proteins has come mainly through the study of dominant or semidominant, gain-of-function *aux/iaa Arabidopsis* mutant lines. This work reports on the downregulation of a tomato *Aux/IAA* gene, *IAA9*, resulting in a pleiotropic morphological and developmental phenotype. Downregulation of *IAA9* altered leaf architecture and vascular venation

patterning, fruit set and development, apical dominance, and many other aspects of vegetative and reproductive growth in the *AS-IAA9* transgenic lines. The pleiotropic phenotypes were consistent with the multiorgan pattern of expression of the *IAA9* gene, and most of them are indicative of a role for the encoded protein in auxin responsiveness. The combined evidence suggests that *IAA9* acts as a negative regulator of auxin responses in tomato.

A number of Aux/IAA proteins have been shown to function as transcriptional repressors through interactions with ARF proteins that bind to auxin-responsive elements in the promoters of auxin-responsive genes. In *Arabidopsis*, mutations in domain II of Aux/IAA proteins result in the relative stabilization of the mutated proteins, leading in most cases to phenotypes consistent with a reduced auxin responsiveness (Reed, 2001). Gain-of-function mutations in individual *Arabidopsis Aux/IAA* genes result in distinctive phenotypes, presumably indicative of the specific expression patterns and/or alterations in the dynamics of



**Figure 9.** NPA-Treated Wild-Type Plants Phenocopy the Simple-Leaf Phenotype of AS-*IAA9* Plants.



**Figure 10.** Expression of Primary Auxin-Responsive Genes in *IAA9*-Inhibited Lines.

(A) RT-PCR analyses of the expression of auxin-responsive genes (*IAA3*, *IAA6*, *IAA2*, *GH3*, *SAUR*, and *DR12*) in the wild type and *IAA9*-inhibited (AS) line (AS213) upon auxin treatment (20  $\mu$ M IAA for 3 h).

(B) Expression pattern of the auxin-inducible *IAA3* promoter (*Pro<sub>IAA3</sub>*) fused to the *GUS* reporter gene in wild-type seedlings. Inset contains enlarged picture showing *GUS* staining limited to small spots corresponding to the root tip and lateral root initiation sites (arrowheads).

(C) Expression pattern of *Pro<sub>IAA3</sub>:GUS* in wild-type seedlings treated with exogenous auxin (20  $\mu$ M IAA). Inset shows *GUS* staining throughout the root.

(D) Expression pattern of *Pro<sub>IAA3</sub>:GUS* in AS-*IAA9* background with no auxin treatment. Inset shows *GUS* staining present in all parts of the root system. Bars = 5 mm in (B) to (D).

homoprotein/heteroprotein interactions between Aux/IAA and ARFs. Intragenic suppressors of gain-of-function alleles (Timpte et al., 1994; Rouse et al., 1998; Tian and Reed, 1999) and T-DNA insertional mutants (Nagpal et al., 2000) have subtle or indiscernible phenotypes (Reed, 2001; Liscum and Reed, 2002). It is generally thought that functional redundancy in the *Aux/IAA* gene family in *Arabidopsis* accounts for the absence of clear phenotypes, and as an attempt to overcome this problem, efforts are now underway to isolate lines with mutations in several *Aux/IAA*s. In this work, the *IAA9* antisense construct was mainly targeted to the 3' untranslated region to prevent cross-inhibition of other *Aux/IAA* genes through a sequence homology-based posttranscriptional mechanism. The efficiency of this strategy was validated by assessing the expression of nine *Aux/IAA* genes, including *IAA6*, which shares the highest sequence similarity with *IAA9*, and none of them were downregulated in the AS-*IAA9*

transgenic plants (Figure 10A; see Supplemental Figure 5 online). These data support the hypothesis that the observed phenotypes were directly attributable to the downregulation of the *IAA9* gene. Although the phylogenetic data clearly indicate that *IAA9* protein bears specific domains that are not found in other *Aux/IAA* subfamilies, it remains possible, however, that the expression of a yet unidentified tomato *Aux/IAA* might be directly affected by the transgene. We have generated a number of transgenic lines downregulated in other *Aux/IAA* genes, and none of them have reproduced the dramatic leaf and fruit phenotypes displayed by the AS-*IAA9* lines. In this context, the AS-*IAA9* tomato lines present a valuable means to determine *Aux/IAA* dynamics in auxin response and to reveal distinct roles for *Aux/IAA* genes that have not been described previously, including their role in fruit set and leaf morphogenesis.

**Figure 9.** (continued).

(A) Mock treatments are presented in the top row, and NPA treatments (1  $\mu$ M) are presented in the bottom row. NPA-treated wild-type seedlings reproduced the simple-leaf and leaf-fusion phenotypes (inset).

(B) Effect of NPA treatment on primary root elongation in light-grown wild-type and AS-*IAA9* plants.

(C) Inhibition of lateral root formation in wild-type and AS-*IAA9* lines (AS213 and AS250) upon NPA application. Error bars represent mean  $\pm$  SE ( $n \geq 63$ ).

(D) Leaf vascular patterns in wild-type (left panels), AS-*IAA9* (middle panels), and NPA-treated wild-type plants (right panels). Top row, overall views of venation pattern; middle row, close-up pictures of the leaf tips; bottom row, proximal region of midvein. Bars = 1 mm (top and bottom rows) and 100  $\mu$ m (middle row).

*IAA9* was isolated from a tomato fruit cDNA library and found to code for a putative member of the Aux/IAA protein family (Jones et al., 2002). The strong sequence similarity of the highly conserved Aux/IAA domains I to IV, the nuclear localization of the encoded protein, and the phenotypes of the *AS-IAA9* tomato lines strongly suggest that *IAA9* encodes a functional Aux/IAA protein. Aux/IAA multigene families have been isolated from *Arabidopsis* (Abel et al., 1995), tobacco (Dargeviciute et al., 1998), tomato (Nebenfuhr et al., 2000), soybean (*Glycine max*; Ainley et al., 1988), and a number of other dicot and monocot species. Tomato *IAA9* and its putative *Arabidopsis* orthologs, *IAA9* and *IAA8*, belong to Aux/IAA subfamily IV (Figure 2), which includes members encoding proteins 50 to 89% larger than Aux/IAs from all other subfamilies. In addition to the domains I to IV common to all Aux/IAs, subfamily IV proteins have a highly conserved, extended N-terminal region upstream of domain I and between domains I and II, comprising up to 150 extra amino acids. This includes a (Pxx)<sub>2</sub> motif located between domains I and II and several other amino acid domains perfectly conserved across plant species (shaded in gray in Figure 1) that are indicative of a conserved function for these regions.

Auxin responsiveness of Aux/IAA genes differs considerably from gene to gene. Depending on the gene and concentration of the exogenous hormone, responses range from strong and rapid induction (Abel et al., 1995) to a slight decrease in transcript abundance (Rogg et al., 2001). *IAA9* exhibits several features in common with its *Arabidopsis* orthologs *IAA9* and *IAA8*, including widespread high-level basal expression compared with other Aux/IAA genes and a relatively weak auxin inducibility. No phenotypes have been reported to date for knockout or gain-of-function mutations in any member of subfamily IV in *Arabidopsis* (*IAA9* and *IAA8*) or in any other plant species.

Fruit set is triggered by pollination and fertilization, and auxin signaling is thought to play a dynamic role in the regulation of fruit set and early growth. It has been postulated that auxin is first produced by elongating pollen tubes and then by the embryo and endosperm in the developing seeds. Subsequent development of the fruit appears to depend on these sources of auxin. Supporting this hypothesis, the auxin-resistant tomato mutant, *diageotropica*, has reduced fruit set, fruit weight, and seed production (Balbi and Lomax, 2003) and the application of either auxin or auxin transport inhibitors that cause an increase in auxin in the ovary stimulate fruit set and the development of parthenocarpic fruit (Gustafson, 1937; Beyer and Quebedeaux, 1974). Parthenocarpy has also been induced in tomato, eggplant (*Solanum melongena*), and tobacco by ovary-targeted ectopic expression of *Agrobacterium iaaM* and *rolB* genes, which confer higher auxin production and increased auxin sensitivity, respectively (Ficcadenti et al., 1999; Carmi et al., 2003). Because *AS-IAA9* plants show precocious fruit set and marked parthenocarpy, it appears that in wild-type plants the presence of the *IAA9* protein prevents ovary development prior to pollination, potentially by acting as a negative regulator of auxin response pathways. Based on our data supporting a transcriptional repressor activity for *IAA9*, its downregulation in the antisense lines may release the expression of target auxin-responsive genes, thus mimicking a burst of auxin produced during pollination leading to fruit set and development independent of pollination and fertilization.

In the *pat-2* and *pat-3/pat-4* tomato mutants, it has been postulated that elevated levels of gibberellins (GAs) account for parthenocarpic fruit set and growth (Fos et al., 2000, 2001). In these mutants, fruit set can be inhibited by paclobutrazol, an inhibitor of GA synthesis, and this inhibition can be reversed by exogenous GA application. Because auxin has been shown to promote GA biosynthesis (Ross et al., 2000) and to modulate GA responses (Fu and Harberd, 2003), it is possible that the precocious fruit set and development in *AS-IAA9* lines is mediated by modified GA responses through auxin/GA crosstalk. Treatment of *AS-IAA9* flowers with paclobutrazol or crosses between *AS-IAA9* lines and GA-deficient mutants (*gib-1*, *gib-2*, and *gib-3*) may reveal whether auxin induces parthenocarpy via GA signaling or whether the two hormones induce fruit set and development by separate, parallel pathways.

Once fruit set has occurred, two peaks of auxin production are seen during normal tomato fruit development (Gillaspy et al., 1993), the first at 10 d after anthesis and the second coinciding with the final phase of seed development. Parthenocarpic fruit are generally smaller than seeded fruit, suggesting that seed-derived auxin is required for full pericarp cell expansion (Mapelli et al., 1978). *AS-IAA9* parthenocarpic fruit are of a similar size to seeded wild-type fruit (Figures 5F and 6C) and have the same flesh consistency and fresh weight, suggesting that downregulation of *IAA9* can also compensate for the lack of seed-influenced pericarp cell expansion.

One of the most striking phenotypes exhibited by the *AS-IAA9* plants is the simple leaf architecture, which stands in a marked contrast to the wild-type compound leaf. In some cases, changes in leaf shape are coupled with changes in vascular patterning, suggesting a link between the controls on cell growth, division, and differentiation (Dengler and Kang, 2001). Although, the molecular mechanisms controlling vascular differentiation and leaf morphogenesis during plant organ ontogeny are yet to be fully elucidated, a number of observations have implicated auxin in the formation of vascular tissues in plant organs. These include vascular strand formation in response to local auxin application (Sachs, 2000) and the effects of impaired auxin transport on vascular patterning (Mattsson et al., 1999; Koizumi et al., 2005). In *Arabidopsis* leaf primordia, auxin response patterns presage sites of procambial differentiation, and auxin is instrumental in patterning *Arabidopsis* leaf vasculature (Mattsson et al., 2003). The *Arabidopsis* gain-of-function *aux/iaa* mutant *bd1* and auxin-resistant mutant *axr6* exhibit severely disrupted vascular networks (Hamann et al., 1999; Hobbie et al., 2000), and in the loss-of-function *arf* mutant, *monopteros*, marginal leaf veins are missing or interrupted (Przemeck et al., 1996). The increased vascular network in *AS-IAA9* leaves indicates that downregulation of *IAA9* resulted in more cells undergoing vascular differentiation and that *IAA9* is a key mediator in the auxin-dependent regulation of vascular vein patterning and leaf morphogenesis. Interestingly, treatment of wild-type tomato plants with NPA resulted in a simple leaf shape and enhanced vascular differentiation that mimic the *AS-IAA9* phenotype in an exaggerated manner. It appears, therefore, that in tomato and *Arabidopsis*, disruption of correct pattern of either auxin distribution or response can affect the correct establishment of leaf architecture. Our data bring direct evidence for the role of an Aux/IAA gene in leaf morphogenesis and indicate that

enhanced auxin responsiveness leads to altered vascular network patterning and leaf ontogeny. To investigate the possible involvement of IAA9 in other tomato leaf morphology mutants, we examined transcript accumulation of the gene in *entire* and *Lanceolate*, two natural monogenic mutants with a simple-leaf phenotype, and in the *Petroselinum* mutant that has supercompound leaves. IAA9 transcript accumulation was found to be strongly reduced in the *entire* mutant (Figure 5C) but not in the *Lanceolate* and *Petroselinum* mutants (see Supplemental Figure 1 online). However, in the case of *Lanceolate*, the expression data should be regarded cautiously, as only semidominant mutants are available and the presence of one copy of the wild-type allele may be sufficient to drive a normal expression of the IAA9 gene. Our data suggest either that LANCEOLATE and PETROSELINUM act downstream of IAA9 or that other signaling pathways not involving IAA9 operate in the making of compound leaves. Ectopic expression of the maize homeobox gene *KNOTTED-1* in tomato resulted in a supercompound leaf architecture, with each leaf having up to 2000 leaflets (Hareven et al., 1996). However, in *AS-IAA9* lines, transcript accumulation of two tomato orthologs of *KNOTTED-1* was not affected in the leaves or in other organs (see Supplemental Figure 2 online). These data suggest the coexistence of multiple components controlling leaf ontogeny and vascular vein system patterning. Moreover, these pathways seem not to compensate for one another, as the presence of all these components seem to be required for proper leaf morphogenesis.

The phenotypes of *AS-IAA9* lines confirm the existence of both overlapping and distinct roles for *Aux/IAA* genes in plant developmental processes. The phenotypes associated with gain-of-function mutations in *Aux/IAA* genes have been attributed in most cases to a reduced auxin responsiveness, and in the *Arabidopsis axr2-1* and *shy2-2* mutants, this was clearly correlated with a lower expression level of some auxin-induced genes (Reed, 2001). Most of the gain-of-function mutants, such as *iaa3/shy2-2*, *iaa6/shy1-1*, *iaa7/axr2-1*, and *iaa17/axr3-1*, have short hypocotyls, and *AS-IAA9* seedlings have longer hypocotyls than the wild type, suggesting that these *Aux/IAA* proteins may act as inhibitors of hypocotyl elongation. *Arabidopsis* gain-of-function *aux/iaa* mutants display distinct and specific phenotypes. For example, *iaa12/bdl-1* lacks an embryonic root (Hamann et al., 1999), *iaa14/slr* has no lateral roots (Fukaki et al., 2002), and *iaa19/mgs2* mutants have a phototropic hypocotyls (Tatematsu et al., 2004). This not only reveals different roles for the *Aux/IAA* members, but also emphasizes the high level of complexity permitted by the large number of *Aux/IAA* genes. However, it is yet unclear whether these differences result from cell-specific expression or functional differences among the corresponding *Aux/IAA* proteins. Using an antisense strategy to downregulate other members of the *Aux/IAA* gene family (*IAA1*, *IAA3*, and *IAA8*) in tomato revealed that transgenic lines altered in different gene members share some common phenotypes that related to altered auxin responses, but they also displayed specific, reproducible phenotypes (H. Wang and M. Bouzayen, unpublished data). It is particularly noteworthy that none of these downregulated lines displayed precocious fruit set and parthenocarpy, clearly indicating distinct roles for tomato *Aux/IAAs* during plant growth and development.

Most of the phenotypes exhibited by *AS-IAA9* plants were consistent with the IAA9 protein being a negative regulator of

auxin responses. Expression of the auxin-responsive *IAA3* promoter-driven *GUS* reporter gene in the *AS-IAA9* genetic background clearly indicates that IAA9 can act as a negative regulator of auxin-responsive gene expression. Moreover, the promoter becomes active in tissues and organs where it is normally silent, mimicking the expression pattern obtained upon exogenous auxin treatment.

It is striking that all *Aux/IAA* knockout mutants examined so far have no or subtle phenotypes, while partially silenced *AS-IAA9* lines exhibited clear phenotypes. It is possible that the complete knockout of these genes triggers a compensation mechanism through functionally redundant genes and that this mechanism is not activated when a residual expression of the gene is still present, like in the *AS-IAA9* lines. Strikingly, while a number of *aux/iaa* gain-of-function mutants have been described, our attempts to generate *IAA9*-overexpressing lines under the control of the 35S constitutive promoter were unsuccessful. It is possible that the presence of high levels of IAA9 protein without discrimination between tissue types is incompatible with organ differentiation that is vital to normal plant development. Indeed, in the case of *Arabidopsis aux/iaa* gain-of-function mutants, although the proteins are more stable, the expression of the gene is still under the tight control of the endogenous promoter that continues to finely tune the expression of the mutated gene in a tissue-specific manner. These observations sustain the idea that the expression of *Aux/IAA* genes is under a complex and highly coordinated control mechanism.

## METHODS

### Plant Material and Growth Conditions

Tomato (*Solanum lycopersicum* Mill. cv *MicroTom* and *Ailsa Craig*) plants were grown under standard greenhouse conditions. The conditions for the culture chamber room are as follows: 14-h-day/10-h-night cycle, 25/20°C day/night temperature, 80% humidity, 250  $\mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$  intense luminosity. Seeds were first surface-sterilized in 50% bleach solution for 10 min, rinsed seven to nine times in sterile distilled water, and dried on Whatman paper and then sown in recipient Magenta vessels containing 50 mL of 50% Murashige and Skoog (MS) culture medium added with  $\text{R}_3$  vitamin (0.5 mg  $\text{L}^{-1}$  thiamine, 0.25 mg  $\text{L}^{-1}$  nicotinic acid, and 0.5 mg  $\text{L}^{-1}$  pyridoxine), 1.5% (w/v) sucrose, and 0.8% (w/v) agar, pH 5.9. The tomato mutants, *entire* LA2922, *Lanceolate* LA335, and *Petroselinum* LA2532, were provided by C.M. Rick Tomato Genetics Resource Center (<http://trgc.ucdavis.edu/>).

### Plant Transformation

To generate *AS-IAA9* transgenic plants, forward 5'-TGGCCACCCATTC-GATCTTTTAG-3' and reverse 5'-AGACAACTCCAATATCAAACGG-3' primers encompassing the 3' untranslated region and part of the 3' terminal coding region of the IAA9 cDNA were used to amplify a partial IAA9 clone. This fragment was then cloned into pGA643 binary vector in antisense orientation under the transcriptional control of the cauliflower mosaic virus 35S promoter and the *Nos* terminator. Transgenic plants were generated by *Agrobacterium tumefaciens*-mediated transformation according to Jones et al. (2002), and transformed lines were first selected on kanamycin (70 mg  $\text{L}^{-1}$ ) and then analyzed by both PCR and DNA gel blot analysis to check the presence and the number of T-DNA insertions and to discriminate between different transformation events in the various transgenic lines obtained.

### Phenotypical and Physiological Characterizations of *AS-IAA9*

Phenotypical characterization was performed on homozygote lines. For auxin dose–response experiments, cotyledon explants from wild-type and *AS-IAA9* 9-d-old seedlings were incubated on MS medium containing the indicated auxin (NAA) concentrations in growth chamber conditions described as above for 10 d. For auxin dose–response experiments performed with hypocotyl segments, hypocotyl fragments, 8 mm long, were isolated just below the cotyledons nodes from 5-d-old light-grown seedlings and then immediately floated in sucrose/MES buffer (1% sucrose [w/v] and 5 mM MES/KOH, pH 6.0) in 8-cm Petri dishes. After 1 to 2 h pre-incubation, the hypocotyl segments were randomly distributed to fresh buffer solutions with or without NAA and measured following 23 h of incubation with gentle agitation at room temperature. For NPA treatment, the seeds were sown on MS medium containing 1  $\mu$ M NPA, and the phenotypes affecting root and leaf growth were observed on young 19-d-old seedlings. For auxin treatment, 21-d-old tomato seedlings were harvested and incubated in 50% MS buffer containing 20  $\mu$ M IAA or not (mock treatment) for the time indicated (30 min, 3 h). Thereafter, tissues were immediately frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until RNA extraction. For light microscopy used for determination of venation patterning, plant materials were prepared as follows. The third leaves were taken from 28-d-old light-grown seedlings and fixed overnight in a solution (acetic acid 14%, ethanol 84%). The samples were then dehydrated through a graded ethanol series and cleared in a solution of chloral hydrate (200 g chloral hydrate; 20 g glycerol; 50 g distilled water). Stereomicroscopic photographs were taken under light-field conditions. Mature xylem cells appear dark under light-field optics due to the refraction properties of their thicker secondary cell walls.

### Flower Emasculation and Cross Assay

Flower buds of wild-type or transgenic plants were emasculated before dehiscence of anthers (closed flowers) to avoid accidental self-pollination. Cross-pollination was performed on emasculated flowers one day prior to anthesis. For emasculation and cross-fertilization experiments, 8 to 10 flowers were kept per plant to ensure equivalent growth conditions for all fruit.

### Gene Expression Analysis

Total RNA was extracted according to Jones et al. (2002). The RT-PCR analysis was performed as described previously (Zegzouti et al., 1999), and in each PCR reaction, the internal reference *ubi3* was coamplified with the target gene. Forward (F) and reverse (R) primers used for RT-PCR amplification of the target genes in each RNA sample are the following: for *IAA9* (F 5'-TGGCCACCCATTCGATCTTTAG-3' and R 5'-CGCAACACACATTAGTTGCAG-3'), for *IAA3* (F 5'-AACCAAGACTCAGCTCCTGCACC-3' and R 5'-CATCACAACAAGCATCCAATC-3'), for *IAA8* (F 5'-ATGACTGAGCTAACTCTCGGCTTA-3' and R 5'-ACTCGACGATCCCCCAGGTGTTCT-3'), for *IAA2* (F 5'-AAGCGAGCTATGTTAAAGTGAGCA-3' and R 5'-CCGTTGTATCCATCTGTTTCTGAA-3'), for *IAA6* (F 5'-AGGAGACTGAGCTGAGACTTGGGTT-3' and R 5'-CAACTGAACCTGTTCTCCTTCAT-3'), for *GH3* (F 5'-AGCTCGTCATCAACATACGC-3' and R 5'-CAACTCGCCCTTGTGATAAAC-3'), for *SAUR* (F 5'-GGCTATCCG-TATGCCTCGTA and R 5'-CCACCCATCGGATGATTA-3'), for *DR12* (F 5'-CATGCTGATTTGTTGTACCTTAC-3' and R 5'-GTCTAAAAGAGCACTCCCTC-3'), and for *ubi3* (F 5'-AGAAGAAGACCTACACCAAGCC-3' and R 5'-TCCAAGGTTGTACATACATC-3').

### Protoplast Isolation and Transient Expression of *IAA9-GFP* Fusion Gene

The coding sequence of *IAA9* was cloned as a C-terminal fusion in frame with the GFP into the pGreen vector (Hellens et al., 2000) and expressed

under the control of the 35S promoter. Protoplasts used for transfection were obtained from suspension-cultured tobacco (*Nicotiana tabacum*) BY-2 cells according to the method described previously (Leclercq et al., 2005). Protoplasts were transfected by a modified polyethylene glycol method as described by Abel and Theologis (1994). Typically, 0.2 mL of protoplast suspension ( $0.5 \times 10^6$ ) was transfected with 50  $\mu$ g of shared salmon sperm carrier DNA and 30  $\mu$ g of either *35S:IAA9-GFP* or *35S:GFP* (control) plasmid DNA. Transfected protoplasts were incubated 16 h at  $25^{\circ}\text{C}$  and analyzed for GFP fluorescence by confocal microscopy as by Leclercq et al. (2005). All transient expression assays were repeated at least three times.

### Histochemical GUS Analysis

Transgenic lines bearing the *IAA3* promoter/*GUS* fusion construct (*Pro<sub>IAA3</sub>-GUS*) were incubated at  $37^{\circ}\text{C}$  overnight with GUS staining solution (100 mM sodium phosphate buffer, pH 7.2, 10 mM EDTA, 0.1% Triton, and 1 mM 5-bromo-4-chloro-3-indolyl- $\beta$ -D-glucuronic acid) to reveal GUS activity. Following GUS staining, samples were washed several times to extract chlorophyll using a graded ethanol series.

### Sequence Analysis

Amino acid sequence alignments were performed using ClustalW (<http://bioweb.pasteur.fr/seqanal/interfaces/clustalw-simple.html>) assisted by manual adjustment. Phylogenetic analyses were performed with PAUP (Phylogenetic Analysis Using Parsimony, version 4.0 b10). All heuristic searches for optimal trees were performed by TBR (tree bisection reconnection) branch swapping with option MULPARS in effect. Starting trees were obtained by random addition (10 replicates). For each of the 5000 bootstrap replicates, 10 heuristic searches were performed with random addition of protein samples. The phylogenetic tree was displayed with TreeView (<http://taxonomy.zoology.gla.ac.uk/rod/treeview.html>).

### Accession Numbers

Sequence data from this article can be found in the GenBank/EMBL data libraries under accession number AJ937282. GenBank accession numbers for the sequences analyzed in the phylogenetic analysis are as follows: AtIAA1 (P49677), AtIAA2 (P49678), AtIAA3 (Q38822), AtIAA4 (P33077), AtIAA5 (P33078), AtIAA6 (Q38824), AtIAA7 (Q38825), AtIAA8 (Q38826), AtIAA9 (Q38827), AtIAA10 (Q38828), AtIAA11 (Q38829), AtIAA12 (Q38830), AtIAA13 (Q38831), AtIAA14 (Q38832), AtIAA16 (O24407), AtIAA17 (P93830), AtIAA18 (O24408), AtIAA19 (O24409), AtIAA20 (O24410), AtIAA26 (Q8LAL2), AtIAA27 (Q9ZSY8), AtIAA28 (Q9XFM0), AtIAA29 (Q93WC4), AtIAA30 (Q9M1R4), AtIAA31 (Q8H174), AtIAA32 (Q8RYC6), AtIAA33 (Q9FKM7), AtIAA34 (Q9C5X0), ZeARP (AAM12952), NtIAA9 (CAD10639), VvIAA (AAL92850), StIAA (AAM29182), CsIAA2 (BAA85821), OsIAAa (BAD81331.1), OsIAAb (BAD67992.1), OsIAAc (BAD22024.1), OsIAA1 (CAC80823.1), OsIAA2 (AAK98708.1), OsIAA8 (CAF28457.1), and OsIAA18 (BAA99424.1).

### ACKNOWLEDGMENTS

This work was supported by the Midi-Pyrénées Regional Council (Grants 01008920 and 03001146) and forms part of the requirement for the degree of PhD for H.W. We thank Alain Jauneau for technical support in confocal microscopy analysis, Luc Legal for his helpful advice in performing phylogenetic analysis, and Lydie Tessarotto, Héliène Mondières, and Dominique Saint-Martin for tomato genetic transformation and plant growth. H.W. was supported by a scholarship from the Association Franco-Chinoise pour la Recherche Scientifique et Technique and Z.L. by a scholarship from the Institut National de la Recherche Agronomique.



Received April 12, 2005; revised June 13, 2005; accepted August 3, 2005; published August 26, 2005.

## REFERENCES

- Abel, S., Nguyen, M.D., and Theologis, A. (1995). The *PS-IAA4/5*-like family of early auxin-inducible mRNAs in *Arabidopsis thaliana*. *J. Mol. Biol.* **251**, 533–549.
- Abel, S., and Theologis, A. (1994). Transient transformation of *Arabidopsis* leaf protoplasts: A versatile experimental system to study gene expression. *Plant J.* **5**, 421–427.
- Ainley, W.M., Walker, J.C., Nagao, R.T., and Key, J.L. (1988). Sequence and characterization of two auxin-regulated genes from soybean. *J. Biol. Chem.* **263**, 10658–10666.
- Balbi, V., and Lomax, T.L. (2003). Regulation of early tomato fruit development by the *diageotropica* gene. *Plant Physiol.* **131**, 186–197.
- Beyer, E.M., and Quebedeaux, B. (1974). Parthenocarpy in cucumber: Mechanism of action of auxin transport inhibitors. *J. Am. Soc. Hortic. Sci.* **99**, 385–390.
- Carmi, N., Salts, Y., Dedicova, B., Shabtai, S., and Barg, R. (2003). Induction of parthenocarpy in tomato via specific expression of the *rolB* gene in the ovary. *Planta* **217**, 726–735.
- Dargeviciute, A., Roux, C., Decreux, A., Sitbon, F., and Perrot-Rechenmann, C. (1998). Molecular cloning and expression of the early auxin-responsive *Aux/IAA* gene family in *Nicotiana tabacum*. *Plant Cell Physiol.* **39**, 993–1002.
- Dengler, N., and Kang, J. (2001). Vascular patterning and leaf shape. *Curr. Opin. Plant Biol.* **4**, 50–56.
- Dharmasiri, N., Dharmasiri, S., and Estelle, M. (2005). The F-box protein TIR1 is an auxin receptor. *Nature* **435**, 441–445.
- Ficcadenti, N., Sestili, S., Pandolfini, T., Cirillo, C., Rotino, G., and Spena, A. (1999). Genetic engineering of parthenocarpic fruit development in tomato. *Mol. Breed.* **5**, 463–470.
- Fos, M., Nuez, F., and Garcia-Martinez, J.L. (2000). The gene *pat-2*, which induces natural parthenocarpy, alters the gibberellin content in unpollinated tomato ovaries. *Plant Physiol.* **122**, 471–480.
- Fos, M., Proano, K., Nuez, F., and Garcia-Martinez, J.L. (2001). Role of gibberellins in parthenocarpic fruit development induced by the genetic system *pat-3/pat-4* in tomato. *Physiol. Plant* **111**, 545–550.
- Friml, J. (2003). Auxin transport—Shaping the plant. *Curr. Opin. Plant Biol.* **6**, 7–12.
- Fu, X., and Harberd, N.P. (2003). Auxin promotes *Arabidopsis* root growth by modulating gibberellin response. *Nature* **421**, 740–743.
- Fukaki, H., Tameda, S., Masuda, H., and Tasaka, M. (2002). Lateral root formation is blocked by a gain-of-function mutation in the *SOLITARY-ROOT/IAA14* gene of *Arabidopsis*. *Plant J.* **29**, 153–168.
- Galweiler, L., Guan, C., Muller, A., Wisman, E., Mendgen, K., Yephremov, A., and Palme, K. (1998). Regulation of polar auxin transport by AtPIN1 in *Arabidopsis* vascular tissue. *Science* **282**, 2226–2230.
- Gillaspy, G., Ben-David, H., and Grissem, W. (1993). Fruit: A developmental perspective. *Plant Cell* **5**, 1439–1451.
- Gustafson, F.G. (1937). Parthenocarpy induced by pollen extracts. *Am. J. Bot.* **24**, 102–107.
- Hamann, T., Benkova, E., Baurle, I., Kientz, M., and Jurgens, G. (2002). The *Arabidopsis* *BODENLOS* gene encodes an auxin response protein inhibiting *MONOPTEROS*-mediated embryo patterning. *Genes Dev.* **16**, 1610–1615.
- Hamann, T., Mayer, U., and Jurgens, G. (1999). The auxin-insensitive *bodenlos* mutation affects primary root formation and apical-basal patterning in the *Arabidopsis* embryo. *Development* **126**, 1387–1395.
- Hareven, D., Gutfinger, T., Parnis, A., Eshed, Y., and Lifschitz, E. (1996). The making of a compound leaf: Genetic manipulation of leaf architecture in tomato. *Cell* **84**, 735–744.
- Hellens, R.P., Edwards, A.E., Leyland, N.R., Bean, S., and Mullineaux, P. (2000). pGreen: A versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation. *Plant Mol. Biol.* **42**, 819–832.
- Hobbie, L., McGovern, M., Hurwitz, L.R., Pierro, A., Liu, N.Y., Bandyopadhyay, A., and Estelle, M. (2000). The *axr6* mutants of *Arabidopsis thaliana* define a gene involved in auxin response and early development. *Development* **127**, 23–32.
- Jones, B., Frasse, P., Olmos, E., Zegzouti, H., Li, Z.G., Latche, A., Pech, J.C., and Bouzayen, M. (2002). Down-regulation of DR12, an auxin-response-factor homolog, in the tomato results in a pleiotropic phenotype including dark green and blotchy ripening fruit. *Plant J.* **32**, 603–613.
- Kepinski, S., and Leyser, O. (2005). The *Arabidopsis* F-box protein TIR1 is an auxin receptor. *Nature* **435**, 446–451.
- Kessler, S., and Sinha, N. (2004). Shaping up: The genetic control of leaf shape. *Curr. Opin. Plant Biol.* **7**, 65–72.
- Kim, B.C., Soh, M.S., Kang, B.G., Furuya, M., and Nam, H.G. (1996). Two dominant photomorphogenic mutations of *Arabidopsis thaliana* identified as suppressor mutations of *hy2*. *Plant J.* **15**, 441–456.
- Kim, J., Harter, K., and Theologis, A. (1997). Protein-protein interactions among the Aux/IAA proteins. *Proc. Natl. Acad. Sci. USA* **94**, 11786–11791.
- Koizumi, K., Naramoto, S., Sawa, S., Yahara, N., Ueda, T., Nakano, A., Sugiyama, M., and Fukuda, H. (2005). VAN3 ARF-GAP-mediated vesicle transport is involved in leaf vascular network formation. *Development* **132**, 1699–1711.
- Leclercq, J., Ranty, B., Sanchez-Ballesta, M.T., Li, Z., Jones, B., Jauneau, A., Pech, J.C., Latche, A., Ranjeva, R., and Bouzayen, M. (2005). Molecular and biochemical characterization of LeCRK1, a ripening-associated tomato CDPK-related kinase. *J. Exp. Bot.* **56**, 25–35.
- Liscum, E., and Reed, J.W. (2002). Genetics of Aux/IAA and ARF action in plant growth and development. *Plant Mol. Biol.* **49**, 387–400.
- Mapelli, S.C., Frova, G., Torti, G., and Soressi, G. (1978). Relationship between set, development and activities of growth regulators in tomato fruits. *Plant Cell Physiol.* **19**, 1281–1288.
- Mattsson, J., Ckurshumova, W., and Berleth, T. (2003). Auxin signaling in *Arabidopsis* leaf vascular development. *Plant Physiol.* **131**, 1327–1339.
- Mattsson, J., Sung, Z.R., and Berleth, T. (1999). Responses of plant vascular systems to auxin transport inhibition. *Development* **126**, 2979–2991.
- Nagpal, P., Walker, L.M., Young, J.C., Sonawala, A., Timpte, C., Estelle, M., and Reed, J.W. (2000). *AXR2* encodes a member of the Aux/IAA protein family. *Plant Physiol.* **123**, 563–574.
- Nebenfuhr, A., White, T.J., and Lomax, T.L. (2000). The *diageotropica* mutation alters auxin induction of a subset of the *Aux/IAA* gene family in tomato. *Plant Mol. Biol.* **44**, 73–84.
- Ouellet, F., Overvoorde, P.J., and Theologis, A. (2001). IAA17/AXR3: Biochemical insight into an auxin mutant phenotype. *Plant Cell* **13**, 829–841.
- Park, J.Y., Kim, H.J., and Kim, J. (2002). Mutation in domain II of IAA1 confers diverse auxin-related phenotypes and represses auxin-activated expression of *Aux/IAA* genes in steroid regulator-inducible system. *Plant J.* **32**, 669–683.
- Przemeck, G.K., Mattsson, J., Hardtke, C.S., Sung, Z.R., and Berleth, T. (1996). Studies on the role of the *Arabidopsis* gene *MONOPTEROS* in vascular development and plant cell axialization. *Planta* **200**, 229–237.

- Reed, J.W.** (2001). Roles and activities of Aux/IAA proteins in *Arabidopsis*. *Trends Plant Sci.* **6**, 420–425.
- Rogg, L.E., Lasswell, J., and Bartel, B.** (2001). A gain-of-function mutation in *IAA28* suppresses lateral root development. *Plant Cell* **13**, 465–480.
- Ross, J.J., O'Neill, D.P., Smith, J.J., Kerckhoffs, L.H., and Elliott, R.C.** (2000). Evidence that auxin promotes gibberellin A1 biosynthesis in pea. *Plant J.* **21**, 547–552.
- Rouse, D., Mackay, P., Stirnberg, P., Estelle, M., and Leyser, O.** (1998). Changes in auxin response from mutations in an *AUX/IAA* gene. *Science* **279**, 1371–1373.
- Sachs, T.** (2000). Integrating cellular and organismic aspects of vascular differentiation. *Plant Cell Physiol.* **41**, 649–656.
- Servant, F., Bru, C., Carrere, S., Courcelle, E., Gouzy, J., Peyruc, D., and Kahn, D.** (2002). ProDom: Automated clustering of homologous domains. *Brief. Bioinform.* **3**, 246–251.
- Tatematsu, K., Kumagai, S., Muto, H., Sato, A., Watahiki, M.K., Harper, R.M., Liscum, E., and Yamamoto, K.T.** (2004). *MASSUGU2* encodes Aux/IAA19, an auxin-regulated protein that functions together with the transcriptional activator NPH4/ARF7 to regulate differential growth responses of hypocotyl and formation of lateral roots in *Arabidopsis thaliana*. *Plant Cell* **16**, 379–393.
- Tian, Q., and Reed, J.W.** (1999). Control of auxin-regulated root development by the *Arabidopsis thaliana* *SHY2/IAA3* gene. *Development* **126**, 711–721.
- Timpte, C., Wilson, A.K., and Estelle, M.** (1994). The *axr2-1* mutation of *Arabidopsis thaliana* is a gain-of-function mutation that disrupts an early step in auxin response. *Genetics* **138**, 1239–1249.
- Tiwari, S.B., Hagen, G., and Guilfoyle, T.J.** (2004). Aux/IAA proteins contain a potent transcriptional repression domain. *Plant Cell* **16**, 533–543.
- Ulmasov, T., Hagen, G., and Guilfoyle, T.J.** (1997). ARF1, a transcription factor that binds to auxin response elements. *Science* **276**, 1865–1868.
- Vogler, H., and Kuhlemeier, C.** (2003). Simple hormones but complex signalling. *Curr. Opin. Plant Biol.* **6**, 51–56.
- Yang, X., Lee, S., So, J.H., Dharmasiri, S., Dharmasiri, N., Ge, L., Jensen, C., Hangarter, R., Hobbie, L., and Estelle, M.** (2004). The IAA1 protein is encoded by *AXR5* and is a substrate of SC<sup>F<sup>TIR1</sup></sup>. *Plant J.* **40**, 772–782.
- Zegzouti, H., Jones, B., Frasse, P., Marty, C., Maitre, B., Latch, A., Pech, J.C., and Bouzayen, M.** (1999). Ethylene-regulated gene expression in tomato fruit: Characterization of novel ethylene-responsive and ripening-related genes isolated by differential display. *Plant J.* **18**, 589–600.



**The Tomato Aux/IAA Transcription Factor IAA9 Is Involved in Fruit Development and Leaf Morphogenesis**

Hua Wang, Brian Jones, Zhengguo Li, Pierre Frasse, Corinne Delalande, Farid Regad, Salma Chaabouni, Alain Latché, Jean-Claude Pech and Mondher Bouzayen

*PLANT CELL* 2005;17;2676-2692; originally published online Aug 26, 2005;

DOI: 10.1105/tpc.105.033415

This information is current as of December 10, 2008

<b>Supplemental Data</b>	<a href="http://www.plantcell.org/cgi/content/full/tpc.105.033415/DC1">http://www.plantcell.org/cgi/content/full/tpc.105.033415/DC1</a>
<b>References</b>	This article cites 51 articles, 25 of which you can access for free at: <a href="http://www.plantcell.org/cgi/content/full/17/10/2676#BIBL">http://www.plantcell.org/cgi/content/full/17/10/2676#BIBL</a>
<b>Permissions</b>	<a href="https://www.copyright.com/ccc/openurl.do?sid=pd_hw1532298X&amp;issn=1532298X&amp;WT.mc_id=pd_hw1532298X">https://www.copyright.com/ccc/openurl.do?sid=pd_hw1532298X&amp;issn=1532298X&amp;WT.mc_id=pd_hw1532298X</a>
<b>eTOCs</b>	Sign up for eTOCs for <i>THE PLANT CELL</i> at: <a href="http://www.plantcell.org/subscriptions/etoc.shtml">http://www.plantcell.org/subscriptions/etoc.shtml</a>
<b>CiteTrack Alerts</b>	Sign up for CiteTrack Alerts for <i>Plant Cell</i> at: <a href="http://www.plantcell.org/cgi/alerts/ctmain">http://www.plantcell.org/cgi/alerts/ctmain</a>
<b>Subscription Information</b>	Subscription information for <i>The Plant Cell</i> and <i>Plant Physiology</i> is available at: <a href="http://www.aspb.org/publications/subscriptions.cfm">http://www.aspb.org/publications/subscriptions.cfm</a>

## II. Premier post-doctorat (2004-2008) : Étude de génomique fonctionnelle du développement du fruit de tomate par des approches combinées de transcriptomique et de métabolomique.

### Résumé de Postdoc I

À l'issue de ma thèse, grâce à l'obtention d'un financement d'un projet européen trilatéral et d'un autre projet européen FP6 Eu-SOL, j'ai pu poursuivre mes travaux de recherche sur **un nouvel axe du laboratoire** : « le développement **précoce des fruits** », sous la direction du Professeur Mondher Bouzayen. Mes travaux consistaient à évaluer le rôle de différentes hormones dans la fructification, la croissance et la qualité finale du fruit. La stratégie consisterait à étudier l'effet soit de la modification de l'état hormonal par traitement exogène, soit par transgénèse affectant la signalisation hormonale. Cette étude visait à mettre en évidence l'influence des phytohormones sur la fructification, le développement et la maturation des fruits. La démarche suivie combinait des approches de génomiques fonctionnelles, de biologie moléculaire, de physiologie et de biochimie. Ces travaux ont conduit à ma **deuxième publication** « **The Plant Cell** » (Wang et al., 2009). Grâce à la réalisation de ces projets, j'ai ainsi développé des outils génétiques et des méthodes qui continuent d'être exploités.

### Travaux de Postdoc I

#### *1. Analyse cytologique de l'expression de IAA9 au cours de la fructification*

J'ai entrepris d'analyser l'évolution de l'expression du gène IAA9 au cours des phases précoces du développement du fruit (stades bouton floral à 5 jours après anthèse) par la technique de l'hybridation *in situ*. Les résultats indiquent que le gène IAA9 présente une expression dynamique au cours de la fructification. En effet, pendant les phases qui précèdent l'anthèse, IAA9 s'exprime fortement et spécifiquement dans certains tissus tels que les ovules, le placenta et le style mais suite à la pollinisation son expression baisse fortement dans ces tissus et apparaît parallèlement à un faible niveau partout dans le jeune fruit. Par contre, dans les fleurs émasculées, le profil d'expression d'IAA9 reste inchangé par rapport au stade avant anthèse. Dans les fruits induits par les hormones, j'ai pu mettre en évidence une forte induction du gène IAA9 par le traitement aux gibbérellines dans les jeunes fruits, surtout dans le style, le stigmate et l'ovule. Le traitement NPA (inhibiteur du transport de l'auxine) entraîne aussi l'induction du gène IAA9 mais beaucoup moins fortement que celle obtenue par les gibbérellines et avec un profil spatial différent puisque l'expression se concentre majoritairement dans le placenta et le péricarpe.

#### *2. Etude des facteurs hormonaux impliqués dans le développement du fruit de tomate et leur incidence sur la qualité du fruit*

Diverses hormones ont été appliquées sur les ovaires de tomate après émasculation. La fructification et les effets sur les composantes de la qualité des fruits sont mesurés à différents stades de développement du fruit. Cette étude démontre que l'acide gibbérellique, l'AIA et le NPA induisent la fructification en absence de pollinisation. La caractérisation physicochimique montre que les fruits

obtenus sont différents de ceux issus de pollinisation naturelle en termes de teneur en sucre, d'acidité, de taille, de forme, de durée de développement et de maturation, et d'organisation des tissus intérieurs. Les études de microscopie indiquent que le profil d'ontogenèse des fruits sont spécifiques de chaque traitement hormonal. L'étude comprend également l'analyse par PCR en temps réel de l'expression des gènes, telle que la famille des Aux/IAA, qui sont associées aussi à la réponse à l'auxine afin d'évaluer leur rôle dans le processus de fructification des jeunes ovaires. Le profil transcriptomique est réalisé par puce à ADN pour chaque traitement hormonal et comparé à ceux obtenus pour les fruits naturellement autofécondés. En collaboration avec l'équipe de « Métabolite Profiling » (Max Planck Institute, Golm, Allemagne) une analyse des variations métaboliques est réalisée qui vise à révéler des corrélations entre certaines voies métaboliques et la régulation hormonale. Elle permet également d'établir des corrélations entre des réseaux métaboliques et des réseaux de gènes. Ces travaux sont réalisés dans le cadre d'un programme Trilateral associant des équipes françaises (INRA, Toulouse, Avignon et Bordeaux), espagnoles (CSIC Valencia et Malaga) et allemandes (Max Planck, Golm).

### *3. Identification des facteurs clés de la fructification indépendante de la pollinisation chez les plantes AS-IAA9*

J'ai entrepris une étude comparative de l'évolution de l'expression génique globale chez les *DR4AS* et sauvages à trois stades différents du développement précoce du fruit : bouton floral, anthèse et post-anthèse (4 jours après anthèse). L'utilisation des puces à oligo de tomate incluant 12000 unigènes produits au laboratoire a permis d'identifier plusieurs groupes de gènes mobilisés différemment au cours des processus de fructification dépendant et indépendant de la pollinisation. La validation des profils d'expression spécifique des gènes les plus intéressants est actuellement réalisée par PCR quantitative. Une étude similaire est également poursuivie pour comparer l'évolution de l'expression génique globale des deux types de fruits au cours des différents stades de maturation.

### *4. La régulation transcriptionnelle et post-traductionnelle de l'expression du gène IAA9*

Il est bien démontré que les membres de la famille des Aux/IAA à laquelle appartient le gène IAA9 sont régulés au niveau post-traductionnel. On sait depuis très récemment que cette régulation qui conduit à la dégradation de protéines Aux/IAA est à la base du mode d'action de l'auxine. Les travaux que j'ai réalisés dans la période récente montrent que le gène IAA9 répond à une régulation transcriptionnelle. Dans le but de comprendre la part de régulation post-transcriptionnelle de ce gène, j'ai entrepris maintenant de réaliser deux types de constructions avec le gène rapporteur GUS fusionné soit au promoteur IAA9 seul soit au promoteur IAA9 incluant la partie codante du gène qui correspond à la région N-terminale responsable de la dégradation de la protéine par le système d'ubiquitination. L'expression de ces deux constructions de façon stable dans des lignées de tomate transgéniques permettra d'évaluer les situations (ou bien les tissus) où les deux niveaux de régulation interviennent simultanément et celles où seule la régulation post-traductionnelle contrôle l'expression du gène IAA9. Ces constructions permettront également de connaître les signaux hormonaux et les stimuli qui induisent chacun des mécanismes de régulations envisagés.

Annexe 2 Article de Postdoc I



# Regulatory Features Underlying Pollination-Dependent and -Independent Tomato Fruit Set Revealed by Transcript and Primary Metabolite Profiling <sup>W</sup>

Hua Wang,<sup>a,b</sup> Nicolas Schauer,<sup>c</sup> Bjoern Usadel,<sup>c</sup> Pierre Frasse,<sup>a,b</sup> Mohamed Zouine,<sup>a,b</sup> Michel Hernould,<sup>d</sup> Alain Latché,<sup>a,b</sup> Jean-Claude Pech,<sup>a,b</sup> Alisdair R. Fernie,<sup>c</sup> and Mondher Bouzayan<sup>a,b,1</sup>

<sup>a</sup> Université de Toulouse, Institut National Polytechnique - Ecole Nationale Supérieure Agronomique Toulouse, Génomique et Biotechnologie des Fruits, Castanet-Tolosan F-31326, France

<sup>b</sup> Institut National de la Recherche Agronomique, Génomique et Biotechnologie des Fruits, Chemin de Borde Rouge, Castanet-Tolosan, F-31326, France

<sup>c</sup> Max-Planck-Institut für Molekulare Pflanzenphysiologie, 14476 Potsdam-Golm, Germany

<sup>d</sup> Unité Mixte de Recherche 619 Biologie du Fruit, Institut National de la Recherche Agronomique, Université de Bordeaux, Institut de Biologie Végétale Moléculaire, 33883 Villenave d'Ornon, France

**Indole Acetic Acid 9 (IAA9) is a negative auxin response regulator belonging to the *Aux/IAA* transcription factor gene family whose downregulation triggers fruit set before pollination, thus giving rise to parthenocarpy. In situ hybridization experiments revealed that a tissue-specific gradient of *IAA9* expression is established during flower development, the release of which upon pollination triggers the initiation of fruit development. Comparative transcriptome and targeted metabolome analysis uncovered important features of the molecular events underlying pollination-induced and pollination-independent fruit set. Comprehensive transcriptomic profiling identified a high number of genes common to both types of fruit set, among which only a small subset are dependent on *IAA9* regulation. The fine-tuning of *Aux/IAA* and *ARF* genes and the downregulation of *TAG1* and *TAGL6* MADS box genes are instrumental in triggering the fruit set program. Auxin and ethylene emerged as the most active signaling hormones involved in the flower-to-fruit transition. However, while these hormones affected only a small number of transcriptional events, dramatic shifts were observed at the metabolic and developmental levels. The activation of photosynthesis and sucrose metabolism-related genes is an integral regulatory component of fruit set process. The combined results allow a far greater comprehension of the regulatory and metabolic events controlling early fruit development both in the presence and absence of pollination/fertilization.**

## INTRODUCTION

Fruit development is a genetically programmed process, unique to flowering plants, which provides a suitable environment for seed maturation and seed dispersal. Given the fundamental nature of both the dietary and biological significance of fruit, it is unsurprising that the molecular dissection of fruit development has recently received considerable interest (Giovannoni, 2001; Pandolfini et al., 2007; Serrani et al., 2008). The fruit is the result of the development of the ovary, and fruit organogenesis originates from a flower primordium. The mature flower can either be fertilized and develop into a fruit or, in the absence of successful pollination, can enter the abscission process. The onset of the development of an ovary into fruit, the so-called fruit set, thus constitutes an essential mechanism for fruit production. Despite the importance of the fruit set process, most research effort to

date has been dedicated to the processes of fruit maturation (Alba et al., 2005) and fruit growth (Carrari et al., 2006), while the molecular mechanisms underlying the onset of ovary development itself remain relatively poorly defined.

In normal fruit development, the initiation of fruit set depends on the successful completion of pollination and fertilization; this suggests that both pollination and seed-derived signals are required for fruit initiation and subsequent development. Fruit growth and shape are known to be modified by differences in seed genotype and seed number (Sedgley and Griffin, 1989). However, the signals that trigger fruit growth after fertilization remain unknown. Unfavorable conditions, such as extreme temperatures, may prevent pollination and hence also fruit set. Mild temperature stress, leading to loss of pollen viability, can result in the production of underfertilized puffy fruits of poor quality, while severe temperature stress can completely abolish fruit set. In tomato (*Solanum lycopersicum*) and many other species, a major limiting factor for fruit set is the extreme sensitivity of microsporogenesis and pollination to moderate fluctuations in temperature and humidity (Picken, 1984). Hence, parthenocarpy, the growth of the ovary into a seedless fruit in the absence of pollination and fertilization, has long been recognized as an important trait that circumvents the problems of low fruit set in

<sup>1</sup> Address correspondence to bouzayan@ensat.fr.

The author responsible for distribution of materials integral to the finding presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Mondher Bouzayan (bouzayan@ensat.fr).

<sup>W</sup> Online version contains Web-only data.

www.plantcell.org/cgi/doi/10.1105/tpc.108.060830

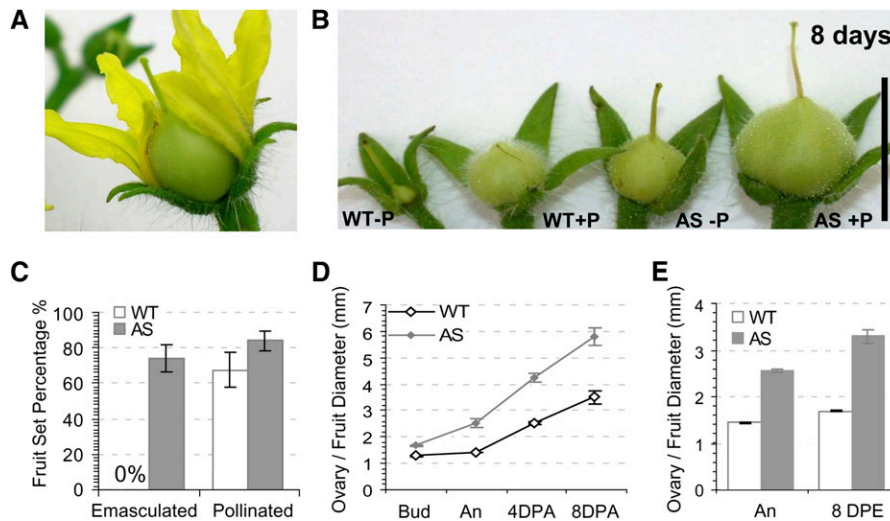
harsh environmental conditions (Gorguet et al., 2005). Fertilization-independent fruit set can occur either naturally (genetic parthenocarpy) or by induction via exogenous application of phytohormones, such as auxin and gibberellins (GAs), to the flower (Gustafson, 1936; Coombe, 1960). The presence of both germinated pollen and developing seeds appears to be essential for fruit growth and development, since they probably serve as sources of phytohormones that are, most likely, continuously required throughout seed and fruit formation (Nitsch, 1970; Talon et al., 1992; Ozga et al., 2002). In keeping with this hypothesis, elevated endogenous phytohormone levels have been observed during parthenocarpic fruit set (George et al., 1984; Talon et al., 1992). Accordingly, parthenocarpy can be induced in diverse agricultural species not only by the exogenous application of auxins, cytokinins, or GAs (Gillaspy et al., 1993; Vivian-Smith and Koltunow, 1999; Serrani et al., 2007) but also by increasing either auxin levels or auxin response in ovaries and ovules (Carmi et al., 2003; Rotino et al., 2005) or by increasing the GA response (Potts et al., 1985). Recent molecular analyses of fertilization-independent fruit formation in both tomato and *Arabidopsis thaliana* have, however, identified auxin signaling as one of the early events in the fruit initiation cascade. Furthermore, components of the auxin signaling pathway are also involved in repressing fruit initiation until the fertilization cue (Vivian-Smith et al., 2001; Wang et al., 2005; Goetz et al., 2006). When considered together, these data demonstrate that hormones, such as auxin, play an important role in fruit initiation. However, the exact mechanism by which auxin promotes fruit initiation and the nature of the genes that control fruit set remain an open question.

We previously reported that downregulation of *IAA9*, a member of the *Aux/IAA* transcription factor gene family encoding a negative auxin response regulator, in tomato resulted in pollination-independent fruit set giving rise to parthenocarpy (Wang et al., 2005). In this study, taking advantage of this unique biological tool, we sought to uncover the molecular events underlying the process of fruit set via a combined transcriptomic and metabolomic approach. Our study reveals that *IAA9* closely regulates the initiation of fruit set by establishing a spatial expression gradient whose release triggers the flower-to-fruit transition. The comparative analysis at the transcriptomic and metabolic levels of pollination-induced natural fruit set and fertilization-independent fruit set identifies auxin and ethylene signaling, as well as photosynthesis and sugar metabolism, as major events of the fruit set program and potential components of the regulatory mechanism underlying this developmental process. These data are discussed in comparison to current models of fruit set as well as within the context of the development of future strategies for the biotechnological exploitation of parthenocarpy.

## RESULTS

### Fertilization-Independent Fruit Set in *AS-IAA9*

MicroTom tomato lines downregulated in the expression of the *IAA9* gene (*AS-IAA9*) exhibit precocious fruit set prior to anthesis, resulting in parallel fruit and flower development (Figure 1A) and uncoupling fruit set from pollination, leading to parthenocarpy



**Figure 1.** Comparison of Wild-Type and *AS-IAA9* Ovary/Fruit Development during Fruit Set.

(A) *AS-IAA9* lines exhibit precocious fruit set prior to anthesis, resulting in abnormal parallel development of fruit and flower at anthesis stage.

(B) Impact of pollination on fruit size at 8 DPA. +P, pollinated; -P, nonpollinated. Bar = 8 mm.

(C) Percentage of fruit set (the transition from flower to fruit and subsequent fruit development) from emasculated and pollinated flowers in wild-type and *AS-IAA9* lines.

(D) Ovary/fruit diameter in wild-type and *AS-IAA9* lines at anthesis (An) and 8 d postemasculation (DPE).

(E) Accelerated ovary-fruit enlargement in *AS-IAA9* (AS) compared with the wild type.

In (C) to (E), error bars represent  $\pm$  SE of three independent trials, for each trial  $n \geq 20$ . AS, *IAA9*-antisense lines; Bud, flower bud; An, anthesis.

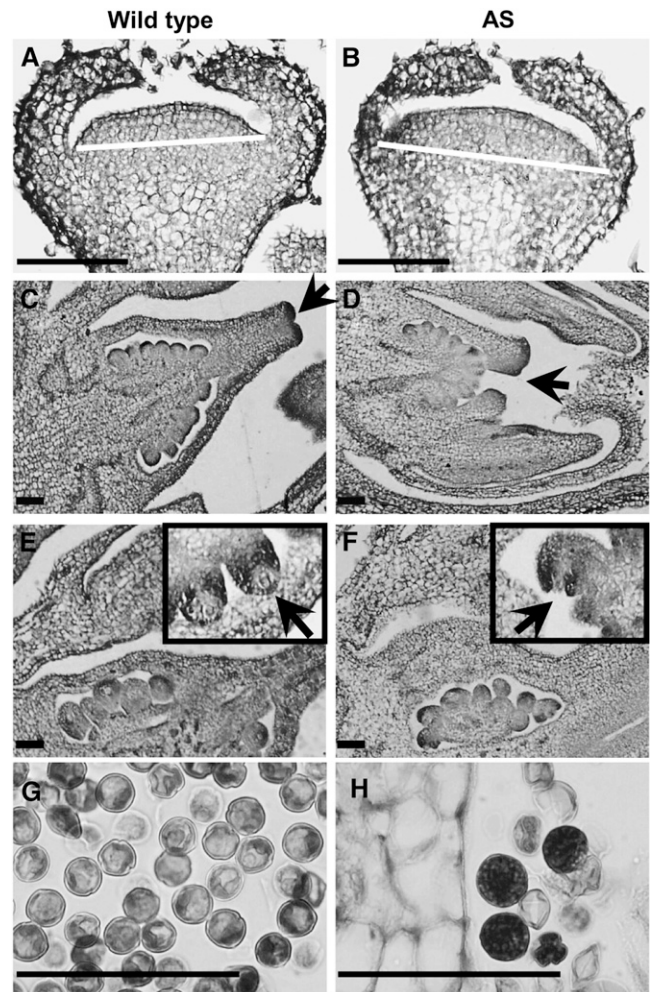
(Wang et al., 2005). The rate of successful fruit set of emasculated *AS-IAA9* lines (75%) was similar to that observed in the wild type (68%) under natural pollination conditions (Figure 1C). At 8 d postanthesis (DPA), unpollinated *AS-IAA9* young fruits and naturally pollinated wild type had similar size (Figure 1B). Manual pollination stimulated the development of *AS-IAA9* fruit, leading to a fruit size at 8 DPA that were on average 76% bigger than *AS-IAA9* emasculated fruit or wild-type naturally pollinated ones (Figures 1D and 1E). Moreover, Figure 1D indicates that from bud stage onward, *AS-IAA9* displayed accelerated ovary fruit enlargement compared with the wild type. Taken together, these data indicated that pollination of *AS-IAA9* has no significant impact on the rate of fruit set, whereas it substantially promotes fruit enlargement.

### Downregulation of *IAA9* Affects Flower/Ovary Development

Given that fruit organogenesis is initiated from the floral primordia, light microscopy analysis was performed to investigate early fruit organogenesis in the wild type and *AS-IAA9*. Figures 2A and 2B reveal that the floral meristem is larger in *AS-IAA9* compared with the wild type, and comparison of 2-mm-long flower buds (Figures 2C and 2D) revealed that the wild type exhibits fused carpels while, at the same size, carpels are still growing toward fusion in *AS-IAA9*. When flower buds are 4 mm long, the nucellus in the wild type is completely embedded in the integument, while, in *AS-IAA9*, the integument is still in the process of surrounding the nucellus (Figures 2E and 2F). These observations indicate that the *AS-IAA9* flower organ corresponds to a more juvenile stage than the wild type and that organ differentiation is not more advanced in the *IAA9*-downregulated lines but rather that the organ exhibits larger growth. Mature *AS-IAA9* flowers (anthesis stage) contained a higher proportion of aborted pollen grains (Figure 2H) than the wild type (Figure 2G), while those not aborted were significantly larger and more intensely colored than in the wild type. However, the latter observation is largely in keeping with our previous observation that *AS-IAA9* pollen retains the capacity to fertilize emasculated wild-type flowers (Wang et al., 2005).

### Spatial and Temporal Regulation of *IAA9* Expression during Natural Fruit Set

In situ hybridization during wild-type flower/ovary development and fruit set revealed that *IAA9* mRNA distribution was detectable in the whole floral meristem but was more abundant in emerging organs, such as stamen and carpel (Figure 3B). In the emerging petals, the signal intensity was higher on the adaxial sides. A control hybridization using the sense probe showed no signal (Figure 3A). Though unevenly distributed, the *IAA9* mRNA signal was detected across all flower tissues but exhibited a steady increase throughout flower development peaking at anthesis (Figures 3C to 3G). At these stages, accumulation of the *IAA9* transcript formed a gradient wherein the signal is strongest in ovule, sporogenous tissue, tapetum, petals, vascular bundles, developing style, placenta, and funiculus, but, by contrast, the signal was weak in sepals, ovary wall, and the



**Figure 2.** Histological Analysis of Organ Differentiation Program in Wild-Type and *AS-IAA9* Flowers.

(A) and (B) Floral meristem is larger in *AS-IAA9* line (AS) than in the wild type as indicated by white bars.

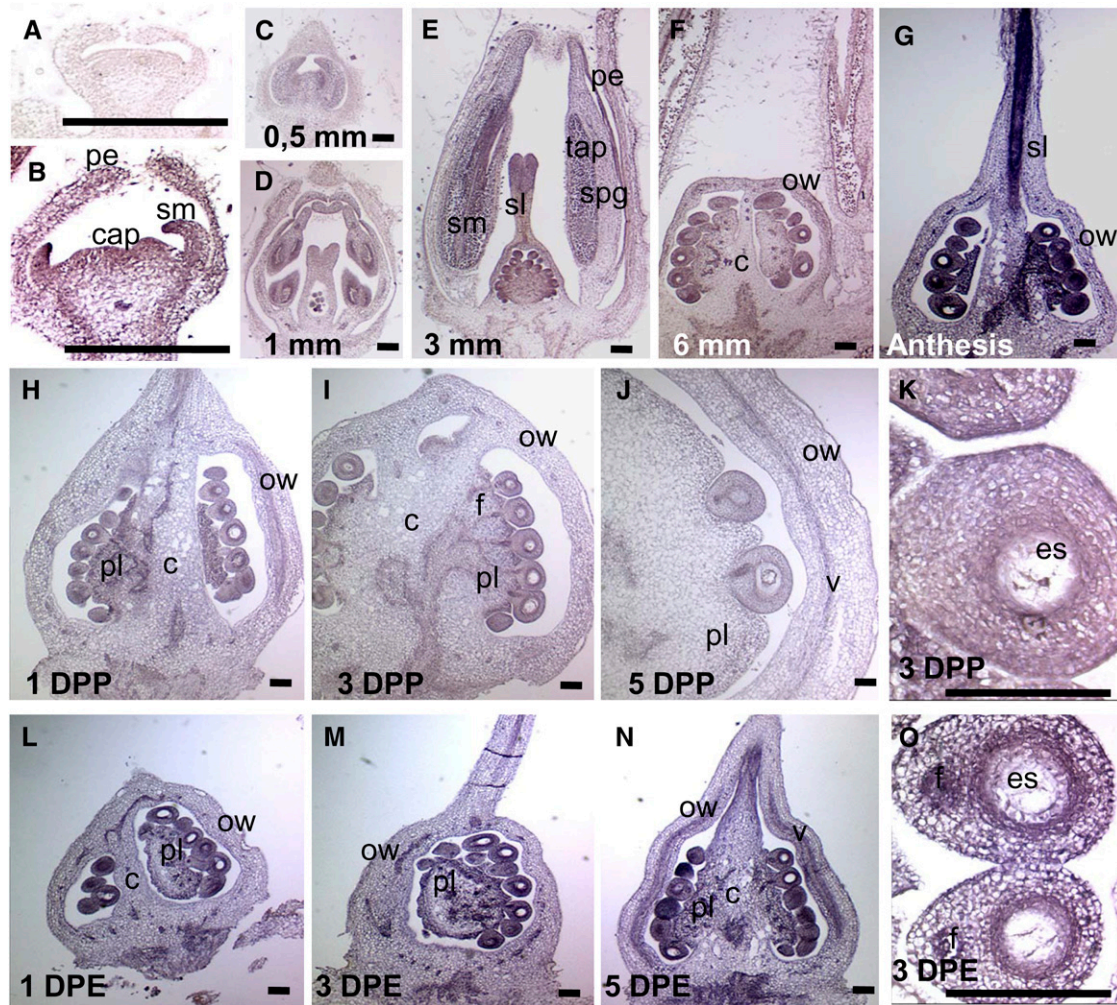
(C) and (D) In 2-mm-long flower buds, wild-type lines exhibit fully fused carpels, whereas at the same size, carpels are still growing toward fusion (arrows) in *AS-IAA9*.

(E) and (F) In 4-mm-long flower buds, the nucellus is already completely enveloped by the integuments in the wild type, whereas it is still not fully covered by the integuments in *AS-IAA9* lines (insets).

(G) and (H) In mature flowers (anthesis stage), a high proportion of pollen grains are aborted in antisense lines compared with the wild type. Nonaborted pollens are significantly bigger and intensely colored in antisense lines compared with the wild type. Bars = 100  $\mu$ m.

columella. Successful pollination and fertilization triggered the dissipation of the *IAA9* transcript gradient, which tended to spread across the developing fruit leading to a net decrease in transcript abundance in the placenta, funiculus, and inner integument of embryonic sac (Figures 3H to 3K). Conversely, in the absence of pollination, emasculated flowers retained the original transcript gradient associated with arrest of fruit development (Figures 3L to 3O). A close examination of the ovule confirmed





**Figure 3.** In Situ Hybridization Reveals That Pollination Triggers the Release of the *IAA9* Transcript Gradient.

(A) Low background signal detected in control hybridization experiment performed with *IAA9* sense probe.

(B) *IAA9* mRNAs are distributed all over the floral meristem, with higher accumulation in emerging organs, such as petals, stamen, and carpel. Signal intensity is higher in the adaxial sides of the emerging stamen.

(C) to (G) *IAA9* mRNA signal increases throughout flower development with uneven distribution leading to the formation of a gradient peaking at anthesis stage. Flowers from all the stages analyzed were put in the same slide and were therefore developed for the same amount of time.

(H) to (J) Pollinated ovaries at 1, 3, and 5 d after pollination (DPP). Pollination results in a rapid release of the *IAA9* gradient leading to a net decrease in *IAA9* mRNA accumulation in the placenta, funiculus, and inner integument of embryonic sac.

(K) Close-up examination of fertilized ovule. Successful fertilization releases the *IAA9* mRNA gradient, resulting in a spreading of the *IAA9* signal all over the developing ovule.

(L) to (N) Emasculated ovaries at 1, 3, and 5 DPE. In the absence of pollination, emasculated flowers retain the *IAA9* expression gradient and display an arrest of ovary development.

(O) Close-up examination of unfertilized ovule shows that 3 d after emasculature, a strong *IAA9* mRNA gradient is maintained with a high signal detected in cell layers of inner integuments and funiculus tissues.

Magnification is  $\times 5$  in (C) to (J) and (L) to (N) and  $\times 40$  in (A), (B), (K), and (O). Bars = 100  $\mu\text{m}$ . sp, sepal; sm, stamen; cap, carpels; sl, style; spg, sporogenous tissue; tap, tapetum; c, columella; ow, ovary wall; pl, placenta; f, funiculus; v, vascular bundles; es, embryo sac.

that after flower emasculature, unfertilized ovules displayed a strong gradient in *IAA9* transcript abundance, which maintains a high level of *IAA9* transcripts in cell layers of inner integuments and funiculus tissues (Figure 3O). Successful fertilization released this gradient, leading to a spreading of the *IAA9* transcript across the developing ovule at 3 DPA (Figure 3K).

### Experimental Design for Combined Transcriptomic and Metabolomic Analysis of Fruit Set

In an attempt to identify important regulatory genes and metabolic pathways involved in tomato fruit set, we used transcriptomic and metabolomic approaches as well as real-time PCR analyses of target genes. The complete experimental design

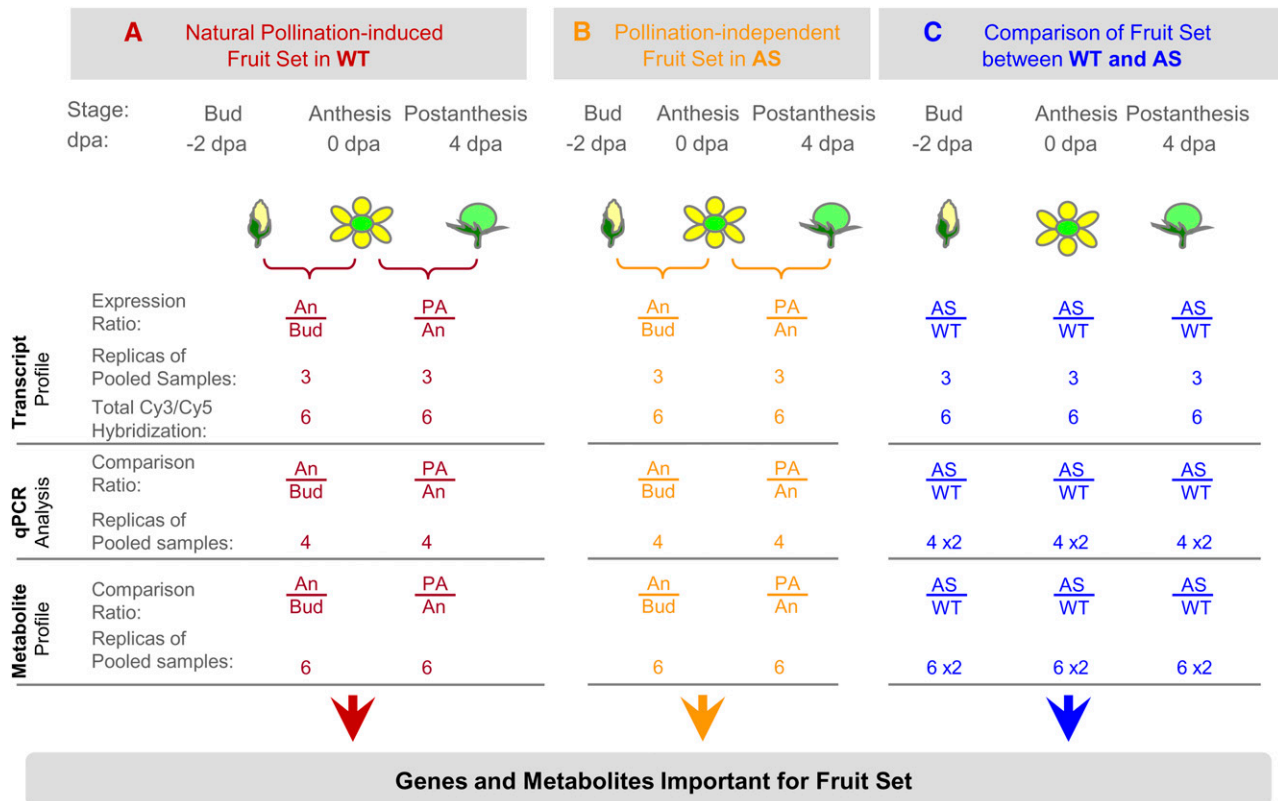
included three parallel experiments (Figure 4). The first experiment aimed at defining the genes whose expression is associated with the transitions from bud through anthesis to the postanthesis stages and the corresponding metabolic shifts during natural fertilization-induced fruit set. The second experiment was designed to define the genes differentially expressed during pollination-independent fruit set in *AS-IAA9* plants. The third experiment aimed at identifying *IAA9* regulated genes (directly or indirectly regulated) during fruit set in ovary/fruit tissues. For this purpose, we conducted comprehensive analysis of gene expression in *AS-IAA9* and the wild type at three developmental stages during fruit set: flower bud, anthesis, and postanthesis. For each experiment, metabolomic changes associated with both types of fruit set were examined in wild-type and two independent *AS-IAA9* lines (Figure 4).

Oligonucleotide microarrays (EU-TOM1) containing 11,860 tomato unigenes were used to compare transcript profiles corresponding to preanthesis flower buds (−2 DPA), anthesis

mature flower (0 DPA), and postanthesis fertilized flower (4 DPA) in the wild type and *AS-IAA9*. For each point, six Cy-labeled cDNAs corresponding to three independent biological samples with a dye-swap were hybridized to six independent array slides. The statistical analysis of the data was performed with R/MAANOVA package (see Methods).

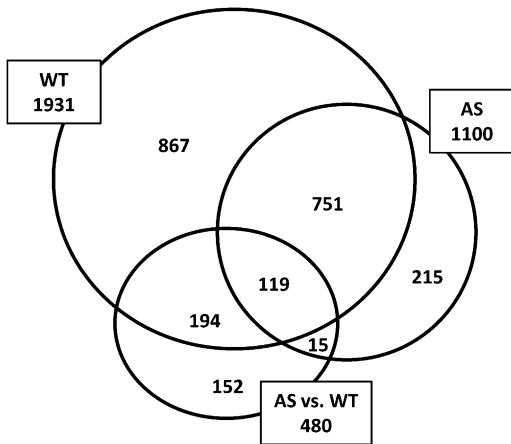
### General Feature of Fruit Set Expression Profiles

To identify new candidate genes that are essential for the fruit set process, we performed a comprehensive analysis of gene expression associated with the flower-to-fruit transition in wild-type pollination-dependent fruit set and in *AS-IAA9* pollination-independent fruit set. We compared the sets of differentially expressed genes between the two genotypes (Figure 5) and found, notably, that the overwhelming majority of the differentially expressed genes were common to both processes of fruit set (see Supplemental Table 4 online). In a further step, putative



**Figure 4.** Experimental Design for Analyzing Transcriptomic and Metabolomic Changes Associated with Fruit Set in the Wild Type and *AS-IAA9*.

Combined transcriptomic and metabolomic approaches were used to investigate the molecular events associated with pollination-induced natural fruit set in the wild type (A), pollination-independent fruit set in *AS-IAA9* (AS) (B), and to identify differentially expressed genes or altered metabolites in antisense (AS) fertilization-free fruit set versus pollination-dependent fruit set in the wild type (C). Bud (equivalent to 2 d before anthesis, −2 DPA), anthesis (An), and postanthesis (PA; 4 DPA) stages were considered. For transcriptome analyses, a direct comparison of *AS-IAA9* lines to their wild-type counterpart was employed. At each of the three developmental stages, Cy-labeled cDNAs were hybridized to six independent 12k-oligomicroarrays (EU-TOM1) using a triple dye-swap design. The abundance of a broad range of metabolites was quantified using GC-MS. Fruit/ovary tissues were pooled from wild-type (cv MicroTom) and *AS-IAA9* homozygous lines, and the samples were split into two parts, one to be used in microarray hybridization and the other in metabolite quantification. Cy3, cyanine3 fluor; Cy5, cyanine5 fluor; AS, *IAA9* antisense lines; qPCR, quantitative real-time PCR.



**Figure 5.** Venn Diagram of Fruit Set-Associated Genes.

Venn diagram showing genes differentially expressed during the three stages of flower-to-fruit transition in the wild type (red circle) and AS-*IAA9* (AS; yellow circle) tomato and differentially expressed between pollination-dependent and -independent fruit set (AS versus WT, blue circle). Overlap between each set of genes is indicated by different colors. Total numbers of differentially expressed genes are indicated in white boxes; the number of genes in each subset is indicated within the appropriate colored domain.

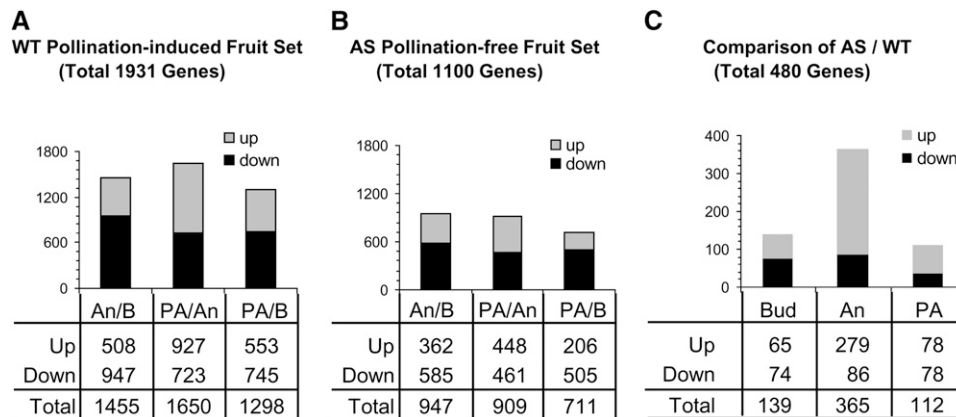
correlations between the genes that are differentially expressed during the transition from bud through anthesis to postanthesis in the wild type and AS-*IAA9* and those differentially expressed between the two genotypes were identified. In total, this study identified 870 genes (751 + 119) that are differentially expressed in both pollination-independent and pollination-dependent fruit set, among which only 119 display *IAA9*-dependent regulation (Figure 5; see Supplemental Table 5 online). Moreover, among

the genes exhibiting differential expression during pollination-dependent fruit set (1931), the vast majority of genes were independent of *IAA9* regulation. Similarly, among the genes that are common to both types of fruit set, the majority were independent of *IAA9* regulation. This reflects the fact that only a small proportion of fruit set-associated genes is dependent on normal *IAA9* regulation.

**Transcriptome Analysis of Pollination-Induced Fruit Set in the Wild Type**

Statistical analysis indicated that among 1931 differentially expressed genes in the wild type, 1455 genes exhibited differential expression during the transition from bud to anthesis stages and 1650 genes between anthesis and postanthesis stages (Figure 6; see Supplemental Table 1 online). The high number of genes displaying altered expression during fruit set is indicative of dynamic regulatory mechanisms underlying this developmental process. Functional categorization of the differentially expressed genes revealed that the major groups affected belong to gene categories encoding components of signal transduction, proteins involved in metabolic pathways, or proteins associated with cell division (Table 1). These results suggest that the fruit set process in tomato requires complex regulatory control and intensive cell division as well as reconfiguration of metabolic processes.

Surprisingly, the number of differentially regulated genes between bud and postanthesis was lower than during the transitions from bud to anthesis and from anthesis to postanthesis, consistent with a mechanism wherein a large number of genes undergo a transient up- or downregulation at the anthesis stage. Moreover, the majority of differentially expressed genes are downregulated during the transition from flower bud to anthesis, while the reverse occurs during the transition from anthesis to postanthesis (Figure 6A), indicating that there is no continuum of ovary development from bud to postanthesis.



**Figure 6.** Global Gene Expression Pattern in Pollination-Dependent and -Independent Fruit Set.

(A) and (B) Differentially expressed genes (either up- or downregulated) during the transition from flower bud to anthesis and from anthesis to postanthesis in natural pollination-induced fruit set of wild-type plants (A) and pollination-free fruit set in antisense (AS) lines (B). (C) Differentially expressed genes between AS and the wild type during fruit set at three developmental stages: bud (B), anthesis (An), and postanthesis (PA). Upregulated (Up), downregulated (Down), and total altered genes (total) refer to changes in transcript accumulation displaying log<sub>2</sub> (ratio) higher than 0.5 or lower than -0.5 (P value = 0.05).

**Table 1.** Functional Classification of Differentially Expressed Genes during Natural Pollinated Wild-Type Fruit Set

Assignment	An/B (Wild Type)		PA/An (Wild Type)		PA/B (Wild Type)	
	No. of Genes	%	No. of Genes	%	No. of Genes	%
Hormone responses	67	5	75	5	61	5
Signal transduction <sup>a</sup>	349	24	365	22	295	23
Metabolism	286	20	357	22	269	21
Cell division	116	8	161	10	136	10
Photosynthesis	52	4	54	3	37	3
Cell wall	39	3	45	3	36	3
Cell structure	24	2	29	2	16	1
Stress/defense responses	97	7	100	6	85	7
Transporter	72	5	92	6	73	6
Others	38	3	44	3	43	3
Unknown	315	22	328	20	247	19
Total	1455	100	1650	100	1298	100

<sup>a</sup>Including transcription factor.

### Transcriptome Analysis of Fertilization-Independent Fruit Set in AS-*IAA9*

Statistical analysis of the global gene expression associated with pollination-independent fruit set in AS-*IAA9* indicated that 947 genes exhibited differential expression during the transition from bud to anthesis (Figure 6B; see Supplemental Table 2 online) and 909 from anthesis to postanthesis (Figure 6B). In contrast with pollination-induced fruit set where the highest transition occurs between anthesis and postanthesis, the major shift in terms of change in transcript accumulation in AS-*IAA9* occurs at the transition from bud to anthesis.

Functional categorization of the differentially expressed genes (Table 2) revealed that the major groups affected during the process of fruit set in antisense lines belong to gene categories encoding components of signal transduction, metabolic pathways, and cell division. Of particular note, the percentage of cell division-associated genes changing on the developmental transition is considerably lower in AS-*IAA9* compared with that found in the wild type.

### Comparative Transcriptome Analysis of Wild-Type and AS-*IAA9* Fruit Set

We next sought to identify genes that are differentially expressed between the two genotypes during fruit set. When all three developmental stages are considered, a total of 480 genes showed differential expression between wild-type and AS-*IAA9* lines (Figure 5), and the difference between the two lines was greatest at anthesis (Figure 6C). Functional classification of these differentially expressed genes (Table 3) revealed that the major discriminating factors between the two lines are dependent on the developmental stage. That is, at flower bud stage, the major groups of differential genes belong to the category of signal transduction, whereas at anthesis stage, the most prevalent groups encode proteins involved in cell division and metabolism.

At postanthesis stage, the most discriminating groups are those associated with photosynthesis (35%) and metabolism. Functional classification of these genes (Table 3) also indicates that among the differentially expressed genes, those exhibiting potential regulatory function (such as those involved in hormone response or signal transduction) decrease steadily from bud to postanthesis stage, while those associated with potential executor function (such as those involved in photosynthesis, cell division, and metabolism), increase from bud to postanthesis.

### Validation of Microarray Expression Data by Quantitative RT-PCR

Quantitative real-time-PCR (qRT-PCR) was used to validate the microarray analysis data and to further define the expression pattern of selected genes, relating to our biological focus, during fruit set. Representatives from 11 functional groups of differentially expressed genes between AS-*IAA9* and the wild type (Table 1) were selected for qRT-PCR. In total, 25 genes were analyzed and the resultant qRT-PCR profiles were compared with microarray expression profiles and scored as matching if they agreed in the expression pattern. By this criterion, 90% of the values have the same pattern of expression in the microarray experiment as in the qRT-PCR experiment, whereas for the remaining 10%, only one or two biological repetition out of four showed the same pattern of expression in the microarray experiment (see Supplemental Table 7 online).

### Downregulation of *IAA9* Strongly Promotes Photosynthesis-Related Genes at Fruit Set

During pollination-induced fruit set, photosynthesis-related genes were downregulated during the transition from bud to anthesis, whereas they were generally upregulated from anthesis to postanthesis (see Supplemental Table 1 online). By contrast, photosynthesis-related genes were strongly activated in AS-*IAA9* throughout fruit set (Table 3), and, notably, all these genes without exception (43 in total) were upregulated in the AS-*IAA9*

**Table 2.** Functional Classification of Differentially Expressed Genes during Pollinated-Independent Fruit Set in AS-*IAA9* Lines

Assignment	An/B (AS)		PA/An (AS)		PA/B (AS)	
	No. of Genes	%	No. of Genes	%	No. of Genes	%
Hormone responses	50	5	48	5	39	5
Signal transduction <sup>a</sup>	242	26	212	23	170	24
Metabolism	180	19	190	21	135	19
Cell division	54	6	50	6	42	6
Photosynthesis	26	3	28	3	25	4
Cell wall	23	2	24	3	20	3
Cell structure	18	2	17	2	14	2
Stress/defense responses	66	7	59	6	55	8
Transporter	53	6	55	6	42	6
Others	22	2	24	3	22	3
Unknown	213	22	202	22	147	21
Total	947	100	909	100	711	100

<sup>a</sup>Including transcription factor.



**Table 3.** Functional Classification of Differentially Expressed Genes between Antisense and the Wild Type during Fruit Set

Assignment	Bud (AS versus WT)		An (AS versus WT)		PA (AS versus WT)	
	No. of Genes	%	No. of Genes	%	No. of Genes	%
Hormone responses	16	12	17	5	4	4
Signal transduction <sup>a</sup>	23	17	36	10	11	10
Metabolism	21	15	74	20	22	20
Cell division	22	16	108	30	13	12
Photosynthesis	2	1	29	8	39	35
Cell wall	7	5	18	5	2	2
Cell structure	1	1	8	2	5	4
Stress/defense responses	19	14	27	7	8	7
Transporter	8	6	14	4	2	2
Others	4	3	11	3	0	0
Unknown	16	12	23	6	6	5
Total	139	100	365	100	112	100

<sup>a</sup>Including transcription factor.

lines compared with the wild type (see Supplemental Table 3 online). qRT-PCR analysis (Figures 7A and 7B) further confirmed the expression profile revealed by microarray for two photosynthesis-associated genes encoding chlorophyll *a/b* binding protein 4 (*CAB4*) and chlorophyll *a/b* binding protein 1D (*CAB-1D*). The two genes were upregulated at postanthesis stages in the wild type, and their upregulation continued during subsequent cell division and enlargement phases of fruit development (Figure 7C). In line with the upregulation of photosynthesis-related genes, downregulation of *IAA9* was associated with a dramatic upregulation of genes involved in sucrose catabolism (Figure 7D), particularly at the postanthesis stage and to a lesser extent at the anthesis stage, suggesting the importance of these genes in the processes of the flower-to-fruit transition. Notably, all sugar metabolism genes that are differentially expressed between *AS-IAA9* and the wild type displayed upregulation during the fruit set developmental process, suggesting that activation of sugar metabolism may be an important event coupled with the initiation of pollination-independent fruit set.

### The Fruit Set Process Recruits a High Number of Transcriptional Regulators

Up to 199 transcription factors displayed differential expression during natural fruit set. Such a high number of transcriptional regulators is indicative of the massive and complex regulation required for the fruit set process (see Supplemental Table 1 online). Comparing the transcriptomic profiles associated with the fruit set process in the wild type and *AS-IAA9* revealed that 18 transcription factors showed differential expression between the two genotypes (see Supplemental Table 3 online). qRT-PCR analysis showed that during natural pollination-induced fruit set in the wild type, the expression of two MADS box genes, *Tomato Agamous1* (*TAG1*) and *Tomato Agamous-like 6* (*TAGL6*), underwent dramatic downregulation following fertilization exhibiting seven- and eightfold decreases in the level

of transcript accumulation, respectively (Figure 8A). Interestingly, during later stages of fruit development, the expression level of the two genes continued to display sharp and steady decreases with *TAG1* and *TAGL6* transcripts dropping at 10 DPA to ~10 and 1% of their levels at 1 DPA, respectively (Figure 8C). Moreover, compared with the wild type, the two MADS box genes underwent a net downregulation during fruit set in *AS-IAA9* lines, which was most prominent at the anthesis stage (Figure 8B). Since both natural and pollination-independent fruit set were both characterized by a net downregulation of *TAG1* and *TAGL6*, it can be speculated that these two MADS box genes are key players in the process of fruit set and that *AS-IAA9* may promote fertilization-independent fruit set by downregulating their expression.

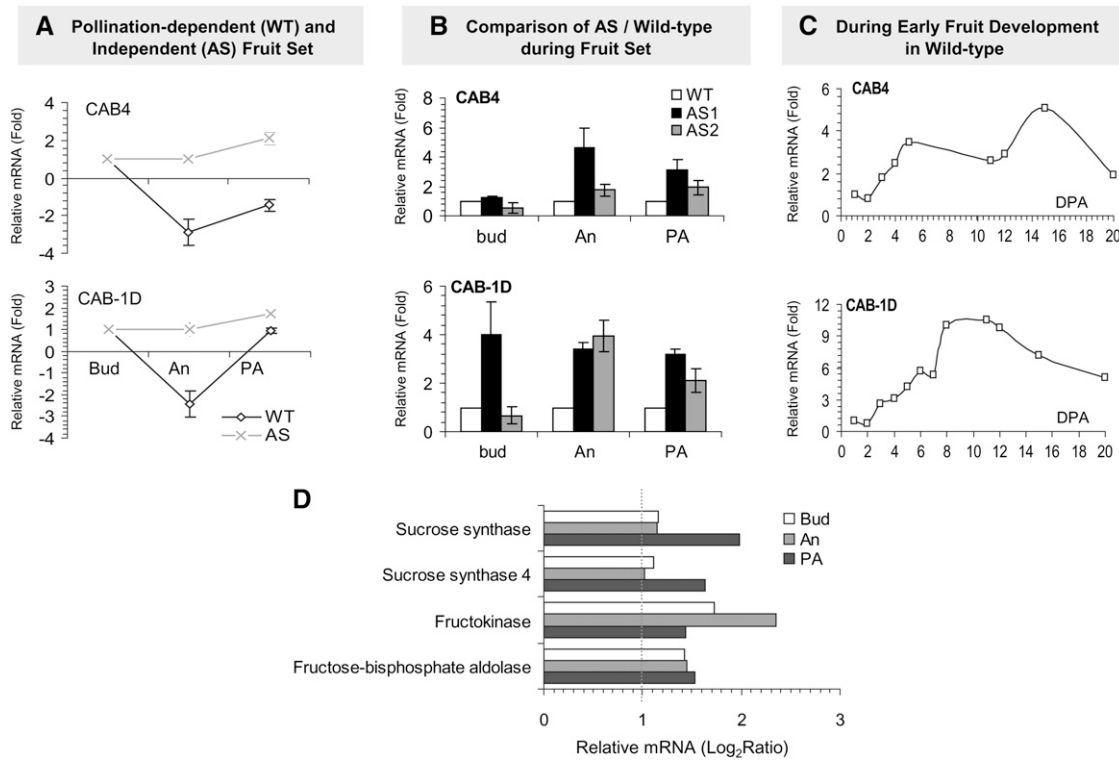
### Upregulation of Cell Division-Related Genes Is Advanced in *AS-IAA9* Lines

A number of cell division, protein biosynthesis, and cell wall-related genes were upregulated at the postanthesis stage in the wild type, while their activation occurred earlier in *AS-IAA9* at the anthesis stage (see Supplemental Tables 1 and 2 online). Remarkably, when comparing pollination-dependent and -independent fruit set, these genes were found to be upregulated in *AS-IAA9* at the anthesis stage but downregulated at postanthesis (see Supplemental Table 3 online). These included *cyclinA3*, *cyclinb*, and 19 unigenes coding for histones that all displayed higher expression in *AS-IAA9* at anthesis stage. A large group of 67 ribosomal protein genes was also all upregulated at anthesis stage in *AS-IAA9*. A total of 24 cell wall-related genes were differentially expressed in *AS-IAA9*, with a large majority of them being upregulated, including *cellulose synthases*, *expansins*, *extensins*, *polygalacturonase*, and *glucanases*. In particular, the polygalacturonase inhibitor protein (*PGIP1*) gene, putatively involved in fruit enlargement, was very strongly upregulated (sevenfold) at both bud and anthesis stages in *AS-IAA9*, but only at postanthesis in naturally pollinated wild-type young fruit (see Supplemental Tables 1 and 3 online).

The expression of two histone protein genes and two ribosomal protein genes during the flower-to-fruit transition in wild-type lines displayed a strong activation at 4 DPA (Figure 9A) and then a sharp decline after 10 DPA, coinciding with the shift from cell division to cell enlargement phases (Figure 9C). Interestingly, the upregulation of these genes occurred at anthesis stage in *AS-IAA9*. Consistently, when compared with the wild type, the expression of all four genes was upregulated in *AS-IAA9* at both bud and anthesis stages but was clearly downregulated at the postanthesis stage (Figure 9B). By contrast, the *expansin10* gene showed higher expression in *AS-IAA9* compared with the wild type throughout the three stages of fruit set, and its transcript levels remained high after the shift from cell division to fruit enlargement during natural fruit development (Figures 9B and 9C).

### Expression of Phytohormone-Related Genes during the Process of Fruit Set

Taking into account the prominent role played by phytohormones in triggering and coordinating the flower-to-fruit transition



**Figure 7.** Expression Analysis of Genes Involved in Photosynthetic Processes and Sugar Metabolism.

**(A)** Expression kinetics of selected photosynthetic genes during pollination-dependent fruit set in wild-type (black line) and pollination-free fruit set in *AS-IAA9* (gray line) assessed by qRT-PCR. cDNA were prepared from the same RNA samples used in the microarray experiment. Data are expressed as relative values, based on the values of bud ovary taken as reference sample set to 1.

**(B)** Comparison of transcript levels between the wild type (white bar) and two independent *AS-IAA9* lines *AS1* (black bar) and *AS2* (gray bar) assessed by qRT-PCR at three developmental stages. Data are expressed as relative values, based on the reference wild-type samples set to 1.0 at each stage considered.

**(C)** Transcript accumulation during early stage of natural pollination-induced fruit set in the wild type. Relative expression levels for stages 2 to 20 DPA were determined based on the reference 1 DPA sample set to 1.0. *CAB4*, chlorophyll *a/b* binding protein 4; *CAB-1D*, chlorophyll *a/b* Binding Protein 1D. Error bars represent  $\pm$  SE of four biological repetitions.

**(D)** Comparative transcript accumulation in the wild type at different stages of genes involved in sugar metabolism identified by microarray analysis as differentially expressed between wild-type and antisense lines.

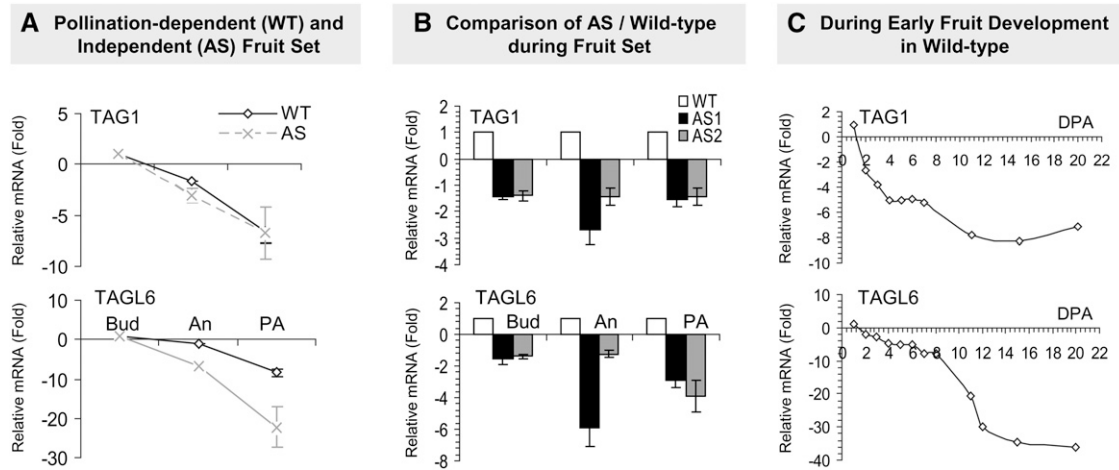
developmental process, we closely examined transcript accumulation of genes associated with hormone responses and metabolism. During natural pollinated fruit set, almost 5% of the differentially expressed genes was related to hormone metabolism and signaling (Table 1). Functional category classification of these genes indicated that auxin, ethylene, and polyamine-associated genes were the most prominent in this developmental process (Figure 10A, left panel) as well as in the pollination-free fruit set of *AS-IAA9* (Figure 10A, right panel). The high number of ethylene-related genes observed to change suggests that in addition to the well-established role of auxin and GAs (Pandolfini et al., 2007; Serrani et al., 2007), ethylene must also play an active role in fruit set. Detailed evaluation of the ethylene-associated genes reveals that they included both key ethylene signal transduction pathway and ethylene biosynthesis genes (see Supplemental Table 1 online).

qRT-PCR analysis (Figures 10B to 10D) indicated that, in both the wild type and *AS-IAA9*, the expression of all hormone-related

genes selected underwent dramatic downregulation following the anthesis stage (Figures 10B to 10D, left panel), which continued until at least 14 DPA in the wild type (Figures 10B to and 10D, right panel). Notably, the downregulation of auxin-related genes occurred earlier (2 DPA) than that of ethylene-related genes (6 DPA). As a general feature, the hormone-related genes displayed lower expression in *AS-IAA9* lines compared with the wild type (Figures 10B to 10D, middle panel). According to these data, auxin and ethylene are the two hormones most prominently involved in the control of the flower-to-fruit transition developmental process.

#### **Aux/IAA and ARF Genes Are Strongly Regulated during Fruit Set**

Considering the role of *IAA9* in mediating auxin responses and that generally ascribed to auxin in the process of fruit set, we sought to study the expression profile of all members of the *Aux/*



**Figure 8.** Expression Analysis of Two MADS Box Genes Assessed by qRT-PCR.

**(A)** Expression kinetics of two MADS box genes (*TAG1* and *TAGL6*) during pollination-dependent (wild-type, black line) and pollination-free (*AS-IAA9*, gray line) fruit set. Data are expressed as relative values based on bud ovary taken as reference sample set to 1.0.

**(B)** Comparison of transcript levels between wild-type (white bar) and two independent *AS-IAA9* lines *AS1* (black bar) and *AS2* (gray bar) at three developmental stages. Data are expressed as relative values, based on the reference wild-type samples set to 1.0.

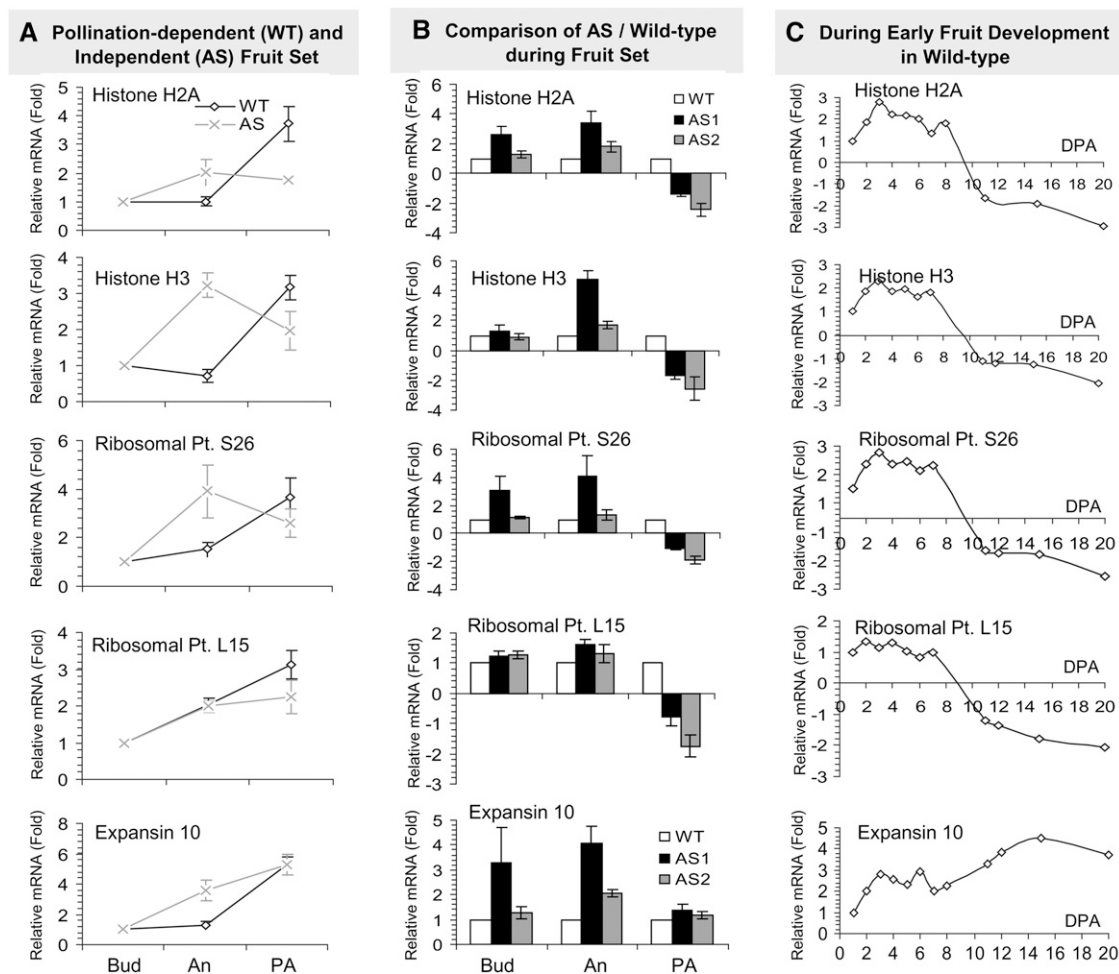
**(C)** Transcript accumulation during early stage of natural pollination-induced fruit set in the wild type. Relative expression levels for stages 2 to 20 DPA were determined based on the reference value of 1 DPA sample set to 1.0. *TAG1*, *Tomato Agamous 1*; *TAGL6*, *Tomato Agamous-like6*. Error bars represent  $\pm$  SE of four biological repetitions.

*IAA* and auxin transcription factors (*ARF*) gene families available in tomato. The expression data (Figure 11) indicate that throughout the process of natural fruit set, both *ARFs* and *Aux/IAAs* display a dramatic shift in their expression, suggesting an active role of these transcriptional regulators in this developmental process. Moreover, downregulation of *IAA9* resulted in a feedback regulation of a number of *ARF* and *Aux/IAA* genes. Of particular note is the fact that half of the *ARF* genes tested (7 out of 14) undergo opposite regulation during the transitions from bud to anthesis and from anthesis to postanthesis (Figure 11A). The expression of the overwhelming majority of *ARF* genes at the bud stage was strongly upregulated in *AS-IAA9* lines compared with the wild type (Figure 11B). Notably, *ARF4*, *ARF8*, and *ARF18* showed strong upregulation upon pollination in the wild type, and their expression levels were higher in the *AS-IAA9* line at bud stage, suggesting their potential involvement in the process of pollination-independent fruit set. Most *Aux/IAA* genes (13 out of 18) were downregulated during the transition from bud to anthesis, whereas a number of them (six genes) displayed clear upregulation during the transition from anthesis to postanthesis. The *Aux/IAAs* genes that were upregulated upon pollination in the wild type (*IAA1*, *IAA3*, *IAA11*, *IAA13*, *IAA14*, and *IAA30*) displayed higher expression in *AS-IAA9* than in wild-type lines at bud stage, suggesting that the upregulation of these genes is integral to the fruit set process.

#### Metabolite Accumulation in Ovary/Fruit during Natural Pollinated Fruit Set

Information related to metabolic changes underlying the flower-to-fruit transition are still lacking, and the studies of flower

metabolism in various species generally tend to be confined to specific pathways of importance with respect to color or fragrance (Dudareva et al., 1996; Moyano et al., 1996; Dunphy, 2006). To identify the changes in primary metabolism that underlie the flower-to-fruit transition, an established gas chromatography-mass spectrometry (GC-MS) method was applied to extracts from the materials described in the experimental design (Figure 4). Comparing the metabolite contents between the samples harvested at anthesis and bud stages in the wild type revealed little difference in the metabolism of these developmental stages. By contrast, many changes were observed between the samples harvested at anthesis and postanthesis. The differences detected are summarized in Figure 12 (while the absolute values are presented in Supplemental Table 6 online). In the wild type, differences observed between bud and anthesis stages were confined to increases in raffinose, glycerol 3-phosphate, fructose 6-phosphate, Val, Phe, Tyr, and His and a decrease in the level of galactonate-1,4-lactone (Figure 11, white bars). When comparing anthesis and postanthesis stages, a total of 22 of the 73 compounds measured were found at significantly different levels (see Supplemental Table 6 online). In a few instances, the direction of these changes was reversed from that observed between bud and anthesis. Raffinose, mannose, L-ascorbate, glucose-6-phosphate, maltitol, glycerol-3-phosphate, Leu, Val, Ala, Phe, Trp, isocitrate, Gln, Arg, and putrescine all significantly decrease at postanthesis to values as low as those recorded at the bud stage (see Supplemental Table 6 online). By contrast, few increases in metabolite level were observed with galactonate-1,4-lactone, gluconate-1,5-lactone, threonate, pyruvate, Asp, and 5-oxoproline being markedly elevated at this time point.



**Figure 9.** Expression Analysis of Cell Division-Related Genes Assessed by qRT-PCR.

**(A)** Expression kinetics of two histone genes (*Histone H2A* and *Histone H3*), two ribosomal protein genes (*Ribosomal Pt. S26* and *Ribosomal Pt. L15*), and *Expansin 10* gene during pollination-dependent (wild-type, black line) and pollination-free (*AS-IAA9*; gray line) fruit set. Data are expressed as relative values, based on bud ovary taken as reference sample set to 1.

**(B)** Comparison of transcript levels between wild-type (white bar) and two independent *AS-IAA9* lines *AS1* (black bar) and *AS2* (gray bar) at three developmental stages. Data are expressed as relative values, based on the reference wild-type sample set to 1.0.

**(C)** Transcript accumulation during early stage of natural pollination-induced fruit set in the wild type. Relative expression levels for stages 2 to 20 DPA were determined based on the reference 1 DPA sample set to 1.0. Error bars represent  $\pm$  SE of four biological repetitions.

### Metabolite Accumulation in Ovary/Fruit Revealed Considerable Differences between Fertilization-Dependent and -Independent Fruit Set

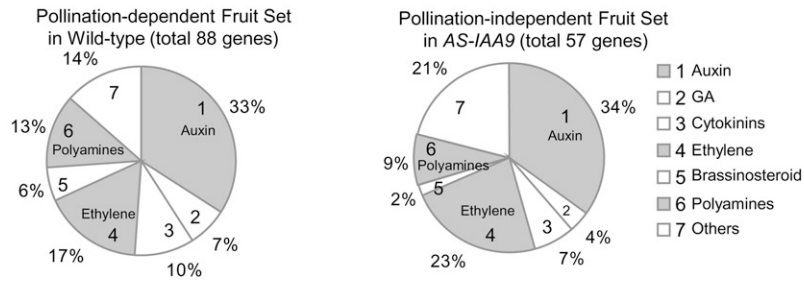
The scale of change in metabolite contents between the floral bud and anthesis samples was dramatically enhanced in the transgenic lines (Figure 12, gray bars). These changes included increased levels of glucose, fructose, sucrose, xylose, inositol, gluconate-1,5-lactone, galactonate-1,4-lactone, ribose-5-phosphate, amino acids derived from pyruvate (Leu, Ile, and Ala), oxaloacetate (Asp, Asn, and Lys), and 2-oxoglutarate (Glu and Gln), as well as increased levels of Gly, threonate, and citrate at the anthesis stage. In addition, the shikimate pathway was upregulated in the antisense lines, as evidenced by the higher levels of both Tyr and the large increase in shikimate itself (Figure 12, gray bars; see

Supplemental Table 6 online). The transgenic genotype also displayed large changes with many metabolites returning to levels comparable to those observed at the bud stage. Particularly dramatic changes in the levels of fructose, glucose, inositol, citrate, glycerol 3-phosphate, fructose 6-phosphate, Ala, Asn, Glu, Leu, Tyr, and shikimate can be seen in Figure 12 (gray bars). These combined changes indicate that far fewer of the metabolites displayed significant differences at postanthesis stage in comparison to the control (bud stage).

Fertilization-independent fruit set is characterized by a massive increase in sugars and sugar derivatives at the anthesis stage, with fructose, glucose, sucrose, inositol, and trehalose all being dramatically increased. However, the levels of glucose, inositol, and trehalose fell substantially at subsequent stages.

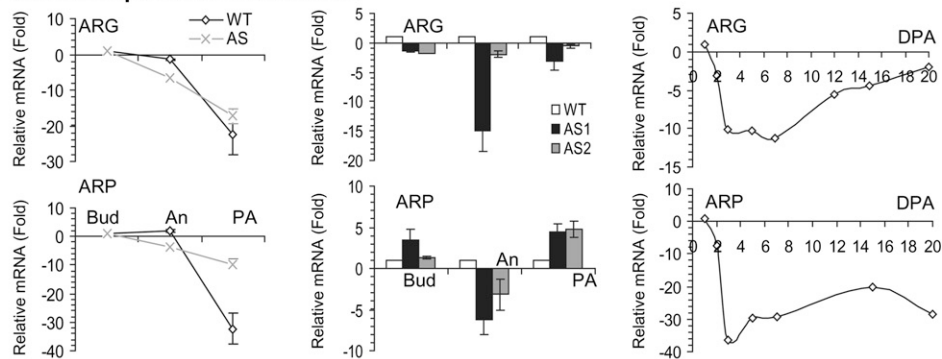


**A Percentage of Genes Involved in Hormones Response during Pollination-dependent and Independent Fruit Set**

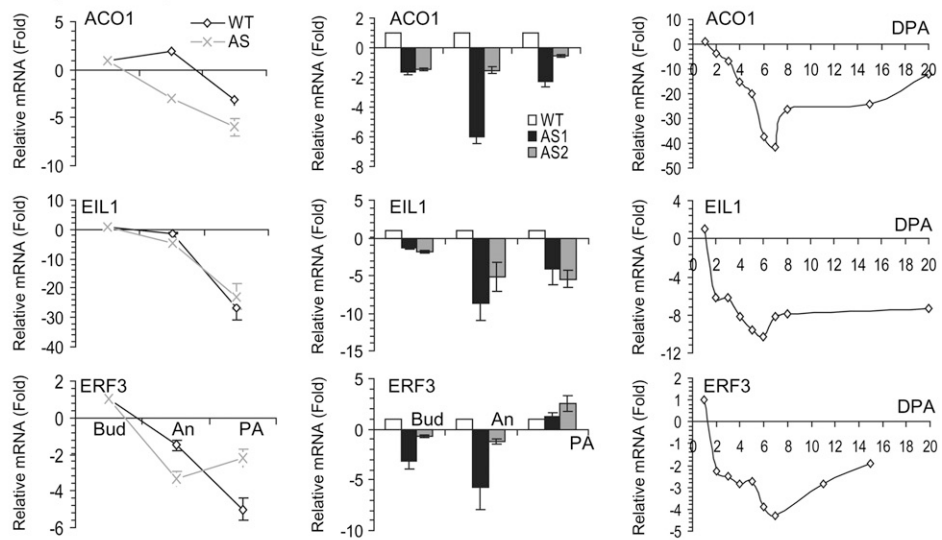


**Pollination-dependent (WT) and Independent (AS) Fruit Set      Comparison of AS / Wild-type during Fruit Set      During Early Fruit Development in Wild-type**

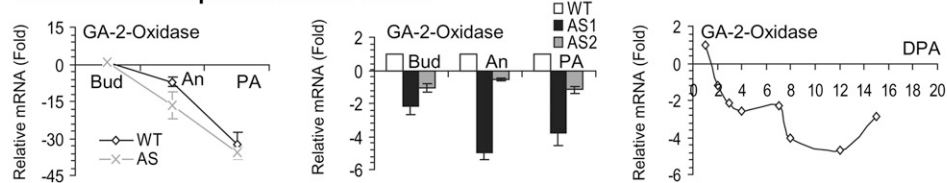
**B Auxin Response Related Genes**



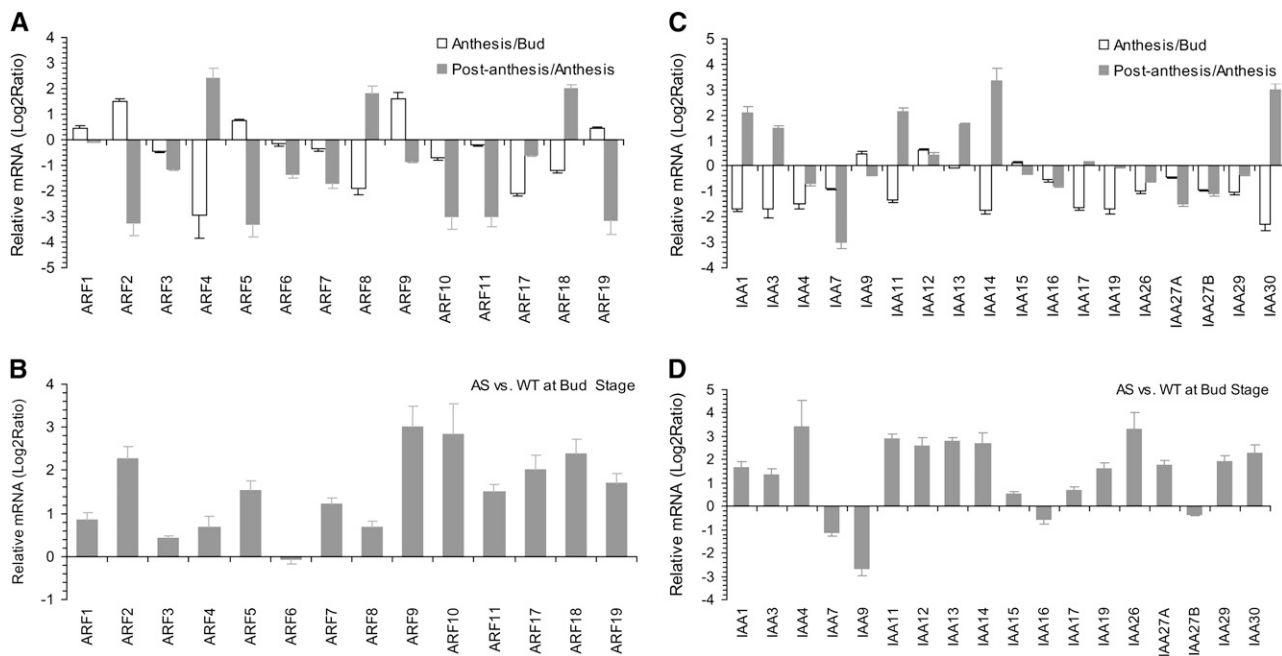
**C Ethylene Response Related Genes**



**D Gibberellins Response Related Genes**



**Figure 10.** Expression Analysis of Hormone-Related Genes.



**Figure 11.** Expression Profile of Auxin Response Transcription Factors during Fruit Set.

**(A)** and **(C)** Expression profile of *ARF* **(A)** and *Aux/IAA* **(C)** during pollination-induced fruit set in wild-type lines. Data are expressed as relative values, white bars indicate the relative transcripts level in anthesis stage compared with bud stage, and dark-gray bars indicate level in postanthesis stage compared with anthesis stage.

**(B)** and **(D)** Expression profile of *ARFs* **(B)** and *Aux/IAAs* **(D)** in *AS-IAA9* pollination-free fruit set lines. Data are expressed as relative values, which indicate the relative transcript levels in *AS-IAA9* lines compared with the wild type at bud stage. Expression analysis was assessed by real-time PCR, and data are expressed as relative value  $\log_2$  ratio.

Interestingly, this effect is not mirrored by the levels of fructose 6-phosphate or glycerol 3-phosphate, which exhibited similar dynamic behavior across development in both types of fruit set. Organic acids were also generally present at higher levels throughout development during fertilization-independent fruit set. Notably, citrate was at significantly higher levels in the transgenics during anthesis, as was the ascorbate precursor galactonate-1, 4-lactone; however, both compounds returned to levels observed in the wild type by 4 DPA. Consistent changes in the levels of some amino acids were also observed (Figure 12, black bars). Ala is present at considerably higher levels at both anthesis and postanthesis stages in the transgenic lines, as is Leu, perhaps indicating a higher metabolic rate associated with

fertilization-independent fruit set. Glu and Gln as well as Tyr and shikimate are all at significantly enhanced levels in the *AS-IAA9* lines at the bud stage. In addition, the levels of Trp are clearly decreased at all three developmental stages, perhaps indicating an important role for this metabolite in fertilization-dependent fruit set.

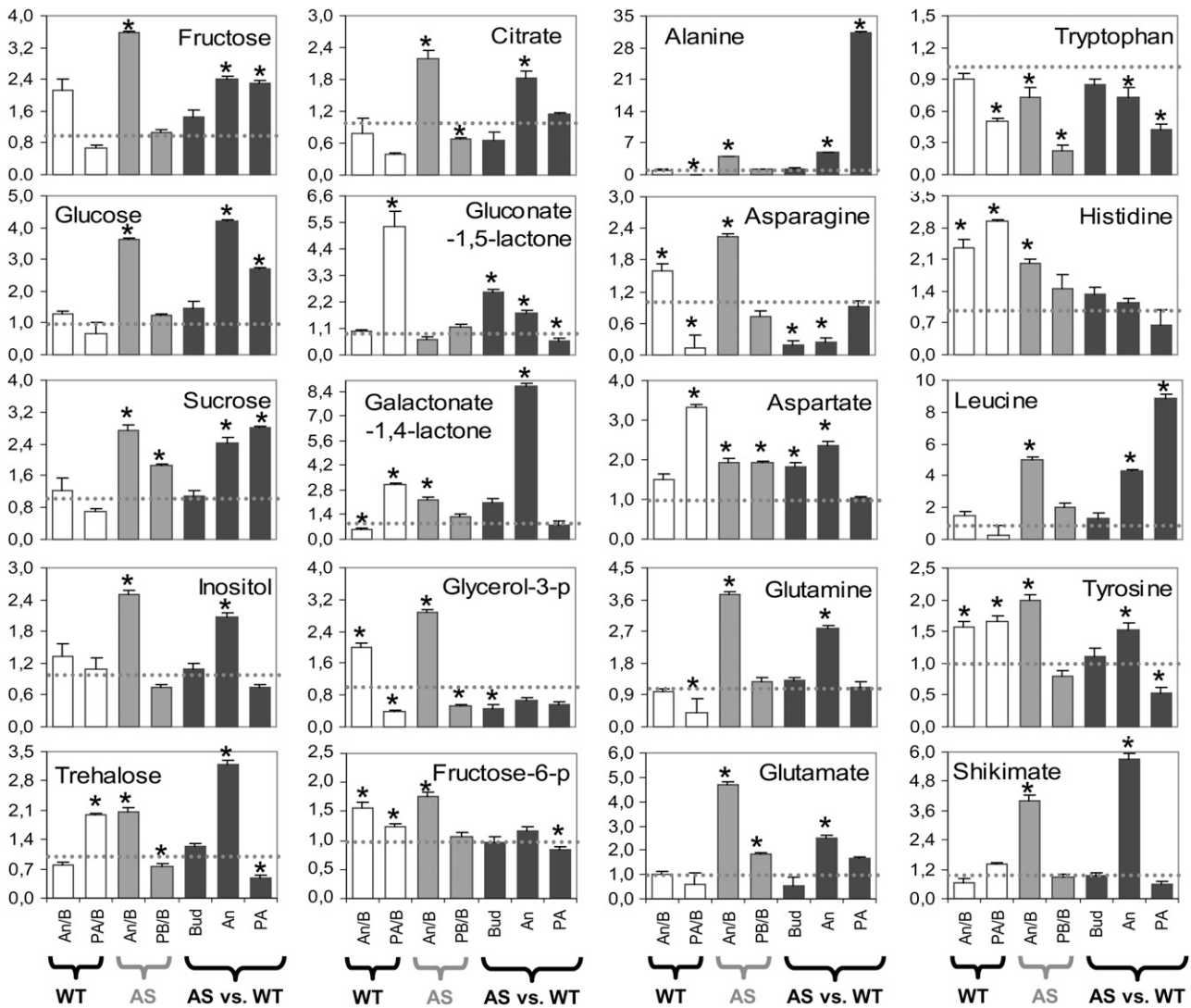
#### Integration of Metabolite and Transcript Profiling Data Obtained during Pollination-Dependent Fruit Set

Given that the determination of metabolite and transcript levels were performed in the exact same samples, we then investigated whether further biological insights were accessible following

**Figure 10.** (continued).

**(A)** Expression of hormone-related genes differentially expressed during pollination-dependent fruit set in the wild type (left panel) and pollination-free fruit set in *AS-IAA9* (right panel). The most prominent groups are shaded gray. The category called "Others" includes abscisic acid, jamonic acid, and salicylic acid.

**(B)** to **(D)** Expression analysis assessed by qRT-PCR of auxin response-related genes **(B)**, ethylene response-related genes **(C)**, and GA response-related genes **(D)**. Left panel, expression kinetics of selected hormone-related genes during pollination-dependent fruit set in the wild type (black line) and pollination-free fruit set in *AS-IAA9* (gray line); data are expressed as relative values, based on the reference bud ovary sample set to 1.0. Middle panel, comparison of transcript levels between wild-type (white bar) and two independent *AS-IAA9* lines *AS1* (black bar) and *AS2* (gray bar) at three developmental stages. Data are expressed as relative values, based on the reference wild-type samples set to 1.0. Right panel, transcript accumulation during early stage of natural pollination-induced fruit set in the wild type. Relative expression levels for stages 2 to 20 DPA were determined based on the reference 1 DPA sample set to 1.0. ARG, auxin-repressed gene; ARP, auxin-regulated gene; ACO1, ACC oxidase1. Error bars represent  $\pm$  SE of four biological repetitions.

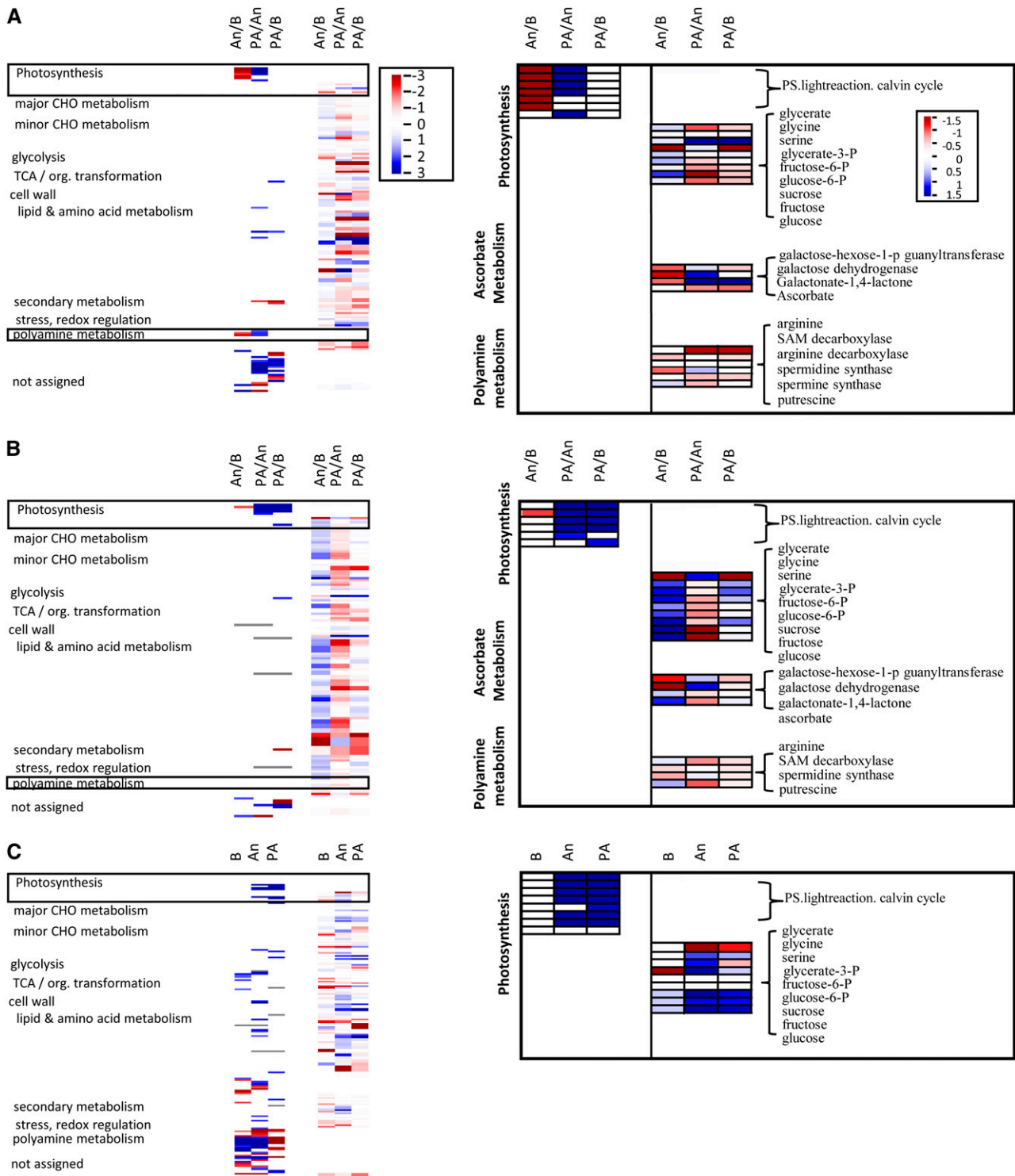


**Figure 12.** Comparative Analyses of Metabolic Changes Associated with Pollination-Dependent and -Independent Fruit Set.

White bars represent metabolic changes during pollination-dependent fruit set in the wild type. The relative metabolite level of anthesis (An/B) and postanthesis (PA/B) were determined compared with that of wild-type flower bud, which was set to 1.0. Gray bars show the metabolic change during pollination-independent fruit set in AS-*IAA9*. The relative metabolite level of anthesis and postanthesis were determined compared with that of AS-*IAA9* flower bud, which was set to 1.0. Black bars present changes in the metabolite profile in antisense line AS1 compared with wild-type lines during flower-to-fruit transition at three different developmental stages of bud, An (anthesis), and PA (postanthesis). The relative metabolite level was determined compared with that of the wild type at each stage. The dotted lines in the diagrams reflect the normalized control (1.0). The changes are represented as means  $\pm$  SE of determinations of six individual pooled samples. An asterisk indicates changes deemed by the Student's *t* test ( $P < 0.05$ ) to be statistically significant.

integration of these data using two different approaches. First, we uploaded the combined data sets into the MapMan software framework to allow full access to our data for further interrogation (<http://mapman.mpimp-golm.mpg.de/supplement/wang/>). A visualization of the connectivity between functionally classified transcripts and metabolites can be seen in the PageMan (Usadel et al., 2005) representation of Figure 13. To comprehensively investigate this, we compared changes in either the wild type (Figure 13A), the *IAA9* antisense lines (Figure 13B), or the ratios

between the two genotypes during each transition and across the whole developmental process (Figure 13C). Interestingly, given that there were proportionally far more changes in the levels of metabolites than transcripts, the general picture is that most metabolite levels do not seem to be directly related to the levels of transcripts associated with their metabolism. However, certain observations, such as the similar trend in the levels of transcripts associated with photosynthesis and the levels of major carbohydrates, suggest that regulatory modules can be



**Figure 13.** Integrated Analysis of Transcriptional and Metabolic Changes.

The left-hand side the figure shows the color-coded results of a Wilcoxon test for a consistent upregulation (blue) or downregulation (red) of individual processes for the whole data set (columns 1 to 3). Average levels of metabolites as well as averages of metabolites per process are also color coded in the same manner (columns 4 to 6). In the panel on the right-hand side, individual processes are magnified, and processes or metabolites are labeled individually. The results of the Wilcoxon test are shown in the first columns. The next columns show individual metabolites or the averages of individual enzyme classes. Processes or enzyme classes not flagged as significant have been omitted for clarity. The three independent graphs represent changes in the wild type (**A**), *IAA9* antisense line (**B**), and the comparison of these genotypes obtained by dividing the  $\log_2$  transformed values of the antisense by similarly transformed values of the wild type (**C**).

identified within the data set (Figure 13A). By contrast, the *IAA9* antisense line displayed clearer correlative changes between transcript and metabolite levels with strong upregulation of transcripts associated with both the photosynthetic apparatus and enzymes of the Calvin cycle.

To further explore correlative changes between metabolites and transcripts, we performed a manual consistency check between those metabolites that display differences in abundance and transcripts associated to the corresponding metabolic pathways (<http://mapman.mpimp-golm.mpg.de/supplement/wang/>). There were many cases in which the changes in metabolites were, at least partially, reflected by changes in the levels of transcripts associated with either their synthesis or degradation. The levels of His, Ala, and Asp provide good examples of such instances (see Supplemental Figures 1 to 3 online). Changes in the levels of His (which increased) were similar to those observed in the transcript level of histidinol-phosphate aminotransferase. In addition, changes documented in Ala levels were similar to those for the transcript levels of Ala glyoxylate transferase. Similarly, Asp displayed a similar upregulation as was observed in one of the isoforms of Asp aminotransferase (see online pathway visualizations and supplemental data for details). In two occasions, further potential transcriptional pathway regulation became apparent following the application of consistency analysis. The first of these is in the pathway of polyamine biosynthesis; at the anthesis stage, several S-adenosylmethionine (SAM) decarboxylases as well as several spermidine synthases were downregulated. However, the levels of the spermidine synthase genes were upregulated again in the postanthesis stage, which might explain the lowered levels of putrescine found during this stage. This metabolite is potentially depleted by its flux into spermidine (Figure 13A, right panel), which is in keeping with previous reports of a rise of free spermidine and spermine levels during early tomato fruit development (Egea-Cortines et al., 1993). Furthermore, a depletion of Arg was also identified, which feeds into the polyamine biosynthesis pathway through Arg decarboxylase. The identification of transcriptional control of this pathway also may explain its identification as an overrepresented functional category. Secondly, the ascorbate pathway exhibited similar behavior (Figure 13A, right panel). While, ascorbate levels fell during the transition from anthesis to the postanthesis stage, the levels of its direct precursor, galactonate 1,4-lactone, conversely decreased during anthesis before rising again at postanthesis. Interestingly, the transcript levels of both galactose-dehydrogenase (L-Gal-DH) and GDP-L-galactose-hexose-1-phosphate guanyltransferase were repressed during the anthesis stage with L-Gal-DH recovering during postanthesis and was thus able to account for the change in levels of galactonate 1,4-lactone. Furthermore, several transcripts coding for enzymes potentially involved in the production of glucuronate (an alternative precursor for ascorbate) were upregulated at later developmental stages.

By contrast, there was no correlation between transcript and metabolite changes for some of the metabolites displaying change throughout the fruit set process, such as raffinose and Leu, no consistent transcript responses could be detected, indicating that these metabolites are unlikely to be controlled at the transcriptional level. However, it is important to note that

caution must be taken when interpreting these data since the EU-TOM1 does not offer genome-wide coverage, and several genes, such as galactinol synthases, which catalyze the first step in the pathway of raffinose synthesis, are not represented on the chip. In the case of Leu, the amino acid was not significantly altered during the anthesis stage but increased thereafter. This was somewhat in contrast with the behavior of transcripts associated with its synthesis, for example, the gene encoding 2-isopropylmalate synthase increases constantly during development, while several branched-chain amino acid transaminases were downregulated. This likely reflects a multilevel regulation of the branched-chain amino acids and suggests that the control resident in this metabolic network shifts among the constituent enzymes during the developmental period.

### Integration of Metabolite and Transcript Profiling Data Obtained during Pollination-Independent Fruit Set

We next searched for consistent changes occurring when comparing the metabolites and transcripts of the transgenic line to that of the wild type. In this instance, not only the major sugars but also the minor sugars altered during development. An interesting example is the increased production of xylose in the antisense line during the later stages of development, which might be controlled by the upregulation of both dual-activity UDP-D-apiose/UDP-D-xylose synthase and UDP-D-xylose synthase isoforms in the transgenic line, since these enzymes represent both known synthetic pathways for xylose. As opposed to many other nucleotide sugar converting enzymes, it is anticipated that these enzymes catalyze irreversible reactions that would likely require tight regulatory control (Reiter, 2008). Another example in which the transgenic situation was somewhat different from that of the wild type is that the sharp rise in Leu synthesis, particular at the anthesis stage, could be explained in the antisense line by a concerted upregulation of 3-isopropylmalate dehydrogenase and 2-isopropylmalate synthase and a downregulation of Leu degrading enzymes, such as 3-methylcrotonyl-CoA carboxylase. These changes thus corroborate the hypothesis we made on the basis of the wild-type data alone, that Leu metabolism is controlled by multilevel regulation.

As observed for pollination-dependent fruit set, there were several changes in transcripts associated with ascorbate metabolism. While the metabolite L-galactono-lactone was increased in the antisense line at anthesis, the transcripts involved in the initial stages of the ascorbate synthesis pathway as well as the potential myo-inositol pathway were upregulated in the transgenic lines. However, in this instance, these transcriptional changes were not reflected in changing ascorbate levels. Moreover, it was observed that Arg decarboxylases were downregulated, while spermidine synthases were upregulated at this time point when comparing to the pollination dependent fruit set. Once again, this was not reflected in changes in the metabolites themselves. Figure 13C clearly highlights the differences between the two genotypes, revealing the dramatic upregulation of photosynthesis in the *IAA9* transgenics as well as an upregulation of glycerate 3-phosphate and the sugars sucrose, glucose, and fructose.

### Differences between the Pollination-Dependent and -Independent Fruit Set Are More Discriminative at the Metabolomic Than the Transcriptomic Level

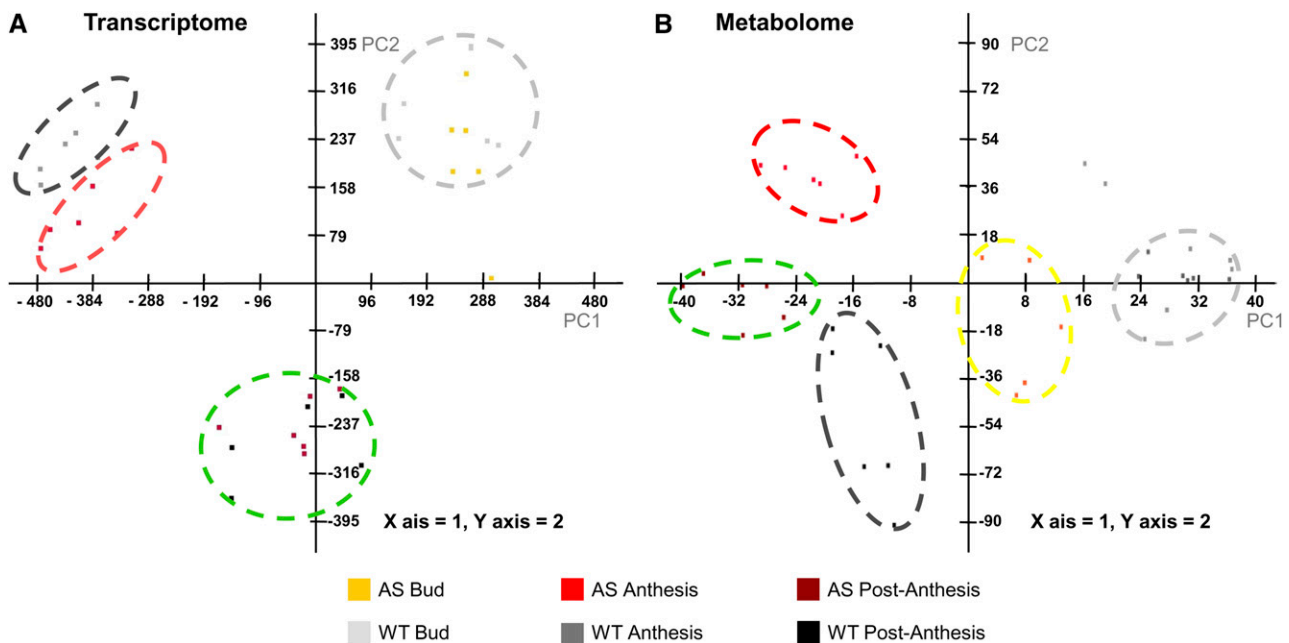
Having performed an integrative analysis, we next decided to compare the discriminatory power of the metabolite and transcript data sets independently. Principal component analysis of the 480 transcripts whose expression is significantly altered between pollination-dependent and -independent fruit set (Figure 14A) allowed clear discrimination of developmental stages along both the first and second principal component, leading to three separated groups corresponding to bud (light-gray circle), anthesis (red and dark-gray circles), and postanthesis stages (green circle). However, they could not discriminate between the wild type and *AS-IAA9* (AS) at bud and postanthesis stages, while at anthesis stage, wild-type (dark-gray circle) and *AS-IAA9* (red circle) were clearly separated along the two principal components. It indicates that at the global transcriptional level, there are no significant differences between fertilization-dependent and -independent fruit set but rather specific functional changes that are sufficient to mediate the observed phenotypic responses.

Principle component analysis (PCA) of metabolite data (Figure 14B) allowed clear discrimination both between the two genotypes and between the various stages of fruit set process. Wild-

type samples (gray and black spots) could be discriminated from that of *AS-IAA9* (colored spots) along both the first and second principal component, highlighting the global differences in terms of metabolic changes between fertilization-dependent and fertilization-independent fruit set. Along the first principal component, the postanthesis stage separate from the earlier time points in both genotypes. In the wild type, they could not however discriminate anthesis from bud stages (gray circle); by contrast, *AS-IAA9* displayed a clear separation of anthesis (red circle) from bud stage (yellow circle), displaying a precocious shift of metabolite levels in *AS-IAA9*, consistent with the precocious development previously observed in these lines. Evaluation of the metabolites that made the major contribution to the discrimination observed in this graph revealed that they were components of cell wall and energy metabolism.

### DISCUSSION

This study was designed to harness transcript and metabolite profiling technologies to evaluate changes occurring within the developing ovary and fruits of both wild-type and *AS-IAA9* plants to discriminate between fertilization-dependent and -independent fruit set. The underlying rationale was that such a study



**Figure 14.** PCA of Transcripts and Metabolites during the Flower-to-Fruit Transition.

Score plot of two principle components of transcripts and metabolites showing that developmental stages as well as different genotypes (wild type and *AS-IAA9*) cluster separately.

**(A)** PCA of the 480 significantly altered transcript levels between *AS1* and wild-type lines at all three developmental stages (see Venn diagram in Figure 5). Changes during development allow discrimination of developmental stages, leading to three separated groups corresponding to bud (light-gray circle), anthesis (red and dark-gray circle), and postanthesis stages (green circle).

**(B)** PCA of metabolite data. Wild-type samples clustered in two groups corresponding to bud/anthesis (gray) and postanthesis (black), while *AS-IAA9* samples clustered into three groups corresponding to bud (yellow), anthesis (red), and postanthesis (green).

could provide a valuable resource for understanding how IAA9 exerts its function on gene transcription and subsequently on primary metabolite accumulation and developmental processes. The data published so far on *Aux/IAA* genes have provided relatively little information concerning their *in vivo* function and the *AS-IAA9* transgenic lines and therefore represent a unique genetic resource in which the IAA9-mediated auxin response is effectively constitutive. It should be noted that this study has been performed with the MicroTom cultivar of tomato. This cultivar is developing into a useful research tool with many studies, documenting that it is essentially equivalent to larger nonmutated cultivars (Obiadalla-Ali et al., 2004a; Dan et al., 2005; Tsugane et al., 2005). The MicroTom background harbors several genetic mutations, including a floral meristem determinancy mutation and a brassinosteroid mutation (Meissner et al., 1997). However, given that our previous work showed that the altered fruit set phenotype in the antisense lines was fully reproducible in another genetic background (Wang et al., 2005), the use of MicroTom should not greatly complicate the interpretation of the results presented here.

The nature of the signals and sequence of events that stimulate or limit the processes involved in fruit initiation remain largely unknown, and the genes and mechanisms involved in translating the signal for fruit initiation evoked either by natural pollination or application of hormones, such as auxin and GAs, are still obscure. The vast majority of studies on fruit initiation focused on hormonal regulation and parthenocarpic fruit set (Vivian-Smith and Koltunow, 1999; Pandolfini et al., 2002; Wang et al., 2005; Goetz et al., 2006; Marti et al., 2007; Serrani et al., 2007, 2008; de Jong et al., 2009), though the effect of altering the flavonoid pathway has also been reported to lead to parthenocarpic fruit (Schijlen et al., 2006). However, to date, relatively few studies have used genomic tools to look at the global transcriptomic changes associated with flower-to-fruit transition (Vriezen et al., 2008). Other studies implementing a transcriptomic approach focused on the events occurring from 8 d after pollination onward and addressed more specifically either the expansion phase (Lemaire-Chamley et al., 2005) or later fruit development and ripening phases (Alba et al., 2005; Carrari et al., 2006).

Recent work in tomato revealed that downregulation of *IAA9* results in the uncoupling of fruit set from pollination and fertilization, thus giving rise to seedless (parthenocarpic) fruit. These data indicated that *IAA9* is likely to participate within a regulatory complex that negatively regulates fruit initiation (Wang et al., 2005). *IAA9* is a member of the *Aux/IAA* family constituting short-lived transcription factors involved in auxin response (Abel et al., 1995; Gray et al., 2001). The rapid turnover of these proteins allows exquisite control of the auxin response in plants (Dharmasiri and Estelle, 2002; Dreher et al., 2006), with their levels being mediated at both transcriptional and posttranslational levels. In this study, we used *in situ* mRNA hybridization to allow spatial resolution of the expression gradient of *IAA9* during fruit organogenesis. This expression was elevated in the ovary and peaked in specific tissues, such as the placenta, funiculus, and the inner layer of integument in the embryo sac. Within the mature flower, a distinctive expression gradient was apparent, and upon successful pollination and fertilization, this gradient was released with *IAA9* expression being spread across all

tissues of the developing ovary/fruit. However, in the absence of fertilization, this gradient persisted, providing evidence that the initiation of normal fruit development and enlargement requires the dissipation of the *IAA9* mRNA gradient, thus suggesting an important regulatory role for *IAA9* in the process of fruit set. Although we have yet to examine *IAA9* protein levels, we speculate that prior to pollination and fertilization, *IAA9* transcripts (and proteins) are kept at a normal level but that following perception of the fertilization signal the *IAA9* protein is rapidly and extensively degraded. While the rapid postfertilization decline in transcript levels is in keeping with this hypothesis, further experiments will be necessary to directly delineate the mechanism linking the fertilization signal to alterations in *IAA9* transcript abundance and ultimately to fruit set and growth. The observation that *IAA9* gradients persist in unfertilized ovules indicates that the negative regulation of fruit set by *IAA9* is active throughout the flower-to-fruit transition, indicating a central role for the ovule in the mediation of fruit development.

The data documented both here and in our previous work (Wang et al., 2005) provide compelling arguments on the important role of auxin in the early events of tomato fruit development. It has been recently demonstrated that the interaction between *Aux/IAA* proteins and ARFs is instrumental in auxin-dependent transcriptional regulation (Tiwari et al., 2004; Dharmasiri et al., 2005). The identity of the exact ARF(s) putatively interacting with *IAA9* is currently unknown, but a likely candidate is the protein encoded by the tomato ortholog of *Arabidopsis* ARF8. Indeed, recent studies have revealed that the parthenocarpic mutant *fwf*, which harbors a lesion in *ARF8*, exhibits an uncoupling of fruit set and growth from pollination and fertilization events (Goetz et al., 2006). Notably, ectopic expression in the tomato of an aberrant form of the *Arabidopsis* *ARF8* induces parthenocarpic fruit set (Goetz et al., 2007). Under normal circumstances, pollination is known to induce increases in the levels of both auxin and ethylene in floral organs, correlating with the observation of subsequent growth (O'Neill, 1997; Llop-Tous et al., 2000). It is thus conceivable that in wild-type flowers, an auxin burst induced by pollination and fertilization leads to the degradation of *IAA9* protein via a proteolytic pathway such as that described for *Arabidopsis* (Woodward and Bartel, 2005). This scenario would thus be anticipated to abolish the repression of crucial auxin-responsive fruit initiation genes by the ARF-*IAA9* protein complex. In the *AS-IAA9* lines, however, such a mechanism of repression is clearly impaired. It seems likely that the very low abundance of *IAA9* in these lines is not capable of forming the inhibitory complex of ARF-*IAA9* and thus leads to constitutive activation of auxin-responsive and fruit initiation genes. The ultimate consequence of this lack of control is fruit set in the absence of fertilization and subsequent parthenocarpy. In support of an active role of auxin during the flower-to-fruit transition, our data reveal that many transcriptional regulators from both ARF and *Aux/IAA* type undergo dramatic shifts in their expression throughout natural fruit set, suggesting that the coordinated regulation of these genes is integral to this developmental process.

Comprehensive transcriptomic profiling of the flower-to-fruit transition identified a large number of genes that are common to both pollination-induced and pollination-independent fruit set, among which only a small subset are *IAA9* dependent. This

highlights the fact that a minimal change in gene expression can have a dramatic effect on the developmental fate of the flower organ. In contrast with pollination-induced fruit set, where the highest transition occurs between anthesis and postanthesis, the major shift in terms of changes in both transcript and metabolite accumulation in *AS-IAA9* occurs at the bud-to-anthesis transition. In accordance with this, functional categorization of the differentially expressed genes in both pollination-induced and pollination-independent fruit set reveals that the major groups of genes affected are related to signal transduction, metabolic pathways, and cell division. These results suggest that the fruit set process in tomato requires complex regulatory control and reconfiguration of metabolic processes. Evaluating the genes that are differentially expressed between wild-type and *AS-IAA9* lines during the flower-to-fruit transition revealed that the differences between the two lines are greatest at the anthesis stage. These data may explain why *AS-IAA9* enter the fruit differentiation process earlier than the wild type and with no requirement for pollination/fertilization signal.

Among the large number of transcription factors recruited during the process of fruit set in the wild type, the category of MADS box genes is the most striking regarding their sharp downregulation. Notably, *TAG1* and *TAGL6* genes undergo dramatic downregulation following fertilization, and the levels of their transcripts remain low throughout the early stages of fruit development. It has been reported that the *Agamous* gene takes part in the control of fruit organogenesis in *Arabidopsis* (Mizukami and Ma, 1992), and downexpression of *TM29*, another MADS box gene in the tomato, results in parthenocarpic fruit development (Ampomah-Dwamena et al., 2002). Moreover, downregulation of *TAG1* expression in tomato via antisense strategy results in dramatic alteration of flower development with conversion of stamens into petaloid organs and indeterminate growth of the floral meristem as well as in strong reduction of seed production (Pnueli et al., 1994). Given their dramatic downregulation during both pollination-dependent and -independent fruit set, *TAG1* and *TAGL6* genes emerge as potential active players in the process of fruit set. Moreover, since the transcript levels of both MADS box genes is low in *AS-IAA9* compared with the wild type, it is possible that this low expression promotes fruit set even in the absence of flower fertilization. Whether *IAA9* acts in concert with these MADS box genes to control fruit set initiation and whether auxin directly regulates their expression remains to be clarified.

In accordance with the precocious development of the ovary in *AS-IAA9*, a number of cell division-related genes show upregulation at anthesis stage in antisense lines, while their activation only occurs at postanthesis in wild-type lines. Following the same lines of evidence, the expression of two histone and two ribosomal protein genes was upregulated at both bud and anthesis stages in *AS-IAA9* but only at postanthesis in the wild type, again indicating that cell division may well have been boosted in the pollination-independent process.

The changes in transcript accumulation of a number of *ARF* and *Aux/IAA* genes are clearly indicative of active auxin signaling throughout the process of fruit set. Our data also reveal that auxin and ethylene-related genes are the most predominantly altered in terms of change in transcript accumulation with comparatively

only few changes concerning GA. This brings ethylene into the field as a potential major player in coordinating the process of fruit set. It has long been reported that flower pollination induces a burst of ethylene production (Pech et al., 1987), though the role of this in fruit set has not been resolved. More recently, a transcriptomic profiling of fruit set identified ethylene-related genes among the most altered during fruit set (Vriezen et al., 2008). Downregulation of auxin-related genes occurs earlier (2 DPA) than that of ethylene-related genes (6 DPA), suggesting the existence of a temporal dependence on hormone action for the flower-to-fruit transition to proceed properly.

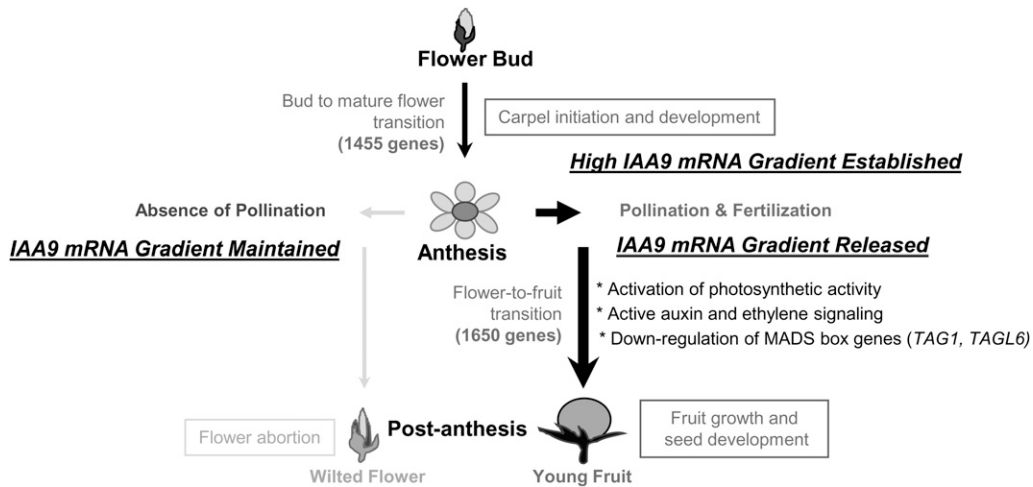
Another striking feature revealed by both metabolomic and transcriptomic profiling is related to aspects of photosynthesis and sugar metabolism during the flower-to-fruit transition. Given the widely documented crosstalk between hormone and sugar signaling (Arenas-Huertero et al., 2000; Leon and Sheen, 2003), it is highly interesting that the antisense lines were additionally characterized as being upregulated in the sucrose synthase pathway of sucrose degradation. This pathway is more energy efficient than that mediated by invertase (Bologa et al., 2003) and is likely the prominent pathway of sucrose degradation in later stages of development of many heterotrophic tissues, including tomato (Hackel et al., 2006). Thus, the observed elevation of the transcripts of this pathway could, at least partially, explain the precocious nature of fruit development in the transgenics. Two possible explanations could account for the observed elevation in the levels of sucrose, glucose, and fructose in the antisense lines: (1) it could be the consequence of a more efficient downloading of photoassimilate from the fruit, as observed on the introgression of a wild species allele of cell wall invertase (Fridman et al., 2004), or (2) it could be due to an increase in the fruits own photoassimilate production. The strong activation of photosynthesis-related genes during fruit set in *IAA9* downregulated lines is a major phenomenon. Strikingly, the activation of photosynthesis-related genes is delayed to the postanthesis stage in the wild type but takes place at the bud stage in *AS-IAA9*. Many recent studies have endorsed the prevailing opinion that fruit growth and metabolism are predominantly supported by photoassimilate supply from source tissues (Farrar et al., 2000; Nunes-Nesi et al., 2005; Schauer et al., 2006). However, it is important to note that the carpel of the fruit is essentially a modified leaf that has folded into a tubular structure enclosing the ovules (Bowman et al., 1991; Gillaspay et al., 1993) and that cells in developing fruit contain photosynthetically active chloroplasts and express both nuclear-encoded and plastidially encoded genes for photosynthetic proteins (Piechulla et al., 1987). Furthermore, combination of indirect evidence provides support for the hypothesis that during early developmental fruit, photosynthesis itself may provide a considerable contribution to both metabolism and growth of the organ (Tanaka et al., 1974; Obiadalla-Ali et al., 2004b). Recent studies have indicated that in tomato pericarp cells, the induction of genes related to photosynthesis and chloroplast biogenesis positively correlate with chloroplast numbers and cell size (Kolotilin et al., 2007). The results of this study provide compelling evidence that fruit photosynthesis plays an important role in fruit establishment. It is likely that the end phenotype of the lines is additionally influenced by the upregulation of the sucrose synthase pathway



of sucrose degradation since alterations of sucrose metabolism have often been correlated with changes in yield parameters (Lytovchenko et al., 2007). The combined data provide compelling evidence that photosynthesis is upregulated at the level of gene expression and that this is concurrent with a precocious fruit set and growth advantage in the antisense lines. It is dangerous to conclude causality on the basis of consistency analysis alone; therefore, we are generally not inclined to speculate as to whether the transcriptional change causes the metabolic one or vice versa. However, since it is extensively documented that high sugar contents repress photosynthesis (for review, see Rolland et al., 2006), in this instance, we can effectively exclude that the observed increases in sugars cause the observed increases in transcription of genes associated with photosynthesis. When taken alongside the altered growth rate, these data convincingly suggest that the transcriptional upregulation of photosynthesis is a key event required to sustain the fruit set developmental process. Consistent with this hypothesis is the fact that other features of the metabolite profiles of the transgenics reflect conditions that could be anticipated to promote photosynthesis and growth; for example, the observed changes in ascorbate and carbon nitrogen metabolism are diagnostic of conditions favoring efficient photosynthesis (for example, see Nunes-Nesi et al., 2005), while the elevation of intermediates of both ascorbate and shikimate pathways in the transformants are additionally consistent with the exhibited elevated growth. Ascorbate metabolism is one of the few pathways in which a clear association between changes at the levels of transcripts correlates with changes in metabolites. While the structural features of plant ascorbate biosynthesis are currently being clarified (for reviews, see Smirnov et al., 2000; Ishikawa

et al. 2006), the results from this study could well shed light on the regulation of this important pathway. At the transcriptional level, both galactose-dehydrogenase (L-Gal-DH) and GDP-L-galactose-hexose-1-phosphate guanyltransferase displayed clear differences that correlated well with the absolute ascorbate level, implying that these enzymes are likely highly important in determining the final level of this metabolite. The latter enzyme was recently cloned and suggested to be very important in determining the final ascorbate content (Laing et al., 2007). The other clear example of transcriptional regulation of metabolite content is provided by the polyamines, which seem to be particularly important in pollination-dependent fruit set. The inclusion of this class of metabolites as important regulators of fruit development is not without precedent since they have long been described to play an ethylene-dependent (Saftner and Baldi, 1990) role in flowering and fruit ripening (Kakkar and Rai, 1993). While more recent studies have revealed that engineering high fruit polyamine content enhances both the vine life and the nutritional content of tomato (Mehta et al., 2002), it has been recognized that pollination-independent fruit set in tomato is accompanied by a change in polyamine metabolism and that polyamine metabolism in general proves to be an interesting example for metabolite-enzyme activity correlation (Egea-Cortines et al., 1993). Despite the identification of these interesting correlations between transcript and metabolite level, it is clear that understanding the mechanisms of regulation of both these pathways and assessing their functional importance with respect to the fruit set developmental process will require substantial further work.

So far, only few molecular components of the flower-to-fruit transition have been described (Wang et al., 2005; Goetz et al.,



**Figure 15.** Model for Regulatory Events Underlying the Fruit Set Process in Tomato.

*IAA9* regulates the initiation of fruit set by establishing a spatial expression gradient whose release triggers the flower-to-fruit transition. Transcriptomic and metabolomic analysis reveal novel regulatory points in pathways associated with natural pollination fruit set and pollination-independent fruit set. The comparative analysis at the transcriptomic and metabolic levels of pollination-induced and fertilization-free fruit set identifies auxin and ethylene signaling as well as photosynthesis and sugar metabolism as major events of the fruit set program and potential components of the regulatory mechanism underlying this developmental process. The downregulation of MADS box transcription factors (*TAG1* and *TAGL6*) also emerges as a key event of the fruit set process.

2006; Pandolfini et al., 2007), and little is known concerning the overarching regulatory mechanisms that underpin this transition. Our study brings new insights not only on global biochemical and molecular events underlying this process, but also on their hormonal regulation. The model presented in Figure 15 highlights the role of *IAA9* expression and distribution in triggering ovary development into young fruit. Pollination induces a net release of the *IAA9* mRNA gradient that leads to the activation of the fruit set program, including changes in gene expression and metabolite accumulation. Of particular note were the concerted changes at transcript and metabolite levels of photosynthetic and sugar metabolism. The coupling of these changes to increased growth suggests that they may also be components of the control of the initiation of fruit development. In parallel, both auxin and ethylene signaling emerge as the main hormonal regulation events involved in the flower-to-fruit transition; however, the contribution albeit to a lesser extent of other phytohormones cannot be excluded. When taken together the combined data reported here suggest that *IAA9* exerts a profound influence on the flower-to-fruit transition by, either directly or indirectly, affecting a substantial number of transcription factors belonging to ARF and Aux/IAA families and by downregulating the expression of *TAG1* and *TAGL6* MADS box genes. However, although *IAA9* only affects a small number of transcriptional events, these changes provoke dramatic changes both at the levels of development and primary metabolism during the period of fruit set and subsequent early fruit growth. As such, these results allow a far greater comprehension of the molecular events controlling early fruit development both in the presence and absence of fertilization.

## METHODS

### Plant Material and Experimental Design

Tomato plants (*Solanum lycopersicum* cv MicroTom) and two independent homozygote lines of AS-*IAA9* were grown under standard greenhouse conditions (14-h-day/10-h-night cycle, 25/20°C day/night temperature, 80% humidity, and 250  $\mu\text{mol m}^{-2} \text{s}^{-1}$  light intensity). Flower emasculation was performed before dehiscence of anthers (closed flowers) to avoid accidental self-pollination. At all three stages (bud, anthesis, and postanthesis), only the ovary, including style and stigma, or developing young fruits were collected as samples. At the flower bud stage, the ovaries were collected just before anthesis (closed flowers). Anthesis stage samples were collected on the first day of flower opening. Stage-specific pooled samples (>50 ovary/fruit) were divided into two groups, with one group being used for transcriptome profiling experiments and the other one for metabolite profiling experiments.

### Histological Analysis

For histological analysis, flower buds of 0.5 to 8 mm in length were fixed in FAA solution (4% paraformaldehyde, 50% ethanol, and 5% acetic acid in 1 $\times$  PBS), placed under vacuum for 10 min, and incubated overnight at 4°C before being dehydrated in alcohol and embedded in paraffin (Paraplast plus; Sigma-Aldrich). At least 10 buds were sampled and checked for each stage of development. Histological preparations were performed according to Baldet et al. (2006). For histological analysis, 80- $\mu\text{m}$ -thick sections were stained with 0.05% toluidine blue. Slides were observed under a microscope (Zeiss-Axioplan).

### In Situ Hybridization

A 246-bp fragment from the 3' untranslated region of *IAA9* was amplified with primers *IAA9\_inF* (5'-AGGGCTATGGAAAAGTGTCGGAGCAGAAAT-3') and *IAA9\_inR* (5'-AACTTAAAGAGGACATATATTACGCA-3') and cloned into pGEM-T easy. Plasmids were sequenced to verify identity and orientation of inserts. Sense (control) and antisense digoxigenin-labeled probes were generated by run-off transcription using SP6 RNA polymerases according to the manufacturer's protocol (Roche Diagnostics). For in situ hybridization, tomato flower buds were sampled and processed as described by Baldet et al. (2006).

### RNA Extraction

Total RNA was isolated from ovary/young fruits using an RNA extraction kit (RNeasy plant mini kit; Qiagen) and was DNase-treated according to the manufacturer's protocol (RNase-Free DNase Set; Qiagen).

### Microarray Experiment Analysis

For each developmental stage collected, we did direct comparison of AS-*IAA9* line AS1 with their normal counterpart using two-color hybridizations onto EU-TOM1 microarray slides. The EU-TOM1 chip contains 11,860 different 70-mers oligonucleotides. The majority of the probes were designed from gene sequences gathered from the Lycopersicon Combined Build #3 Unigene database at Cornell University ([http://www.sgn.cornell.edu/unigene\\_builds/](http://www.sgn.cornell.edu/unigene_builds/)).

Labeled cDNA was prepared according to the Pronto Plus direct system protocol (Promega). Five micrograms of total RNA was used for each labeling reaction. Labeled probes were annealed to the target oligonucleotides for 14 to 16 h at 42°C. Arrays were washed twice at 42°C (90 s and 5 min, respectively) in 2 $\times$  SSC/0.1% SDS and then twice for 2 min in 1 $\times$  SSC and twice for 1 min in 0.1 $\times$  SSC at ambient temperature. Slides were then dried by centrifugation (5 min to 800 rpm). For each of the three stages, three biological repeats were performed in dye-swap experiments (total of 18 hybridizations). Dye bias was minimized by conducting half of the replicate hybridizations with wild-type cDNAs labeled with Cy3 and half with Cy5. Images were acquired using a Genepix 4000A microarray scanner (Molecular Devices) at 10- $\mu\text{m}$  resolution per pixel, adjusting the photomultiplier tension to achieve optimal distribution of the signal and few number of saturated spots. Quantification was performed with Genepix Pro 3 software (Molecular Devices).

To identify genes that are differentially expressed between wild-type and AS-*IAA9* lines, analysis was done in R (<http://www.r-project.org/>) using modules from the BioConductor R/MAANOVA package (Kerr et al., 2000). Microarray intensities were  $\log_2$  transformed and normalized using LOWESS algorithm to remove intraslide dye bias. This preprocessed data set was then analyzed using the R/MAANOVA package with a fixed model and array, dye, and sample (ASvsWT) as covariables. For selection of differentially expressed genes, false discovery rate corrected permutation P values (500 permutations) were employed (Benjamini and Hochberg, 2001). Among these genes, only those displaying a  $\log_2$  (ratio) higher than 0.5 or lower than -0.5 were retained, which corresponds to fold difference >1.4. R/MAANOVA analysis was selected as the statistical method of choice since it permits the separate consideration of array and dye effects as experimental covariables, omitting the requirement for interslide normalization approaches, notoriously known to introduce artificial bias (Woo et al., 2005). In the case, where the development of only the wild-type or only the antisense lines were considered, the background was subtracted from the foreground signal; afterwards, the red and green channels were decoupled and either all channels describing antisense line or all channels describing wild-type expression were  $\log_2$  transformed and these values were quantile normalized. To these quantile normalized values a linear model was fitted using the limma package

taking technical replicates into account, and P values were determined using the empirical Bayes procedure as implemented in the limma package (Smyth, 2004).

### Real-Time qRT-PCR Reactions

The same total RNA samples extracted for the microarray experiment were used to synthesize cDNA templates required for qRT-PCR analysis. cDNA synthesis was performed using a qRT-PCR cDNA synthesis kit (Ominiscript Reverse Transcription; Qiagen). Gene-specific primers were designed using Primer Express software Applied Biosystems (see Supplemental Table 8 online). qRT-PCRs were performed using SYBR GREEN PCR master mix (Applied Biosystems) on an ABI PRISM 7900HT sequence detection system. For all real-time PCR experiments, at least three biological replicates were performed, and each reaction was run in triplicate. Relative fold differences were calculated based on the comparative Ct (threshold constant) method using Sl-Actin-51 (accession number Q96483) as an internal standard with primers Actin\_F (5'-TGT-CCCTATTACGAGGGTTATGC-3') and Actin\_R (5'-AGTTAAATCAC-GACCAGCAAGAT-3'). To determine relative fold differences for each sample in each experiment, the Ct value for each gene was normalized to the Ct value for Sl-Actin-51 and was calculated relative to a calibrator using the equation  $2^{-\Delta\Delta C_t}$ .

### Metabolite Analysis

Metabolite extraction was performed as described previously (Roessner et al., 2001; Schauer et al., 2006). One hundred milligrams of tomato tissue from the three developmental stages were homogenized using a ball mill precooled with liquid nitrogen. Derivatization and GC-MS analysis were performed as described previously (Lisec et al., 2006). The GC-MS system was comprised of a CTC CombiPAL autosampler, an Agilent 6890N gas chromatograph, and a LECO Pegasus III TOF-MS running in EI+ mode. Metabolites were identified in comparison to database entries of authentic standards (Schauer et al., 2005b). Recovery experiments concerning the validity of the extraction and processing have been documented previously (Schauer et al., 2005a). If two observations from the metabolite data set are described in the text as different, this means that their difference was determined to be statistically significant ( $P < 0.05$ ) by the performance of Student's *t* test using the algorithm incorporated into Microsoft Excel 7.0. PCA was performed using the informatic program MeV 4.0 (Saeed et al., 2003)

### Metabolite-Transcript Consistency Analyses

To find links between metabolites and transcripts, a consistency analysis was performed on the basis of biological pathway knowledge essentially as detailed by Gibon et al. (2006). For this purpose, metabolites that were differentially abundant under the experimental conditions evaluated were identified, and transcripts coding for enzymes involved in the metabolism of the compound in question were extracted from the entire data sets using MapMan (Usadel et al., 2005). Subsequently, the tomato mapping files for MapMan (Urbanczyk-Wochniak et al., 2005) were used to visualize relevant pathways and to judge consistency between changes in metabolite and transcript levels by displaying both levels of information at the same time. In brief, the data sets were loaded into PageMan. First, the whole data set, including metabolites and transcripts, was analyzed for a consistent up- or downregulation. This was done Bin and subbin-wise using a Wilcoxon test with Benjamini Hochberg type false discovery rate control using the built-in PageMan function. The metabolite data were then added separately by averaging bin-wise, thus enabling the display of metabolites individually but also the visualization of their contribution to individual bins.

### Accession Number

Microarray data have been deposited in the Array-Express public repository with the access number E-MEXP-1617.

### Supplemental Data

The following materials are available in the online version of this article.

**Supplemental Figure 1.** Visualization of Ala Synthesis.

**Supplemental Figure 2.** Visualization of His Synthesis.

**Supplemental Figure 3.** Visualization of Asp Metabolism.

**Supplemental Table 1.** Differentially Expressed Genes during Natural Pollination-Dependent Fruit Set in the Wild Type.

**Supplemental Table 2.** Differentially Expressed Genes during Pollination-Independent Fruit Set in AS-IAA9 Lines.

**Supplemental Table 3.** Differentially Expressed Genes between AS-IAA9 Lines and the Wild Type in Ovary/Young Fruit.

**Supplemental Table 4.** Differentially Expressed Genes during Both Pollination-Dependent and -Independent Fruit Set.

**Supplemental Table 5.** IAA9 Regulated and Differentially Expressed Genes during both Pollination-Dependent and -Independent Fruit Set.

**Supplemental Table 6.** Relative Metabolites Contents during Pollination-Dependent and -Independent Fruit Set.

**Supplemental Table 7.** Validation of Microarray Expression Data by Quantitative RT-PCR.

**Supplemental Table 8.** Primers Used for qRT-PCR and the Accession Numbers of the Relevant Genes.

### ACKNOWLEDGMENTS

This research was supported by the European Commission (EU-SOL Project PI 016214) and by the Midi-Pyrénées Regional Council (Grants 06003789 and 07003760). A.R.F. and M.B. were partially supported by the France-Germany-Spain Trilateral, and A.R.F. and N.S. were supported by a Deutsche-Israeli Project grant. We thank L. Tessarotto, H. Mondès, O. Berseille, and D. Saint-Martin (Institut National Polytechnique Toulouse) for tomato genetic transformation and plant growth, L. Trouilhet (Toulouse Midi Pyrénées Genopole) for the production of tomato oligomicroarray slides, and S. Sokol (Toulouse Midi Pyrénées Genopole) for helping in microarray data analysis.

Received May 16, 2008; revised March 16, 2009; accepted April 24, 2009; published May 12, 2009.

### REFERENCES

- Abel, S., Nguyen, M.D., and Theologis, A.** (1995). The PS-IAA4/5-like family of early auxin-inducible mRNAs in *Arabidopsis thaliana*. *J. Mol. Biol.* **251**: 533–549.
- Alba, R., Payton, P., Fei, Z., McQuinn, R., Debbie, P., Martin, G.B., Tanksley, S.D., and Giovannoni, J.J.** (2005). Transcriptome and selected metabolite analyses reveal multiple points of ethylene control during tomato fruit development. *Plant Cell* **17**: 2954–2965.
- Ampomah-Dwamena, C., Morris, B.A., Sutherland, P., Veit, B., and Yao, J.L.** (2002). Down-regulation of TM29, a tomato SEPALLATA

- homolog, causes parthenocarpic fruit development and floral reversion. *Plant Physiol.* **130**: 605–617.
- Arenas-Huertero, F., Arroyo, A., Zhou, L., Sheen, J., and Leon, P.** (2000). Analysis of *Arabidopsis* glucose insensitive mutants, *gin5* and *gin6*, reveals a central role of the plant hormone ABA in the regulation of plant vegetative development by sugar. *Genes Dev.* **14**: 2085–2096.
- Baldet, P., Hernould, M., Laporte, F., Mounet, F., Just, D., Mouras, A., Chevalier, C., and Rothan, C.** (2006). The expression of cell proliferation-related genes in early developing flowers is affected by a fruit load reduction in tomato plants. *J. Exp. Bot.* **57**: 961–970.
- Benjamini, Y., and Hochberg, Y.** (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* **29**: 1185–1188.
- Bologna, K.L., Fernie, A.R., Leisse, A., Loureiro, M.E., and Geigenberger, P.** (2003). A bypass of sucrose synthase leads to low internal oxygen and impaired metabolic performance in growing potato tubers. *Plant Physiol.* **132**: 2058–2072.
- Bowman, J.L., Smyth, D.R., and Meyerowitz, M.** (1991). Genetic interactions among floral homeotic genes of *Arabidopsis*. *Development* **112**: 1–20.
- Carmi, N., Salts, Y., Dedicova, B., Shabtai, S., and Barg, R.** (2003). Induction of parthenocarpy in tomato via specific expression of the *rolB* gene in the ovary. *Planta* **217**: 726–735.
- Carrari, F., Baxter, C., Usadel, B., Urbanczyk-Wochniak, E., Zanor, M.I., Nunes-Nesi, A., Nikiforova, V., Centero, D., Ratzka, A., Pauly, M., Sweetlove, L.J., and Fernie, A.R.** (2006). Integrated analysis of metabolite and transcript levels reveals the metabolic shifts that underlie tomato fruit development and highlight regulatory aspects of metabolic network behavior. *Plant Physiol.* **142**: 1380–1396.
- Coombe, B.G.** (1960). Relationship of growth and development to changes in sugars, auxins, and gibberellins in fruit of seeded and seedless varieties of *Vitis vinifera*. *Plant Physiol.* **35**: 241–250.
- Dan, Y.H., Yan, H., Muniyikwa, T., Dong, J., Zhang, Y.L., and Armstrong, C.L.** (2005). MicroTom a high throughput model transformation system for functional genomics. *Plant Cell Rep.* **25**: 432–441.
- de Jong, M., Wolters-Arts, M., Feron, R., Mariani, C., and Vriezen, W.H.** (2009). The *Solanum lycopersicum* auxin response factor 7 (SIARF7) regulates auxin signaling during tomato fruit set and development. *Plant J.* **57**: 160–170.
- Dharmasiri, N., Dharmasiri, S., and Estelle, M.** (2005). The F-box protein TIR1 is an auxin receptor. *Nature* **435**: 441–445.
- Dharmasiri, S., and Estelle, M.** (2002). The role of regulated protein degradation in auxin response. *Plant Mol. Biol.* **49**: 401–409.
- Dreher, K.A., Brown, J., Saw, R.E., and Callis, J.** (2006). The *Arabidopsis* Aux/IAA protein family has diversified in degradation and auxin responsiveness. *Plant Cell* **18**: 699–714.
- Dudareva, N., Cseke, L., Blanc, V.M., and Pichersky, E.** (1996). Evolution of floral scent in *Clarkia*: Novel patterns of S-linalool synthase gene expression in the *C. breweri* flower. *Plant Cell* **8**: 1137–1148.
- Dunphy, P.J.** (2006). Location and biosynthesis of monoterpenyl fatty acyl esters in rose petals. *Phytochemistry* **67**: 1110–1119.
- Egea-Cortines, M., Cohen, E., Arad, S., Bagni, N., and Mizrahi, Y.** (1993). Polyamine levels in pollinated and auxin-induced fruit of the tomato (*Lycopersicon esculentum*) during development. *Physiol. Plant.* **87**: 14–20.
- Farrar, J.F., Pollock, C., and Gallagher, J.** (2000). Sucrose and the integration of metabolism in vascular plants. *Plant Sci.* **154**: 1–11.
- Fridman, E., Carrari, F., Lui, Y.S., Fernie, A.R., and Zamir, D.** (2004). Zooming in on a quantitative trait for tomato yield using interspecific introgressions. *Science* **305**: 1786–1789.
- George, W.L., Scott, J.W., and Splittstoesser, W.E.** (1984). Parthenocarpy in tomato. *Hortic. Rev. (Am. Soc. Hortic. Sci.)* **6**: 65–84.
- Gibon, Y., Usadel, B., Blaessing, O.E., Kamlage, B., Hoehne, M., Trethewey, R.N., and Stitt, M.** (2006). Integration of metabolite with transcript and enzyme activity profiling during diurnal cycles in *Arabidopsis* rosettes. *Genome Biol.* **7**: R76.
- Gillaspy, G., Ben-David, H., and Grissem, W.** (1993). Fruits: A developmental perspectives. *Plant Cell* **5**: 1439–1451.
- Giovannoni, J.** (2001). Molecular biology of fruit maturation and ripening. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **52**: 725–749.
- Goetz, M., Hooper, L.C., Johnson, S.D., Rodrigues, J.C., Vivian-Smith, A., and Koltunow, A.M.** (2007). Expression of aberrant forms of AUXIN RESPONSE FACTOR8 stimulates parthenocarpy in *Arabidopsis* and tomato. *Plant Physiol.* **145**: 351–366.
- Goetz, M., Vivian-Smith, A., Johnson, S.D., and Koltunow, A.M.** (2006). AUXIN RESPONSE FACTOR8 is a negative regulator of fruit initiation in *Arabidopsis*. *Plant Cell* **18**: 1873–1886.
- Gorguet, B., van Heusden, A.W., and Lindhout, P.** (2005). Parthenocarpic fruit development in tomato. *Plant Biol (Stuttg.)* **7**: 131–139.
- Gray, W.M., Kepinski, S., Rouse, D., Leyser, O., and Estelle, M.** (2001). Auxin regulates SCF(TIR1)-dependent degradation of AUX/IAA proteins. *Nature* **414**: 271–276.
- Gustafson, F.G.** (1936). Inducement of fruit development by growth promoting chemicals. *Proc. Natl. Acad. Sci. USA* **22**: 628–636.
- Hackel, A., Schauer, N., Carrari, F., Fernie, A.R., Grimm, B., and Kuhn, C.** (2006). Sucrose transporter LeSUT1 and LeSUT2 inhibition affects tomato fruit development in different ways. *Plant J.* **45**: 180–192.
- Ishikawa, T., Dowdle, J., and Smirnoff, N.** (2006). Progress in manipulating ascorbic acid and accumulation in plants. *Physiol. Plant.* **126**: 343–355.
- Kakkar, R.K., and Rai, V.K.** (1993). Plant polyamines in flowering and fruit ripening. *Phytochemistry* **33**: 1281–1288.
- Kerr, M.K., Martin, M., and Churchill, G.A.** (2000). Analysis of variance for gene expression microarray data. *J. Comput. Biol.* **7**: 819–837.
- Kolotilin, I., Koltai, H., Tadmor, Y., Bar-Or, C., Reuveni, M., Meir, A., Nahon, S., Shlomo, H., Chen, L., and Levin, I.** (2007). Transcriptional profiling of high pigment-2dg tomato mutant links early fruit plastid biogenesis with its overproduction of phytonutrients. *Plant Physiol.* **145**: 389–401.
- Laing, W.A., Wright, M.A., Cooney, J., and Bulley, S.M.** (2007). The missing step of the L-galactose pathway of ascorbate biosynthesis in plants, an L-galactose guanylyltransferase, increases leaf ascorbate content. *Proc. Natl. Acad. Sci. USA* **104**: 9534–9539.
- Lemaire-Chamley, M., Petit, J., Garcia, V., Just, D., Baldet, P., Germain, V., Fagard, M., Mouassite, M., Cheniclet, C., and Rothan, C.** (2005). Changes in transcriptional profiles are associated with early fruit tissue specialization in tomato. *Plant Physiol.* **139**: 750–769.
- Leon, P., and Sheen, J.** (2003). Sugar and hormone connections. *Trends Plant Sci.* **8**: 110–116.
- Lisee, J., Schauer, N., Kopka, J., Willmitzer, L., and Fernie, A.R.** (2006). Gas chromatography mass spectrometry-based metabolite profiling in plants. *Nat. Protocols* **1**: 387–396.
- Llop-Tous, I., Barry, C.S., and Grierson, D.** (2000). Regulation of ethylene biosynthesis in response to pollination in tomato flowers. *Plant Physiol.* **123**: 971–978.
- Lytovchenko, A., Sonnewald, U., and Fernie, A.R.** (2007). The complex network of non-cellulosic carbohydrate metabolism. *Curr. Opin. Plant Biol.* **10**: 227–235.
- Marti, C., Orzaez, D., Ellul, P., Moreno, V., Carbonell, J., and Granell, A.** (2007). Silencing of DELLA induces facultative parthenocarpy in tomato fruits. *Plant J.* **52**: 865–876.

- Mehta, R.A., Cassol, T., Li, N., Ali, N., Hanada, A.K., and Mattoo, A.K.** (2002). Engineered polamine accumulation in tomato enhances phytonutrient content, juice quality, and vine life. *Nat. Biotechnol.* **20**: 613–618.
- Meissner, R., Jacobson, Y., Melame, S., Levyatuv, S., Shalev, G., Ashri, A., Elkind, Y., and Levy, A.** (1997). A new model system for tomato genetics. *Plant J.* **12**: 1465–1472.
- Mizukami, Y., and Ma, H.** (1992). Ectopic expression of the floral homeotic gene AGAMOUS in transgenic Arabidopsis plants alters floral organ identity. *Cell* **71**: 119–131.
- Moyano, E., Martinez-Garcia, J.F., and Martin, C.** (1996). Apparent redundancy in myb gene function provides gearing for the control of flavonoid biosynthesis in antirrhinum flowers. *Plant Cell* **8**: 1519–1532.
- Nitsch, J.P.** (1970). Hormonal factors in growth and development. In *The Biochemistry of Fruits and Their Products*, A.C. Hulme, ed (New York: Academic Press), pp. 427–472.
- Nunes-Nesi, A., Carrari, F., Lytovchenko, A., Smith, A.M.O., Loureiro, M.E., Ratcliffe, R.G., Sweetlove, L.J., and Fernie, A.R.** (2005). Enhanced photosynthetic performance and growth as a consequence of decreasing mitochondrial malate dehydrogenase activity in transgenic tomato plants. *Plant Physiol.* **137**: 611–622.
- Obiadalla-Ali, H., Fernie, A.R., Kossmann, J., and Lloyd, J.R.** (2004a). Developmental analysis of carbohydrate metabolism in tomato (*Lycopersicon esculentum* cv. Micro-Tom) fruits. *Physiol. Plant.* **120**: 196–204.
- Obiadalla-Ali, H., Fernie, A.R., Lytovchenko, A., Kossmann, J., and Lloyd, J.R.** (2004b). Inhibition of chloroplastic fructose 1,6-bisphosphate in tomato fruits leads to decreased fruit size, but only small changes in carbohydrate metabolism. *Planta* **219**: 533–540.
- O'Neill, S.D.** (1997). Pollination regulation of flower development. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **48**: 547–574.
- Ozga, J.A., van Huizen, R., and Reinecke, D.M.** (2002). Hormone and seed-specific regulation of pea fruit growth. *Plant Physiol.* **128**: 1379–1389.
- Pandolfini, T., Molesini, B., and Spena, A.** (2007). Molecular dissection of the role of auxin in fruit initiation. *Trends Plant Sci.* **12**: 327–329.
- Pandolfini, T., Rotino, G.L., Camerini, S., Defez, R., and Spena, A.** (2002). Optimisation of transgene action at the post-transcriptional level: high quality parthenocarpic fruits in industrial tomatoes. *BMC Biotechnol.* **2**: 1.
- Pech, J.C., Latche, A., Larrigaudière, C., and Reid, M.S.** (1987). Control of early ethylene synthesis in pollinated petunia flowers. *Plant Physiol. Biochem.* **25**: 431–437.
- Picken, A.J.F.** (1984). A review of pollination and fruit set in the tomato (*Lycopersicon esculentum* Mill). *J. Hort. Sci.* **59**: 1–13.
- Piechulla, B., Glick, R.E., Bahl, H., Melis, A., and Grisse, W.** (1987). Changes in photosynthetic capacity and photosynthetic protein pattern during tomato fruit ripening. *Plant Physiol.* **84**: 911–917.
- Pnueli, L., Hareven, A.D., Rounsley, S.D., Yanofsky, M.F., and Lifschitz, E.** (1994). Isolation of the tomato agamous gene TAG1 and analysis of its homeotic role in transgenic plants. *Plant Cell* **6**: 163–173.
- Potts, W.C., Reid, J.B., and Murfet, I.C.** (1985). Internode length in *Pisum*: Gibberellins and the slender phenotype. *Physiol. Plant.* **63**: 357–364.
- Reiter, W.D.** (2008). Biochemical genetics of nucleotide sugar interconversion reactions. *Curr. Opin. Plant Biol.* **11**: 236–243.
- Roessner, U., Luedemann, A., Brust, D., Fiehn, O., Linke, T., Willmitzer, L., and Fernie, A.** (2001). Metabolic profiling allows comprehensive phenotyping of genetically or environmentally modified plant systems. *Plant Cell* **13**: 11–29.
- Rolland, F., Baena-Gonzalez, E., and Sheen, J.** (2006). Sugar sensing and signalling in plants: Conserved and novel mechanisms. *Annu. Rev. Plant Biol.* **57**: 675–709.
- Rotino, G.L., Acciarri, N., Sabatini, E., Mennella, G., Lo Scalzo, R., Maestrelli, A., Molesini, B., Pandolfini, T., Scalzo, J., Mezzetti, B., and Spena, A.** (2005). Open field trial of genetically modified parthenocarpic tomato: Seedlessness and fruit quality. *BMC Biotechnol.* **5**: 32.
- Saeed, A., et al.** (2003). TM4: A free, open-source system for microarray data management and analysis. *Biotechniques* **34**: 374–378.
- Saftner, B.A., and Baldi, B.G.** (1990). Polyamine levels and tomato fruit development - possible interaction with ethylene. *Plant Physiol.* **92**: 547–550.
- Schauer, N., et al.** (2006). Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nat. Biotechnol.* **24**: 447–454.
- Schauer, N., Steinhauser, D., Strelkov, S., Schomburg, D., Allison, G., Moritz, T., Lundgren, K., Roessner-Tunali, U., Forbes, M.G., Willmitzer, L., Fernie, A.R., and Kopka, J.** (2005b). GC-MS libraries for the rapid identification of metabolites in complex biological samples. *FEBS Lett.* **579**: 1332–1337.
- Schauer, N., Zamir, D., and Fernie, A.R.** (2005a). Metabolic profiling of leaves and fruit of wild species tomato: a survey of the *Solanum lycopersicum* complex. *J. Exp. Bot.* **56**: 297–307.
- Schijlen, E., Ric de Vos, C.H., Jonker, H., van den Broeck, H., Molthoff, J., van Tunen, A., Martens, S., and Bovy, A.** (2006). Pathway engineering for healthy phytochemicals leading to the production of novel flavonoids in tomato fruit. *Plant Biotechnol. J.* **4**: 433–444.
- Sedgley, M., and Griffin, A.R.** (1989). *Sexual Reproduction of Tree Crops*. (London: Academic Press).
- Serrani, J.C., Ruiz-Rivero, O., Fos, M., and Garcia-Martinez, J.L.** (2008). Auxin-induced fruit-set in tomato is mediated in part by gibberellins. *Plant J.* **56**: 922–934.
- Serrani, J.C., Sanjuan, R., Ruiz-Rivero, O., Fos, M., and Garcia-Martinez, J.L.** (2007). Gibberellin regulation of fruit set and growth in tomato. *Plant Physiol.* **145**: 246–257.
- Smirnoff, N., Conklin, P.L., and Loewus, F.A.** (2000). Biosynthesis of ascorbate acid in plants: A renaissance. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **52**: 437–459.
- Smyth, G.K.** (2004). Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3** (online) doi/10.2202/1544-6115.1027.
- Talon, M., Zacarias, L., and Primo-Millo, E.** (1992). Gibberellins and parthenocarpic ability in developing ovaries of seedless mandarins. *Plant Physiol.* **99**: 1575–1581.
- Tanaka, A., Fujita, K., and Kikuchi, K.** (1974). Nutriophysiological studies on the tomato plant, III. Photosynthetic rate of individual leaves in relation to the dry matter production of plants. *Soil Sci. Plant Nutr.* **20**: 173–183.
- Tiwari, S.B., Hagen, G., and Guilfoyle, T.J.** (2004). Aux/IAA proteins contain a potent transcriptional repression domain. *Plant Cell* **16**: 533–543.
- Tsugane, T., Watanabe, M., Yano, K., Sakurai, N., Suzuki, H., and Shibata, D.** (2005). Expressed sequence tags of full-length cDNA clones from the miniature tomato (*Lycopersicon esculentum*) cultivar Micro-Tom. *Plant Biotechnol.* **22**: 161–165.
- Urbanczyk-Wochniak, E., et al.** (2005). Conversion of MapMan to allow the analysis of transcript data from Solanaceous species: Effects of genetic and environmental alterations in energy metabolism in the leaf. *Plant Mol. Biol.* **60**: 773–792.
- Usadel, B., et al.** (2005). Extension of the visualization tool MapMan to allow statistical analysis of arrays, display of corresponding genes, and comparison with known responses. *Plant Physiol.* **138**: 1195–1204.

- Vivian-Smith, A., and Koltunow, A.M.** (1999). Genetic analysis of growth-regulator-induced parthenocarpy in *Arabidopsis*. *Plant Physiol.* **121**: 437–451.
- Vivian-Smith, A., Luo, M., Chaudhury, A., and Koltunow, A.** (2001). Fruit development is actively restricted in the absence of fertilization in *Arabidopsis*. *Development* **128**: 2321–2331.
- Vriezen, W.H., Feron, R., Maretto, F., Keijman, J., and Mariani, C.** (2008). Changes in tomato ovary transcriptome demonstrate complex hormonal regulation of fruit set. *New Phytol.* **177**: 60–76.
- Wang, H., Jones, B., Li, Z., Frasse, P., Delalande, C., Regad, F., Chaabouni, S., Latche, A., Pech, J.C., and Bouzayen, M.** (2005). The tomato Aux/IAA transcription factor IAA9 is involved in fruit development and leaf morphogenesis. *Plant Cell* **17**: 2676–2692.
- Woo, Y., Krueger, W., Kaur, A., and Churchill, G.** (2005). Experimental design for three-color and four-color gene expression microarrays. *Bioinformatics* **21**: 459–467.
- Woodward, A.W., and Bartel, B.** (2005). Auxin: Regulation, action, and interaction. *Ann. Bot. (Lond.)* **95**: 707–735.

**Regulatory Features Underlying Pollination-Dependent and -Independent Tomato Fruit Set Revealed by Transcript and Primary Metabolite Profiling**

Hua Wang, Nicolas Schauer, Bjoern Usadel, Pierre Frasse, Mohamed Zouine, Michel Hernould, Alain Latché, Jean-Claude Pech, Alisdair R. Fernie and Mondher Bouzayen  
*PLANT CELL* 2009;21;1428-1452; originally published online May 12, 2009;  
DOI: 10.1105/tpc.108.060830

This information is current as of July 20, 2009

<b>Supplemental Data</b>	<a href="http://www.plantcell.org/cgi/content/full/tpc.108.060830/DC1">http://www.plantcell.org/cgi/content/full/tpc.108.060830/DC1</a>
<b>References</b>	This article cites 82 articles, 35 of which you can access for free at: <a href="http://www.plantcell.org/cgi/content/full/21/5/1428#BIBL">http://www.plantcell.org/cgi/content/full/21/5/1428#BIBL</a>
<b>Permissions</b>	<a href="https://www.copyright.com/ccc/openurl.do?sid=pd_hw1532298X&amp;issn=1532298X&amp;WT.mc_id=pd_hw1532298X">https://www.copyright.com/ccc/openurl.do?sid=pd_hw1532298X&amp;issn=1532298X&amp;WT.mc_id=pd_hw1532298X</a>
<b>eTOCs</b>	Sign up for eTOCs for <i>THE PLANT CELL</i> at: <a href="http://www.plantcell.org/subscriptions/etoc.shtml">http://www.plantcell.org/subscriptions/etoc.shtml</a>
<b>CiteTrack Alerts</b>	Sign up for CiteTrack Alerts for <i>Plant Cell</i> at: <a href="http://www.plantcell.org/cgi/alerts/ctmain">http://www.plantcell.org/cgi/alerts/ctmain</a>
<b>Subscription Information</b>	Subscription information for <i>The Plant Cell</i> and <i>Plant Physiology</i> is available at: <a href="http://www.aspb.org/publications/subscriptions.cfm">http://www.aspb.org/publications/subscriptions.cfm</a>

### III. Deuxième post-doctorat (10.2008-08.2009): Identification de gènes régulateurs impliqués dans la formation de la paroi chez *Arabidopsis* par des approches post-génomiques.

Pendant mon deuxième Post-doc, j'ai réalisé une mobilité thématique afin d'étudier les mécanismes de mise en place de la paroi secondaire chez la plante modèle *Arabidopsis*, à des fins de production de biocarburant. Ces recherches s'inscrivent dans le cadre du projet FP7 : « Improving plant cell walls for use as a renewable industrial feedstock », à l'UMR5546 LRSV (Laboratoire de Recherche en Sciences Végétales), CNRS / Université Toulouse III – Paul Sabatier, INP Toulouse, France, à l'interface des équipes du Dr. Déborah Goffner et Dr. Jacqueline Grima-Pettenati. Au bout de 10 mois j'ai obtenu un poste de Maître des conférences à l'université Toulouse III. Ces travaux ont conduit à un article en premier auteur (Cassan-Wang, Goue, et al., 2013).

#### Résumé de Post-doctorat II

L'acquisition du système vasculaire, il y a environ 400 millions d'années a été un facteur décisif de l'adaptation des végétaux à la vie terrestre. Il est composé de tissus spécialisés dans la conduction et le soutien, le xylème et le phloème. Chez les arbres, le bois ou xylème secondaire est responsable de la croissance en diamètre des troncs. Il dérive de l'activité d'un tissu méristématique, le cambium vasculaire. La formation du xylème est un exemple fascinant de différenciation qui nécessite la coordination spatiale et temporelle de l'expression de plusieurs centaines de gènes impliqués dans la division cellulaire, l'élongation, l'acquisition d'une paroi secondaire lignifiée ainsi que dans la mort cellulaire programmée. Les parois secondaires des cellules du xylème possèdent des caractéristiques uniques (composition biochimique et association tridimensionnelle des polymères) dont dépendent les propriétés intrinsèques du bois chez les arbres forestiers. Elles se caractérisent notamment par l'abondance de lignines, composés phénoliques hydrophobes, qui représentent un obstacle à l'utilisation optimale des espèces destinées à l'industrie papetière.

Mon projet avait pour but d'identifier et de caractériser des facteurs de transcription d'*Arabidopsis* impliqués dans le développement vasculaire, par des approches de génétique inverse, de biologie cellulaire et de biochimie. Je devais notamment étudier l'importance de certains de ces gènes dans la formation du (pro)cambium, dans l'initiation de la différenciation du xylème, dans le dépôt des lignines et lors de la mort cellulaire programmée. Un criblage à grande échelle d'une collection de mutants ADN-T d'*Arabidopsis*, a été adopté pour identifier des plantes affectées dans la formation de leurs parois cellulaires. Ce criblage reposait sur la recherche de lignées présentant des altérations d'organisation vasculaire, telles que l'hypo-lignification, l'auto-fluorescence modifiée des tissus vasculaires, un dépôt ectopique de lignines, ou encore une perturbation des propriétés mécaniques de leurs tiges florales. Ces recherches fondamentales ont pour objectif d'identifier des facteurs génétiques importants en vue d'optimiser l'exploitation d'espèces végétales modèles pour la production de biomasse.



Annexe 3 Article de Post-doctorat II



# Identification of novel transcription factors regulating secondary cell wall formation in *Arabidopsis*

Hua Cassan-Wang, Nadia Goué, Mohammed N. Saidi<sup>†</sup>, Sylvain Legay<sup>†</sup>, Pierre Sivadon<sup>†</sup>, Deborah Goffner<sup>†</sup> and Jacqueline Grima-Pettenati\*

Laboratoire de Recherche en Sciences Végétales, UMR5546, Centre National de la Recherche Scientifique, Université Toulouse III, UPS, Toulouse, France

## Edited by:

Maurice Bosch, Aberystwyth University, UK

## Reviewed by:

Akiyoshi Kawaoka, Nippon Paper Industries Co., Japan  
Christian Dubos, Institut National de la Recherche Agronomique, France

## \*Correspondence:

Jacqueline Grima-Pettenati, Laboratoire de Recherche en Sciences Végétales, UMR5546, Centre National de la Recherche Scientifique, Université Toulouse III, UPS, 24 Chemin de Borde Rouge, BP 42617 Auzeville, 31326 Castanet-Tolosan, Toulouse, France  
e-mail: grima@lrsv.ups-tlse.fr

## <sup>†</sup>Present Address:

Mohammed N. Saidi, Laboratoire des Biotechnologies Végétales Appliquées à l'Amélioration des Cultures, Ecole Nationale d'Ingénieurs de Sfax, Sfax, Tunisie;  
Sylvain Legay, IJPB, UMR1318 INRA-AgroParisTech, Versailles, France;  
Pierre Sivadon, UMR 5254 IPREM, CNRS, Université de Pau et des Pays de l'Adour, Pau, France;  
Deborah Goffner, UMI 3189 Faculté de Médecine secteur Nord, Marseille, France

The presence of lignin in secondary cell walls (SCW) is a major factor preventing hydrolytic enzymes from gaining access to cellulose, thereby limiting the saccharification potential of plant biomass. To understand how lignification is regulated is a prerequisite for selecting plant biomass better adapted to bioethanol production. Because transcriptional regulation is a major mechanism controlling the expression of genes involved in lignin biosynthesis, our aim was to identify novel transcription factors (TFs) dictating lignin profiles in the model plant *Arabidopsis*. To this end, we have developed a post-genomic approach by combining four independent *in-house* SCW-related transcriptome datasets obtained from (1) the fiber cell wall-deficient *wat1 Arabidopsis* mutant, (2) *Arabidopsis* lines over-expressing either the master regulatory activator *EgMYB2* or (3) the repressor *EgMYB1* and finally (4) *Arabidopsis* orthologs of *Eucalyptus* xylem-expressed genes. This allowed us to identify 502 up- or down-regulated TFs. We preferentially selected those present in more than one dataset and further analyzed their *in silico* expression patterns as an additional selection criteria. This selection process led to 80 candidates. Notably, 16 of them were already proven to regulate SCW formation, thereby validating the overall strategy. Then, we phenotyped 43 corresponding mutant lines focusing on histological observations of xylem and interfascicular fibers. This phenotypic screen revealed six mutant lines exhibiting altered lignification patterns. Two of them [Bel-like HomeoBox6 (*blh6*) and a zinc finger TF] presented hypolignified SCW. Three others (*myb52*, *myb-like TF*, *hb5*) showed hyperlignified SCW whereas the last one (*hb15*) showed ectopic lignification. In addition, our meta-analyses highlighted a reservoir of new potential regulators adding to the gene network regulating SCW but also opening new avenues to ultimately improve SCW composition for biofuel production.

**Keywords:** secondary cell wall, xylem, fibers, transcription factor, *Arabidopsis*, lignin, co-expression, biofuels

## INTRODUCTION

Plant cells are enclosed in cell walls, which provide them with structural support and regulate growth and differentiation. There are two main types of cell walls: primary cell walls and secondary cell walls (SCWs). Primary cell walls are formed in all plant cells and are composed mainly of cellulose, hemicellulose, and pectin. SCWs are much thicker and are deposited in the inner side of primary cell walls only in some highly specialized tissues and cell types such as xylem vessels and fiber cells. Lignified SCW are the most abundant source of renewable biomass on earth, and are widely used for construction, paper, and energy. In the context of the energetic crisis, lignocellulosic biomass has received growing attention as raw material for the production of second-generation biofuels.

SCWs are composed of cellulose, hemicelluloses, and lignin. The impregnation with lignin renders SCWs waterproof and resistant, allowing water conduction through xylem vessels as well

as mechanical support. On the other hand, for lignocellulosic biofuel production, lignin is a major negative factor preventing hydrolytic enzymes from gaining access to cellulose and, as a result limits the saccharification potential. The biosynthetic pathways leading to SCWs formation are highly regulated at the transcriptional level. Tremendous progress has been made during the last decade supporting the existence of a complex hierarchical regulatory network of transcription factors (TFs). Most of those belong to two large TF families: R2R3-MYB and NAC (NAM/ATAF/CUC) (Demura and Fukuda, 2007; Zhong and Ye, 2007; Grima-Pettenati et al., 2012; Wang and Dixon, 2012; Zhong et al., 2012). Some members of the NAC TF family are key regulators of SCW formation in fibers and/or in vessels. This particular SCW related subgroup of NACs include the NAC SECONDARY WALL THICKENING PROMOTING FACTOR 1 (NST1), SECONDARY WALL-ASSOCIATED NAC DOMAIN PROTEIN1 (SND1/NST3), NST2, and the VASCULAR-RELATED NAC

DOMAIN (VND6 and VND7) (Kubo et al., 2005; Mitsuda et al., 2005, 2007; Zhong et al., 2006, 2012; Ko et al., 2007). Over expression of any of these *NACs* led to ectopic lignification in cells that normally contain only primary cell walls (for review, see Grima-Pettenati et al., 2012). A double mutation of *SND1* and *NST1* resulted in loss of SCW in fibers, whereas the simultaneous repression of *VND6* and *VND7* led to a defect in vessel SCW thickenings (Kubo et al., 2005; Zhong et al., 2007b). In the regulatory hierarchical network SDN1, NST1/2, and VND6/7 are first-level master switches controlling downstream TF regulators (Zhong and Ye, 2007; Wang and Dixon, 2012; Zhong et al., 2012). The second layer of regulators includes many MYB TFs (MYB20, MYB42, MYB43, MYB46, MYB52, MYB54, MYB58, MYB69, MYB61, MYB63, MYB83, MYB85, and MYB103) (Zhong et al., 2008; Ko et al., 2009; McCarthy et al., 2009; Zhou et al., 2009; Romano et al., 2012) as well as several other TFs like SND2, SND3, KNAT7, AtC3H14 Zinc finger TF (Zhong et al., 2008; Ko et al., 2009). Some of these are also master regulators since they control the biosynthesis of the three main components of SCW i.e., cellulose, xylan, and lignin. The discovery of this multi-leveled hierarchical regulatory network has been a breakthrough in our understanding of the regulation of the lignified SCW, although it is far from being complete. For instance, only a few TFs characterized hitherto are regulating specifically one of the SCW components, although three MYBs (MYB85, MYB58, and MYB63) were reported to be lignin-specific. In addition, our knowledge of the molecular mechanisms determining the heterogeneous SCW deposition in different cell types, as well as those governing the various patterns of SCW deposition is still very poor. More efforts are needed to get a comprehensive picture of the transcriptional regulation of the SCWs both from a fundamental and an applied perspective.

As a step toward this goal, we searched for novel TFs potentially implicated in the control of lignin deposition. To do this, we set up a post-genomic approach combining four original in-house SCW-related transcriptomic data sets that enabled us to identify 80 candidates belonging to major plant TFs families (i.e., *NAC*, *MYB*, *bHLH*, *Zinc finger*, *HomeoBox*, and *AP2/ERF*). Most of them have not yet been functionally characterized. Histochemical analyses of the corresponding mutants revealed six strong candidates regulating the biosynthesis of lignin and/or the whole SCW biosynthetic program: BLH6 (Bel-like HomeoBox6; AT4G34610), HB5 (AT5G65310), HB15 (AT1G52150), MYB-like TF (AT3G11280), MYB52 (AT1G17950), and Zinc finger TF (AT3G46620).

## RESULTS

### A POST-GENOMIC APPROACH TO IDENTIFY NOVEL REGULATORY GENES INVOLVED IN SCW FORMATION

In order to identify novel regulatory genes involved in SCW formation, we took advantage of four large scale *in house* transcriptomic data sets and developed a post-genomic approach. A flow chart of the main steps of this original strategy is described in **Figure 1**. The first SCW-related *in house* transcriptome dataset came from the *Arabidopsis* mutant, *wat1* (*walls are thin 1*), which has little to no SCW in fibers (Ranocha et al., 2010). In this mutant, the transcript levels of many genes associated with the

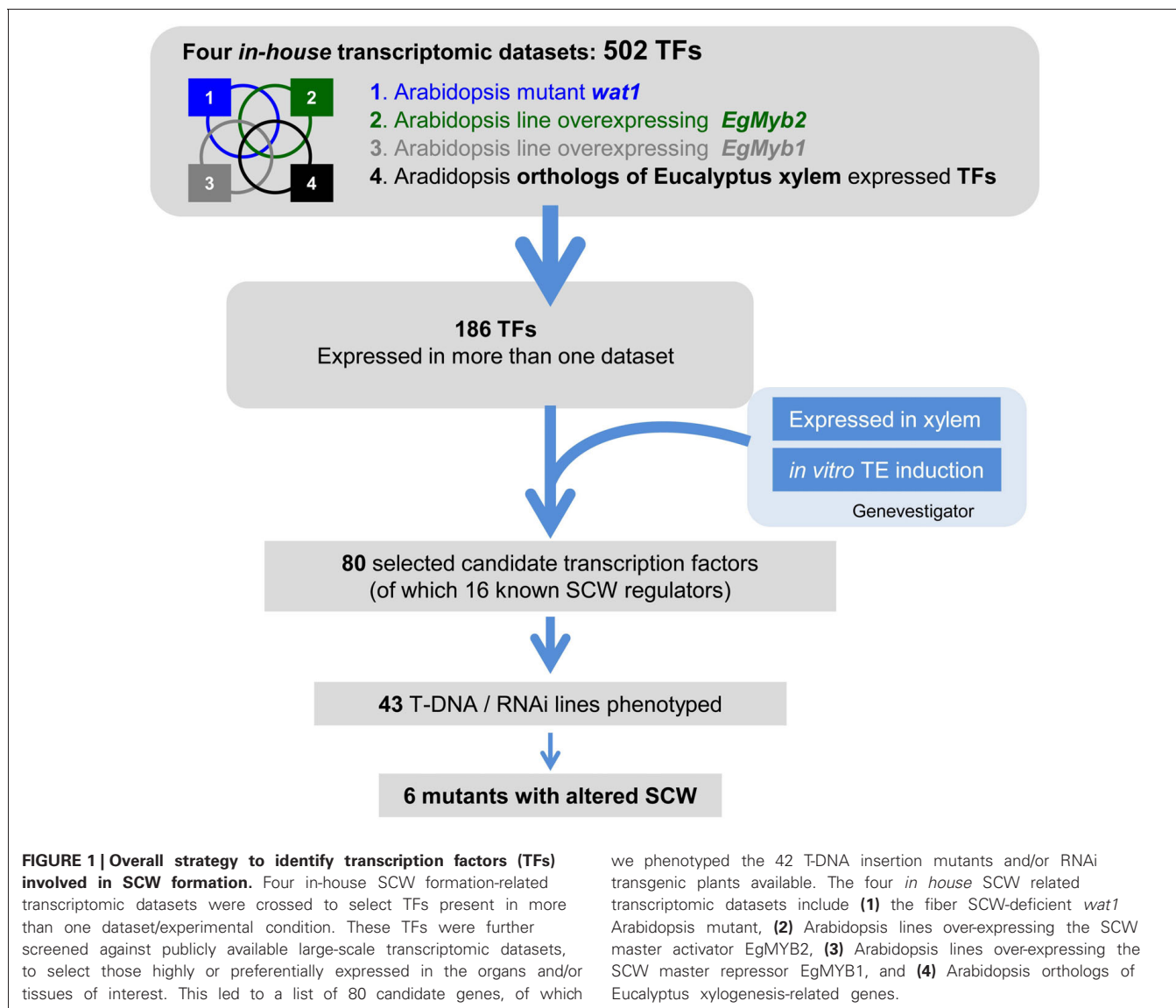
regulation and/or the biosynthesis of the different SCW wall polymers were dramatically reduced in keeping with the mutant phenotype. Within the genes up/or down-regulated in the mutant background, we identified 97 TFs including some well-known SCW-regulating TFs such as *SND1*, *SNT1*, and *MYB46* (Table S1). The second transcriptome dataset was comprised of 240 TFs (Table S2) that exhibited de-regulated expression in *Arabidopsis* lines over-expressed the SCW-master activator *EgMYB2*, which is a *Eucalyptus* R2R3 MYB TF highly expressing in xylem cells undergoing SCW thickening (Goicoechea et al., 2005). *EgMYB2* is able to activate the promoters of lignin (Goicoechea et al., 2005), cellulose, and xylan biosynthetic genes (Zhong and Ye, 2009), leading to thicker SCW in *EgMYB2* over-expressing lines in tobacco (Goicoechea et al., 2005; De Micco et al., 2012). Moreover, the closest orthologs of *EgMYB2* in *Arabidopsis* *AtMYB46* and *AtMYB83* encode for master regulators capable of activating the whole SCW biosynthetic program (Zhong et al., 2007a; McCarthy et al., 2009), and *EgMYB2* was able to complement the *myb46-myb83* double mutant (Zhong et al., 2010). The third transcriptome dataset included 309 TFs (Table S3) deregulated in *Arabidopsis* lines over-expressing the SCW-repressor, *EgMYB1* (Legay et al., 2010). *EgMYB1* over-expressors exhibited fewer lignified fibers particularly in the interfascicular zones and reduced SCW thickenings. Klason lignin content was moderately but significantly reduced and decreased transcript accumulation was observed for genes involved in the biosynthesis of lignins, cellulose, and xylan (Legay et al., 2010). Finally, the fourth dataset was composed of 87 TFs (Table S4) that were the *Arabidopsis* orthologs of *Eucalyptus* TFs preferentially expressed in differentiating xylem (Rengel et al., 2009), a tissue that is particularly rich in cells undergoing SCW deposition and lignification. Altogether, these four transcriptomic datasets allowed us to identify a total of 502 candidate TFs. To narrow down the number of candidates for functional validation, we selected 186 that were identified in two datasets (Table S5). It should be noted that 43 of those were found in three data sets and only three were common to the four datasets *bHLH5* (AT5G46760), *IAA9* (AT5G65670), and *AP2 TF RAP2.2* (AT3G14230).

### CROSS-COMPARISON WITH PUBLICLY AVAILABLE MICROARRAY DATA

To further narrow down the selection of the 186 candidate TFs for further functional analysis, we examined their *in silico* expression patterns using Genevestigator (Hruz et al., 2008). We restricted our list to genes that were preferentially and/or highly expressed in situations in which SCW formation is prevalent i.e., in xylem, the basal part of the inflorescence stem, and/or in cell suspension cultures undergoing *in vitro* SCW formation (Kubo et al., 2005). This *in silico* expression screen allowed us to obtain a final list of 80 candidate SCW TFs (**Table 1**).

The most represented families in decreasing order were the MYB TF family (containing 19 members), the AP2 ERF (Ethylene Response Factor) TF family (12), the HB (HomeoBox protein) TF family (9), and the Aux/IAA family (7) (**Figure 2**). The other TFs belonged to the WRKY (5), *NAC* (5), *bHLH* (5), Zinc finger TF (5), *bZip* (3), and ARF (2) families, respectively.

It is noteworthy that 16 of the 80 candidates were already shown to regulate SCW formation. They include eight MYB



TF (*MYB20*, *MYB42*, *MYB46*, *MYB52*, *MYB61*, *MYB63*, *MYB75*, *MYB85*, and *MYB103*) (Zhong et al., 2007a, 2008; Ko et al., 2009; Zhou et al., 2009; Ohman et al., 2013), four NAC TF (*SND1*, *NST1* *XND1*, and *VNI2*) (Mitsuda et al., 2005; Zhong et al., 2006; Zhao et al., 2008; Yamaguchi et al., 2010), three Homeodomain containing TF (*HB14*, *HB15*, and *KNAT7*) (McConnell et al., 2001; Emery et al., 2003; Kim et al., 2005; Zhong et al., 2008) and one WRKY TF (*WRKY12*) (Wang et al., 2010) (complete list in Table 1). This significant proportion of characterized SCW related TF in our final list validates well the strategy used in this study. For example, *MYB46* (Zhong et al., 2007a), *MYB85* (Zhong et al., 2008), and *WRKY12* (Wang et al., 2010) were present in three of the four transcriptomic datasets.

#### PHENOTYPES OF TF T-DNA MUTANT OR RNAi TRANSGENIC LINES

We then collected and characterized publicly available T-DNA mutant lines or RNAi transgenic lines that corresponded to 43 of the 80 candidate genes (Table S6). The information concerning

the different lines including T-DNA insertion position and *in-house* databases source is presented in Table S6. Phenotyping was performed on 20 cm-high mutant stems grown in short-day growth conditions. Under these conditions, the basal part of the stem abundantly develops cells undergoing SCW thickening (xylem vessel cells, xylary fiber cells, and interfascicular fiber cells). Histological analyses of SCW were performed using the natural auto fluorescence of phenolic compounds under UV-light as well as phloroglucinol-HCl staining, which is indicative of the lignin content. We found significant alteration of lignin profiles in six mutant lines corresponding to two MYB TFs: *MYB like TF* (*AT3G11280*) and *MYB52* (*AT1G17950*), three HomeoBox TF *HB5* (*AT5G65310*), *BLH6* (*AT4G34610*), and *HB15* (*AT1G52150*) and a *Zinc finger TF* (*AT3G46620*), although the overall organization of vascular bundles and interfascicular fibers was not altered in these six mutant lines (Figures 3, 4).

Under UV-light the intensity of auto-fluorescence was lower in *zinc finger TF* (Figure 3B) and in *blh6* (Figure 3C) mutant lines in

**Table 1 | List of 80 candidate TFs obtained after cross-comparison of transcriptomic datasets.**

ID	Annotation	<i>in house data sets (Log<sub>2</sub>Ratio)</i>				<i>Publicly available data</i>		No. of ordered mutants
		wat1	OE EgMYB2	OE EgMYB1	Eg Xylem	TE induction 6d*	Xylem**	
AT3G50260	AP2 TF DREB, CEJ1 (subfamily A-5)	2.7	-1.6	-2.44		0.3	Low_ns	0
AT5G61590	AP2 TF ERF (subfamily B-3)	3.4	-1.1	2.1		3.7	High_p	1
AT5G07580	AP2 TF ERF (subfamily B-3)		2.2	2.4		0.9	High_p	1
AT5G51190	AP2 TF ERF (subfamily B-3)	-3.2	1.8	2.4		1.5	High_p	1
AT5G13330	AP2 TF ERF (subfamily B-4)	7.3	-0.6			1.4	High_p	0
AT3G14230	AP2 TF ERF RAP2.2 (subfamily B-2)	2.0	1.4	1.1	x	0.2	High_p	2
AT1G43160	AP2 TF ERF RAP2.6 (subfamily B-4)	2.2	-0.7	-1.5		-0.2	Low_ns	0
AT1G68840	AP2 TF ERF RAP2.8 (RAV2)	1.9	-2.0	-2.3		-2.0	Medium_ns	0
AT3G25890	AP2 TF ERF, CRF11 (subfamily B-6)	2.0	1.0	1.2		0.1	Medium_p	0
AT5G25190	AP2 TF ERF, ESE3 (subfamily B-6)	-2.3			x	2.0	Low_ns	0
AT2G44840	AP2 TF ERF13 (subfamily B-3)	-2.8		-0.8	x	0.0	Medium_u	1
AT5G47230	AP2 TF ERF5 (subfamily B-3)	-2.7	1.3			-0.4	Medium_u	0
AT5G60450	ARF4	-1.7	2.0	1.9		-0.4	High_s	2
AT1G30330	ARF6		1.3	1.6	x	-0.8	Medium_u	1
AT5G54680	bHLH105	-3.7	0.6		x	-0.3	High_u	1
AT2G46510	bHLH17		1.1		x	-0.2	Medium_u	0
AT1G68810	bHLH30	-2.4	2.0	1.8		2.4	High_s	3
AT3G25710	bHLH32	-2.1	1.2	1.6		0.0	High_s	2
AT5G46760	bHLH5	1.8	1.5	1.2	x	0.7	High_p	2
AT1G75390	bZIP TF	-2.1	0.6			-2.9	High_p	0
AT2G18160	bZIP TF	-1.8	0.9			2.4	High_p	1
AT3G51960	bZIP TF	2.3	-0.8	-0.9		0.5	Low_u	0
AT3G56850	bZIP TF (ABA_REB3)		0.6	0.6	x	0.8	Medium_u	0
AT1G72830	CCAAT-binding TF	2.8	1.6	1.5		-0.6	High_s	0
AT3G28730	HMG		1.1	0.7	x	-0.4	High_p	0
<b>AT2G34710</b>	<b>homeobox-leucine zipper TF (HB14)</b>			0.8	x	4.1	High_p	0
<b>AT1G52150</b>	<b>homeobox-leucine zipper TF (HB15)</b>		1.3		x	4.2	High_s	3
AT5G65310	homeobox-leucine zipper TF (HB5)	-2.1		0.9		-2.5	Low_u	3
AT5G41410	homeodomain TF (BEL1)	2.2	-1.2	-0.6		3.0	Low_ns	0
AT5G02030	homeodomain TF (BELLRINGER)		1.4	1.6		0.2	High_p	0
AT4G34610	homeodomain TF (BLH6)	-2.0	0.7	0.7		1.4	Low_p	2
AT1G70510	homeodomain TF (KNAT2)		1.4	2.1		-0.1	Medium_s	0
<b>AT1G62990</b>	<b>homeodomain TF (KNAT7)</b>	-7.9	2.8	2.4		2.6	High_s	0
AT1G62360	homeodomain TF (SHOOT MERISTEMLESS)		3.9	3.7	x	-0.1	High_s	0
AT4G28640	IAA11	-2.6	1.6	1.8		1.4	Medium_s	1
AT2G33310	IAA13		2.3	1.7	x	1.1	High_p	0
AT3G04730	IAA16		1.7	1.4		1.2	High_u	0
AT2G46990	IAA20		1.8	1.5		-0.3	Low_u	0
AT5G25890	IAA28	1.4	2.3	2.1		3.3	High_s	1
AT2G22670	IAA8		1.2	1.3		0.6	High_p	0
AT5G65670	IAA9	-1.8	0.9	1.2	x	0.8	High_p	0
AT1G10200	LIM TF			0.7	x	1.4	Medium_u	0
AT1G01060	myb related TF (LHY)	-1.8	2.6		x	0.4	High_u	1
<b>AT1G63910</b>	<b>MYB103</b>	-11.8		2.4		3.2	Medium_s	1
AT1G48000	MYB112	2.8	-1.7	-1.5		-0.1	Low_u	0
<b>AT1G66230</b>	<b>MYB20</b>	-3.1				1.0	Medium_u	1
AT5G07690	MYB29	-4.4	1.9	2.3		-0.2	Low_u	0
AT4G38620	MYB4		1.4		x	-1.3	Medium_p	3
<b>AT4G12350</b>	<b>MYB42</b>		1.3	1.0		0.6	Low_u	0
<b>AT5G12870</b>	<b>MYB46</b>	-2.8	0.9	1.4		5.3	Medium_s	2

(Continued)



Table 1 | Continued

ID	Annotation	<i>in house data sets (Log<sub>2</sub>Ratio)</i>			Publicly available data		No. of ordered mutants	
		wat1	OE EgMYB2	OE EgMYB1	Eg Xylem	TE induction 6d*		Xylem**
AT3G46130	MYB48	3.0	1.4	1.8		0.1	High_s	0
<b>AT1G17950</b>	<b>MYB52</b>	-2.1	2.3	2.4		3.3	Medium_s	2
AT5G59780	MYB59	4.7	1.0	0.8		2.7	High_p	1
AT1G09540	MYB61	-2.4		1.8		no probe	no probe	0
<b>AT1G79180</b>	<b>MYB63</b>	-5.9	1.3			2.3	Medium_s	0
<b>AT2.36650</b>	<b>MYB75</b>		-1.3	-1.9		0.1	Low_u	0
AT3G50060	MYB77	-3.2	1.4	0.8		-0.5	Medium_u	0
<b>AT4G22680</b>	<b>MYB85</b>	-2.1	2.2	1.6		2.1	Low_u	1
AT5G05790	myb-like TF	-1.9	0.7	0.9		1.7	High_s	0
AT5G17300	myb-like TF	-1.8	0.8			1.2	Medium_u	0
AT3G11280	myb-like TF	1.9		0.6		4.5	High_p	3
AT1G12260	NAC007, VND4	-1.9				5.0	Low_u	1
<b>AT1G32770</b>	<b>NAC012, SND1, NST3</b>	-2.6		1.8		0.0	Medium_s	0
<b>AT2G46770</b>	<b>NAC043, NST1</b>	-0.8			x	no probe	no probe	0
<b>AT5G64530</b>	<b>NAC104, XND1</b>			0.9	x	3.6	High_s	0
<b>AT5G13180</b>	<b>NAC83, VND-INTERACTING2 (VNI2)</b>		-0.7		x	3.3	High_u	0
AT4G37750	AINTEGUMENTA (ANT)	-5.4		1.0		0.1	Low_p	0
AT1G21450	scarecrow-like TF 1 (SCL1)		1.2		x	2.2	High_p	1
AT2G47070	squamosa TF-like 1 (SPL1)		0.8		x	0.7	High_u	0
<b>AT2G44745</b>	<b>WRKY12</b>	-2.9	2.3	2.7		-0.1	High_s	1
AT2G30250	WRKY25	2.6	-0.7	-0.6		-1.2	Medium_u	2
AT4G23550	WRKY29	-7.7	2.1	1.6		no probe	no probe	1
AT1G80840	WRKY40	-3.0			x	-0.4	High_u	1
AT3G01970	WRKY45	3.4	-0.8	-1.2		-0.5	Medium_u	0
AT5G13080	WRKY75	3.5	-0.8	-0.9		-0.3	Low_ns	0
AT2G28510	zinc finger Dof-type	-1.7	-0.8	-0.9		0.2	Low_ns	0
AT1G68360	zinc finger TF	-3.7	1.8	2.0		1.1	High_s	0
AT2G40140	zinc finger TF (CCCH-type)	-1.8	0.6			0.3	High_u	0
AT3G55980	zinc finger TF (CCCH-type)	-3.6			x	0.5	High_u	2
AT3G46620	zinc finger TF-n129	-3.4				1.6	High_u	2

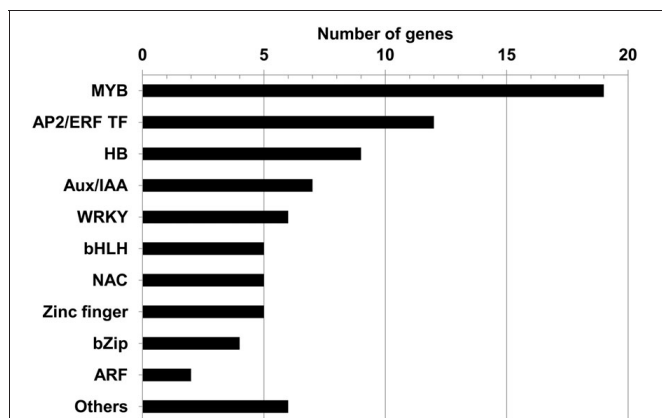
\*TE induction 6d, 6 days after induction (1  $\mu$ M brassinolide and 10 mM H<sub>3</sub>BO<sub>3</sub>) when tracheary element were actively forming.

\*\*Xylem in hypocotyl of adult plant; High, Medium, and Low for their expression level; \_s, specific expression in xylem; \_p, preferential expression; \_u, ubiquitous expression; \_ns, higher expression in non-xylem cells. Bold indicates known genotypes, and corresponding secondary wall or xylem cell identity phenotypes.

both vascular bundles and interfascicular regions as compared to the control (Figure 3A), suggesting a global decrease in phenolic compound deposition. The SCW in xylem vessels, xylary fibers, and interfascicular fibers were observed in more detail using phloroglucinol-HCl staining. Little to no SCW was deposited in xylary fibers (Figures 3H,I) and moreover, SCW thickness was largely reduced in interfascicular fibers (thin and weak phloroglucinol staining) as compared to wild-type (Figure 3G) suggesting that these lines were hypolignified.

Auto-fluorescence under UV light was more intense in *myb like TF* (Figure 3D), *hb5* (Figure 3E), and *myb52* (Figure 3F) lines than in controls (Figure 3A), especially in the interfascicular region, suggesting an increased deposition of phenolic compounds and possibly lignins. This was further confirmed by a massive and intense phloroglucinol staining indicating an enhanced lignin deposition in the interfascicular fiber and xylary

fiber cells of these mutants (Figures 3J–L). Extra-layers of cells with lignified SCW were detected in the external layers of both interfascicular fibers and metaxylem vessels in two lines *myb-like TF* (Figure 3J, green arrows) and *myb52* (Figure 3L, green arrows) as compared to the control (Figure 3A). This observation suggests that secondary xylem formation was enhanced and appeared earlier than in wild-type. A strong fluorescent signal was also detected in the phloem cap cells (Figures 3D–F, blue arrows) in all three highly auto-fluorescent lines suggesting a transition of phloem cap cells to phloem sclereids (highly lignified) which was further confirmed by phloroglucinol staining (Figures 3J–L, blue arrows). Similarly, auto-fluorescent signals (Figure 3F) and strong phloroglucinol-HCl staining (Figure 3L, pink arrow) were detected in the epidermal cells of some *myb52* T-DNA insertion lines revealing an ectopic deposition of lignin.



**FIGURE 2 | Distribution of the 80 SCW candidate TFs according to family.** Amongst them, four families have been shown previously to regulate SCW formation: MYB, NAC, HB (Homeodomain containing TF), and WRKY; three TF families are involved in hormone signaling including ARF, Aux/IAA, and AP2 ERF TF families. “Others” include the following TF families: LIM, CCAAT binding, HMG (High Mobility Group, belonging to a transcription complex), AINTEGUMENTA (ANT; AP2 like TF), Scarecrow-like TF1 (SCL1), and Squamosa-like1 (SPL1).

Both auto-fluorescence and phloroglucinol staining of stem sections of *hb15* (Figures 4B,D) showed that large parenchyma cells adjacent to the inner side of the interfascicular fibers (red arrow), as well as smaller xylem parenchyma cells surrounding the protoxylem (yellow arrow) exhibited lignified SCW. The corresponding cells in the control have non-lignified primary walls (Figures 4A,C). As compared to the control, extra layers of cells with lignified SCW were present in the most external rows of the interfascicular fibers and xylem (Figures 4B,D, green arrows) suggesting an enhanced and early formation of secondary xylem. Moreover, both xylary and interfascicular fibers in *hb15* lines exhibited both a more intense auto-fluorescence and staining by phloroglucinol than that of the control suggesting higher lignin content.

The overall growth behavior of the mutants did not differ significantly from the controls, except the bolting and flowering time were altered in three of the mutant lines. Hypolignified *blh6* and *zinc finger* lines bolted and flowered earlier than controls (Figures 5A,C) whereas the hyperlignified *hb15* line exhibited delayed bolting and flowering (Figures 5B,C). In addition, *hb15* mutants exhibited aerial rosettes at the base of the lateral inflorescence branches instead of growing cauline leaves as in wild-type plants (Figure 5D).

#### CO-EXPRESSION ANALYSIS OF CANDIDATE TFs GENES

Since it is known that transcriptionally coordinated genes tend to be functionally related (Ruprecht and Persson, 2012), we performed co-expression analyses for the six candidate genes in order to further validate their role in controlling SCW synthesis and get some clues about their function. The co-expression genes lists were generated using the Genevestigator platform (<https://www.genevestigator.com>), *Arabidopsis* co-expression data mining tools and GeneCAT. All six candidate TFs were co-expressed with genes related to cell wall formation (Tables 2, 3 and Tables S7–S10)

albeit to different extents ranging from 10 to 66% of SCW associated genes amongst the 50 first co-expressed genes. The most remarkably high co-expression profiles were found for *MYB52* and *BLH6*.

Among the top 50 genes co-expressed with *BLH6*, 30 (60%) were related to SCW formation (Table 2). Notably, they included genes involved in the biosynthesis of the three main cell wall polymers i.e., three main SCW cellulose synthases genes (CESAs) (*IRX5/CESA4*, *IRX3/CESA7*, and *IRX1/CESA8*), three genes implicated in xylan biosynthesis [*IRX10/GUT2*, *GUX1* (Glucuronic acid substitution of xylan1), and *XTH32* (xyloglucan endotransglucosylase/hydrolase 32) (Brown et al., 2009; MacMillan et al., 2010)] as well as two laccase genes (*IRX12/LAC4* and *LAC17*) involved in monolignol polymerization (Berthet et al., 2011). Moreover, several SCW master transcriptional regulators were also co-expressed with *BLH6*, such as *MYB46*, *MYB83*, and *SND2* (Zhong et al., 2008; Zhong and Ye, 2012).

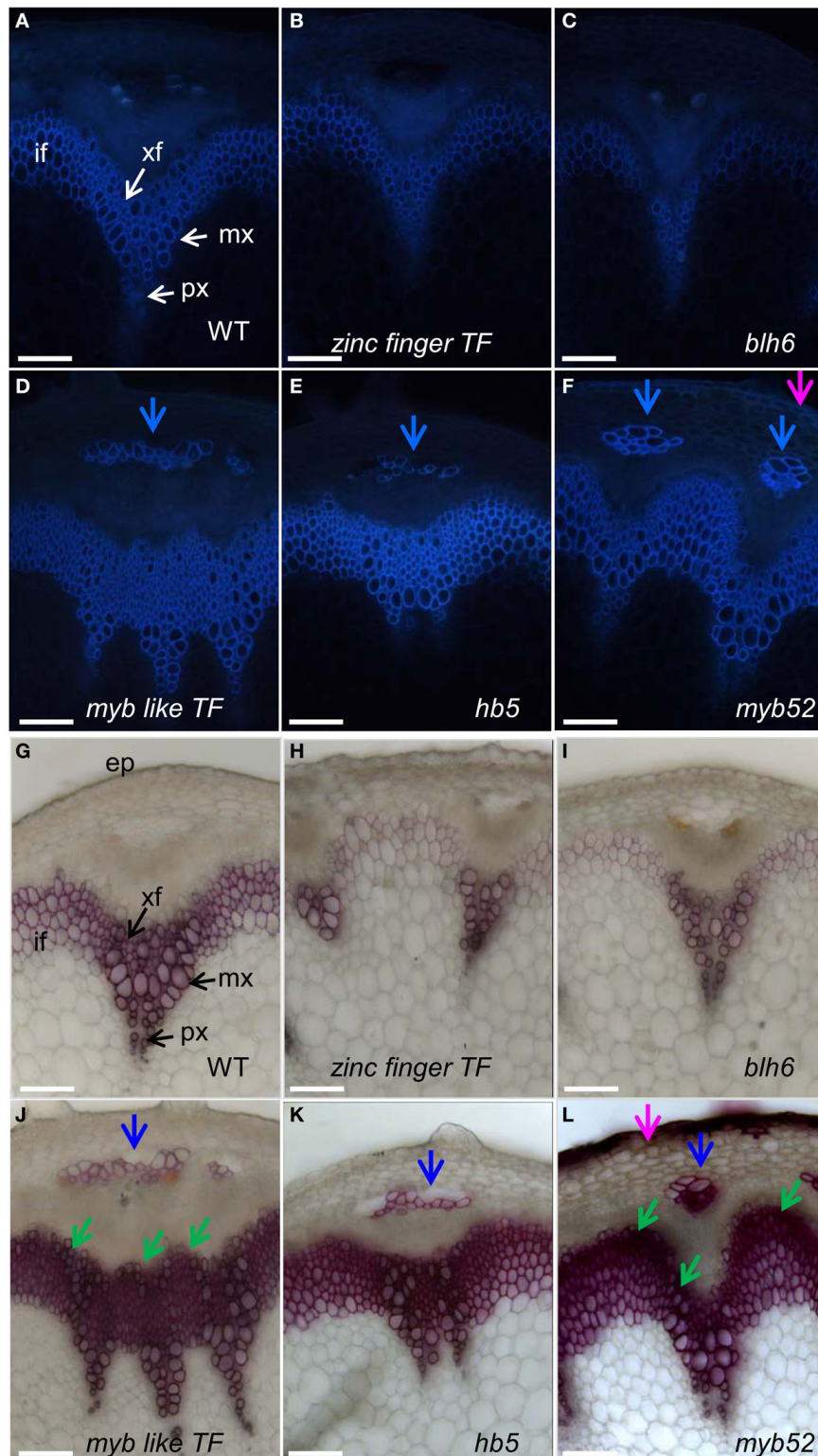
The co-expression analysis of *MYB52* revealed that 33 of the top 50 genes (66%) co-expressed with *MYB52* were related to SCW formation. The list of co-expressed genes included genes encoding key biosynthetic enzymes of the main SCW polymers i.e., cellulose (*IRX3/CESA7*, *IRX1/CES8*, *IRX6/COBL4*, and *CTL2*), hemicelluloses (*IRX10/GUT1*, *IRX8*, *IRX9*, *IRX15*, and *IRX15-L*), and possibly lignins (*LAC2* and *LAC10*). *MYB52* was also co-expressed with many SCW regulators such as *MYB54*, *MYB85*, *MYB69*, *MYB103*, *SND2/NAC073*, and *XND1/NAC104* (Xylem NAC Domain1).

#### DISCUSSION

Functional genomics approaches developed during the last decade have generated numerous candidate genes related to SCW formation in *Arabidopsis* and other plant species. Whereas these large individual gene lists make difficult the choice of the most promising candidates for the further functional validation, meta-analyses combining multiple transcriptomic data sets offer a new way to reveal some core regulators.

By cross-comparing four SCW-related transcriptomic datasets, we selected 186 TFs present in at least two experimental conditions. Since these datasets came from very different backgrounds (mutants and over-expressors of SCW regulatory genes as well as orthologs of *Eucalyptus* xylem expressed genes), the selection of genes appearing in more than one dataset likely helped us to identify “core regulators” of SCW formation but might also have filtered out some more specific regulators. We further restricted the candidate gene list by including *in silico* analyses of their expression making the hypothesis that TFs expressed highly or preferentially in xylem tissues and/or during tracheary elements formation would be the most promising candidates. Indeed this strategy was successful since among the 80 genes that came out, 16 have already been reported to be regulators of the SCW. They included, for instance, the master regulators *SND1* and *MYB46* as well as the lignin-specific *MYB85*.

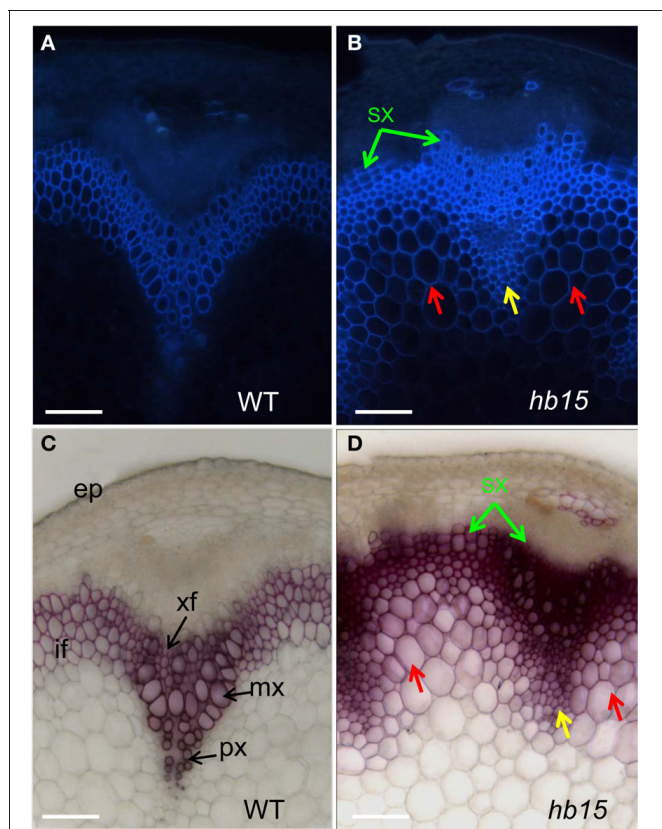
Forty-three mutant lines were phenotyped but only six exhibited a notable cell wall phenotype. This high proportion of mutants without phenotype is not surprising since many mutants targeting only one TF are known to yield mild to no phenotype



**FIGURE 3 | Stem cross sections of five T-DNA mutant lines presenting either hypo or hyperlignified SCWs.** Sections of wild-type plant and T-DNA mutants were observed under UV light (A–F) or stained with phloroglucinol-HCl (G–L). Phloem cap cells and ectopic lignification in epidermal cells are indicated by blue and pink arrows, respectively.

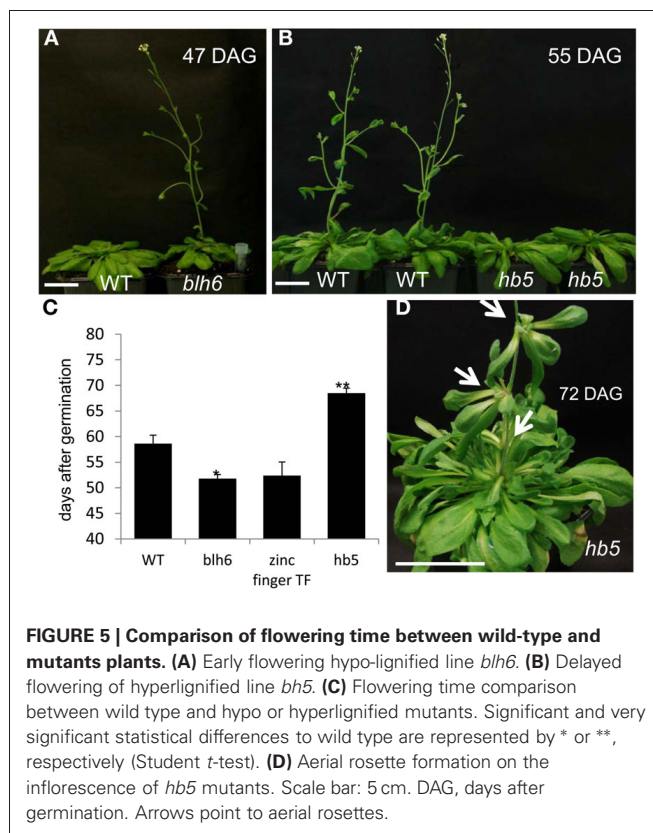
Observations were made at the basal part of inflorescence stems at the stage of newly formed green siliques, about two weeks after bolting, when the inflorescence stems reached 20 cm height. if, interfascicular fiber; xf, xylary fiber; mx, metaxylem; px, protoxylem; sx, secondary xylem; ep, epidermis. Scale bar: 20  $\mu$ m.





**FIGURE 4 | Stem sections of the *hb15* T-DNA mutant.** Auto-fluorescence observed under UV light of wild type (A) and *hb15* mutant (B) Phloroglucinol-HCl staining of lignin in wild type (C) and *hb15* mutant (D). Ectopic lignification in large parenchyma cells and in small parenchyma cells surrounding protoxylem is indicated by red arrows and yellow arrows, respectively; precocious secondary walled secondary xylem formation is indicated by green arrow. if, interfascicular fiber; xf, xylary fiber; mx, metaxylem; px, protoxylem; sx, secondary xylem; ep, epidermis. Scale bar: 20  $\mu$ m.

(Okushima et al., 2005; Overvoorde et al., 2005; Jensen et al., 2011; Ruprecht et al., 2011). This is particularly true for multi-gene families TFs where functional redundancy prevents the observation of distinct phenotypes in knock-out mutants. This is indeed the case for a large proportion of the SCW regulators characterized hitherto including some of the sixteen highlighted here. For instance, whereas a single mutant of the SCW master transcriptional activator *MYB46* did not exhibit any cell wall phenotype, the double knock out mutant *myb46/myb83* with its closest ortholog *MYB83* showed a severe reduction of SCW thickness (Zhong et al., 2007a; McCarthy et al., 2009). Therefore, genes for which the corresponding single mutants exhibited no phenotype in this study may still be interesting candidates taking part in the regulation of SCW formation. Further experiments using over-expressors and/or mutants of two or more paralog genes would increase the probability of obtaining informative phenotypes and insight into their functions. Our *in silico* analyses pointed out some very promising genes which should be further characterized using such approaches.



**FIGURE 5 | Comparison of flowering time between wild-type and mutant plants.** (A) Early flowering hypo-lignified line *blh6*. (B) Delayed flowering of hyperlignified line *hb5*. (C) Flowering time comparison between wild type and hypo or hyperlignified mutants. Significant and very significant statistical differences to wild type are represented by \* or \*\*, respectively (Student *t*-test). (D) Aerial rosette formation on the inflorescence of *hb5* mutants. Scale bar: 5 cm. DAG, days after germination. Arrows point to aerial rosettes.

The most abundantly represented TF family in our list was the MYB family (19 members) of which eight (belonging to the R2R3 subfamily) have already been shown to regulate either the phenylpropanoid pathway and/or the SCW formation. It is the case for *MYB46* (Zhong et al., 2007a), *MYB63* (Zhou et al., 2009), *MYB85* (Zhong et al., 2008), and *MYB103* (Ohman et al., 2013). We phenotyped *myb52* insertion lines that exhibited a strong hyperlignification phenotype, thus suggesting that *MYB52* could be a repressor of the lignin biosynthesis and possibly of the whole SCW formation. This result is in apparent contradiction with a previous study showing that the dominant repression of *MYB52* caused a severe reduction in SCW thickening in both interfascicular fibers and xylary fibers of the inflorescence stem (Zhong et al., 2008). The authors concluded that *MYB52* was an activator of the SCW although no phenotype was detectable when over-expressed. A likely explanation to these apparent discrepancies is that *MYB52* encodes a transcriptional repressor as clearly suggested by our knock-out mutant phenotype and therefore its dominant repression would result in a stronger transcriptional repression. *MYB52* appeared to be tightly co-expressed with *MYB54* and *WAT1*. It is also co-expressed with several cellulose and xylan biosynthetic genes and with *MYB85*, a specific regulator of the lignin biosynthesis (Zhong et al., 2008). Altogether, these results suggest for *MYB52* a repressor role of the whole SCW program although this needs to be supported by further experiments.

Besides these canonic R2R3 MYBs, four MYB-like proteins were present in the candidate list and one mutant was analyzed.

**Table 2 | Top 50 co-expressed genes with *BLH6* using Genevestigator platform.**

Probeset	AGI	Gene symbol	Score	Description
253247_at	AT4G34610	BLH6	1	BEL1-like homeodomain 6
<b>252025_at</b>	<b>AT3G52900</b>		<b>0.75</b>	<b>Family of unknown function (DUF662)*</b>
250600_at	AT5G07800		0.73	Flavin-binding monooxygenase family protein
<b>264573_at</b>	<b>AT1G05310</b>		<b>0.71</b>	<b>Pectin lyase-like superfamily protein</b>
245657_at	AT1G56720		0.71	Protein kinase superfamily protein
<b>251069_at</b>	<b>AT5G01930</b>	<b>MAN6</b>	<b>0.70</b>	<b>Glycosyl hydrolase superfamily protein</b>
263629_at	AT2G04850		0.70	Auxin-responsive family protein
<b>250933_at</b>	<b>AT5G03170</b>	<b>FLA11</b>	<b>0.69</b>	<b>FASCICLIN-like arabinogalactan-protein 11</b>
<b>262796_at</b>	<b>AT1G20850</b>	<b>XCP2</b>	<b>0.68</b>	<b>xylem cysteine peptidase 2</b>
245864_at	AT1G58070		0.68	Unknown protein
258938_at	AT3G10080		0.68	RmlC-like cupins superfamily protein
<b>251131_at</b>	<b>AT5G01190</b>	<b>LAC10</b>	<b>0.68</b>	<b>laccase 10*</b>
<b>246827_at</b>	<b>AT5G26330</b>		<b>0.67</b>	<b>Cupredoxin superfamily protein*</b>
246439_at	AT5G17600		0.67	RING/U-box superfamily protein
251249_at	AT3G62160		0.66	HXXXD-type acyl-transferase family protein
<b>257757_at</b>	<b>AT3G18660</b>	<b>PGSIP1,GUX1</b>	<b>0.66</b>	<b>Glucuronic acid substitution of xylan 1</b>
<b>264184_at</b>	<b>AT1G54790</b>		<b>0.65</b>	<b>GDSL-like Lipase/Acylhydrolase family protein*</b>
253876_at	AT4G27430	CIP7	0.65	COP1-interacting protein 7
259688_at	AT1G63120	RBL2	0.65	RHOMBOID-like 2
<b>267094_at</b>	<b>AT2G38080</b>	<b>LAC4,IRX12,LMCO4</b>	<b>0.65</b>	<b>Laccase/Diphenol oxidase family protein</b>
250664_at	AT5G07080		0.64	HXXXD-type acyl-transferase family protein
<b>250322_at</b>	<b>AT5G12870</b>	<b>MYB46</b>	<b>0.64</b>	<b>myb domain protein 46</b>
250120_at	AT5G16490	RIC4	0.64	ROP-interactive CRIB motif-containing protein 4
251297_at	AT3G62020	GLP10	0.64	germin-like protein 10
256367_at	AT1G66810		0.64	Zinc finger C-x8-C-x5-C-x3-H type family protein
265277_at	AT2G28410		0.63	Unknown protein
<b>246425_at</b>	<b>AT5G17420</b>	<b>IRX3,MUR10,CESA7</b>	<b>0.63</b>	<b>Cellulose synthase family protein</b>
266424_at	AT2G41330		0.62	Glutaredoxin family protein
<b>253798_at</b>	<b>AT4G28500</b>	<b>ANAC073,SND2</b>	<b>0.62</b>	<b>NAC domain containing protein 73</b>
<b>260867_at</b>	<b>AT1G43790</b>	<b>TED6</b>	<b>0.62</b>	<b>Tracheary element differentiation-related 6</b>
251050_at	AT5G02440		0.61	Unknown protein
<b>264493_at</b>	<b>AT1G27440</b>	<b>GUT1,GUT2,IRX10</b>	<b>0.61</b>	<b>Exostosin family protein</b>
261653_at	AT1G01900	SBT1.1	0.61	Subtilase family protein
262922_at	AT1G79420		0.61	Protein of unknown function (DUF620)
<b>266244_at</b>	<b>AT2G27740</b>		<b>0.61</b>	<b>Family of unknown function (DUF662)*</b>
<b>257896_at</b>	<b>AT3G16920</b>	<b>CTL2</b>	<b>0.60</b>	<b>chitinase-like protein 2</b>
<b>249070_at</b>	<b>AT5G44030</b>	<b>IRX5,NWS2,CESA4</b>	<b>0.60</b>	<b>cellulose synthase A4</b>
<b>247590_at</b>	<b>AT5G60720</b>		<b>0.60</b>	<b>Protein of unknown function, DUF547*</b>
<b>267037_at</b>	<b>AT2G38320</b>	<b>TBL34</b>	<b>0.60</b>	<b>TRICHOME BIREFRINGENCE-LIKE 34*</b>
<b>247648_at</b>	<b>AT5G60020</b>	<b>LAC17</b>	<b>0.60</b>	<b>laccase 17</b>
<b>257151_at</b>	<b>AT3G27200</b>		<b>0.60</b>	<b>Cupredoxin superfamily protein*</b>
<b>256155_at</b>	<b>AT3G08500</b>	<b>MYB83</b>	<b>0.60</b>	<b>myb domain protein 83</b>
<b>254618_at</b>	<b>AT4G18780</b>	<b>IRX1,CESA8,LEW2</b>	<b>0.60</b>	<b>Cellulose synthase family protein</b>
<b>251630_at</b>	<b>AT3G57420</b>		<b>0.60</b>	<b>Protein of unknown function (DUF288)**</b>
<b>260914_at</b>	<b>AT1G02640</b>	<b>BXL2</b>	<b>0.60</b>	<b>beta-xylosidase 2**</b>
258357_at	AT3G14350	SRF7	0.59	STRUBBELIG-receptor family 7
<b>263841_at</b>	<b>AT2G36870</b>	<b>XTH32</b>	<b>0.59</b>	<b>xyloglucan endotransglucosylase/hydrolase 32</b>
259657_at	AT1G55180	PLDEPSILON,PLDALPHA4	0.59	phospholipase D alpha 4
<b>255150_at</b>	<b>AT4G08160</b>		<b>0.59</b>	<b>xylanase, glycosyl hydrolase family 10 protein*</b>
<b>246512_at</b>	<b>AT5G15630</b>	<b>IRX6,COBL4</b>	<b>0.58</b>	<b>COBRA-like extracellular glycosyl-phosphatidyl inositol-anchored protein family</b>
<b>253191_at</b>	<b>AT4G35350</b>	<b>XCP1</b>	<b>0.58</b>	<b>xylem cysteine peptidase 1</b>

\*Annotated involved in "xylan biosynthetic process" and/or "cell wall biogenesis" and/or "cell wall macromolecule metabolic process" in Tair (<http://www.arabidopsis.org>).

\*\*Located in cell wall according to Tair. Bold indicates related to cell wall formation.

**Table 3 | Top 50 co-expressed genes with MYB52 using Genevestigator platform.**

Probeset	AGI	Gene symbol	Score	Description
255903_at	AT1G17950	MYB52,BW52	1	myb domain protein 52
<b>245735_at</b>	<b>AT1G73410</b>	<b>MYB54</b>	<b>0.90</b>	<b>myb domain protein 54</b>
252550_at	AT3G45870		0.82	nodulin MtN21 /EamA-like transporter family protein
<b>264559_at</b>	<b>AT1G09610</b>	<b>GXM3</b>	<b>0.82</b>	<b>GLUCURONOXYLAN METHYLTRANSFERASE 3 (DUF579)</b>
248761_at	AT5G47635		0.81	Pollen Ole e 1 allergen and extensin family protein
<b>252211_at</b>	<b>AT3G50220</b>	<b>IRX15</b>	<b>0.81</b>	<b>Protein of unknown function (DUF579)</b>
<b>266783_at</b>	<b>AT2G29130</b>	<b>LAC2</b>	<b>0.80</b>	<b>laccase 2*</b>
<b>257233_at</b>	<b>AT3G15050</b>	<b>IQD10</b>	<b>0.79</b>	<b>IQ-domain 10*</b>
256054_at	AT1G07120		0.79	Tetratricopeptide repeat (TPR)-like superfamily protein
<b>254277_at</b>	<b>AT4G22680</b>	<b>MYB85</b>	<b>0.78</b>	<b>myb domain protein 85</b>
<b>264493_at</b>	<b>AT1G27440</b>	<b>GUT1,GUT2,IRX10</b>	<b>0.77</b>	<b>Exostosin family protein</b>
<b>250770_at</b>	<b>AT5G05390</b>	<b>LAC12</b>	<b>0.77</b>	<b>laccase 12*</b>
246425_at	<b>AT5G17420</b>	<b>IRX3,MUR10,CESA7</b>	<b>0.76</b>	<b>Cellulose synthase family protein</b>
251009_at	AT5G02640		0.76	Unknown protein
259584_at	AT1G28080		0.76	RING finger protein
<b>248121_at</b>	<b>AT5G54690</b>	<b>GAUT12,IRX8,LGT6</b>	<b>0.75</b>	<b>galacturonosyltransferase 12</b>
261928_at	AT1G22480		0.74	Cupredoxin superfamily protein
254170_at	AT4G24430		0.74	Rhamnogalacturonate lyase family protein
<b>246512_at</b>	<b>AT5G15630</b>	<b>IRX6,COBL4</b>	<b>0.74</b>	<b>COBRA-like4</b>
264495_at	AT1G27380	RIC2	0.74	ROP-interactive CRIB motif-containing protein 2
245105_at	AT2G41610		0.74	Unknown protein
<b>253327_at</b>	<b>AT4G33450</b>	<b>MYB69</b>	<b>0.74</b>	<b>myb domain protein 69</b>
<b>257896_at</b>	<b>AT3G16920</b>	<b>CTL2</b>	<b>0.73</b>	<b>chitinase-like protein 2</b>
<b>255150_at</b>	<b>AT4G08160</b>		<b>0.73</b>	<b>glycosyl hydrolase family 10 protein*</b>
266714_at			0.73	No_match
<b>260326_at</b>	<b>AT1G63910</b>	<b>MYB103</b>	<b>0.73</b>	<b>myb domain protein 103</b>
<b>248907_at</b>	<b>AT5G46340</b>	<b>RWA1</b>	<b>0.72</b>	<b>REDUCED WALL ACETYLATION 1, O-acetyltransferase family</b>
<b>266488_at</b>	<b>AT2G47670</b>		<b>0.72</b>	<b>Plant invertase/pectin methylesterase inhibitor family protein</b>
<b>253379_at</b>	<b>AT4G33330</b>	<b>PGSIP3,GUX2</b>	<b>0.72</b>	<b>plant glycogenin-like starch initiation protein 3</b>
<b>261399_at</b>	<b>AT1G79620</b>		<b>0.72</b>	<b>Leucine-rich repeat protein kinase family protein*</b>
<b>251478_at</b>	<b>AT3G59690</b>	<b>IQD13</b>	<b>0.71</b>	<b>IQ-domain 13*</b>
<b>265463_at</b>	<b>AT2G37090</b>	<b>IRX9</b>	<b>0.71</b>	<b>GT43 glycosyl transferase43</b>
<b>266708_at</b>	<b>AT2G03200</b>		<b>0.71</b>	<b>Eukaryotic aspartyl protease family protein*</b>
248887_at	AT5G46115		0.71	Unknown protein
267414_at	AT2G34790	MEE23,EDA28	0.70	FAD-binding Berberine family protein
246344_at	AT3G56730		0.70	Putative endonuclease or glycosyl hydrolase
<b>263470_at</b>	<b>AT2G31900</b>	<b>ATMYO5,XIF</b>	<b>0.70</b>	<b>myosin-like protein XIF*</b>
264305_at	AT1G78815	LSH7	0.70	LIGHT SENSITIVE HYPOCOTYLS 7, Protein of unknown function (DUF640)
<b>251131_at</b>	<b>AT5G01190</b>	<b>LAC10</b>	<b>0.69</b>	<b>laccase 10*</b>
<b>251093_at</b>	<b>AT5G01360</b>	<b>TBL3</b>	<b>0.69</b>	<b>Plant protein of unknown function (DUF828)*</b>
<b>247264_at</b>	<b>AT5G64530</b>	<b>XND1,ANAC104</b>	<b>0.69</b>	<b>xylem NAC domain 1</b>
260430_at	AT1G68200		0.69	Zinc finger C-x8-C-x5-C-x3-H type family protein
<b>261809_at</b>	<b>AT1G08340</b>		<b>0.69</b>	<b>Rho GTPase activating protein*</b>
<b>253798_at</b>	<b>AT4G28500</b>	<b>SND2,NAC073</b>	<b>0.69</b>	<b>NAC domain containing protein 73</b>
246342_at	AT3G56700	FAR6	0.69	Fatty acid reductase 6
<b>247030_at</b>	<b>AT5G67210</b>	<b>IRX15-L</b>	<b>0.69</b>	<b>Protein of unknown function (DUF579)</b>
<b>254618_at</b>	<b>AT4G18780</b>	<b>IRX1,CESA8,LEW2</b>	<b>0.69</b>	<b>Cellulose synthase family protein</b>
<b>247590_at</b>	<b>AT5G60720</b>		<b>0.68</b>	<b>Protein of unknown function, DUF547*</b>
<b>249439_at</b>	<b>AT5G40020</b>		<b>0.68</b>	<b>Pathogenesis-related thaumatin superfamily protein*</b>
<b>253877_at</b>	<b>AT4G27435</b>		<b>0.68</b>	<b>Protein of unknown function (DUF1218)*</b>
253710_at	AT4G29230	NAC075	0.68	NAC domain containing protein 75

\*Annotated involved in "xylan biosynthetic process" and/or "cell wall biogenesis" and/or "cell wall macromolecule metabolic process" in *Tair* (<http://www.arabidopsis.org>). Bold indicates related to cell wall formation.

Although none of these MYB-like factors has been yet reported as regulators of the SCW, the *myb-like TF* T-DNA mutant had a clear hyperlignification phenotype suggesting a repressor role of the lignin biosynthesis and/or SCW. The *myb-like TF* gene was annotated in TAIR (<http://www.arabidopsis.org>) as a putative MYB domain containing TF able to interact with the gene product of vacuolar ATPase subunit B1 (VHA-B1). Interestingly, it is highly co-expressed with a newly reported gene *XIP1* (*XYLEM INTERMIXED WITH PHLOEM1*), a leucine-rich repeat receptor-like kinase (Table S8). The *XIP1* knock-down mutants shows the accumulation of cells with ectopic lignification in regions of phloem in the vascular bundles of inflorescence stems (Bryan et al., 2012).

The homeodomain containing TFs were well represented in the list of candidate genes with nine members. Members of this family have been shown to regulate procambium cell activities by promoting secondary walled xylem cell differentiation during vascular development. Some HD-ZIP III TF (HB8, PHV/HB9, PHB/HB14, REV/IFL1, and CAN/HB15) and KANADI TF (KAN1-KAN3) were shown to be involved in the secondary walled cell type formation and patterning in roots and stems (Baima et al., 2001; Emery et al., 2003; Kim et al., 2005; Ilegems et al., 2010). Three of the homeodomain TF mutants analyzed in our study exhibited SCW phenotypes. The *blh6* mutant had less lignified SCW mainly in the xylary and interfascicular fibers, whereas in the *hb5* mutant the fibers in both fascicular and interfascicular regions were heavily lignified. In the *hb15* mutant, both regions were also highly lignified but in addition ectopic lignification was observed in the parenchymatous cells adjacent to fiber and xylem cells (Figures 4B,D). This suggests that HB15 represses the SCW formation program rather than only promote the xylem cell differentiation as was concluded from earlier studies where down-regulation of *CAN/HB15* stimulated xylem production, and over-expression of a miR166-resistant *HB15* (gain-of-function mutant) resulted in reduced xylem formation (Kim et al., 2005). Co-expression analyses revealed interesting clues for *BLH6* which was co-expressed with genes involved in the biosynthesis of the three main polymers i.e., cellulose, xylan, and lignin as well as with the master regulator *MYB46* and its closest paralog *MYB83*. Together with the hypolignified phenotype of the mutant and the thinner SCW particularly in the fibers, this further supports a role of BLH6 as an activator of the whole SCW program.

*AP2 ERF* TF (AT3G14230) was identified in all four SCW-related transcriptomic datasets and exhibited high and preferential expression in xylem, but the corresponding mutant had no detectable cell wall phenotype. Twelve members of the AP2 ERF TF family were highlighted by our *in silico* approach, seven of which had high and preferential expression in xylem and another (AT5G61590) was strongly induced during *in vitro* tracheary element formation. Although this family was the second most highly represented TF family just after the MYBs in the 80 candidate list, none of its members have yet been shown to be directly involved in the regulation of SCW formation. This family therefore deserves more attention especially because it was recently reported that ethylene regulates cambium activity and promotes secondary walled xylem formation (Love et al., 2009).

Some members of the auxin-dependent TFs Aux/IAA and ARF families have been shown to be involved in vascular tissue formation. For example, loss-of-function in *ARF5/MP* (Hardtke and Berleth, 1998) and gain-of-function in *IAA12/BDL* (Hamann et al., 2002) resulted in reduced and discontinuous vascular formation. These TF families were also highly represented within the 80 candidates with seven and two members for Aux/IAA and ARF, respectively. *IAA9* was a very promising candidate found in the four transcriptomic datasets, highly and preferentially expressed in xylem and during tracheary elements differentiation. Unfortunately the corresponding mutant was unavailable at the time this work was performed. T-DNA insertion mutants corresponding to *ARF4*, *ARF6* and *IAA28*, and an *IAA11* RNAi transgenic line were analyzed here but did not show any obvious SCW phenotype. This is very likely due to their functional redundancy as reported in previous studies (Okushima et al., 2005; Overvoorde et al., 2005). The creation of double/triple mutants of these paralog genes might be necessary to further assess their involvement in SCW formation.

The hypolignified lines *blh6* and *zinc finger TF* displayed earlier flowering time as compared to control whereas the hyperlignification line *hb5* exhibited delayed flowering time. Two previous studies demonstrated that flowering induction time was determinant for xylem expansion and SCW formation in Arabidopsis hypocotyls and roots. Some major QTLs for SCW thickening during xylem expansion and fiber differentiation correlated tightly with a major flowering time QTL. In addition, transient induction of flowering at the rosette stage promoted SCW thickening and xylem expansion (Sibout et al., 2008). Double mutant of two flowering time genes *soc1 ful* showed a synergistically delayed flowering time and a dramatically increased SCW formation with wood development present throughout all stems and to a much larger extent than any Arabidopsis mutant described to date (Melzer et al., 2008). Collectively these results suggest that the flowering induction is coupled with the SCW thickening program and xylem formation.

In conclusion, we described here a post-genomic approach that enabled us to propose a list of 80 promising candidate genes potentially regulating SCW formation and/or lignification. Many of the available mutants analyzed did not provide any detectable SCW phenotype and complementary approaches (overexpression, using different alleles, dominant repression, or multiple mutants) are now necessary to further characterize their function. However, the six TFs of which mutants exhibited clear lignin phenotypes, further highlight the complexity of the regulatory network controlling SCW formation. Their in depth functional characterization should allow a better understanding of the regulation of lignification and SCW formation which may ultimately be used to improve the saccharification potential.

## MATERIALS AND METHODS

### CROSS-COMPARISON OF MICROARRAY DATASETS

Four *in house* microarray datasets were generated in our laboratory. In brief, datasets are from *wat1* T-DNA Arabidopsis mutant CATMA microarray (Ranocha et al., 2010); EgMYB1 (Legay et al., 2010), EgMYB2 over-expressed in *Arabidopsis* (unpublished), and orthologs of *Eucalyptus* xylem expressed genes



(Rengel et al., 2009). Publicly available microarray datasets were extracted from Genevestigator (<https://www.genevestigator.com>) (Hruz et al., 2008) by using Arabidopsis ATH1 22k array platform (7392 array datasets).

### PLANT MATERIAL AND GROWTH CONDITION

The mutant lines were isolated from the T-DNA mutagenized populations in the SALK collection (Alonso et al., 2003) and from the RNAi transgenic plant populations in the Agrikola collection (<http://www.agrikola.org>). Seeds were obtained from the Nottingham Arabidopsis Stock Center (NASC) (<http://arabidopsis.info/>) and GABI (<http://www.gabi-kat.de/>). Homozygote lines were obtained from NASC or generated in lab and verified by PCR genotyping with gene specific primers and the respective left border primers of the T-DNA listed in supplementary Table S11. The transcript levels of each target gene in the six T-DNA insertion mutant were assessed (Figure S1) and the corresponding primers are listed in supplementary Table S12. Plants were grown in jiffy peat pellets then transferred to standard soil in culture room in short day conditions [9 h light, 200  $\mu\text{mol photons m}^{-1}\text{s}^{-1}$ , 22°C (day)/20°C (night), 70% RH]. The flowering time was considered from sowing day until the flower stem reached 20 cm in height.

### MICROSCOPY

The histological comparative analysis of SCW between wild type and mutants was done at the stage of newly formed green siliques, about 2 weeks after bolting, when the inflorescence stems reach 20 cm in height. At this stage, the basal part of the inflorescence stem abundantly develops cells undergoing secondary wall thickening (xylem vessel cells, fascicular, and interfascicular fiber cells). Lignin polymers are the characteristic components of SCW and are normally absent from primary cell wall, therefore we used lignin deposition detection techniques to screen for SCW phenotype. Two methods were then chosen to detect the lignin polymers in the sections for microscopic observation. Firstly we used the natural auto fluorescence of the aromatic ring moieties on the subunits of the lignin polymer under UV-light exposition. Secondly, we used the phloroglucinol-HCl coloration which stains specifically lignin polymer precursors coniferaldehyde and *p*-coumaraldehyde in the SCW giving a red-purple color when observed under normal light. Cross sections of inflorescence stems at the basal end (100–150  $\mu\text{m}$ ) were either

observed using auto-fluorescence or stained with phloroglucinol-HCl. Auto-fluorescence was observed with a Leica microscope (excitation filter Bp 340–380 nm; suppression filter Lp 430 nm; <http://leica.com>). Phloroglucinol-HCl was directly applied on the slide. Images were recorded with a CCD camera (Photonic Science, <http://www.photonic-science.co.uk>).

### CO-EXPRESSION ANALYSIS

Three co-expression analysis tools were explored using Genevestigator (<https://www.genevestigator.com>), Arabidopsis co-expression data mining tools (<http://www.arabidopsis.leeds.ac.uk/act/>), and GeneCAT (<http://genecat.mpg.de/>). The results were presented using Genevestigator output tables and genes classified according to gene ontology semantic (Berardini et al., 2004). We used Genevestigator Arabidopsis ATH1 22k array platform with *in absentia* parameters that comprise all 7392 qualified datasets and is regardless of the underlying microarray datasets and the bait genes (i.e., all samples, condition-independent, and no-tissues specific bait genes), 50 was as “cut-off” threshold for co-expressed genes list.

### ACKNOWLEDGMENTS

This work was supported by grants from the European FP7 project RENEWALL (FP7-211982), the Centre National pour la Recherche Scientifique (CNRS), and the Université Toulouse III Paul Sabatier (UPS). This work was part of the Laboratoire d'Excellence (LABEX) project entitled TULIP (ANR-10-LABX-41). The authors are grateful to Prof S. Hawkins (Université de Lille, France) for kindly communicating unpublished data on Arabidopsis lines over-expressing *EgMYB2*. We also acknowledge Dr. P. Ranocha (LRSV) for his precious advice and help since the beginning of this work, Y. Martinez (FR3450) for assistance with microscopy analysis. Thanks also to PhD student H. Yu for her help in quantifying the transcript levels of *HB15* and *ZINC FINGER TF* in their corresponding T-DNA insertional mutants and the internship training students C. Lin and R. Kardinskaite for their help with plant growth, genotyping, and phenotyping.

### SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: [http://www.frontiersin.org/Plant\\_Biotechnology/10.3389/fpls.2013.00189/abstract](http://www.frontiersin.org/Plant_Biotechnology/10.3389/fpls.2013.00189/abstract)

### REFERENCES

- Alonso, J. M., Stepanova, A. N., Leisse, T. J., Kim, C. J., Chen, H., Shinn, P., et al. (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301, 653–657. doi: 10.1126/science.1086391
- Baima, S., Possenti, M., Matteucci, A., Wisman, E., Altamura, M. M., Ruberti, I., et al. (2001). The Arabidopsis ATHB-8 HD-zip protein acts as a differentiation-promoting transcription factor of the vascular meristems. *Plant Physiol.* 126, 643–655. doi: 10.1104/pp.126.2.643
- Berardini, T. Z., Mundodi, S., Reiser, L., Huala, E., Garcia-Hernandez, M., Zhang, P., et al. (2004). Functional annotation of the Arabidopsis genome using controlled vocabularies. *Plant Physiol.* 135, 1–11. doi: 10.1104/pp.104.040071
- Berthet, S., Demont-Caulet, N., Pollet, B., Bidzinski, P., Cezard, L., Le Bris, P., et al. (2011). Disruption of LACCASE4 and 17 results in tissue-specific alterations to lignification of *Arabidopsis thaliana* stems. *Plant Cell* 23, 1124–1137. doi: 10.1105/tpc.110.082792
- Brown, D. M., Zhang, Z., Stephens, E., Dupree, P., and Turner, S. R. (2009). Characterization of IRX10 and IRX10-like reveals an essential role in glucuronoxylan biosynthesis in Arabidopsis. *Plant J.* 57, 732–746. doi: 10.1111/j.1365-313X.2008.03729.x
- Bryan, A. C., Obaidi, A., Wierzba, M., and Tax, F. E. (2012). XYLEM INTERMIXED WITH PHLOEM1, a leucine-rich repeat receptor-like kinase required for stem growth and vascular development in *Arabidopsis thaliana*. *Planta* 235, 111–122. doi: 10.1007/s00425-011-1489-6
- De Micco, V., Ruel, K., Joseleau, J. P., Grima-Pettenati, J., and Aronne, G. (2012). Xylem anatomy and cell wall ultrastructure of *Nicotiana Tabacum* after lignin genetic modification through transcriptional activator EGMVYB2. *IAWA J.* 33, 269–286.
- Demura, T., and Fukuda, H. (2007). Transcriptional regulation in wood formation.

- Trends Plant Sci.* 12, 64–70. doi: 10.1016/j.tplants.2006.12.006
- Emery, J. F., Floyd, S. K., Alvarez, J., Eshed, Y., Hawker, N. P., Izhaki, A., et al. (2003). Radial patterning of Arabidopsis shoots by class III HD-ZIP and KANADI genes. *Curr. Biol.* 13, 1768–1774. doi: 10.1016/j.cub.2003.09.035
- Goicoechea, M., Lacombe, E., Legay, S., Mihaljevic, S., Rech, P., Jauneau, A., et al. (2005). EgMYB2, a new transcriptional activator from Eucalyptus xylem, regulates secondary cell wall formation and lignin biosynthesis. *Plant J.* 43, 553–567. doi: 10.1111/j.1365-313X.2005.02480.x
- Grima-Pettenati, J., Soler, M., Camargo, E. L. O., and Wang, H. (2012). Transcriptional regulation of the lignin biosynthetic pathway revisited: new players and insights. *Lignins Biosynth. Biodegrad. Bioeng.* 61, 173–218.
- Hamann, T., Benkova, E., Baurle, I., Kientz, M., and Jurgens, G. (2002). The Arabidopsis BODENLOS gene encodes an auxin response protein inhibiting MONOPTEROS-mediated embryo patterning. *Genes Dev.* 16, 1610–1615. doi: 10.1101/gad.229402
- Hardtke, C. S., and Berleth, T. (1998). The Arabidopsis gene MONOPTEROS encodes a transcription factor mediating embryo axis formation and vascular development. *EMBO J.* 17, 1405–1411. doi: 10.1093/emboj/17.5.1405
- Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., et al. (2008). Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Adv. Bioinformatics* 2008, 420747. doi: 10.1155/2008/420747
- Ilegems, M., Douet, V., Meylan-Bettex, M., Uyttewaal, M., Brand, L., Bowman, J. L., et al. (2010). Interplay of auxin, KANADI and Class III HD-ZIP transcription factors in vascular tissue formation. *Development* 137, 975–984. doi: 10.1242/dev.047662
- Jensen, J. K., Kim, H., Cocuron, J. C., Orler, R., Ralph, J., and Wilkerson, C. G. (2011). The DUF579 domain containing proteins IRX15 and IRX15-L affect xylan synthesis in Arabidopsis. *Plant J.* 66, 387–400. doi: 10.1111/j.1365-313X.2010.04475.x
- Kim, J., Jung, J. H., Reyes, J. L., Kim, Y. S., Kim, S. Y., Chung, K. S., et al. (2005). microRNA-directed cleavage of ATHB15 mRNA regulates vascular development in Arabidopsis inflorescence stems. *Plant J.* 42, 84–94. doi: 10.1111/j.1365-313X.2005.02354.x
- Ko, J. H., Kim, W. C., and Han, K. H. (2009). Ectopic expression of MYB46 identifies transcriptional regulatory genes involved in secondary wall biosynthesis in Arabidopsis. *Plant J.* 60, 649–665. doi: 10.1111/j.1365-313X.2009.03989.x
- Ko, J. H., Yang, S. H., Park, A. H., Lerouxel, O., and Han, K. H. (2007). ANAC012, a member of the plant-specific NAC transcription factor family, negatively regulates xylary fiber development in Arabidopsis thaliana. *Plant J.* 50, 1035–1048. doi: 10.1111/j.1365-313X.2007.03109.x
- Kubo, M., Udagawa, M., Nishikubo, N., Horiguchi, G., Yamaguchi, M., Ito, J., et al. (2005). Transcription switches for protoxylem and metaxylem vessel formation. *Genes Dev.* 19, 1855–1860. doi: 10.1101/gad.1331305
- Legay, S., Sivadon, P., Blervacq, A. S., Pavy, N., Baghdady, A., Tremblay, L., et al. (2010). EgMYB1, an R2R3 MYB transcription factor from eucalyptus negatively regulates secondary cell wall formation in Arabidopsis and poplar. *New Phytol.* 188, 774–786. doi: 10.1111/j.1469-8137.2010.03432.x
- Love, J., Bjorklund, S., Vahala, J., Hertzberg, M., Kangasjarvi, J., and Sundberg, B. (2009). Ethylene is an endogenous stimulator of cell division in the cambial meristem of Populus. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5984–5989. doi: 10.1073/pnas.0811660106
- MacMillan, C. P., Mansfield, S. D., Stachurski, Z. H., Evans, R., and Southerton, S. G. (2010). Fasciclin-like arabinogalactan proteins: specialization for stem biomechanics and cell wall architecture in Arabidopsis and Eucalyptus. *Plant J.* 62, 689–703. doi: 10.1111/j.1365-313X.2010.04181.x
- McCarthy, R. L., Zhong, R., and Ye, Z. H. (2009). MYB83 is a direct target of SND1 and acts redundantly with MYB46 in the regulation of secondary cell wall biosynthesis in Arabidopsis. *Plant Cell Physiol.* 50, 1950–1964. doi: 10.1093/pcp/pcp139
- McConnell, J. R., Emery, J., Eshed, Y., Bao, N., Bowman, J., and Barton, M. K. (2001). Role of PHABULOSA and PHAVOLUTA in determining radial patterning in shoots. *Nature* 411, 709–713. doi: 10.1038/35079635
- Melzer, S., Lens, F., Gennen, J., Vanneste, S., Rohde, A., and Beeckman, T. (2008). Flowering-time genes modulate meristem determinacy and growth form in Arabidopsis thaliana. *Nat. Genet.* 40, 1489–1492. doi: 10.1038/ng.253
- Mitsuda, N., Iwase, A., Yamamoto, H., Yoshida, M., Seki, M., Shinozaki, K., et al. (2007). NAC transcription factors, NST1 and NST3, are key regulators of the formation of secondary walls in woody tissues of Arabidopsis. *Plant Cell* 19, 270–280. doi: 10.1105/tpc.106.047043
- Mitsuda, N., Seki, M., Shinozaki, K., and Ohme-Takagi, M. (2005). The NAC transcription factors NST1 and NST2 of Arabidopsis regulate secondary wall thickenings and are required for anther dehiscence. *Plant Cell* 17, 2993–3006. doi: 10.1105/tpc.105.036004
- Ohman, D., Demedts, B., Kumar, M., Gerber, L., Gorzsas, A., Goeminne, G., et al. (2013). MYB103 is required for FERULATE-5-HYDROXYLASE expression and syringyl lignin biosynthesis in Arabidopsis stems. *Plant J.* 73, 63–76. doi: 10.1111/tpj.12018
- Okushima, Y., Overvoorde, P. J., Arima, K., Alonso, J. M., Chan, A., Chang, C., et al. (2005). Functional genomic analysis of the AUXIN RESPONSE FACTOR gene family members in Arabidopsis thaliana: unique and overlapping functions of ARF7 and ARF19. *Plant Cell* 17, 444–463. doi: 10.1105/tpc.104.028316
- Overvoorde, P. J., Okushima, Y., Alonso, J. M., Chan, A., Chang, C., Ecker, J. R., et al. (2005). Functional genomic analysis of the AUXIN/INDOLE-3-ACETIC ACID gene family members in Arabidopsis thaliana. *Plant Cell* 17, 3282–3300. doi: 10.1105/tpc.105.036723
- Ranocha, P., Denance, N., Vanholme, R., Freydrich, A., Martinez, Y., Hoffmann, L., et al. (2010). Walls are thin 1 (WAT1), an Arabidopsis homolog of Medicago truncatula NODULIN21, is a tonoplast-localized protein required for secondary wall formation in fibers. *Plant J.* 63, 469–483. doi: 10.1111/j.1365-313X.2010.04256.x
- Rengel, D., San Clemente, H., Servant, F., Ladouce, N., Paux, E., Wincker, P., et al. (2009). A new genomic resource dedicated to wood formation in Eucalyptus. *BMC Plant Biol.* 9:36. doi: 10.1186/1471-2229-9-36
- Romano, J. M., Dubos, C., Prouse, M. B., Wilkins, O., Hong, H., Poole, M., et al. (2012). AtMYB61, an R2R3-MYB transcription factor, functions as a pleiotropic regulator via a small gene network. *New Phytol.* 195, 774–786. doi: 10.1111/j.1469-8137.2012.04201.x
- Ruprecht, C., Mutwil, M., Saxe, F., Eder, M., Nikoloski, Z., and Persson, S. (2011). Large-scale co-expression approach to dissect secondary cell wall formation across plant species. *Front. Plant Sci.* 2:23. doi: 10.3389/fpls.2011.00023
- Ruprecht, C., and Persson, S. (2012). Co-expression of cell-wall related genes: new tools and insights. *Front. Plant Sci.* 3:83. doi: 10.3389/fpls.2012.00083
- Sibout, R., Plantegenet, S., and Hardtke, C. S. (2008). Flowering as a condition for xylem expansion in Arabidopsis hypocotyl and root. *Curr. Biol.* 18, 458–463. doi: 10.1016/j.cub.2008.02.070
- Wang, H., Avci, U., Nakashima, J., Hahn, M. G., Chen, F., and Dixon, R. A. (2010). Mutation of WRKY transcription factors initiates pith secondary wall formation and increases stem biomass in dicotyledonous plants. *Proc. Natl. Acad. Sci. U.S.A.* 107, 22338–22343. doi: 10.1073/pnas.1016436107
- Wang, H. Z., and Dixon, R. A. (2012). On-off switches for secondary cell wall biosynthesis. *Mol. Plant* 5, 297–303. doi: 10.1093/mp/ssr098
- Yamaguchi, M., Ohtani, M., Mitsuda, N., Kubo, M., Ohme-Takagi, M., Fukuda, H., et al. (2010). VND-INTERACTING2, a NAC domain transcription factor, negatively regulates xylem vessel formation in Arabidopsis. *Plant Cell* 22, 1249–1263. doi: 10.1105/tpc.108.064048
- Zhao, C., Avci, U., Grant, E. H., Haigler, C. H., and Beers, E. P. (2008). XND1, a member of the NAC domain family in Arabidopsis thaliana, negatively regulates lignocellulose synthesis and programmed cell death in xylem. *Plant J.* 53, 425–436. doi: 10.1111/j.1365-313X.2007.03350.x
- Zhong, R., Demura, T., and Ye, Z. H. (2006). SND1, a NAC domain transcription factor, is a key regulator of secondary wall synthesis in fibers of Arabidopsis. *Plant Cell* 18, 3158–3170. doi: 10.1105/tpc.106.047399
- Zhong, R., Lee, C., and Ye, Z. H. (2010). Evolutionary conservation of the transcriptional network regulating secondary cell wall biosynthesis. *Trends Plant Sci.* 15, 625–632. doi: 10.1016/j.tplants.2010.08.007
- Zhong, R., Lee, C., and Ye, Z. H. (2012). Global analysis of direct targets of secondary wall NAC master switches in Arabidopsis.

- Mol. Plant* 3, 1087–1103. doi: 10.1093/mp/ssq062
- Zhong, R., Lee, C., Zhou, J., McCarthy, R. L., and Ye, Z. H. (2008). A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in Arabidopsis. *Plant Cell* 20, 2763–2782. doi: 10.1105/tpc.108.061325
- Zhong, R., Richardson, E. A., and Ye, Z. H. (2007a). The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in Arabidopsis. *Plant Cell* 19, 2776–2792. doi: 10.1105/tpc.107.053678
- Zhong, R., Richardson, E. A., and Ye, Z. H. (2007b). Two NAC domain transcription factors, SND1 and NST1, function redundantly in regulation of secondary wall synthesis in fibers of Arabidopsis. *Planta* 225, 1603–1611. doi: 10.1007/s00425-007-0498-y
- Zhong, R., and Ye, Z. H. (2007). Regulation of cell wall biosynthesis. *Curr. Opin. Plant Biol.* 10, 564–572. doi: 10.1016/j.pbi.2007.09.001
- Zhong, R., and Ye, Z. H. (2009). Transcriptional regulation of lignin biosynthesis. *Plant Signal. Behav.* 4, 1028–1034. doi: 10.4161/psb.4.11.9875
- Zhong, R., and Ye, Z. H. (2012). MYB46 and MYB83 bind to the SMRE sites and directly activate a suite of transcription factors and secondary wall biosynthetic genes. *Plant Cell Physiol.* 53, 368–380. doi: 10.1093/pcp/pcr185
- 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. doi: 10.1105/tpc.108.063321
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 27 March 2013; paper pending published: 10 April 2013; accepted: 23 May 2013; published online: 11 June 2013.
- Citation:* Cassan-Wang H, Goué N, Saidi MN, Legay S, Sivadon P, Goffner D and Grima-Pettenati J (2013) Identification of novel transcription factors regulating secondary cell wall formation in Arabidopsis. *Front. Plant Sci.* 4:189. doi: 10.3389/fpls.2013.00189
- This article was submitted to *Frontiers in Plant Biotechnology*, a specialty of *Frontiers in Plant Science*. Copyright © 2013 Cassan-Wang, Goué, Saidi, Legay, Sivadon, Goffner and Grima-Pettenati. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

## CHAPITRE 3 - Travaux de recherche en tant que Maître de Conférences





En 2009, j'ai été recrutée en tant que Maître de Conférences à l'Université Toulouse III – Paul Sabatier. Mon laboratoire d'accueil est l'UMR 5546 UPS/CNRS « Laboratoire de Recherche en Sciences Végétales » dirigé actuellement par M. Bernard Dumas. J'ai été plus particulièrement rattachée à l'équipe GFE « Génomique fonctionnelle de l'*Eucalyptus* », qui a été dirigée par Mme Jacqueline Grima-Pettenati, cette équipe a évolué en l'équipe ReDyWood « Régulation transcriptionnelle de la formation du bois » dirigée par M. Fabien Mounet actuellement. La thématique de recherche concerne l'identification des régulateurs transcriptionnels clés impliqués dans la formation du bois chez l'*Eucalyptus* aux réponses aux stress abiotiques.

## A. Participation aux projets de l'équipe d'accueil

Ma participation à la thématique de l'équipe d'accueil m'a conduite à m'intégrer dans des projets, et dans les collaborations établies. Ces activités de recherche ont permis d'identifier des nouveaux gènes candidats pour les 'masters' régulateurs EgMYB1 et EgMYB2 par la technique de double hybride chez les levures. Suite aux réponses d'appel d'offre, j'ai obtenu le Projet BQR de l'Université de Toulouse III Paul Sabatier en 2010 avec financement accordé de 10 000 €, ce projet a fait aussi le sujet de stage M2R de M. Matthieu Bensussan. Ces résultats ont pu servir pour initier un nouveau projet européen Plant KBBE « Tree for Joules », financement accordé en 2011, dans lequel j'ai participé au montage du projet et étais impliquée en tant que responsable du Task 2 dans WP1 « Transcript profiling of selected candidates genes ». Lors du déroulement de ce projet j'ai pu identifier et évaluer des gènes de référence pour des études précises de PCR quantitative sur *Eucalyptus*, des dizaines de gènes candidats impliqués dans la régulation de la différenciation et formation des cellules du bois ont ainsi été étudiés (Cassan-Wang et al., 2012; Cassan-Wang, Soler, et al., 2013; Yu et al., 2014; Yu et al., 2015) (annexe 4). Ce projet EU KBBE (2011-2015) m'a aidé à réaliser 7 publications avec une moyenne d'impact factor de 10 (Carocha et al., 2015; Cassan-Wang et al., 2012; Li et al., 2015; Myburg et al., 2014; Soler et al., 2015; Yu et al., 2014; Yu et al., 2015), grâce notamment à ma participation en tant que co-auteur dans l'élaboration d'un article « The genome of *Eucalyptus grandis* » dans le journal « Nature » (IF=42). J'ai ainsi pu établir plusieurs protocoles au laboratoire pour l'étude d'expression génique chez les ligneux (Carocha et al., 2015; Cassan-Wang et al., 2012; Soler et al., 2015; Yu et al., 2014; Yu et al., 2015). J'ai eu le plaisir d'avoir un fils début 2014 et une fille en fin 2015 ce qui explique la diminution du nombre de publication entre 2016 et 2018. Avec le recrutement de Dr. Fabien Mounet en 2011, et la fusion de l'équipe avec une équipe s'intéressant à la réponse aux stress abiotiques dirigée par Prof Marie-Chantal Teulières, j'ai aussi participé à la nouvelle thématique « Stress abiotique » de l'équipe sur « la régulation de la formation du bois en condition de stress abiotiques chez l'*Eucalyptus* » (Hadj Bachir et al., 2022).

### Les résumés des publications réalisées dans cet axe :

1. Cassan-Wang, H., et al. (2012).

**Cassan-Wang H, Soler M, Yu H, Camargo ELO, Carocha V, Ladouce N, Savelli B, Paiva JAP, Lepage JC, Grima-Pettenati J. 2012. Reference Genes for High-Throughput Quantitative Reverse Transcription-PCR Analysis of Gene Expression in Organs and Tissues of *Eucalyptus* Grown in Various Environmental Conditions. *Plant and Cell Physiology* 53(12): 2101-2116. (IF=5.0)**

Interest in the genomics of *Eucalyptus* has skyrocketed thanks to the recent sequencing of the genome of *Eucalyptus grandis* and to a growing number of large-scale transcriptomic studies.

Quantitative reverse transcription-PCR (RT-PCR) is the method of choice for gene expression analysis and can now also be used as a high-throughput method. The selection of appropriate internal controls is becoming of utmost importance to ensure accurate expression results in *Eucalyptus*. To this end, we selected 21 candidate reference genes and used high-throughput microfluidic dynamic arrays to assess their expression among a large panel of developmental and environmental conditions with a special focus on wood-forming tissues. We analyzed the expression stability of these genes by using three distinct statistical algorithms (geNorm, NormFinder and delta Ct), and used principal component analysis to compare methods and rankings. We showed that the most stable genes identified depended not only on the panel of biological samples considered but also on the statistical method used. We then developed a comprehensive integration of the rankings generated by the three methods and identified the optimal reference genes for 17 distinct experimental sets covering 13 organs and tissues, as well as various developmental and environmental conditions. The expression patterns of *Eucalyptus* master genes EgMYB1 and EgMYB2 experimentally validated our selection. Our findings provide an important resource for the selection of appropriate reference genes for accurate and reliable normalization of gene expression data in the organs and tissues of *Eucalyptus* trees grown in a range of conditions including abiotic stress.

2. Grima-Pettenati, J., et al. (2012).

**Grima-Pettenati J, Soler M, Camargo ELO, Wang H. (2012). Transcriptional Regulation of the Lignin Biosynthetic Pathway Revisited: New Players and Insights. Lignin: Biosynthesis, Biodegradation and Bioengineering 61: 173-218. Doi 10.1016/B978-0-12-416023-1.00006-9.**

The discovery that AC elements coordinated the regulation of genes belonging to the entire lignin biosynthetic pathway was the first breakthrough in understanding how lignin biosynthesis is regulated. Since then, tremendous progress has been made in the identification and characterization of many transcription factors (TFs) that regulate the genes of the phenylpropanoid branch pathway leading to lignin. A major breakthrough consisted in the discovery of a hierarchical transcriptional network regulating the biosynthesis of lignified secondary walls (SWs) in *Arabidopsis*. The NAC TFs (VND/NST/SND) work as the first layer of master switches activating the whole SW biosynthetic network through the regulation of a cascade of downstream TFs. Among these, MYB46/83 act as a second layer of master switches. Recent findings, however, reveal that the regulation of SW formation is far more complex than initially thought, involving both positive and negative regulators, dual function regulators, feedback loops, combinatorial complexes and cross talk between pathways. Finally, because of the great potential that lignocellulosic biomass represents for the production of bioenergy, there is a great interest in further elucidating the molecular mechanisms underlying the regulation of lignified SW and subsequently applying this knowledge to improve their saccharification potential for the generation of biofuels.

3. Courtial, A., et al. (2013).

**Courtial A, Soler M, Chateigner-Boutin A, Reymond M, Méchin V, Wang H, Grima-Pettenati J, Barriere Y. 2013. Breeding grasses for capacity to biofuel production or silage feeding value: an updated list of genes involved in maize secondary cell wall biosynthesis and assembly. Myadica 58(1): 67-102.**

In the near future, maize, sorghum, or switchgrass stovers and cereal straws will be a significant source of carbohydrates for sustainable biofuel production, in addition to the current use of grass silage in cattle feeding. However, cell wall properties, including the enzymatic degradability of structural polysaccharides in industrial fermenters or animal rumen, is greatly influenced by the embedding of cell wall carbohydrates in lignin matrix, and the linkages between lignins, p-

hydroxycinnamic acids, and arabinoxylans. Breeding for higher and cheaper biofuel or silage production will thus be based on the discovery of genetic traits involved in each cell wall component biosynthesis and deposition in each lignified tissue. Due to its considerable genetic and genomic backgrounds, maize is the relevant model species for identifying traits underlying cell wall degradability variations in grasses. Maize genes involved or putatively involved in the biosynthesis of cell wall phenolic compounds, cell wall carbohydrates and regulation factors were therefore searched for using data available in grass, *Arabidopsis*, and woody species (mostly poplar and *Eucalyptus*). All maize ortholog genes were searched for using protein sequences and a "blastp" strategy against data available in the [www.maizesequence.org](http://www.maizesequence.org) database. Genes were also mapped in silico considering their physical position in the same database. Finally, 409 candidate genes putatively involved in secondary cell wall biosynthesis and assembly were shown in the maize genome, out of which 130 were related to phenolic compound biosynthesis, 81 were related to cell wall carbohydrate biosynthesis, and 198 were involved in more or less known regulation mechanisms. Most probable candidate genes involved in regulation and assembly of secondary cell wall belonged to the MYB (45 genes) and NAC (38 genes) families, but also included zinc finger and HDZipIII encoding genes. While genes involved in ferulic acid cross-linkages with other cell wall components were little known, several families putatively involved in (arabino)-xylan chain biosynthesis and in feruloyl transfer were shown, including especially arabinosyl-CoA-acyltransferases, feruloyl-AX beta-1,2-xylosyl transferases, and xylan-O-3-arabinosyl transferases. This candidate gene list, which focused on genes and orthologs known to be involved in cell wall component biosynthesis and regulation, cannot be considered as exhaustive. Other genes, whose role in cell wall lignification and deposition have not yet been defined, should very likely be added to the list of candidates required for secondary cell wall assembly. Genes encoding proteins of still unknown function should also be added to the list, as several of the latter are probably involved in lignified tissue biosynthesis and deposition.

#### 4. Myburg, A. A., et al. (2014).

**Myburg, A.A., Grattapaglia, D., Tuskan, G.A., Hellsten, U., Hayes, R.D., Grimwood, J., Jenkins, J., Lindquist, E., Tice, H., Bauer, D., Goodstein, D.M., Dubchak, I., Poliakov, A., Mizrachi, E., Kullam, A.R., Hussey, S.G., Pinard, D., van der Merwe, K., Singh, P., van Jaarsveld, I., Silva-Junior, O.B., Togawa, R.C., Pappas, M.R., Faria, D.A., Sansaloni, C.P., Petroli, C.D., Yang, X., Ranjan, P., Tschaplinski, T.J., Ye, C.Y., Li, T., Sterck, L., Vanneste, K., Murat, F., Soler, M., Clemente, H.S., Saidi, N., Cassan-Wang, H., Dunand, C., Hefer, C.A., Bornberg-Bauer, E., Kersting, A.R., Vining, K., Amarasinghe, V., Ranik, M., Naithani, S., Elser, J., Boyd, A.E., Liston, A., Spatafora, J.W., Dharmwardhana, P., Raja, R., Sullivan, C., Romanel, E., Alves-Ferreira, M., Kulheim, C., Foley, W., Carocha, V., Paiva, J., Kudrna, D., Brommonschenkel, S.H., Pasquali, G., Byrne, M., Rigault, P., Tibbits, J., Spokevicius, A., Jones, R.C., Steane, D.A., Vaillancourt, R.E., Potts, B.M., Joubert, F., Barry, K., Pappas, G.J., Strauss, S.H., Jaiswal, P., Grima-Pettenati, J., Salse, J., Van de Peer, Y., Rokhsar, D.S., and Schmutz, J. (2014). The genome of *Eucalyptus grandis*. *Nature* 510, 356-362.**

Eucalypts are the world's most widely planted hardwood trees. Their outstanding diversity, adaptability and growth have made them a global renewable resource of fibre and energy. We sequenced and assembled >94% of the 640-megabase genome of *Eucalyptus grandis*. Of 36,376 predicted protein-coding genes, 34% occur in tandem duplications, the largest proportion thus far in plant genomes. *Eucalyptus* also shows the highest diversity of genes for specialized metabolites such as terpenes that act as chemical defence and provide unique pharmaceutical oils. Genome sequencing of the *E. grandis* sister species *E. globulus* and a set of inbred *E. grandis* tree genomes reveals dynamic genome evolution and hotspots of inbreeding depression. The *E. grandis* genome is the first reference

for the eudicot order Myrtales and is placed here sister to the eurosids. This resource expands our understanding of the unique biology of large woody perennials and provides a powerful tool to accelerate comparative biology, breeding and biotechnology.

5. Yu, H., et al. (2014).

**Yu H, Soler M, Mila I, San Clemente H, Savelli B, Dunand C, Paiva JA, Myburg AA, Bouzayen M, Grima-Pettenati J, Cassan-Wang H. 2014. Genome-wide characterization and expression profiling of the AUXIN RESPONSE FACTOR (ARF) gene family in *Eucalyptus grandis*. PLoS One 9(9): e108906. (IF=4.2).**

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

6. Carocha, V., et al. (2015).

**Carocha V, Soler M, Hefer C, Cassan-Wang H, Fevereiro P, Myburg AA, Paiva JA, Grima-Pettenati J. 2015. Genome-wide analysis of the lignin toolbox of *Eucalyptus grandis*. New Phytol 206(4): 1297-1313. (IF=6.6)**

Lignin, a major component of secondary cell walls, hinders the optimal processing of wood for industrial uses. The recent availability of the *Eucalyptus grandis* genome sequence allows comprehensive analysis of the genes encoding the 11 protein families specific to the lignin branch of the phenylpropanoid pathway and identification of those mainly involved in xylem developmental lignification. We performed genome-wide identification of putative members of the lignin gene families, followed by comparative phylogenetic studies focusing on bona fide clades inferred from genes functionally characterized in other species. RNA-seq and microfluid real-time quantitative PCR (RT-qPCR) expression data were used to investigate the developmental and environmental responsive expression patterns of the genes. The phylogenetic analysis revealed that 38 *E.grandis* genes are located in bona fide lignification clades. Four multigene families (shikimate O-hydroxycinnamoyltransferase (HCT), p-coumarate 3-hydroxylase (C3H), caffeate/5-hydroxyferulate O-methyltransferase (COMT) and phenylalanine ammonia-lyase (PAL)) are expanded by tandem gene duplication compared with other plant species. Seventeen of the 38 genes exhibited strong, preferential expression in highly lignified tissues, probably representing the *E.grandis* core lignification toolbox. The identification of major genes involved in lignin biosynthesis in *E. grandis*, the most widely planted hardwood crop world-wide, provides the foundation for the development of biotechnology approaches to develop tree varieties with enhanced processing qualities.

7. Li, Q., et al. (2015).

**Li Q, Yu H, Cao PB, Fawal N, Mathe C, Azar S, Cassan-Wang H, Myburg AA, Grima-Pettenati J, Marque C, Teulieres C, Dunand C. 2015. Explosive Tandem and Segmental Duplications of Multigenic Families in *Eucalyptus grandis*. *Genome Biol Evol* 7(4): 1068-1081. (IF=4.5).**

Plant organisms contain a large number of genes belonging to numerous multigenic families whose evolution size reflects some functional constraints. Sequences from eight multigenic families, involved in biotic and abiotic responses, have been analyzed in *Eucalyptus grandis* and compared with *Arabidopsis thaliana*. Two transcription factor families APETALA 2 (AP2)/ethylene responsive factor and GRAS, two auxin transporter families PIN-FORMED and AUX/LAX, two oxidoreductase families (ascorbate peroxidases [APx] and Class III peroxidases [CIII Prx]), and two families of protective molecules late embryogenesis abundant (LEA) and DNAj were annotated in expert and exhaustive manner. Many recent tandem duplications leading to the emergence of species-specific gene clusters and the explosion of the gene numbers have been observed for the AP2, GRAS, LEA, PIN, and CIII Prx in *E. grandis*, while the APx, the AUX/LAX and DNAj are conserved between species. Although no direct evidence has yet demonstrated the roles of these recent duplicated genes observed in *E. grandis*, this could indicate their putative implications in the morphological and physiological characteristics of *E. grandis*, and be the key factor for the survival of this nondormant species. Global analysis of key families would be a good criterion to evaluate the capabilities of some organisms to adapt to environmental variations.

8. Soler, M., et al. (2015).

**Soler M, Camargo EL, Carocha V, Cassan-Wang H, San Clemente H, Savelli B, Hefer CA, Paiva JA, Myburg AA, Grima-Pettenati J. 2015. The *Eucalyptus grandis* R2R3-MYB transcription factor family: evidence for woody growth-related evolution and function. *New Phytol* 206(4): 1364-1377. (IF=6.6)**

The R2R3-MYB family, one of the largest transcription factor families in higher plants, controls a wide variety of plant-specific processes including, notably, phenylpropanoid metabolism and secondary cell wall formation. We performed a genome-wide analysis of this superfamily in *Eucalyptus*, one of the most planted hardwood trees world-wide. A total of 141 predicted R2R3-MYB sequences identified in the *Eucalyptus grandis* genome sequence were subjected to comparative phylogenetic analyses with *Arabidopsis thaliana*, *Oryza sativa*, *Populus trichocarpa* and *Vitis vinifera*. We analysed features such as gene structure, conserved motifs and genome location. Transcript abundance patterns were assessed by RNAseq and validated by high-throughput quantitative PCR. We found some R2R3-MYB subgroups with expanded membership in *E.grandis*, *V.vinifera* and *P.trichocarpa*, and others preferentially found in woody species, suggesting diversification of specific functions in woody plants. By contrast, subgroups containing key genes regulating lignin biosynthesis and secondary cell wall formation are more conserved across all of the species analysed. In *Eucalyptus*, R2R3-MYB tandem gene duplications seem to disproportionately affect woody-preferential and woody-expanded subgroups. Interestingly, some of the genes belonging to woody-preferential subgroups show higher expression in the cambial region, suggesting a putative role in the regulation of secondary growth.

9. Yu, H., et al. (2015).

**Yu H, Soler M, San Clemente H, Mila I, Paiva JA, Myburg AA, Bouzayen M, Grima-Pettenati J, Cassan-Wang H. 2015. Comprehensive Genome-Wide Analysis of the *Aux/IAA* Gene Family in *Eucalyptus*: Evidence for the Role of *EgrIAA4* in Wood Formation. *Plant Cell Physiol* 56(4): 700-714. (IF=5.0)**

Auxin plays a pivotal role in various plant growth and development processes, including vascular differentiation. The modulation of auxin responsiveness through the auxin perception and



signaling machinery is believed to be a major regulatory mechanism controlling cambium activity and wood formation. To gain more insights into the roles of key *Aux/IAA* gene regulators of the auxin response in these processes, we identified and characterized members of the *Aux/IAA* family in the genome of *Eucalyptus grandis*, a tree of worldwide economic importance. We found that the gene family in *Eucalyptus* is slightly smaller than that in *Populus* and *Arabidopsis*, but all phylogenetic groups are represented. High-throughput expression profiling of different organs and tissues highlighted several *Aux/IAA* genes expressed in vascular cambium and/or developing xylem, some showing differential expression in response to developmental (juvenile vs. mature) and/or to environmental (tension stress) cues. Based on the expression profiles, we selected a promising candidate gene, *EgrIAA4*, for functional characterization. We showed that *EgrIAA4* protein is localized in the nucleus and functions as an auxin-responsive repressor. Overexpressing a stabilized version of *EgrIAA4* in *Arabidopsis* dramatically impeded plant growth and fertility and induced auxin-insensitive phenotypes such as inhibition of primary root elongation, lateral root emergence and agravitropism. Interestingly, the lignified secondary walls of the interfascicular fibers appeared very late, whereas those of the xylary fibers were virtually undetectable, suggesting that *EgrIAA4* may play crucial roles in fiber development and secondary cell wall deposition.

10. Camargo, E., et al. (2019).

**ELO Camargo, R Ployet, H Cassan-Wang, F Mounet, J Grima-Pettenati (2018). Digging in wood: New insights in the regulation of wood formation in tree species *Advances in botanical research Molecular Physiology and Biotechnology of Trees*, 89: 201-233.**

Wood, which represents the most abundant lignocellulosic biomass on earth, fulfils key roles in trees, and is also a raw material for multiple end-uses by mankind. The differentiation of this complex vascular tissue starts with cell division in the vascular cambium and is characterized by a massive deposition of lignified secondary cell walls mainly in fibres and vessels. A transcriptional network underlying this differentiation process ensures a tight regulation of the expression of thousands of genes both at the spatial and temporal levels. Most of our current knowledge of this hierarchical network was extrapolated from studies performed in *Arabidopsis*. Here, we review recent findings on the regulation of wood formation in angiosperm trees species highlighting conserved and distinct mechanisms with *Arabidopsis*. We provide examples shedding light on the central role of auxin and its cross-talk with other hormones at different stages of secondary xylem differentiation. Functional studies of trees' wood-associated transcription factors revealed diversified functions as compared to their *Arabidopsis* orthologs. Sophisticated mechanisms of alternative splicing and cross-regulation between the two distinct groups of top-level NAC-domain master regulators were uncovered. These findings underlie the high level of complexity of wood formation in trees and provide a framework for future lines of research in this exiting research field.

11. Dai, Y., et al (2020)

**Implementing the CRISPR/Cas9 Technology in *Eucalyptus* Hairy Roots Using Wood-Related Genes. *Int J Mol Sci* 21: 10 10.3390/ijms21103408.**

*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.

*12. Hadj Bachir, I., et al (2022)*

**Bachir, I., R. Ployet, C. Teulières, H. Cassan-Wang, F. Mounet and J. Grima-Pettenati (2022). Chapter Nine - Regulation of secondary cell wall lignification by abiotic and biotic constraints. *Advances in Botanical Research*. R. Sibout, Academic Press. 104: 363-392. DOI <https://doi.org/10.1016/bs.abr.2022.03.008>**

The appearance of lignin, 400 million-years ago, has been crucial for the successful land colonization by plants. This complex phenolic polymer which confers hydrophobicity and resistance to degradation to secondary cell walls, is also neosynthesized in response to both biotic and abiotic stress. Lignin biosynthesis during development is regulated by a complex hierarchical transcriptional gene regulatory network. Notably, distinct abiotic stress can co-opt different genes of this network, leading to different lignification patterns. Co-optation of this developmental network is a mean to facilitate adaptation to stress by promoting functional adaptation. Supporting this hypothesis, a growing number of transcription factors known to be involved in the regulation of secondary cell wall, were shown to integrate both developmental and environmental signals leading to an increased tolerance to abiotic stress, such as drought or salinity. On the other hand, some transcription factors known to be involved in stress response were shown to impact lignin biosynthesis as part of the functional adaptation performed by plants to cope with natural constraints. In this review, we focus on the most recent findings on the cross-talk between stress-signaling pathways and regulation of lignin biosynthesis in response to abiotic stress and to a lesser extent to biotic stress. We consider the transcriptional regulation level which has received much attention and the emerging roles of posttranscriptional and posttranslational regulations in the control of lignin biosynthesis in response to stress. We end discussing future outlines and challenges to fill gaps in this emerging and exciting research field.





Annexe 4. Un exemple des Articles suite à mon intégration dans l'équipe d'accueil



# Reference Genes for High-Throughput Quantitative Reverse Transcription–PCR Analysis of Gene Expression in Organs and Tissues of *Eucalyptus* Grown in Various Environmental Conditions

Hua Cassan-Wang<sup>1,\*</sup>, Marçal Soler<sup>1</sup>, Hong Yu<sup>1</sup>, Eduardo Leal O. Camargo<sup>1,2</sup>, Victor Carocha<sup>1,3,4</sup>, Nathalie Ladouce<sup>1</sup>, Bruno Savelli<sup>1</sup>, Jorge A. P. Paiva<sup>3,4</sup>, Jean-Charles Leplé<sup>5</sup> and Jacqueline Grima-Pettenati<sup>1,\*</sup>

<sup>1</sup>LRSV, Laboratoire de Recherche en Sciences Végétales, Université Toulouse III, UPS, CNRS, BP 42617, Auzeville, 31326 Castanet Tolosan, France

<sup>2</sup>Universidade Estadual de Campinas, UNICAMP, Instituto de Biologia, Departamento de Genética, Evolução e Bioagentes, Laboratório de Genômica e Expressão, SP, Brasil

<sup>3</sup>Instituto de Investigação Científica e Tropical (IICT/MNE), Palácio Burnay, Rua da Junqueira, 30, 1349-007 Lisboa, Portugal

<sup>4</sup>Laboratório de Biotecnologia de Células Vegetais, Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa (ITQB-UNL), Av. da República, Quinta do Marquês, 2781-901 Oeiras, Portugal

<sup>5</sup>INRA, UR0588 Amélioration Génétique et Physiologie Forestières (AGPF), F-45075 Orléans, France

\*Corresponding authors: Hua Cassan-Wang, E-mail, wang@lrsv.ups-tlse.fr; Jacqueline Grima-Pettenati, E-mail, grima@lrsv.ups-tlse.fr (Received September 21, 2012; Accepted November 1, 2012)

Interest in the genomics of *Eucalyptus* has skyrocketed thanks to the recent sequencing of the genome of *Eucalyptus grandis* and to a growing number of large-scale transcriptomic studies. Quantitative reverse transcription–PCR (RT–PCR) is the method of choice for gene expression analysis and can now also be used as a high-throughput method. The selection of appropriate internal controls is becoming of utmost importance to ensure accurate expression results in *Eucalyptus*. To this end, we selected 21 candidate reference genes and used high-throughput microfluidic dynamic arrays to assess their expression among a large panel of developmental and environmental conditions with a special focus on wood-forming tissues. We analyzed the expression stability of these genes by using three distinct statistical algorithms (geNorm, NormFinder and  $\Delta$ Ct), and used principal component analysis to compare methods and rankings. We showed that the most stable genes identified depended not only on the panel of biological samples considered but also on the statistical method used. We then developed a comprehensive integration of the rankings generated by the three methods and identified the optimal reference genes for 17 distinct experimental sets covering 13 organs and tissues, as well as various developmental and environmental conditions. The expression patterns of *Eucalyptus* master genes *EgMYB1* and *EgMYB2* experimentally validated our selection. Our findings provide an important resource for the selection of appropriate reference genes for accurate and reliable normalization of gene expression data in the organs and tissues of *Eucalyptus* trees grown in a range of conditions including abiotic stresses.

**Keywords:** Abiotic stress • *Eucalyptus* • Gene expression • Normalization • Reference genes • Xylem.

**Abbreviations:** ACT, actin; AS, all samples; Ct, cycle threshold; EF-1 $\alpha$ , elongation factor-1 $\alpha$ ; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; GM, geometric mean; IDH, NADP isocitrate dehydrogenase; L, leaves; M value, gene expression stability value; OT, organs and tissues; PCA, principal component analysis; PP2A, protein phosphatase 2A subunits; qRT–PCR, quantitative reverse transcription–PCR; TUA,  $\alpha$ -tubulin; TUB,  $\beta$ -tubulin; UBQ, polyubiquitin (UBQ); XA, xylem all; XC, xylem–cambium; XES, xylem environmental stimuli.

## Introduction

Wood is the main source of terrestrial biomass; as well as providing fibers and solid wood products, it is a significant renewable and environmentally cost-effective alternative feedstock to biofuel and a major sink for excess atmospheric CO<sub>2</sub> (Plomion et al. 2001, Boudet et al. 2003). Wood formation, or xylogenesis, involves a transition from meristematic cambium cells to highly differentiated xylem cells, and is one of the most remarkable examples of plant cell differentiation. It involves sequential stages of cell division, cell elongation, formation of lignified secondary cell walls and, finally, programmed cell death to produce tracheary elements. This irreversible developmental process requires complex networks of spatial and temporal regulation events in order to coordinate gene expression networks. Exhaustive sequencing of developing xylem tissues in

*Plant Cell Physiol.* 53(12): 2101–2116 (2012) doi:10.1093/pcp/pcs152, available online at [www.pcp.oxfordjournals.org](http://www.pcp.oxfordjournals.org)

© The Author 2012. Published by Oxford University Press on behalf of Japanese Society of Plant Physiologists.

All rights reserved. For permissions, please email: [journals.permissions@oup.com](mailto:journals.permissions@oup.com)

several tree species and in model systems has increased our understanding of wood formation and highlighted the pivotal role of transcriptional regulation (Demura and Fukuda 2007). In *Eucalyptus*, the most planted hardwood tree in the world, mainly for pulp production, and also among the most appealing lignocellulosic feedstocks for bioenergy production (Myburg *et al.* 2007), large-scale transcriptomic studies including microarrays have revealed thousands of genes expressed in differentiating xylem (Paux *et al.* 2004, Foucart *et al.* 2006, Rengel *et al.* 2009). The roles of most of these genes, however, are still to be discovered. To this end, gene expression profiling of various organs and/or under different conditions is a powerful tool to identify those genes relevant to new biological processes, to provide insights into regulatory networks and to select members of multigene families prior to functional characterization. With the sequence of the genome of *E. grandis* publicly accessible since 2011 in the Phytozome database (<http://www.phytozome.net/>) (Goodstein *et al.* 2011), interest in the genomics of *Eucalyptus* has skyrocketed. There are growing numbers of functional genomics studies and whole-transcriptome sequencing studies linked to digital transcript counting (RNaseq) of various *Eucalyptus* species, organs, tissues, developmental stages and environmental conditions (Paiva *et al.* 2011, Villar *et al.* 2011, Camargo *et al.* 2012) and, thus, a pressing need for accurate techniques to validate and mine these high-throughput expression data.

Real-time quantitative reverse transcription-PCR (qRT-PCR) is considered as the most sensitive and specific technique to assess expression patterns of a moderate number of genes and has become the standard method for validating microarray data. One limitation of qRT-PCR was the relatively low number of genes that can be assessed. This limitation has been overcome with the microfluidic technology allowing high-throughput expression measurements using dynamic arrays (Spurgeon *et al.* 2008). This technology, which allows 9,216 simultaneous real-time PCR gene expression measurements in a single run, has been proven to be as reliable as conventional real-time PCR (Spurgeon *et al.* 2008) and was successfully used in several applications, such as, for instance, in animals and humans for single-cell gene expression analysis (Guo *et al.* 2010, Pang *et al.* 2011), but, to the best of our knowledge, so far no study on gene expression measurement in plants has been reported.

When using either low- or high- throughput qRT-PCR, appropriate normalization is essential to obtain an accurate and reliable quantification of the gene expression level. The purpose of normalization is to correct technical or experimental variations in order to reveal the true biological changes in expression. This is especially relevant when the samples come from different individuals, different tissues and/organs, different time courses, different environmental conditions, etc. The success of the normalization strategy in correcting variability between samples is highly dependent on the choice of the appropriate internal control or reference gene since its expression level should be constant among the tissues or cells, the experimental

treatments and the biological conditions tested. No universal reference gene has been found in any plant, animal or medical system. In the pre-genomic era, genes believed to play housekeeping roles in basic cellular processes such as 18S rRNA, *Glyceraldehyde-3-phosphate dehydrogenase* (GAPDH), *Elongation factor-1 $\alpha$*  (EF-1 $\alpha$ ), *Polyubiquitin* (UBQ), *Actin* (ACT),  $\alpha$ -*Tubulin* and  $\beta$ -*Tubulin* (TUA and TUB, respectively) were frequently used as reference genes. The growing use of the very sensitive qRT-PCR technique, however, has shown that the expression of these genes varies under different experimental or biological conditions, and their systematic use without previous validation can lead to the misinterpretation of results (Czechowski *et al.* 2005, Gutierrez *et al.* 2008a). It has become clear that it is necessary to validate the expression stability of a candidate control gene in each experimental system prior to its use for normalization (Gutierrez *et al.* 2008a, Udvardi *et al.* 2008, Bustin *et al.* 2009, Guenin *et al.* 2009).

In their pioneering work, Scheible and colleagues (Czechowski *et al.* 2005) analyzed a very large set of data from Affymetrix ATH1 (whole-genome GeneChip microarrays of RNA from the model plant *Arabidopsis*) and selected hundreds of candidate reference genes that were expressed at similar levels in a wide range of experimental conditions, outperforming traditional housekeeping genes (GAPDH, ACT2, UBQ10, UBC and EF-1 $\alpha$ ) in terms of expression stability. About 20 of these new-generation reference genes were assessed by qRT-PCR throughout development and in a range of environmental conditions, confirming their superior expression stability and lower absolute expression levels when compared with traditional housekeeping genes. These genes included those encoding members of the polyubiquitin family, proteins with potential regulatory functions such as F-box proteins, protein phosphatase 2A subunits (PP2A) and 'expressed proteins' of unknown function. These new reference genes were successfully employed to search for orthologs in unrelated species such as *Vitis vinifera* (Reid *et al.* 2006) and cotton (Artico *et al.* 2010).

Most of the reference gene evaluation studies of plants have been performed on model species and crop species (Czechowski *et al.* 2005, Exposito-Rodriguez *et al.* 2008, Gutierrez *et al.* 2008b, Guenin *et al.* 2009) and fewer have been carried out on woody plants. Strauss and colleagues (Brunner *et al.* 2004) used single-factor analysis of variance (ANOVA) and linear regression analysis to study the expression of 10 potential reference genes in eight tissues from the poplar tree, and Huang and colleagues (Xu *et al.* 2010) tested three algorithms—geNorm, NormFinder and BestKeeper—to assess the stability of nine housekeeping genes during adventitious rooting of poplar. In *Eucalyptus*, traditional housekeeping genes such as 18S rRNA (Foucart *et al.* 2006, Navarro *et al.* 2011), GAPDH (Navarro *et al.* 2011) and *NADP isocitrate dehydrogenase* (IDH) (Paux *et al.* 2004, Goicoechea *et al.* 2005, Gallo de Carvalho *et al.* 2008, Legay *et al.* 2010) have been used frequently as internal controls for gene expression analysis.

Recent studies, however, have reported the selection of reference genes for normalization during cold acclimation (Fernández et al. 2010) and in vitro adventitious rooting in *E. globulus* (de Almeida et al. 2010) as well as in *Eucalyptus* leaves exposed to biotic (*Puccinia psidii*) and abiotic (acibenzolar-*S*-methyl) stresses (Boava et al. 2010). Very recently, Pasquali and colleagues (de Oliveira et al. 2012) selected housekeeping genes based on the 50 genes exhibiting the least variation in microarrays of *E. grandis* leaves and xylem and of *E. globulus* xylem.

In this study, we report the selection and evaluation of 21 candidate reference genes to identify the most suitable internal control gene or gene combinations for normalization of qRT-PCR data from various species of *Eucalyptus* organs/tissues (xylem, cambium, stem, root and leaves), at different developmental stages, and exposed to environmental stimuli such as drought, cold, bending or nitrogen fertilization. We analyze the entire gene expression data set by three different statistical algorithms to compare the ranking of the potential reference genes. In addition, to illustrate the utility of the new reference genes, we provide a detailed expression analysis of two MYB transcription factors in a set of *Eucalyptus* tissues and organs.

## Results and Discussion

### Selection of candidate reference genes

To identify candidate reference genes in *Eucalyptus* over a broad range of developmental and environmental conditions, we first surveyed the *E. grandis* Phytozome database (<http://www.phytozome.net/eucalyptus.php>) for putative orthologs of the top 27 reference genes in Arabidopsis as defined by Scheible and colleagues (Czechowski et al. 2005). Using the best hits in Phytozome, we selected 16 *E. grandis* genes with high orthology probabilities (E-value < 1.00E-60) (Table 1). Despite their weak scores, we also included genes encoding Ubiquitin10 and F-box protein because they are optimal reference genes in Arabidopsis and in cotton (Czechowski et al. 2005, Artico et al. 2010). Also, we added three genes commonly considered to be housekeeping genes in *Eucalyptus*: *GAPDH*, *IDH* and *TUB*. In total, 21 candidate reference genes were evaluated for their expression levels by making use of microfluidic dynamic array technology, which is appropriate and reliable to perform high-throughput gene expression measurements by qRT-PCR (Spurgeon et al. 2008). The cycle threshold (Ct) values (medians and ranges) for each of the 21 candidate genes in 90 samples (including biological triplicates, as described in detail in the Materials and Methods) are shown in Fig. 1. The transcripts encoded by *GAPDH* and *Actin2* (*ACT2*) were the most abundant, whereas those encoded by *F-box* were the least abundant. The variability of Ct values in the 90 samples examined was widest for the *F-box* transcript, whereas the *ACT2* transcripts had the least variable Ct values (Fig. 1).

### Expression profiling of the candidate reference genes

We performed a preliminary analysis of the expression stability of these 21 genes, and decided to discard five of them (*F-box*, *UBQ10*, *UBQ14*, *TUB* and *YLS8*) because their expression varied substantially in different tissues and conditions (data not shown). To obtain an accurate view of the expression profiles of the 16 remaining candidate genes, we organized the 90 samples into five subsets based on their origin: the 'Organs and Tissues' (OT) subset included 13 organs and/or tissues (38 samples); the 'Leaves' (L) subset comprised young and mature leaves from control and cold-treated plants (nine samples); the 'Xylem All' (XA) subset included all xylem samples from trees grown in various environmental conditions and from several species or hybrids (33 samples); the 'Xylem-Cambium' (XC) subset comprised samples from juvenile and mature xylem and cambium-enriched tissues (12 samples), and the 'Xylem Environmental Stimuli' (XES) subset comprised samples from xylem tissues of trees under drought, mechanical stress and/or with different nitrogen fertilization conditions (27 samples). Supplementary Table S2 reports the relative Ct values for 10 genes (*ACT*, *EF-1 $\alpha$* , *GAPDH*, *IDH*, *PP2A-1*, *PP2A-3*, *PPR2*, *PTB*, *SAND* and *UBC2*) in the five subsets indicated above. The mean Ct values (average of three independent biological samples) were in the range of 5.9–19.0, typical of moderately to highly expressed genes as assayed by using dynamic array technology.

### Expression stability analyses

The most widely used statistical algorithm to analyze the stability of gene expression is geNorm (Vandesompele et al. 2002), which calculates the expression stability value (*M*) of a gene defined as the average pairwise variation of that gene relative to all other potential reference genes in a panel of cDNA samples. Genes with the lowest *M* value have the most stable expression, and useful reference genes should have *M* values < 0.5 (Vandesompele et al. 2002). We used geNorm to determine the *M* values of our 16 candidate genes by stepwise exclusion of the least stable control gene in the five sample subset panels described above and in a sixth panel composed of all the samples (AS; 90 samples; Fig. 2). The subset OT including various organs and tissues showed the greatest variation in expression for all the genes tested (with *M* values of 0.85–0.47), when compared with the subsets including samples from only one organ (e.g. subset L had *M* values of 0.57–0.15) or from a single tissue (e.g. subset XA had *M* values of 0.65–0.28). Surprisingly, in the panel containing all samples (AS), gene expression was more variable (*M* values of 0.77–0.34) than in the OT panel. This may be due to the relatively high proportion of xylem samples (33/90). By applying a cut-off *M* value of 0.5, we found that in the OT panel only two genes, *EF-1 $\alpha$*  and *GAPDH*, qualified as stable, and even those had relatively high *M* values (0.47; Table 2). In contrast, many more candidate genes qualified as stable in the other panels: five in AS, nine

**Table 1** Twenty-one potential reference genes and their primer sequences

Gene abbreviation <sup>a</sup>	Gene model	<i>A. thaliana</i> ortholog	<i>A. thaliana</i> annotation	BlastE-value	Identity (%)	Protein size <sup>b</sup>	Blast alignment <sup>c</sup>	Primer fw_start	Primer sequence (forward/reverse 5'-3')
<i>EF-1α</i>	Eucgr.B02473	AT5G60390	Elongation factor 1-alpha	0	95.9%	449	434/449	1,530	ATCCGTCAGACTGCGCTGTTG/ATGCCGTCAGACTGGGCTGTTG
<i>GAPDH<sup>d</sup></i>	Eucgr.H04673	AT1G13440	Glyceraldehyde-3-phosphate dehydrogenase C2	6.50E-165	92.8%	433	320/338	688	TTGTGGGTGTAAACGAGAGGAG/TTGGTAGTGCAACTGGCGGTTGG
<i>UBC9</i>	Eucgr.D01776	AT4G27960	Ubiquitin conjugating enzyme 9	1.90E-82	96.6%	148	148/178	36	ATCTCGAGGTGCGTTCGTTTC/GGACCCGCCCTTCAATATCAACG
<i>Actin2</i>	Eucgr.I00241	AT3G18780	Actin 2	0	92.8%	377	377/377	1,492	AGTTCTTGCAGCCCATAGTCAGG/AGAAAGCACCACCAATCCCAATCCC
<i>CACS</i>	Eucgr.E00287	AT5G46630	Clathrin adaptor complex subunit, medium subunit family protein	0	94.2%	438	413/441	1,384	TGGACAAAGCCACCACCAATTCAGATG/AGCCGACTCGTAATCCAGATGC
<i>YLS8</i>	Eucgr.G02580	AT5G08290	mRNA splicing factor, thioredoxin-like U5 snRNP	2.80E-82	98.6%	317	142/142	650	TGAAGTCTGGCATCAGTTGCGG/AATCGGGCACCTCAGTTATGTCC
<i>Expressed1</i>	Eucgr.F00240	AT4G33380	Expressed protein	1.70E-101	62.2%	348	315/328	737	TGAGCAGCAGCAGATGCTATTG/TTCCAAGGTCACCTCGTTTGGC
<i>UBC2</i>	Eucgr.E03515	AT5G25760	Peroxin-4, ubiquitin conjugating enzyme 2	2.90E-86	93.0%	157	157/157	379	ACTGCTTTATCAAGGGACCACCTCG/TGCTCAGGTATAGCAAAAGGCAAGC
<i>PP2A-3</i>	Eucgr.B03031	AT1G13320	Protein phosphatase 2A subunit A3	0	90.3%	587	587/587	1,720	CAGCGCAAAACAACCTTGAAGCG/ATTATGTGCTGCTTGGCCCAAGTC
<i>SAND</i>	Eucgr.B02502	AT2G28390	SAND family, trafficking protein Mon1	0	66.7%	612	580/607	1,863	TTGATCCACTTGGCGACAAGGC/TCACCCATTGACATACACGATTCG
<i>Expressed3</i>	Eucgr.J01745	AT4G26410	Expressed protein	1.00E-74	59.3%	242	214/263	647	TCGCAAGCAGGAGGTCTAATGG/ACCTCAGCCCGAAGTTTTCAGTG
<i>PTB</i>	Eucgr.J01358	AT3G01150	Polypyrimidine tract-binding protein 1	0	82.4%	459	364/399	1,277	TGTGGAAAGAGAGGTTTGGAAAGG/AGACAAAACCTGAGCCACTGAAGC
<i>Helicase</i>	Eucgr.G00345	AT1G58050	RNA helicase family protein	0	65.1%	852	854/1417	2,775	AGGAGCAAAACCCCTGTCAACTG/ACATGCCCAATTGTGAACAGGAG
<i>UPL7</i>	Eucgr.A01586	AT3G53090	Ubiquitin-protein ligase 7	0	59.7%	1161	1142/1142	3,450	TTGGAGTCAAGGATGTCGAGAGAC/TGCACCGCTTGTATGTGGGAAGC
<i>F-box</i>	Eucgr.E04314	AT5G15710	Galactose oxidase/kelch repeat superfamily protein	3.20E-27	27.8%	445	371/448	688	TTGATGCTGGGGGTTCTTCTC/TCCTGATTCGAGGCTGCAAAAGC

(continued)



Table 1 Continued

Gene abbreviation <sup>a</sup>	Gene model	<i>A. thaliana</i> ortholog	<i>A. thaliana</i> annotation	BlastE-value	Identity (%)	Protein size <sup>b</sup>	Blast alignment <sup>c</sup>	Primer fw_start	Primer sequence (forward/reverse 5'-3')
PPR2	Eucgr.A00809	AT1G62930	Pentatricopeptide repeat (PPR) superfamily protein	2.40E-121	39.1%	589	580/629	1,775	CCATTGTGTGAGATCTGCTCCTC/CTGCTGTTAATCTGGCGTGCTC
PP2A-1	Eucgr.B03386	AT1G59830	Protein phosphatase 2A-2	1.50E-175	98.3%	306	305/306	720	TCGAGCTTTGGACCCCATACAAG/ACCACAAGAGGTACACACATTGGC
UBQ10	Eucgr.H03021	AT4G05320	Polyubiquitin 10	5.60E-46	37.8%	321	296/457	612	TCGGCAGAGGTGATGAGCTTAC/AGGTCTTGGCATCTCTCAAGTTGG
UBQ14	Eucgr.F04448	AT4G02890	Polyubiquitin family gene	1.90E-65	80.9%	154	152/305	436	ATTCACCTGTCCAGCAAAAGGC/TCTCAGCGCAAGCAAAAGATG
Tubulin <sup>d</sup>	Eucgr.K00264	AT5G23860	Tubulin beta 8	0	99.8%	447	430/449	978	CACTCAGCAAATGTGGGATGCG/TAACGGCCATGACGTGGATCTG
IDH <sup>d</sup>	Eucgr.F02901	AT1G65930	NADP-isocitrate dehydrogenase	0	87.6%	416	410/410	1,356	TGCTGTGGCAGCTGAACCTCAAG/ATGTTGTCCCCAGTCACCTAC

<sup>a</sup> Genes were named according to the most similar gene from Arabidopsis.

<sup>b</sup> *Eucalyptus* protein size in amino acids.

<sup>c</sup> Size of the Blastp alignment in amino acids (*Eucalyptus* vs. Arabidopsis).

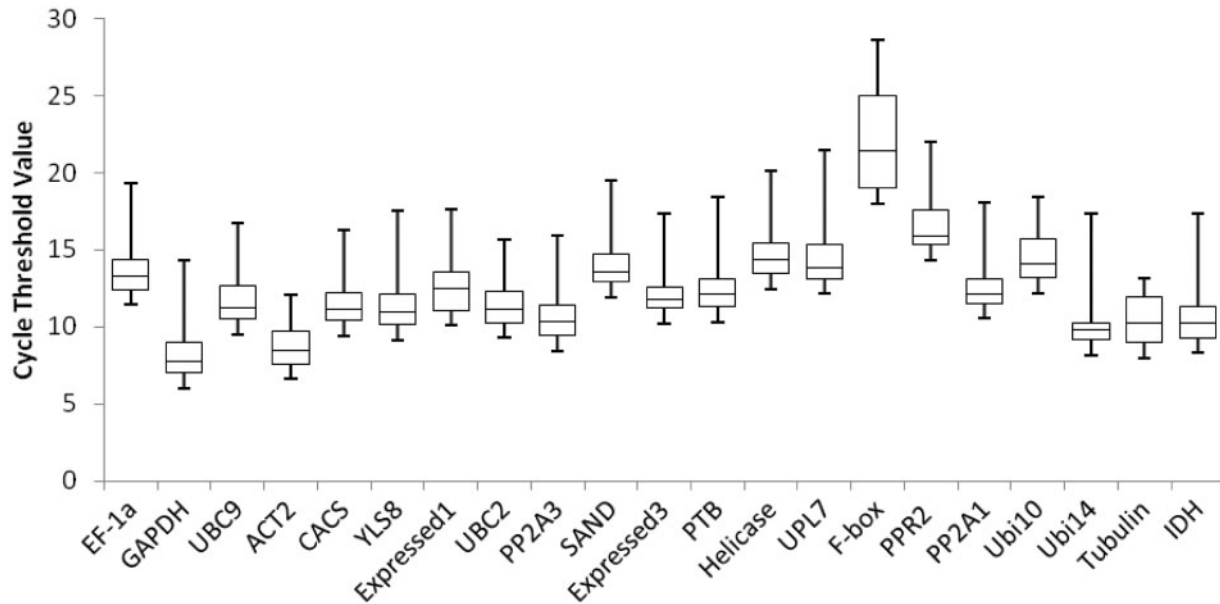
<sup>d</sup> *Eucalyptus* genes used as reference genes in previous studies.

in XA, 10 in XES, 13 in L and 14 in XC. The identity of the most stably expressed genes in a panel varied depending on the panel considered (Table 2). For instance, *EF-1α* was one of the most stably expressed genes in the OT and XC panels, whereas *PP2A-1* and *IDH* were the most stable in the L panel; *PP2A-3* and *SAND* were the most stably expressed genes in the XA, XES and, remarkably, the AS panel, in which all the samples were grouped together. In all panels, *PPR2* was among the least stably expressed genes (ranked consistently 14–16). *Expressed1* ranked poorly in most panels except in L, where it ranked 3 with an M value of 0.2. The commonly used reference gene *GAPDH* ranked well in the OT panel (ranked top but with a relatively high M value of 0.47) and in XC (ranked 5, M value of 0.2); it was also stably expressed in L (ranked 10, M value of 0.35) and in XA (ranked 10, M value of 0.48). *IDH* ranked top in L (M value of 0.15) but less well in OT (ranked 4, M value of 0.54) and even worse in XA (rank 12, M value of 0.52).

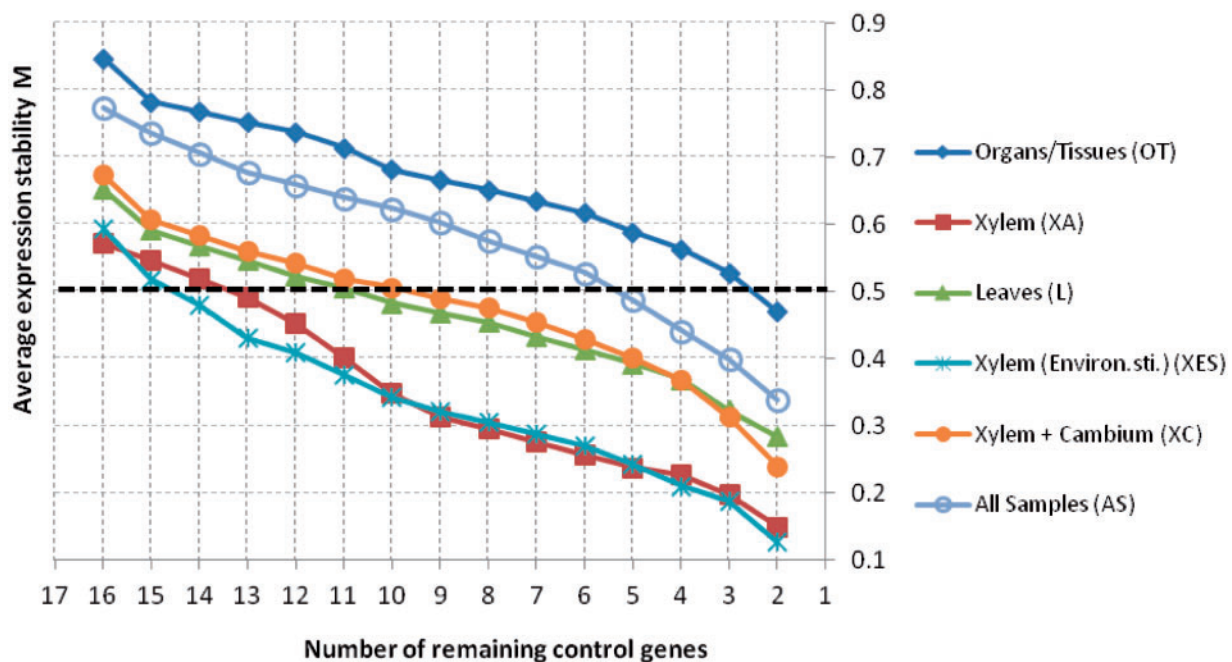
The geNorm program also determines the ideal and/or minimal number of reference genes that would be required to calculate an accurate normalization factor, as the geometric mean (GM) of their relative quantities. A pairwise variation  $V_n/V_{n+1}$  of 0.15 provides a cut-off value below which the inclusion of an additional control gene is not necessary for reliable normalization. This pairwise analysis revealed that the ideal number of reference genes varied among the different data subsets. Use of the two best reference genes was sufficient ( $V_2/V_3 < 0.15$ ) when the subsets were composed of only one type of organ or tissue, and also for the panel containing all samples (AS; Fig. 3). When the panel was composed of less homogeneous samples, such as the OT set, at least three reference genes ( $V_2/V_3 = 0.17$ ,  $V_3/V_4 = 0.13$ ) were necessary for accurate normalization (Fig. 3).

We also analyzed our data with NormFinder, a program taking into account the intra- and intergroup variations for normalization factor calculation and whose results are not affected by occasional co-regulated genes (Andersen et al. 2004; Table 3). Remarkably, in this analysis *PP2A-3* was the first or the second best-ranked gene in each panel, although over all 16 genes there were some differences in the rankings in the different panels. Similarly, *PPR2* was consistently the worst ranked gene in all panels. NormFinder also defines the best combination of two reference genes for each experimental set, which provides a corresponding lower stability value M (Table 3). A comparison of the outcomes using geNorm or NormFinder revealed common features but also significant differences (Table 3 vs. Table 2). The four most stably expressed genes in the AS panel as ranked by geNorm were *PP2A-3*, *SAND*, *UPL7* and *PTB*, whereas when the same data set was ranked by NormFinder the four most stably expressed genes were *PP2A-3*, *CACS*, *GAPDH* and *SAND*. Consistently, *PP2A-3* was ranked top and *SAND* was found among the top four reference genes by both methods. The other genes ranked in the top four by one method (i.e. *UPL7* and *PTB* for geNorm and *CACS* and *GAPDH* for NormFinder) were not well ranked by the other method.





**Fig. 1** Range of cycle threshold (Ct) values of 21 candidate reference genes obtained for 90 *Eucalyptus* samples. Comparisons of the variability of Ct values of the 21 candidate reference genes among all 90 samples are shown as medians (line in box), 25th percentile to the 75th percentile (boxes) and ranges (whiskers).



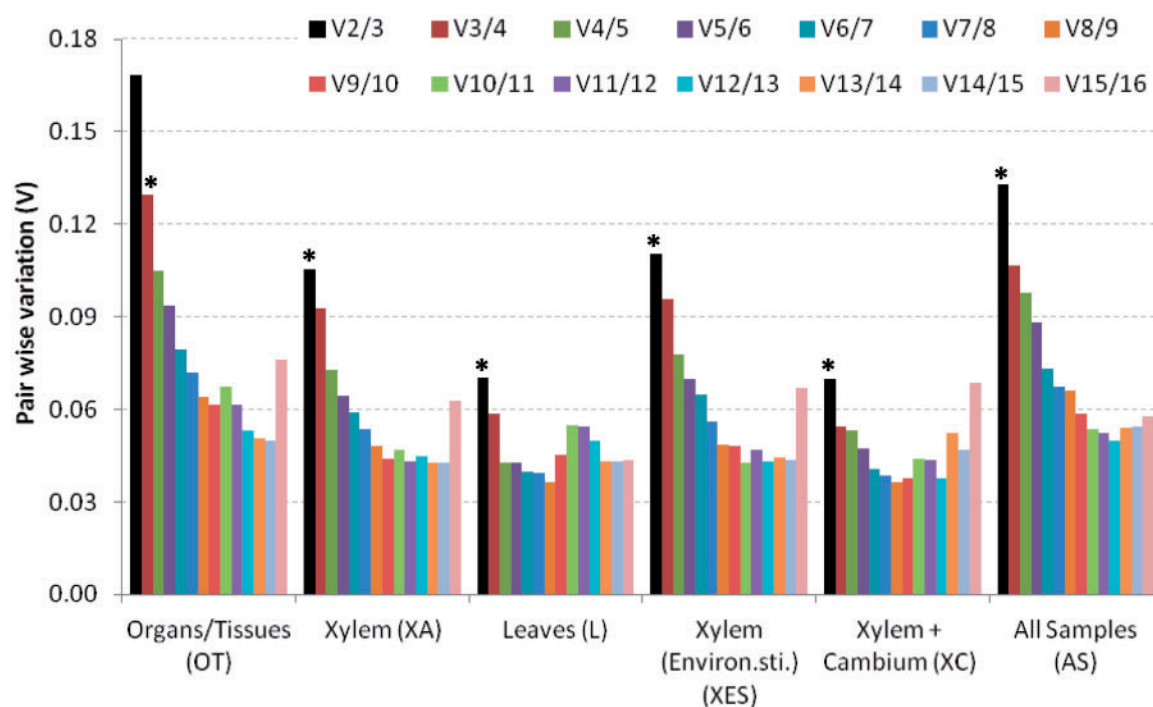
**Fig. 2** Average expression stability values (M) evaluated by geNorm of the remaining control genes during stepwise exclusion of the least stable control gene in the different tissue panels. See also **Table 2** for the ranking of the genes according to their expression stability.

As there were significant differences in the results obtained with geNorm and NormFinder, we employed a third analytical method called  $\Delta Ct$  (Silver et al. 2006). This method compares the relative expression of ‘pairs of genes’ within each sample to identify a reference gene confidently. If the  $\Delta Ct$  value between two genes remains constant when analyzed in different

samples, it means that both genes either are stably expressed or are co-regulated. On this basis, the stability of expression of the candidate genes was ranked according to the reproducibility of the gene expression difference among all tested samples (Silver et al. 2006). Using this method, again, the ranking results differed among the different data panels, but PP2A-3 and SAND

**Table 2** *Eucalyptus* candidate reference genes ranked according to their expression stability evaluated using geNorm algorithm

Rank	Organs/tissues (OT)		Leaves (L)		Xylem (XA) (all samples)		Xylem and cambium (XC) (development)		Xylem (XES) (environmental stimuli)		All samples (AS)	
	Gene	M	Gene	M	Gene	M	Gene	M	Gene	M	Gene	M
1/1	<i>EF-1α/GAPDH</i>	0.47	<i>PP2A-1/IDH</i>	0.15	<i>PP2A-3/SAND</i>	0.28	<i>EF-1α/UBC9</i>	0.13	<i>PP2A-3/SAND</i>	0.24	<i>PP2A-3/SAND</i>	0.34
3	<i>PP2A-1</i>	0.53	<i>Expressed1</i>	0.20	<i>UPL7</i>	0.32	<i>PP2A-1</i>	0.19	<i>UPL7</i>	0.31	<i>UPL7</i>	0.40
4	<i>UBC9</i>	0.56	<i>CACS</i>	0.23	<i>UBC2</i>	0.37	<i>CACS</i>	0.21	<i>UBC2</i>	0.37	<i>PTB</i>	0.44
5	<i>IDH</i>	0.59	<i>Expressed3</i>	0.24	<i>PTB</i>	0.39	<i>GAPDH</i>	0.24	<i>PTB</i>	0.40	<i>UBC2</i>	0.49
6	<i>CACS</i>	0.62	<i>UBC2</i>	0.26	<i>EF-1α</i>	0.41	<i>PTB</i>	0.27	<i>EF-1α</i>	0.43	<i>PP2A-1</i>	0.53
7	<i>Expressed1</i>	0.63	<i>PP2A-3</i>	0.28	<i>UBC9</i>	0.43	<i>PP2A-3</i>	0.29	<i>UBC9</i>	0.45	<i>EF-1α</i>	0.55
8	<i>PP2A-3</i>	0.65	<i>EF-1α</i>	0.30	<i>Expressed3</i>	0.45	<i>UBC2</i>	0.31	<i>Expressed3</i>	0.48	<i>GAPDH</i>	0.58
9	<i>Expressed3</i>	0.67	<i>UBC9</i>	0.31	<i>PP2A-1</i>	0.47	<i>Actin-2</i>	0.32	<i>Helicase</i>	0.49	<i>CACS</i>	0.60
10	<i>Actin-2</i>	0.68	<i>GAPDH</i>	0.35	<i>GAPDH</i>	0.48	<i>Helicase</i>	0.34	<i>PP2A-1</i>	0.51	<i>IDH</i>	0.62
11	<i>UBC2</i>	0.71	<i>Actin-2</i>	0.40	<i>CACS</i>	0.50	<i>IDH</i>	0.38	<i>GAPDH</i>	0.52	<i>UBC9</i>	0.64
12	<i>SAND</i>	0.74	<i>PTB</i>	0.45	<i>IDH</i>	0.52	<i>UPL7</i>	0.41	<i>CACS</i>	0.54	<i>Expressed3</i>	0.66
13	<i>PTB</i>	0.75	<i>UPL7</i>	0.49	<i>Actin-2</i>	0.55	<i>SAND</i>	0.43	<i>IDH</i>	0.56	<i>Helicase</i>	0.68
14	<i>Helicase</i>	0.77	<i>PPR2</i>	0.52	<i>Helicase</i>	0.57	<i>PPR2</i>	0.48	<i>Actin-2</i>	0.58	<i>Actin-2</i>	0.71
15	<i>UPL7</i>	0.78	<i>SAND</i>	0.55	<i>Expressed1</i>	0.59	<i>Expressed1</i>	0.52	<i>Expressed1</i>	0.61	<i>Expressed1</i>	0.74
16	<i>PPR2</i>	0.85	<i>Helicase</i>	0.57	<i>PPR2</i>	0.65	<i>Expressed3</i>	0.59	<i>PPR2</i>	0.67	<i>PPR2</i>	0.77

**Fig. 3** Determination the optimal number of reference genes for normalization according to geNorm. Pairwise variation ( $V_n/V_{n+1}$ ) was analyzed between the normalization factors  $NF_n$  and  $NF_{n+1}$  by the geNorm software. The cut-off value of 0.15, below which the inclusion of an additional reference gene is not required, is indicated by a discontinuous line. \*Optimal number of reference genes for normalization.

were consistently well ranked in most, whereas *PPR2*, *Actin2* and *Expressed1* were the least stably expressed genes (Table 4). Generally, these results were closer to those obtained with NormFinder than with geNorm. This was unexpected because

the  $\Delta C_t$  method is similar to geNorm (Vandesompele et al. 2002), in that it compares pairs of genes (Silver et al. 2006).

Because these three independent methods produced similar results in some cases but discrepancies in others, we decided to

**Table 3** *Eucalyptus* reference genes ranked according to their expression stability evaluated by the NormFinder method

Rank	Organs/tissues (OT)		Leaves (L)		Xylem (XA) (all samples)		Xylem and cambium (XC) (development)		Xylem (XES) (environmental stimuli)		All samples (AS)	
	Gene	SV <sup>a</sup>	Gene	SV	Gene	SV	Gene	SV	Gene	SV	Gene	SV
1	PP2A-3	0.15	PP2A-3	0.04	SAND	0.08	PP2A-3	0.08	SAND	0.09	PP2A-3	0.08
2	PP2A-1	0.23	CACS	0.11	PP2A-3	0.13	GAPDH	0.10	PP2A-3	0.12	CACS	0.11
3	PTB	0.30	EF-1 $\alpha$	0.14	EF-1 $\alpha$	0.13	UBC2	0.13	EF-1 $\alpha$	0.14	GAPDH	0.12
4	SAND	0.30	IDH	0.15	UBC9	0.16	PTB	0.15	UBC9	0.18	SAND	0.13
5	Expressed1	0.30	PP2A-1	0.15	UPL7	0.17	CACS	0.17	UPL7	0.19	UBC9	0.14
6	EF-1 $\alpha$	0.31	UBC2	0.16	Expressed3	0.17	Actin2	0.17	Expressed3	0.19	PP2A-1	0.14
7	IDH	0.31	Expressed3	0.20	UBC2	0.18	PP2A-1	0.20	UBC2	0.21	PTB	0.14
8	CACS	0.32	Expressed1	0.20	GAPDH	0.19	EF-1 $\alpha$	0.21	GAPDH	0.22	IDH	0.14
9	UBC2	0.33	SAND	0.27	PTB	0.20	UBC9	0.23	PTB	0.23	UBC2	0.14
10	Expressed3	0.33	PTB	0.27	PP2A-1	0.26	IDH	0.24	Helicase	0.25	EF-1 $\alpha$	0.14
11	GAPDH	0.36	Actin2	0.32	CACS	0.27	Helicase	0.32	Actin2	0.28	UPL7	0.14
12	Helicase	0.39	Helicase	0.34	Actin2	0.28	UPL7	0.33	PP2A-1	0.28	Helicase	0.15
13	Actin2	0.41	UBC9	0.34	IDH	0.30	SAND	0.33	CACS	0.29	Expressed1	0.15
14	UPL7	0.42	GAPDH	0.35	Helicase	0.30	PPR2	0.52	IDH	0.32	Actin2	0.18
15	UBC9	0.44	UPL7	0.37	Expressed1	0.30	Expressed1	0.54	Expressed1	0.32	Expressed3	0.19
16	PPR2	0.63	PPR2	0.38	PPR2	0.39	Expressed3	0.74	PPR2	0.42	PPR2	0.22
Best combination <sup>a</sup> and PP2A-1	PP2A-3 and CACS	0.09	PP2A-3 and CACS	0.06	PP2A-3 and Expressed3	0.06	PP2A-3 and GAPDH	0.05	PP2A-3 and Expressed3	0.07	SAND and PP2A-3	0.05

<sup>a</sup> SV, stability value.

<sup>b</sup> The best combination of two reference genes in each experimental set.

**Table 4** *Eucalyptus* reference genes ranked according to their expression stability evaluated by the  $\Delta$ Ct method

Rank	Organs/tissues (OT)		Leaves (L)		Xylem (XA) (all samples)		Xylem and cambium (XC) (development)		Xylem (XES) (environmental stimuli)		All samples (AS)	
	Ranking	SD	Ranking	SD	Ranking	SD	Ranking	SD	Ranking	SD	Ranking	SD
1	PP2A-3	0.67	PP2A-3	0.43	PP2A-3	0.52	PP2A-3	0.44	PP2A-3	0.53	PP2A-3	0.62
2	PP2A-1	0.72	EF-1 $\alpha$	0.47	SAND	0.52	EF-1 $\alpha$	0.44	SAND	0.54	SAND	0.66
3	SAND	0.73	Expressed1	0.47	UBC9	0.54	GAPDH	0.45	UBC9	0.57	PP2A-1	0.68
4	EF-1 $\alpha$	0.75	CACS	0.49	Expressed3	0.57	Actin2	0.45	Expressed3	0.6	EF-1 $\alpha$	0.71
5	GAPDH	0.78	UBC2	0.5	PP2A-1	0.57	UBC2	0.46	EF-1 $\alpha$	0.6	GAPDH	0.71
6	PTB	0.79	PP2A-1	0.5	UBC2	0.59	PTB	0.46	PP2A-1	0.61	PTB	0.73
7	UBC2	0.81	IDH	0.5	EF-1 $\alpha$	0.59	CACS	0.47	GAPDH	0.61	CACS	0.73
8	IDH	0.82	Expressed3	0.51	GAPDH	0.6	SAND	0.5	UBC2	0.62	UBC2	0.73
9	CACS	0.82	SAND	0.59	UPL7	0.6	PP2A-1	0.5	UPL7	0.62	UBC9	0.76
10	Expressed1	0.83	UBC9	0.59	PTB	0.6	UBC9	0.51	PTB	0.63	IDH	0.8
11	Expressed3	0.84	GAPDH	0.6	CACS	0.65	UPL7	0.57	Helicase	0.68	Helicase	0.82
12	UPL7	0.86	PTB	0.61	IDH	0.67	IDH	0.58	CACS	0.68	UPL7	0.83
13	Helicase	0.87	UPL7	0.64	Actin2	0.71	Helicase	0.59	IDH	0.71	Expressed3	0.83
14	UBC9	0.88	Actin2	0.65	Helicase	0.74	Expressed1	0.79	Actin2	0.74	Expressed1	0.93
15	Actin2	0.89	PPR2	0.69	Expressed1	0.76	PPR2	0.85	Expressed1	0.78	Actin2	0.97
16	PPR2	1.32	Helicase	0.76	PPR2	1.08	Expressed3	1.11	PPR2	1.15	PPR2	1.06

SD, average of standard deviation.

**Table 5** Comprehensive ranking of *Eucalyptus* reference genes according to the GM method<sup>a</sup>

Rank	Organs/tissues (OT)		Leaves (L)		Xylem (XA) (all samples)		Xylem and cambium (XC) (development)		Xylem (XES) (environmental stimuli)		All samples (AS)	
	Ranking	GM <sup>b</sup>	Ranking	GM	Ranking	GM	Ranking	GM	Ranking	GM	Ranking	GM
1	PP2A-3	4.01	PP2A-3	1.42	SAND	2.04	EF-1 $\alpha$	1.52	SAND	2.05	PP2A-3	2.53
2	PP2A-1	4.33	IDH	1.74	PP2A-3	2.77	PP2A-3	1.71	PP2A-3	2.14	SAND	3.00
3	EF-1 $\alpha$	4.63	PP2A-1	1.74	UPL7	2.94	GAPDH	1.74	UPL7	3.04	PTB	3.57
4	GAPDH	4.95	CACS	1.90	UBC9	2.96	UBC9	1.84	EF-1 $\alpha$	3.10	PP2A-1	3.64
5	IDH	5.33	Expressed1	2.02	UBC2	3.11	CACS	1.98	UBC9	3.25	GAPDH	3.70
6	SAND	5.34	EF-1 $\alpha$	2.16	EF-1 $\alpha$	3.20	PTB	2.06	UBC2	3.35	UBC2	3.72
7	Expressed1	5.46	UBC2	2.26	PTB	3.29	UBC2	2.09	Expressed3	3.50	UPL7	3.75
8	CACS	5.46	Expressed3	2.36	GAPDH	3.46	PP2A-1	2.10	PTB	3.58	CACS	3.81
9	PTB	5.63	UBC9	3.49	Expressed3	3.56	Actin2	2.25	GAPDH	3.81	EF-1 $\alpha$	3.84
10	Expressed3	5.75	GAPDH	3.70	Actin2	3.60	IDH	3.32	PP2A-1	4.06	UBC9	4.16
11	UBC2	5.76	PTB	3.86	PP2A-1	3.81	SAND	3.39	Helicase	4.19	IDH	4.28
12	UBC9	6.04	Actin2	4.03	CACS	4.29	Helicase	3.54	CACS	4.52	Helicase	4.58
13	Actin2	6.37	SAND	4.03	Expressed1	4.44	UPL7	3.72	IDH	4.82	Expressed3	4.83
14	Helicase	6.46	UPL7	4.50	IDH	4.59	Expressed1	5.87	Actin2	4.84	Expressed1	5.05
15	UPL7	6.60	PPR2	4.87	Helicase	4.62	PPR2	5.88	Expressed1	5.27	Actin2	5.33
16	PPR2	9.49	Helicase	5.16	PPR2	6.79	Expressed3	8.13	PPR2	7.28	PPR2	6.04

<sup>a</sup> Calculated by the integration of geNorm, NormFinder and  $\Delta$ Ct methods.<sup>b</sup> Geometric mean of ranking values.

make a comprehensive ranking by integrating the ranking obtained with the three methods. A simple algorithm was designed based on the GM of each gene weight across the three methods (as detailed in the Materials and Methods). The gene

with the lowest GM was considered to be the most stable gene. **Table 5** displays the results obtained for the six data sets, and **Supplementary Table S3** contains the results for 11 additional data sets including stems, roots and cold stress.

### Comparison of the four statistical methods by principal component analysis

To compare the four methods visually, we used principal component analysis (PCA), which takes data points in a high dimensional space and defines new axes (components) that cut across that space such that the first component captures as much of the variance in the data as possible, the second component (orthogonal to the first) captures as much of the remaining variance as possible, and so on. In this study, the data points are the stability values of the 16 genes in a 4D space, with the coordinate in each dimension being the stability value of each gene calculated by each of the four methods. Importantly, because the components cut across this 4D space, each component has contributions from all 16 genes, which allows the original variables (here the four methods) to be represented in new axes defined by the new components. This leads to a graphical representation of the correlations between variables (the four algorithms used to analyze gene expression stability), and also of the similarity between data points (the gene expression stability values). Applied to the expression stability values of the AS, OT, XA and L subsets, the first principal component (PC1) explained 88.0, 82.1, 91.0 and 90.7% of the observed variance, respectively, whereas the second principal component (PC2) explained only 8.4, 15.9, 5.6 and 7.1%, respectively. Thus, in all cases, >95% of the total information was explained by the first plane, defined by the first two components. The projection of the four 'gene stability' variables onto this plane was performed for each of four sample panels (Fig. 4A). In each case, the four methods were grouped together on the extreme right of the PC1 axis (Fig. 4A), which explains the major part of the variance. The distance between them along the PC2 axis was variable depending on the panel considered. In the most heterogeneous panel, OT, geNorm was somewhat distant (on both PC1 and PC2 axes) from NormFinder and  $\Delta$ Ct, which were close to each other (Fig. 4A). For the XA panel and, to a lesser extent, for the L panel, which are the most homogeneous subsets of samples, all four methods grouped spatially (Fig. 4A). As expected, the position of the values obtained by the comprehensive ranking method reflected the fact that they were the GMs of the three other methods.

Because PCA components consist of contributions of all 16 genes, we plotted the score of the stability of the expression of the 16 genes for the two PCA components and for each panel of samples (Fig. 4B). Nearly all the genes projected on the left part of the central vertical axis were within the top six genes as ranked by at least one of the four methods. PP2A3 was always well ranked if not the top-ranked gene by the four methods, whatever the subset considered. Thus, it appears to be the best reference gene in these samples. In contrast, PPR2 was located in all cases on the extreme right of the plots; it is clearly unsuitable as a reference gene. In the AS subset, PP2A-3 (which ranked top by the four statistical methods) appeared as a singleton on the extreme left of the *x*-axis, ahead of the second

most stable gene, SAND (which was also well ranked by the four methods). To the right of these two genes on the *x*-axis were several genes with a very similar score for the first PCA component but with a diverse score for the second component on the *y*-axis. Among them, only PP2A1, located on the horizontal axis, was well ranked by all methods. The main difference between the remaining genes was the method used for ranking. For instance, CACS and UPL7, which were located at the two extremes of the *y*-axis, were well ranked by only one method (NormFinder and geNorm, respectively).

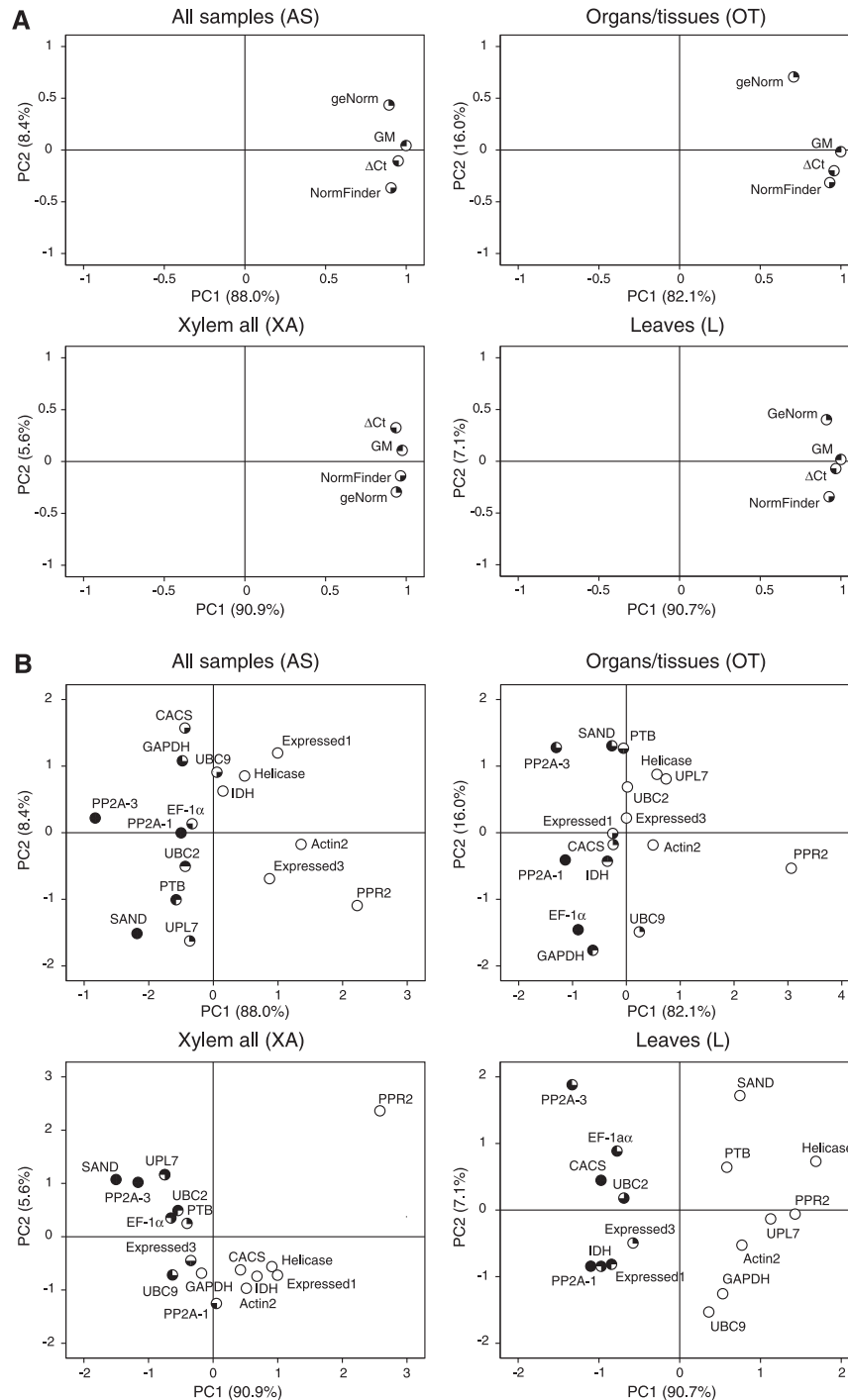
In the very heterogeneous OT subset, PP2A1, EF-1 $\alpha$  and PP2A3 were the three best genes. PP2A1, EF-1 $\alpha$  were well ranked by all methods, whereas PP2A3 was ranked first by NormFinder and  $\Delta$ Ct and only eighth by geNorm (Fig. 4B). The spatial distribution of the values for the genes in the L subset was very different from that of the other three panels, with a clear distinction between the group on the left containing the best reference genes vs. the group on the right, which contained the worst ones (Fig. 4B). There were equal numbers of genes in the two groups and they were distributed evenly along both the *x*- and *y*-axes, in contrast to those in the XA subset, which were distributed more closely on both axes (Fig. 4B). This is consistent with the fact that the three statistical methods were closely grouped for the XA subset (Fig. 4A), and the relative density along the *y*-axis (second PCA component) reflects the fact that there are fewer differences between the stability rankings of the L subset genes by the three methods when compared with the other subsets.

One conclusion from this PCA is that the ranking result depends not only on the nature of the samples and the degree of heterogeneity of the gene panel considered, but also on the analysis method applied to a particular subset of samples: the different rankings obtained by using different methods depend on which sample panel is being considered (Fig. 4A). Interestingly, there is a very good agreement between the ranking of the three best genes obtained by the comprehensive ranking method (reflecting the geometric mean of the ranking of the three other methods; Table 5) and those identified by PCA analysis (Fig. 4B). We therefore used this comprehensive ranking method for 11 further subsets of samples from other tissues, developmental stages and environmental conditions (Supplementary Table S3). The expression of PP2A-3 was consistently stable in all 11 data sets. In roots and stems, UBC2 and CACS were stably expressed, whereas PPR2, Actin2 and Helicase were, as already noticed in the other sample panels, the least stably expressed genes. Taking together the two comprehensive ranking tables (Table 5 and Supplementary Table S3), PP2A-3, PP2A-1, EF-1 $\alpha$ , SAND, IDH, CACS, UBC2 and PTB exhibited relatively stable expression, whereas PPR2, Actin2, Expressed1 and Helicase were the least stable.

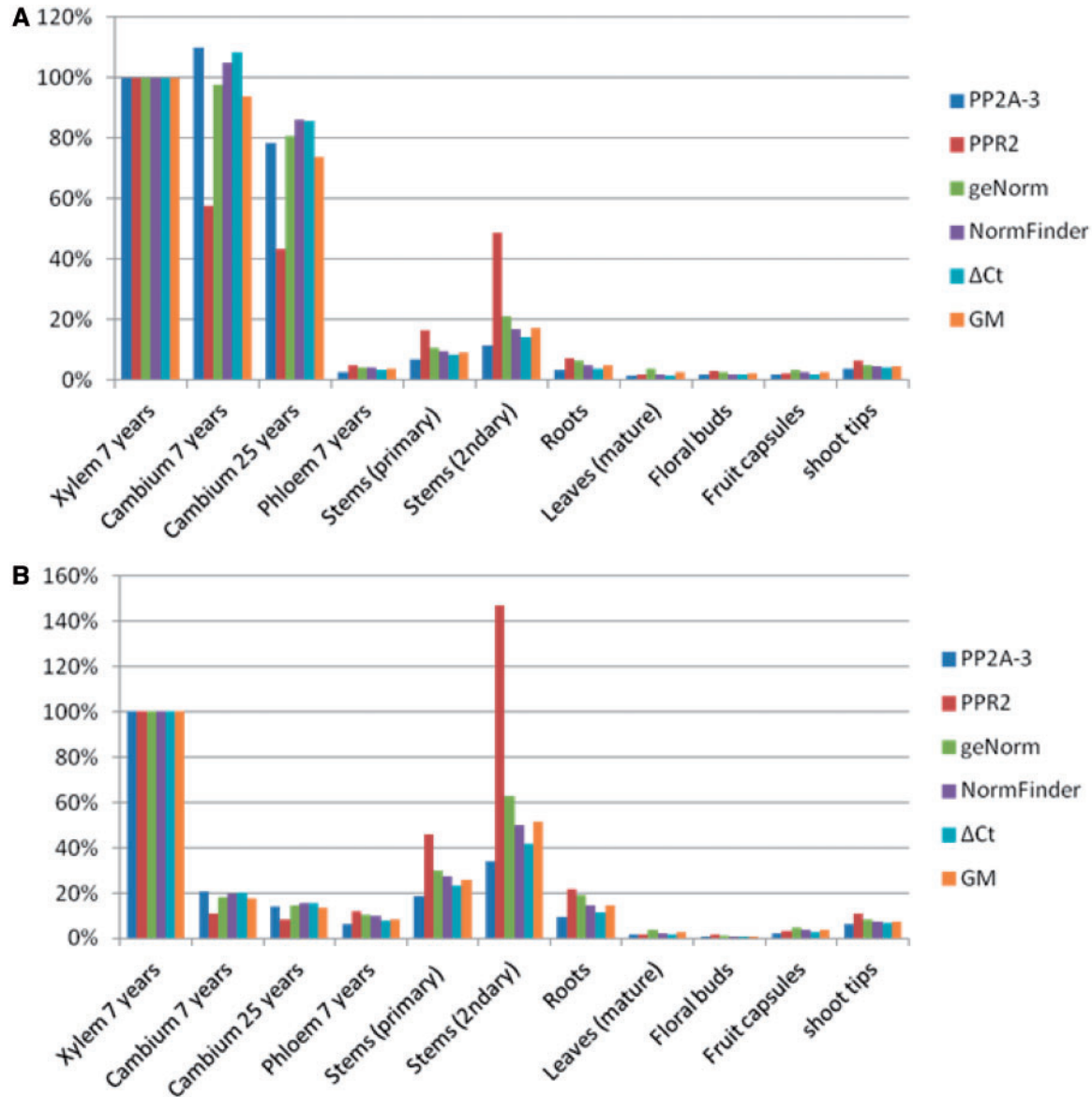
### Expression profiling of EgMYB1 and EgMYB2

To evaluate the effect of the choice of reference gene on the expression profiling of other genes, we analyzed the expression





**Fig. 4** Principal component analysis of four different methods to measure gene stability values. (A) Factor loading of the four methods (geNorm, NormFinder,  $\Delta$ Ct and GM) used to measure gene stability for the two main PCA axes (PC1 and PC2) calculated for four different samples panels (AS, OT, XA and L). Each method is represented by the position of the black quarter in the disc: right upper quarter, geNorm; right lower quarter, NormFinder; left upper quarter, GM (comprehensive method); and left lower quarter,  $\Delta$ Ct. (B) PCA scores of the gene stability values of the 16 reference genes analyzed for the two main PCA axes extracted (PC1 and PC2). Each disc corresponds to a candidate reference gene. Following the same rules as in (A), the position of the black quarter indicates the methods. A black right upper quarter in the disc means that the gene is ranked amongst the six best-classed genes by the geNorm method. Dots with one, two or three black quarters indicate that the gene is ranked amongst the best-classed six genes by one, two or three methods. A black disc (four black quarters) represents genes ranked amongst the first six best-classed genes by all four methods. All the genes well ranked by at least one method are generally found on the left of the vertical axis.



**Fig. 5** Expression profiles of *EgMYB1* and *EgMYB2* in different organs and tissues. Expression ratios of *EgMYB1* and *EgMYB2* for the experimental set of different organs/tissues calculated using (i) the best-ranked reference gene *PP2A-3*; (ii) the worse-ranked reference gene *PPR2*; (iii) combinations of genes recommended by each algorithm, i.e. *PP2A-1*, *EF-1 $\alpha$*  and *GAPDH* for *geNorm*; *PP2A-3* and *PP2A-1* for *NormFinder*; *PP2A-3*, *PP2A-1* and *SAND* for  $\Delta$ Ct; and *PP2A-3*, *PP2A-1* and *EF-1 $\alpha$*  for the comprehensive ranking method (*GM*). Results are expressed as a percentage of the value in xylem which was arbitrarily set to 100%. The mature leaves sample was used as the calibrator.

of two genes, *EgMYB1* and *EgMYB2* (which are master regulators of xylem formation in *Eucalyptus* (Goicoechea *et al.* 2005, Legay *et al.* 2010), in several tissues and organs by using the various combinations of top reference genes indicated by each of the four methods: *EF-1 $\alpha$* , *GAPDH* and *PP2A-1* as indicated by *geNorm*; *PP2A-3* and *PP2A-1* as indicated by *NormFinder*; *PP2A-3*, *PP2A-1* and *SAND* as indicated by  $\Delta$ Ct, and *PP2A-3*, *PP2A-1* and *EF-1 $\alpha$*  as indicated by the comprehensive ranking method. We also used the overall best (*PP2A3*) and the worst (*PPR2*) ranked genes (Fig. 5). The relative expression data we obtained by using the various reference gene combinations

were consistent for both *EgMYB1* (Fig. 5A) and *EgMYB2* (Fig. 5B) expression in all tissues, with the exception of the data obtained by using *PPR2* as a reference gene. This was particularly obvious for the expression of *EgMYB1* and *EgMYB2* in secondary stems, which was largely overestimated by normalizing to *PPR2* expression when comparing with the data obtained with the other reference genes. Using *PPR2* also led to underestimation of *EgMYB1* expression in the cambium samples as well as to an overestimation of *EgMYB2* expression in primary stems. *PPR2* was consistently poorly ranked by the four statistical methods independently of the experimental data set

considered; thus, it should not be used as a reference gene for expression studies in *Eucalyptus*. Conversely, the use of *PP2A3* (best ranked by at least three algorithms) produced relative expression data very similar to those obtained by using combinations of two or three reference genes from all sample subsets. Speleman and colleagues (Vandesompele et al. 2002) recommended the use of three reference genes for pairwise analysis; however, we found that this did not improve normalization when compared with the use of one or two carefully selected reference genes, even in the case of a heterogeneous sample set such as the OT subset in this study (Fig. 3).

By making use of a high-throughput gene expression platform based on microfluidic dynamic arrays, this in-depth study has identified and validated the optimal reference genes for gene expression analysis of a large panel of 90 samples corresponding to 30 distinct organs, tissues or developmental stages and/or environmental conditions (nitrogen fertilization, mechanical, cold and drought stress) in five different species of *Eucalyptus*. The stability of expression of 16 *Eucalyptus* genes was evaluated in 17 experimental data sets, and the data were analyzed by using three statistical algorithms: geNorm (Vandesompele et al. 2002), NormFinder (Andersen et al. 2004) and  $\Delta Ct$  (Silver et al. 2006). As illustrated by PCA, the rankings of gene expression stability depended on the statistical method, on the nature of the biological samples and on the degree of heterogeneity of the panel considered; also, the differences between the methods varied with the sample data set considered. There was very good agreement between the ranking obtained by comprehensive integration ranking, which reflects the GM of the ranking outcomes of the three statistical algorithms, and those identified by PCA analysis. This integrative ranking applied to the 17 experimental panels enabled us to identify, among the 16 candidate genes, the most stable genes: *PP2A-3*, *PP2A-1*, *EF-1 $\alpha$* , *SAND*, *IDH*, *CACS*, *UBC2* and *PTB*, as well as the least stable genes: *PPR2*, *Actin2*, *Expressed1* and *Helicase*. The use of *PPR2* as an internal reference gene was shown to be unsatisfactory because it led to erroneous patterns of expression of two *Eucalyptus* master genes *EgMYB1* and *EgMYB2*. Interestingly, *EF-1 $\alpha$*  and *IDH* have been used traditionally as internal controls for studies of *Eucalyptus* gene expression (Paux et al. 2004, Paux et al. 2005, Legay et al. 2010). Previously, Fett-Netto and colleagues using geNorm identified *IDH* and *SAND* as the most stable genes during in vitro adventitious rooting in *Eucalyptus* (de Almeida et al. 2010). Among the stable genes, *PP2A-3* was the most stable, being among the four best-ranked genes in all the 17 sample panels. We therefore tested *PP2A-3* alone as a reference gene in the most heterogeneous sample panel, OT, for which geNorm recommended at least three reference genes. It is worth noting that for this panel of samples, *PP2A-3* was not in the first three top-ranked genes for geNorm, whereas it was ranked first by NormFinder,  $\Delta Ct$  and by the comprehensive ranking. Interestingly, the use of *PP2A-3* as the only reference gene gave a consistent pattern of expression for both *EgMYB1* and *EgMYB2*. Thus, in our hands, a well-chosen single reference gene is able to provide an

accurate and reliable expression pattern even in a complex and heterogeneous sample panel. This finding is interesting since the systematic use of multiple reference genes is costly and laborious and sometimes even impossible, as in the case of limited available amount of RNA such as samples harvested by laser micro-dissection. In conclusion, this study provided useful clues for accurate and reliable normalization of gene expression in a wide panel of organs, tissues and conditions for *Eucalyptus*, a tree of great economic importance.

## Materials and Methods

### Plant material

Shoot tips and vascular tissue samples (cambium-enriched fraction, secondary phloem and differentiating xylem) were harvested from 7-year-old *Eucalyptus* Gundal hybrids (*gunnii*  $\times$  *darympleana*, genotype 850645) grown in south-west France (Longages) by the Institut Technologique FCBA (Forêt Cellulose Bois-construction Ameublement). Cambium-enriched fractions were also collected from 25-year-old Gundal hybrids (genotype 821290). Vascular tissue sampling was performed as previously described (Paux et al. 2004). Juvenile and mature xylem samples (kindly provided by the RAIZ Institute of Forest and Paper Research, Portugal) were harvested in Herdade do Zambujal from, respectively, 4- and 10-year-old *E. globulus* trees (genotype VC9). Tension and opposite xylem samples were collected at Quinta do Furadouro (Portugal) from 2-year-old trees of three distinct genotypes of *E. globulus* (*GM52*, *BB3* and *MB43* kindly provided by Altri-Florestal, Portugal) after 3 weeks of bending (at 45°). Fruit capsules were harvested from *E. globulus*, genotype C33 (provided by Altri-Florestal, Portugal). Drought-stressed xylem samples were collected from non-irrigated 16-month-old trees (a clonal plantation of *E. urophylla*  $\times$  *E. grandis*, hybrid 1850) after 4 months without rainfall during the dry season in Yanika, Republic of Congo (Villar et al. 2011). Control xylem samples were taken from irrigated trees. These samples were kindly provided by the CIRAD Forêt (France). For plants grown under conditions of nitrogen fertilization, 3-month-old rooted cuttings of a *Eucalyptus* hybrid (*E. urophylla*  $\times$  *E. grandis*, clone *IPB2-H15*, kindly provided by International Paper do Brasil) grown in greenhouse conditions were submitted for 30 d to three different fertilization treatments: 7.5 mM  $\text{NH}_4\text{NO}_3$  for limiting N (N-), 15 mM  $\text{NH}_4\text{NO}_3$  for optimal N (CT) and 30 mM  $\text{NH}_4\text{NO}_3$  for luxuriant N (N+) (Camargo et al. 2012). Each xylem sample consisted of a pool of 20 debarked stems. For cold experiments, 1-year-old *E. globulus* trees (genotype *GM258* provided by Altri-Florestal, Portugal) were submitted to cold (7°C) for 16 h in the dark. In parallel, control plants were maintained for 16 h in the dark in greenhouse conditions. Expanding leaves, fully expanded leaves, primary stems, secondary stems and roots were harvested for each condition. Each sample consisted of pooled tissues from two trees. *Eucalyptus grandis* calli were obtained from in vitro cultures.



For each tissue or organ, three independent biological repetitions were collected, except the fruit capsules samples for which we had only two biological repetitions. Following harvesting, all samples were immediately frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until extraction.

### Total RNA extraction, cDNA synthesis and quality controls

All the procedures used for the qRT-PCR, from the RNA extraction to the calculation of transcript abundance, were performed as described by Udvardi *et al.* (2008), and Derveaux *et al.* (2010). Total RNAs were extracted from 1–5 g of frozen material as described by Southerton *et al.* (1998). RNA concentration and purity were determined by using a NanoDrop spectrophotometer ND-1000 (Thermo Scientific). RNA samples were then treated with a Turbo DNA-free™ kit (Ambion). The absence of remaining genomic DNA was confirmed by PCR using *ubiquitin* primers (EgUbi1\_F, GCGGCTTTT AAGTCTCTTGCGAA; and EgUbi1\_R, TTCGAAGCATAGCTTCCCATATG). The integrity of RNAs was assessed by using the Agilent 2100 Bioanalyzer, and only samples with an RNA integrity number (RIN)  $>7$  were retained for reverse transcription performed by using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems), according to the manufacturer's instructions and using up to 1  $\mu\text{g}$  of total RNA. The quality of each cDNA was assessed by using two pairs of primers located approximately 1.2 kb apart from each other in the 5' and in the 3' regions of the genes encoding *IDH* (5' end primers, F\_AATCGACCTGCTTCGACCCTTC and R\_TCGACCTTGATC TTCTCGAAACCC; 3' end primers, F\_TGCTGTGGCAGCTGA ACTCAAG and R\_ATGTTGTCCGCCAGTCACCTAC) and *PP2A3* (5' end primers, F\_CGGAAGAAGTGGGTGTGTTT and R\_CACAGAGGGTCTCCAATGGT; and 3' end primers, F\_CAG CGGCAAACAAGTGAAGCG and R\_ATTATGTGCTGCATTG CCCAGTC)]. The majority of our cDNA samples showed a Ct value of the 5'-end pair that did not exceed that of the 3'-end pair by more than one cycle number for both genes. These good quality cDNA samples were diluted 5-fold and stored at  $-20^{\circ}\text{C}$  until used for qPCR.

### PCR primer design

The sequences of the putative *E. grandis* orthologs of *Arabidopsis thaliana* genes (best hits) were identified by BLASTP from the Phytozome database (<http://www.phytozome.net/eucalyptus.php>) based on an E-value of  $<1.00\text{E}-60$ . Only *F-box* and *Ubiquitin10* did not meet this criterion, as shown in **Table 1**. Primer pairs were designed using the software QuantPrime (qPCR primer design tool: <http://www.quantprime.de/>; Arvidsson *et al.* 2008) benefiting from exon-intron border, splice variant information of the *E. grandis* genome annotations (Phytozome version 6). Primers were preferentially selected to be as close to the 3' end of the transcripts as possible, and their sequences are shown in **Table 1**.

Oligonucleotides were synthesized by Sigma Life Science (France).

### High-throughput quantitative qRT-PCR

High-throughput qRT-PCR was performed by the Genotoul service in Toulouse (<http://genomique.genotoul.fr/>) using the BioMark® 96:96 Dynamic Array integrated fluidic circuits (Fluidigm Corporation) according to the manufacturer's protocol. Briefly, each cDNA sample was pre-amplified with a pool of primers specific to the target genes by using the following program:  $95^{\circ}\text{C}$  for 10 min, then 14 cycles of  $95^{\circ}\text{C}$  for 15 s and  $60^{\circ}\text{C}$  for 4 min. The pre-amplified products were diluted 1:5 in 10 mM Tris-HCl, pH 8; 0.1 mM EDTA and analyzed by qRT-PCR using the following conditions:  $95^{\circ}\text{C}$  for 10 min, then 35 cycles of  $95^{\circ}\text{C}$  for 15 s and  $60^{\circ}\text{C}$  for 30 s. The specificity of the PCR products was confirmed by analyzing melting curves, following the final PCR cycle. Only primers that produced a linear amplification and qPCR products with a single-peak melting curve were used for further analysis. The efficiency of each pair of primers was determined by plotting the Ct values obtained for serial dilutions of a mixture of all cDNAs and the equation  $\text{Efficiency} = 10^{(-1/\text{slope})} - 1$ . Primer efficiencies for 16 candidate reference genes were higher than 95% and lower than 110% (**Supplementary Table S1**), except for *EF-1 $\alpha$*  (89%) and *SAND* (115%). For all primer pairs, amplicon sizes were around 70 bp, and annealing temperatures around  $63^{\circ}\text{C}$  as indicated in **Supplementary Table S1**. For all reference gene candidates, the amplification plots and melting curves are presented in **Supplementary Fig. S1**.

### Statistical analyses

To analyze the stability of the candidate reference genes, we used three different tools: geNorm, NormFinder and  $\Delta\text{Ct}$ . The geNorm software v3.4 is a Visual Basic Application tool for Microsoft Excel that relies on the principle that the expression ratio of two ideal reference genes should be constant in samples from different experimental conditions or cell types (Vandesompele *et al.* 2002). The NormFinder program uses a model-based approach for identifying the optimal normalization genes among a set of candidates (Andersen *et al.* 2004). It is rooted in a mathematical model of gene expression that enables estimation of both the overall variation of candidate genes and the variation between samples subgroups. For technical reasons, the sample set should contain a minimum of eight samples per group and at least three, but ideally five or more, candidate genes. In this study, the subgroups were defined by either organ, tissue type or experimental condition. The  $\Delta\text{Ct}$  method uses a similar strategy to that of geNorm by comparing relative expression of 'pairs of genes' within each sample to identify reference genes (Silver *et al.* 2006). If the  $\Delta\text{Ct}$  value between two genes remains constant when analyzed in different samples, it means that both genes are either stably expressed or co-regulated. Introduction of a third, fourth or more genes into comparisons will provide information on

which pairs show less variability and hence which gene(s) are stably expressed among samples tested. Finally, we made our own comprehensive ranking of the best reference genes by integrating the ranking obtained by the three methods (geNorm, NormFinder and  $\Delta\text{Ct}$ ). For this, we developed a simple algorithm to present an overall ranking of the best reference genes. Briefly, we first made a linear transformation of the stability values by assigning a series of weight values from 1 to 10 to each reference gene according to the ranking obtained by each algorithm from the most stable gene to the least stable gene; then we calculated the GM of each gene weight across the three methods and finally re-ranked these reference genes. The gene with the lowest GM was considered to be the most stably expressed reference gene.

### Determination of *EgMYB1* and *EgMYB2* expression profiles

We investigated the expression of *Eucalyptus* transcription factor *EgMYB1* (Eucgr.G01774; F\_ACCATGACGAGCCACCA TTTC, R\_ TCAGGTCAGGACACCTTTCTCG) and *EgMYB2* (Eucgr.G03385; F\_AGGCATTGCACCGTCAAGTATG, R\_ TTC TCCTCGGTGGTGGTTGTG) in 90 samples as described under the subheading 'Plant material' above. The primer design and the qRT-PCR conditions were carried out following the same parameters used for the analysis of reference genes. The analyses of the relative expression profiles were obtained through the  $E^{-\Delta\Delta\text{CT}}$  method (Pfaffl 2001) using the efficiency of each MYB primer and each reference gene. The 'leaves' sample was adopted as the calibrator reference tissue.

### Supplementary data

Supplementary data are available at PCP online.

### Funding

This work was supported by grants from the Agence Nationale pour la Recherche (ANR) [Project Tree For Joules ANR-2010-KBBE-007-01]; Centre National pour la Recherche Scientifique (CNRS); University Paul Sabatier Toulouse III (UPS); Fundação para a Ciência e Tecnologia (FCT) [Projects KBBE-TREEFORJOULES P-KBBE/AGR\_GPL/0001/2010, microEGO PTDC/AGR-GPL/098179/2008, and PEst-OE/EQB/LA0004/2011]; and the INTEREG IVB SudoE project Interbio. This work is part of the Laboratoire d'Excellence (LABEX) project entitled TULIP (ANR-10-LABX-41). H.Y., E.L.O.C. and V.C. were supported by PhD grants from the China Scholarship Council [to H.Y.]; the Fundação de Amparo à Pesquisa do Estado de São Paulo [FAPESP, 2008/53520-3; to E.L.O.C.]; the FCT [SFRH/BD/72982/2010; to V.C.]. The Departament d'Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya [a post-doctoral fellowship 'Beatriu de Pinós' to M.S.]; Ciència 2008 program (FCT)/POPH (QREN) [a research contract to J.P.P.].

### Acknowledgments

We are grateful to J.M. Gion and E. Villar (CIRAD, FR), F. Melun and L. Harvenget (FCBA, France), C. Araujo and L. Neves (AltriFlorestal, Portugal) and C. Marques (RAIZ, Portugal) for kindly providing and/or allowing collection of *Eucalyptus* organ and tissue samples, and C. Graça (IICT) and M.N. Saidi (LRSV) for help with sample collection and RNA extraction. We also warmly acknowledge the advice of C. Brière (LRSV) regarding the statistical analyses of our results, and S. Arvidsson for his kind help in linking the *E. grandis* genome to the Quantprime software. Thanks to the Plateforme Génomique Génopole Toulouse/Midi-Pyrénées (Genotoul) for advice and technical assistance with high-throughput Biomark Fluidigm qRT-PCR amplifications. Finally, the authors acknowledge the Eucagene consortium led by A. Myburg and the Department of the Energy (USA) for making available the *E. grandis* genome.

### References

- Andersen, C.L., Jensen, J.L. and Orntoft, T.F. (2004) Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res.* 64: 5245–5250.
- Artico, S., Nardeli, S.M., Brillhante, O., Grossi-de-Sa, M.F. and Alves-Ferreira, M. (2010) Identification and evaluation of new reference genes in *Gossypium hirsutum* for accurate normalization of real-time quantitative RT-PCR data. *BMC Plant Biol.* 10: 49.
- Arvidsson, S., Kwasniewski, M., Riano-Pachon, D.M. and Mueller-Roeber, B. (2008) QuantPrime—a flexible tool for reliable high-throughput primer design for quantitative PCR. *BMC Bioinformatics* 9: 465.
- Boava, L.P., Laia, M.L., Jacob, T.R., Dabbas, K.M., Goncalves, J.F., Ferro, J.A. et al. (2010) Selection of endogenous genes for gene expression studies in *Eucalyptus* under biotic (*Puccinia psidii*) and abiotic (acibenzolar-S-methyl) stresses using RT-qPCR. *BMC Res. Notes* 3: 43.
- Boudet, A.M., Kajita, S., Grima-Pettenati, J. and Goffner, D. (2003) Lignins and lignocelluloses: a better control of synthesis for new and improved uses. *Trends Plant Sci.* 8: 576–581.
- Brunner, A.M., Yakovlev, I.A. and Strauss, S.H. (2004) Validating internal controls for quantitative plant gene expression studies. *BMC Plant Biol.* 4: 14.
- Bustin, S.A., Benes, V., Garson, J.A., Hellems, J., Huggett, J., Kubista, M. et al. (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* 55: 611–622.
- Camargo, E., Costa, L., Soler, M., Salazar, M., Lepikson, J., Gonçalves, D. et al. (2012) Effects of nitrogen fertilization on global xylem transcript profiling of *Eucalyptus urophylla* × *grandis* evaluated by RNA-seq technology. *BMC Proc.* 5(Suppl 7): P106.
- Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K. and Scheible, W.R. (2005) Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. *Plant Physiol.* 139: 5–17.

- de Almeida, M.R., Ruedell, C.M., Ricachenevsky, F.K., Sperotto, R.A., Pasquali, G. and Fett-Neto, A.G. (2010) Reference gene selection for quantitative reverse transcription–polymerase chain reaction normalization during *in vitro* adventitious rooting in *Eucalyptus globulus* Labill. *BMC Mol. Biol.* 11: 73.
- de Oliveira, L.A., Breton, M.C., Bastolla, F.M., Camargo Sda, S., Margis, R., Frazzon, J. *et al.* (2012) Reference genes for the normalization of gene expression in eucalyptus species. *Plant Cell Physiol.* 53: 405–422.
- Demura, T. and Fukuda, H. (2007) Transcriptional regulation in wood formation. *Trends Plant Sci.* 12: 64–70.
- Derveaux, S., Vandesompele, J. and Hellemans, J. (2010) How to do successful gene expression analysis using real-time PCR. *Methods* 50: 227–230.
- Exposito-Rodríguez, M., Borges, A.A., Borges-Perez, A. and Perez, J.A. (2008) Selection of internal control genes for quantitative real-time RT–PCR studies during tomato development process. *BMC Plant Biol.* 8: 131.
- Fernández, M., Villarroel, C., Balbontín, C. and Valenzuela, S. (2010) Validation of reference genes for real-time qRT–PCR normalization during cold acclimation in *Eucalyptus globulus*. *Trees - Struct Funct.* 24: 1109–1116.
- Foucart, C., Paux, E., Ladouce, N., San-Clemente, H., Grima-Pettenati, J. and Sivadon, P. (2006) Transcript profiling of a xylem vs phloem cDNA subtractive library identifies new genes expressed during xylogenesis in *Eucalyptus*. *New Phytol.* 170: 739–752.
- Gallo de Carvalho, M.C., Caldas, D.G., Carneiro, R.T., Moon, D.H., Salvatierra, G.R., Franceschini, L.M. *et al.* (2008) SAGE transcript profiling of the juvenile cambial region of *Eucalyptus grandis*. *Tree Physiol.* 28: 905–919.
- Goicoechea, M., Lacombe, E., Legay, S., Mihaljevic, S., Rech, P., Jauneau, A. *et al.* (2005) EgMYB2, a new transcriptional activator from *Eucalyptus* xylem, regulates secondary cell wall formation and lignin biosynthesis. *Plant J.* 43: 553–567.
- Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J. *et al.* (2011) Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40: D1178–D1186.
- Guenin, S., Mauriat, M., Pelloux, J., Van Wuytswinkel, O., Bellini, C. and Gutierrez, L. (2009) Normalization of qRT–PCR data: the necessity of adopting a systematic, experimental conditions-specific, validation of references. *J. Exp. Bot.* 60: 487–493.
- Guo, G., Huss, M., Tong, G.Q., Wang, C., Li Sun, L., Clarke, N.D. *et al.* (2010) Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. *Dev. Cell* 18: 675–685.
- Gutierrez, L., Mauriat, M., Guenin, S., Pelloux, J., Lefebvre, J.F., Louvet, R. *et al.* (2008a) The lack of a systematic validation of reference genes: a serious pitfall undervalued in reverse transcription–polymerase chain reaction (RT–PCR) analysis in plants. *Plant Biotechnol. J.* 6: 609–618.
- Gutierrez, L., Mauriat, M., Pelloux, J., Bellini, C. and Van Wuytswinkel, O. (2008b) Towards a systematic validation of references in real-time RT–PCR. *Plant Cell* 20: 1734–1735.
- Legay, S., Sivadon, P., Blervacq, A.S., Pavy, N., Baghdady, A., Tremblay, L. *et al.* (2010) EgMYB1, an R2R3 MYB transcription factor from *eucalyptus* negatively regulates secondary cell wall formation in *Arabidopsis* and poplar. *New Phytol.* 188: 774–786.
- Myburg, A.A., Potts, B.M., Marques, C.M.P., Kirst, M., Gion, J.M., Grattapaglia, D. *et al.* (2007) *Eucalyptus*. In *Genome Mapping & Molecular Breeding in Plants*. Edited by Kole, C.R. Springer, Heidelberg.
- Navarro, M., Ayax, C., Martinez, Y., Laur, J., El Kayal, W., Marque, C. *et al.* (2011) Two EguCBF1 genes overexpressed in *Eucalyptus* display a different impact on stress tolerance and plant development. *Plant Biotechnol. J.* 9: 50–63.
- Paiva, J.A.P., Prat, E., Vautrin, S., Santos, M.D., San-Clemente, H., Brommonschenkel, S. *et al.* (2011) Advancing *Eucalyptus* genomics: identification and sequencing of lignin biosynthesis genes from deep-coverage BAC libraries. *BMC Genomics* 12: 137.
- Pang, Z.P., Yang, N., Vierbuchen, T., Ostermeier, A., Fuentes, D.R., Yang, T.Q. *et al.* (2011) Induction of human neuronal cells by defined transcription factors. *Nature* 476: 220–223.
- Paux, E., Carocha, V., Marques, C., Mendes de Sousa, A., Borralho, N., Sivadon, P. *et al.* (2005) Transcript profiling of *Eucalyptus* xylem genes during tension wood formation. *New Phytol.* 167: 89–100.
- 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.
- Pfaffl, M.W. (2001) A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Res.* 29: e45.
- Plomion, C., Leprovost, G. and Stokes, A. (2001) Wood formation in trees. *Plant Physiol* 127: 1513–1523.
- Reid, K.E., Olsson, N., Schlosser, J., Peng, F. and Lund, S.T. (2006) An optimized grapevine RNA isolation procedure and statistical determination of reference genes for real-time RT–PCR during berry development. *BMC Plant Biol.* 6: 27.
- Rengel, D., San Clemente, H., Servant, F., Ladouce, N., Paux, E., Wincker, P. *et al.* (2009) A new genomic resource dedicated to wood formation in *Eucalyptus*. *BMC Plant Biol.* 9: 36.
- Silver, N., Best, S., Jiang, J. and Thein, S.L. (2006) Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. *BMC Mol. Biol.* 7: 33.
- 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.
- Spurgeon, S.L., Jones, R.C. and Ramakrishnan, R. (2008) High throughput gene expression measurement with real time PCR in a microfluidic dynamic array. *PLoS One* 3: e1662.
- Udvardi, M.K., Czechowski, T. and Scheible, W.R. (2008) Eleven golden rules of quantitative RT–PCR. *Plant Cell* 20: 1736–1737.
- Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A. *et al.* (2002) Accurate normalization of real-time quantitative RT–PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3: RESEARCH0034.
- Villar, E., Klopp, C., Noirot, C., Novaes, E., Kirst, M., Plomion, C. *et al.* (2011) RNA-Seq reveals genotype-specific molecular responses to water deficit in *eucalyptus*. *BMC Genomics* 12: 538.
- Xu, M., Zhang, B., Su, X., Zhang, S. and Huang, M. (2010) Reference gene selection for quantitative real-time polymerase chain reaction in *Populus*. *Anal. Biochem.* 408: 337–339.

## B. Développement d'un nouvel axe de recherche visant à identifier des régulateurs de la voie de l'auxine impliqués dans la formation du bois d'*Eucalyptus*.

Avec la publication du génome complet de l'*Eucalyptus* fin 2010 (un an après mon intégration dans l'équipe), j'ai pu saisir cette opportunité en profitant de mon expérience et de mes connaissances sur l'étude de l'auxine. J'ai commencé à développer dans l'équipe un nouvel axe de recherche visant à identifier des régulateurs de la voie de l'auxine impliqués dans la formation du bois d'*Eucalyptus*. Ces travaux de recherche m'ont conduit à encadrer la thèse de Hong Yu (oct. 2010-2014) dont le financement de 4 ans de thèse a été obtenu sur appel d'offre ministériel du gouvernement Chinois. Nous avons établi une procédure d'identification de tous les membres de la famille des gènes d'intérêt (*Aux/IAA* et *ARF*) à l'échelle du génome entier chez l'*Eucalyptus*. Puis, nous avons réalisé des études de séquences, phylogénie et d'expression spatio-temporelle pour sélectionner des gènes candidats pour les études de validation fonctionnelle (Yu et al., 2014; Yu et al., 2015) (Annexe 5 & Annexe 6). Une partie de ces résultats ont été publiés sous forme de poster lors de la conférence internationale « Bioenergy Trees » 2011, à Nancy, qui a été récompensés par un prix « Best Poster ». Avec cette doctorante, nous avons développé ensemble ce nouvel axe de recherche et publié 7 articles dont 5 en premiers auteurs (Cassan-Wang et al., 2012; Li et al., 2015; Wang et al., 2011; Xu et al., 2019; Yu et al., 2022; Yu et al., 2014; Yu et al., 2015) (Annexe 5, Annexe 6, Annexe 8 et Annexe 9). Après la thèse, elle a obtenu un poste permanent de maître de conférences à l'Université de Médecine de Sichuan en Chine. Ces travaux de recherche nous ont permis d'identifier des régulateurs clés auxine dépendants impliqués dans la formation des cellules du bois tels que les membres de la famille *Aux/IAA* (*EgrIAA4*, *EgrIAA9A*, *EgrIAA13* et *EgrIAA20*) et les membres de la famille *ARF* (*EgrARF4*, *EgrARF5*, *EgrARF10* et *EgrARF19*). En particulier, des études fonctionnelles du gène *EgrIAA20* ont montré que la surexpression d'*EgrIAA20* découple spécifiquement la lignification des différents types de cellules dans le xylème secondaire d'*Arabidopsis* (Yu et al., 2022) (Annexe 8). Actuellement, la caractérisation fonctionnelle du gène *EgrIAA9A* est toujours en cours, nous avons obtenu des résultats prometteurs montrant que la surexpression de la protéine *EgrIAA9A* stabilisée chez *Arabidopsis* stimule la différenciation des vaisseaux au sein de cellules du bois et réprime la maturation et la lignification des fibres. La validation fonctionnelle dans un système homologue chez *Eucalyptus* par la technique « hairy root » est en cours et a fait une partie du stage M2 de M. Hugo Génie (Jan-Juin.2023).

Ces activités de recherche m'ont aussi permis d'obtenir un prix "Earl J. Scherago Travel Grant Award" lors de la conférence « International Plant and Animal genome XX » (2738 participants), 2012, San Diego, USA.

En 2017, suite à l'appel d'offre des bourses doctorales du ministère du gouvernement chinois, nous avons obtenu une bourse doctorale de 3 ans pour Mlle Dai Ying (2017-2020), sur le sujet « Mise au point de la technologie CRISPR/Cas9 dans des racines « hairy root » d'*Eucalyptus grandis* et caractérisation fonctionnelle de facteurs de transcription de la signalisation auxinique potentiellement impliqués dans la formation du bois ». Thèse co-encadrée avec Dr. Jacqueline Grima-Pettenati. Malgré des difficultés liées à la crise sanitaire de COVID-19 (nous avons perdu de nombreux matériels biologiques spécifiques (>400 plantes transgéniques) pendant les périodes de confinement), la thèse a pu arriver à son terme et a été soutenue comme prévu en décembre 2020. Ces travaux de recherche font l'objet d'un article en 2020 présentant le premier *Eucalyptus* édité par la technique de pointe de genome editing CRISPR/Cas9 (Dai et al., 2020) (Annexe 7).





Annexe 5. Article *ARF Eucalyptus* à l'échelle du génome

Annexe 6. Article *Aux/IAA Eucalyptus* à l'échelle du génome et études fonctionnelles du gène *EgrIAA4*

Annexe 7. CRISPR/Cas9 *Eucalyptus*

Annexe 8. Article *EgrIAA20*







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

Hong Yu<sup>1</sup>, Marçal Soler<sup>1</sup>, Isabelle Mila<sup>2</sup>, H el ene San Clemente<sup>1</sup>, Bruno Savelli<sup>1</sup>, Christophe Dunand<sup>1</sup>, Jorge A. P. Paiva<sup>3,4</sup>, Alexander A. Myburg<sup>5</sup>, Mondher Bouzayen<sup>2</sup>, Jacqueline Grima-Pettenati<sup>1\*</sup>, Hua Cassan-Wang<sup>1</sup>

**1** LRSV Laboratoire de Recherche en Sciences V eg etales, UMR5546, Universit e Toulouse III/CNRS, Castanet Tolosan, France, **2** Universit e de Toulouse, Institut National Polytechnique-Ecole Nationale Sup erieure Agronomique de Toulouse, Laboratoire de G enomique et Biotechnologie des Fruits, Castanet-Tolosan, France, **3** Instituto de Investigac o Cient fica e Tropical (IICT/MCTES), Lisboa, Portugal, **4** Instituto de Biologia Experimental e Tecnol gica (IBET), Oeiras, Portugal, **5** Department of Genetics, Forestry and Agricultural Biotechnology Institute (FABI), Genomics Research Institute (GRI), University of Pretoria, Pretoria, South Africa

## Abstract

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

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

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

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

**Copyright:**   2014 Yu et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

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

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

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

\* Email: [yu@lrsv.ups-tlse.fr](mailto:yu@lrsv.ups-tlse.fr)

## Introduction

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

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

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

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

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

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

The recent availability of *Eucalyptus grandis* genome [40], the second hardwood forest tree genome fully sequenced, offers new opportunities to get insights into the regulation of secondary growth and cambial activity by ARFs, especially because *Eucalyptus* belongs to evergreen trees that do not present

dormancy in their cambial activity in sharp contrast with deciduous trees like *Populus*. *Eucalyptus* is also the most planted hardwood in the world, mainly for pulp and paper production but is also foreseen as a dedicated energy crop for lignocellulosic biofuel production. Thus, understanding the mechanisms that underlying auxin regulation in *Eucalyptus* wood formation is of interest both in the context of plant development and as a path to improve lignocellulosic biomass production and quality.

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

## Materials and Methods

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

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

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

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

Accession no. <sup>a</sup>	Arabidopsis best hit (blastp) <sup>b</sup>	Amino acid identity % <sup>c</sup>	Short name <sup>d</sup>	Number of predicted alternative transcript <sup>e</sup>	Chromosome <sup>f</sup>	Genome location <sup>g</sup>	ORF (bp) <sup>h</sup>	Deduced polypeptide <sup>i</sup> Length (aa)	MW (kDa)	PI	Exon No.
<i>Eucgr.G00076.1</i>	<i>AtARF1</i>	62.2	<i>EgrARF1</i>	4	7	848,248..860,087	2028	675	75.3	6.06	14
<i>Eucgr.K02197.1</i>	<i>AtARF2</i>	59.3	<i>EgrARF2A</i> **	5	11	29,232,770..29,239,103	2508	835	93.13	6.55	14
<i>Eucgr.803551.1</i>	<i>AtARF2</i>	49.5	<i>EgrARF2B</i>	2	2	60,136,697..60,143,120	2364	787	88.09	6.62	14
<i>Eucgr.D00588.1</i>	<i>AtARF3</i>	46.2	<i>EgrARF3</i>	2	4	10,835,264..10,841,977	2172	723	79.19	6.36	10
<i>Eucgr.802480.1</i>	<i>AtARF4</i>	57.8	<i>EgrARF4</i>	3	2	46,950,723..46,956,747	2394	797	88.53	6.31	12
<i>Eucgr.F02090.1</i>	<i>AtARF5</i>	59.3	<i>EgrARF5*</i>	1	6	28,226,789..28,233,884	2835	944	103.53	5.43	14
<i>Eucgr.D00264.1</i>	<i>AtARF6</i>	69.1	<i>EgrARF6A</i>	2	4	4,224,849..4,232,295	2694	897	99.04	6.04	14
<i>Eucgr.A02065.1</i>	<i>AtARF6</i>	65.6	<i>EgrARF6B</i>	3	1	31,463,716..31,473,722	2613	870	96.64	6.06	14
<i>Eucgr.D01764.1</i>	<i>AtARF9</i>	58.6	<i>EgrARF9A</i>	1	4	31,589,333..31,593,948	1971	656	73.53	6.15	14
<i>Eucgr.E00888.1</i>	<i>AtARF9</i>	58.4	<i>EgrARF9B</i>	2	5	9,303,573..9,308,209	2064	687	76.46	6.09	14
<i>Eucgr.J00923.1</i>	<i>AtARF10</i>	60.1	<i>EgrARF10</i>	1	10	10,070,677..10,074,840	2139	712	77.68	8.32	4
<i>Eucgr.G02838.1</i>	<i>AtARF16</i>	58	<i>EgrARF16A</i>	1	7	46,569,603..46,573,031	2079	692	76.33	6.80	3
<i>Eucgr.K01240.1</i>	<i>AtARF16</i>	52.5	<i>EgrARF16B</i>	1	11	15,584,770..15,588,716	2124	707	77.91	6.91	3
<i>Eucgr.F04380.1</i>	<i>AtARF17</i>	42.1	<i>EgrARF17*</i>	1	6	52,730,940..52,733,856	1881	626	67.75	7.54	3
<i>Eucgr.C03293.1</i>	<i>AtARF19</i>	60.4	<i>EgrARF19A</i>	2	3	62,460,408..62,469,132	3360	1119	124.93	6.30	14
<i>Eucgr.C02178.1</i>	<i>AtARF19</i>	44.8	<i>EgrARF19B</i>	2	3	39,505,080..39,513,066	3360	1119	123.24	5.98	14
<i>Eucgr.K03433.1</i>	-	-	<i>EgrARF24*</i>	1	11	43,352,151..43,357,873	1836	611	68.44	7.24	14

\*Using FGeneSH to complete the complete sequence.

\*\*using specific primers to amplify the genomic DNA to complete the sequence.

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

## Sequence, phylogenetic, gene structure analysis

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

## Plant material

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

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

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

## Transactivation analysis in single cell system

For testing the ability of ARF transcription factors to up or down regulate the expression of auxin responsive promoter DR5, the full-length cDNAs of the ARF transcription factors were cloned in pGreen vector under 35SCaMV promoter to create the effector constructs. The reporter constructs use a synthetic auxin-responsive promoter DR5 fused with the GFP reporter gene. Tobacco BY-2 protoplasts were co-transfected with the reporter

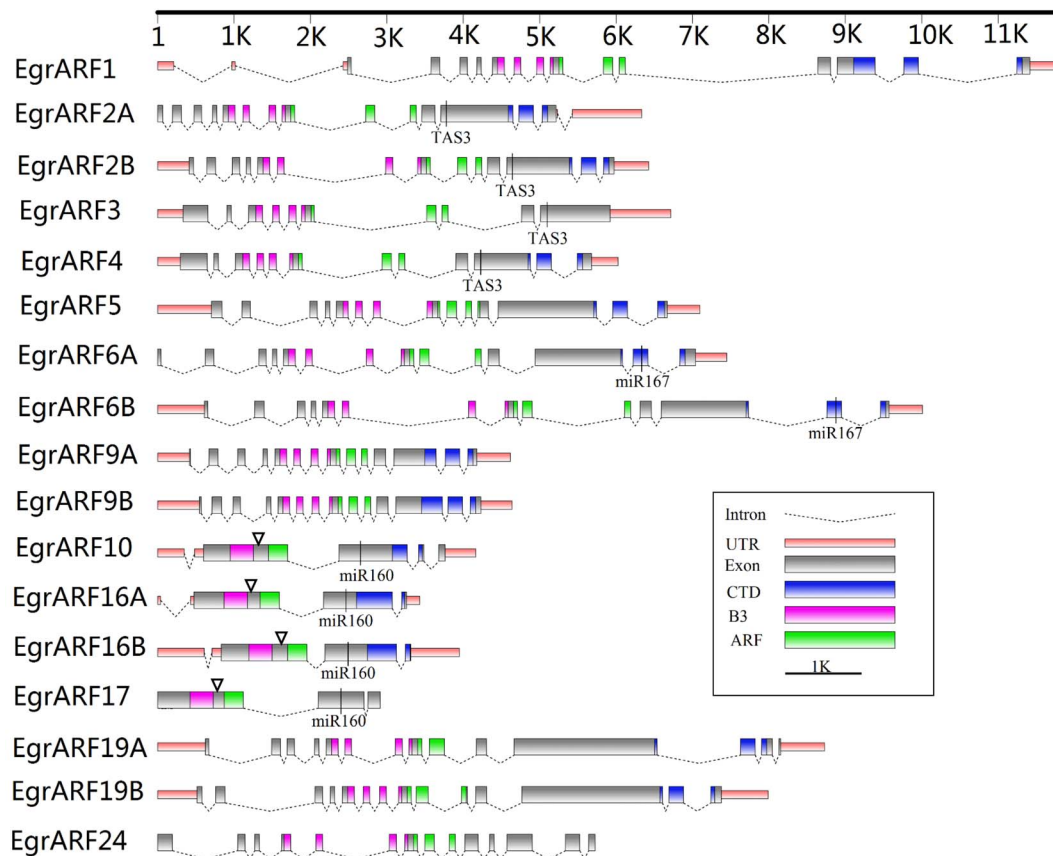
and effector constructs as described in Audran-Delalande et al. [47]. After 16 h incubation, GFP expression was quantified by flow cytometry (LSR Fortessa, BD Biosciences). Data were analysed using BD FACS Diva software. Transfection assays were performed in three independent replicates and 3000–4000 protoplasts were gated for each sample. GFP fluorescence corresponds to the average fluorescence intensity of the protoplasts population after subtraction of auto-fluorescence determined with non-transformed protoplasts. 50  $\mu$ M 2, 4-D was used for auxin treatment. We tested two independent protoplast preparations and for each of them, we performed in three independent transformation replicates. Similar results were obtained with the independent protoplast preparations and the data were represented by one of the preparations. For normalization, protoplasts were transformed with the reporter vector and the effector plasmid lacking the *ARF* gene.

## Results and Discussion

### Identification and chromosomal distribution of *Eucalyptus* ARF genes

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

The predicted proteins encoded by the *EgrARF* genes ranged from 593–1119 amino acid residues (Table 1), with PIs in the range of 5.43–8.32, suggesting that they can work in very different subcellular environments. Sequence analyses of the predicted proteins and Pfam protein motif analysis showed that most of them (14 of the 17 predicted proteins) harbour the typical ARF protein structure comprising a highly conserved DNA-binding domain (DBD) in the N-terminal region composed of a plant specific B3-type subdomain and an ARF subdomain, a variable middle region (MR) that functions as an activation or repression domain, and a carboxy-terminal dimerization (CTD) domain consisting of two highly conserved dimerization subdomains III and IV, similar to those found in Aux/IAAs (Fig. 1). We analysed and aligned the predicted amino acid sequences of the *EgrARFs* (Fig. 1 and Fig. S3). Four out of the 17 *EgrARFs* (*10*, *16A*, *16B* and *17*) exhibited an additional short segment of amino acids (between 15 to 43 amino-acids) in their DBD, between the B3 and ARF subdomains (Fig. 1 and Fig. S3). Such a feature has already been reported in *Arabidopsis* and soybean [36]. At the end of the DBD domain, all of the *EgrARFs* except *EgrARF6A*, *6B* and *19A* contain a conserved putative mono-partite nuclear localization signal (NLS) (Fig. S3) shown to direct the proteins into the nucleus [36,48].



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

doi:10.1371/journal.pone.0108906.g001

Five *EgrARF*s (*EgrARF*5, 6A, 6B, 19A and 19B) harbour a glutamine (Q)-rich middle region (Fig. S3) implying that these proteins are likely transcriptional activators since glutamine enrichment seems to be a distinctive feature of ARF activators in all plant lineages [15,49]. The other 12 *EgrARF*s may function as repressors based on their middle regions enriched either in S (serine), SPL (Serine, Proline, Leucine) or SGL (Serine, Glycine, Leucine) [36] (Fig. S3).

The predicted protein structures of *EgrARF*3 and *EgrARF*17 are lacking dimerization domains III and IV like their potential orthologs in *Arabidopsis* (Fig. 1 and Fig. S3). *EgrARF*24, which has no ortholog in *Arabidopsis*, has a truncated CTD since only Aux/IAA subdomain III is present. The percentage of *EgrARF*s displaying a truncated CTD (17.6%) is similar to that in *Arabidopsis* (17.4%), but lower than in rapeseed (22.6%) or tomato (28.6%) [24,37,50]. These truncated *EgrARF*s are predicted to be unable to interact with Aux/IAA, a sequestration mechanism which may regulate their activity, and hence, they are likely to be insensitive to auxin. However, ARF repressors seem to display very limited interactions with Aux/IAA proteins [18], therefore the lack of domains III and/or IV could also have consequences for the interaction of ARFs with other transcriptional regulators [49].

Compared to *Arabidopsis*, the ARF family in *Eucalyptus* is slightly contracted with 17 versus 23 members. It is worth noting

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

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



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

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

doi:10.1371/journal.pone.0108906.t002

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

### Comparative Phylogenetic analysis of the ARF family

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

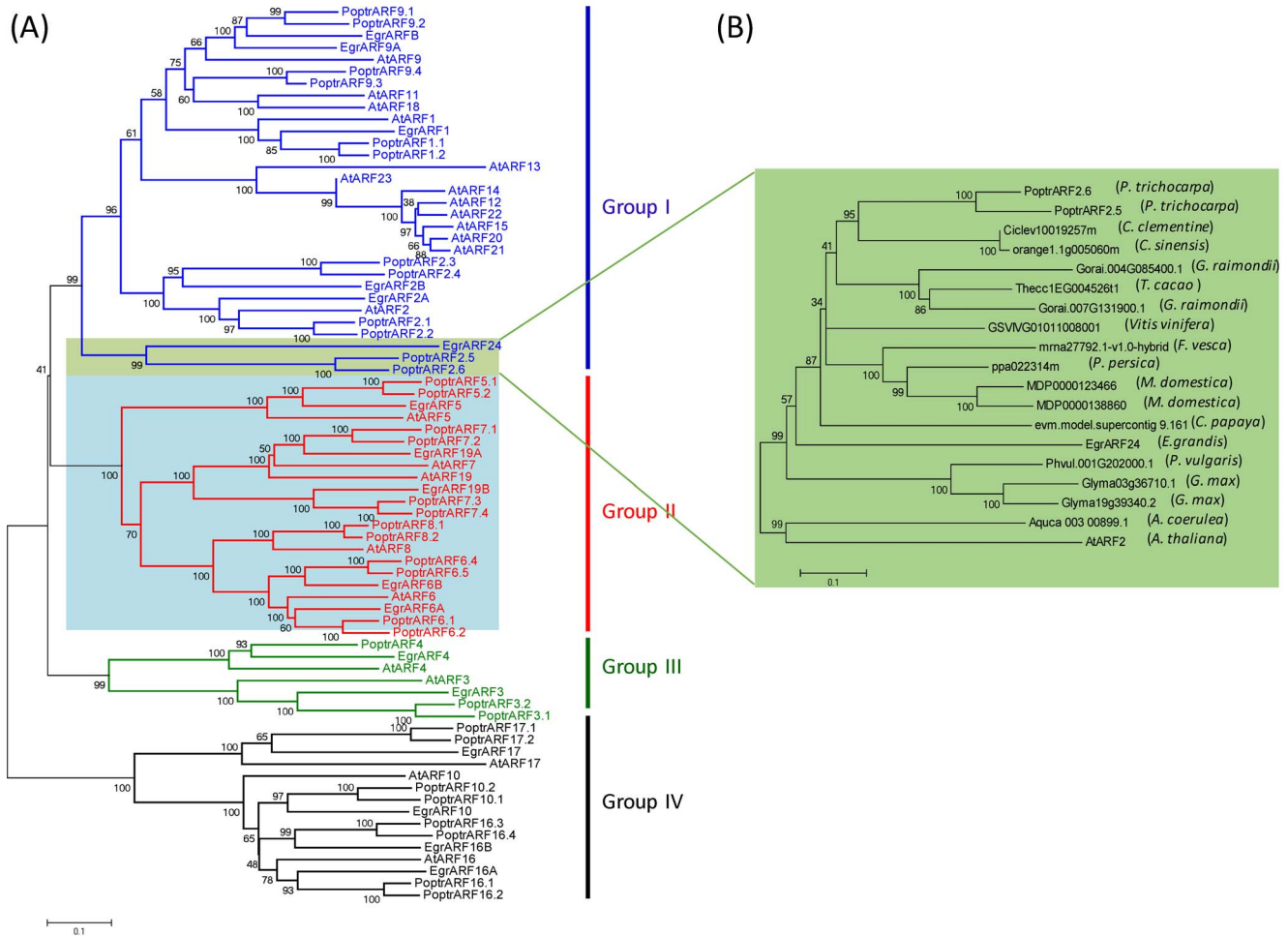
species; the remaining *Egr*ARFs were distributed among the three other groups.

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

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

### Prediction of small regulatory RNAs and their potential ARF targets

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



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

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

We used these newly identified *Eucalyptus* small RNAs as probes to search *in silico* for their target sites in *EgrARF* genes. Ten of the 17 *EgrARF* genes were found to be potential targets of these three small RNAs (Table S7). We identified highly conserved target sites for *EgrmiR160* in *EgrARF10*, *16A*, *16B* and *17*, for *EgrmiR167* in *EgrARF6A* and *B*, and for *EgrTAS3* in *EgrARF2A*, *2B*, *3* and *4* (Table S7). The targeting of three different small RNA to their corresponding target genes was highly conserved between *Arabidopsis* and *Eucalyptus* suggesting common regulation of plant growth and development. For example, *miR160*, a highly conserved miRNA group across the plant

kingdom, is known to target *ARF10*, *ARF16* and *ARF17* to regulate various aspects of plant development [30,52,53]. In *Arabidopsis*, *miR167* regulates lateral root outgrowth [57], adventitious rooting [58], ovule and anther growth, flower maturation [20,29] and jasmonic acid homeostasis [59] by targeting both *AtARF6* and *AtARF8*. Very recently, it has also been shown that *miR167* regulates flower development and female sterility in tomato [60]. Because *Eucalyptus* is a woody perennial plant, one could expect that some small RNAs (for instance *miR160* and *Tasi 3*) could be involved in the regulation of wood formation through targeting of ARF genes preferentially expressed in cambial cells or developing xylem.

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

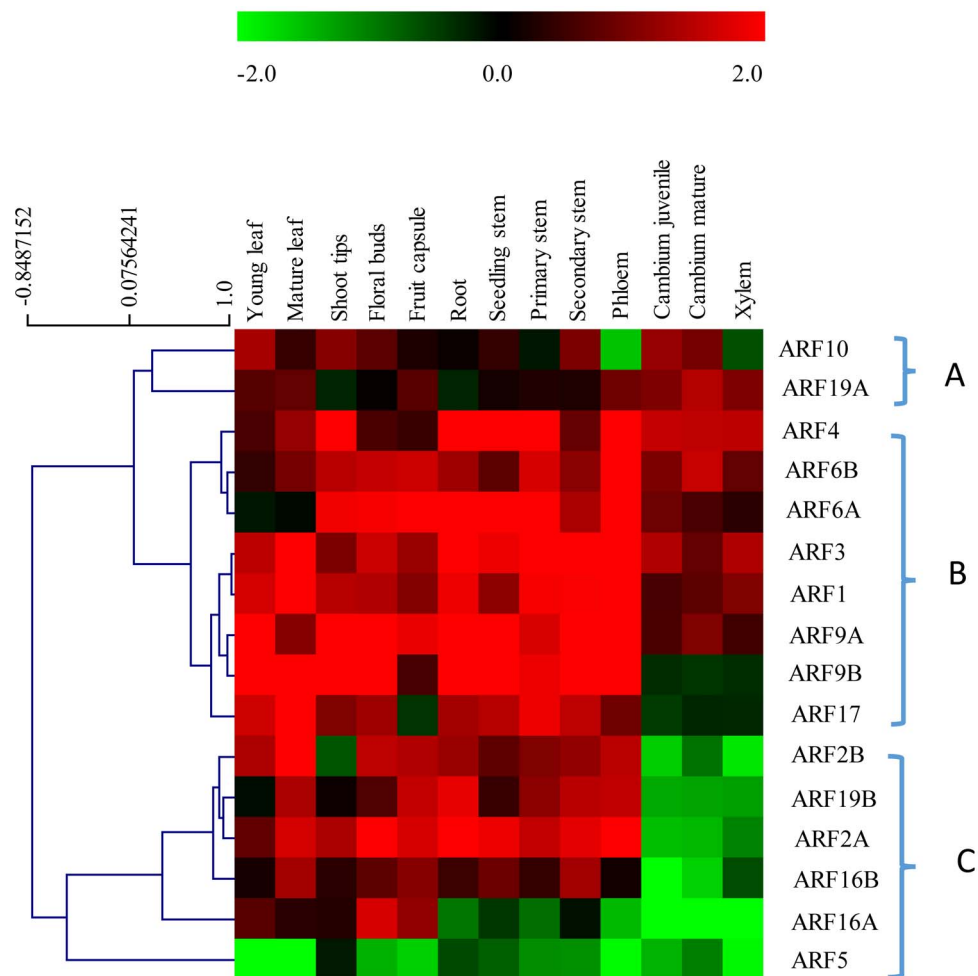
To start investigating the functions of the *EgrARF* genes, we assessed their transcript expression levels in various *Eucalyptus* organs and tissues by qRT-PCR, with special attention to vascular



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

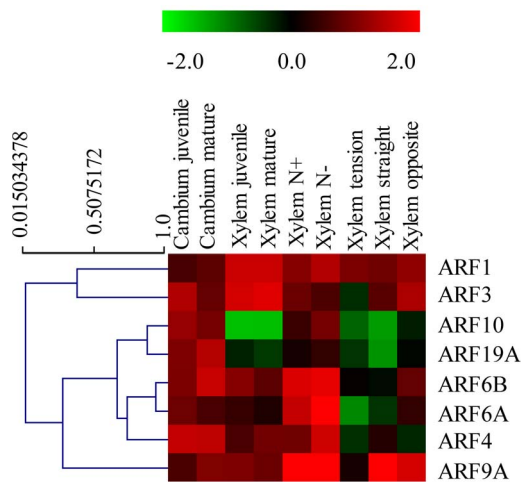
Heatmap representation (Fig. 3) indicated that *EgrARF* genes were expressed across various tissues and organs, but different members displayed preference to particular tissues and/or organs and could therefore be clustered into three main expression groups. Group A is the smallest with only two members *EgrARF10* (predicted repressor) and *EgrARF19A* (predicted activator) showing a relatively higher expression in vascular cambium as compared to other tissues and/or organs. *EgrARF10* was expressed at higher level in cambium (both mature and

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



**Figure 3. Expression profiles of 16 *EgrARF* genes in various organs and tissues.** The heat map was constructed by using the relative expression values determined by qRT-PCR of 16 *EgrARF* genes (indicated on the right) in 13 tissues and organs (indicated at the top) normalized with a control sample (*in vitro* plantlets). In the heat map, red and green indicate relatively high and lower expression ( $\log_2$ ratios) than in the control, respectively. Each measurement is the mean of three independent samples. The heat map and the hierarchical clustering were generated by MultiExperiment Viewer (MEV).

doi:10.1371/journal.pone.0108906.g003

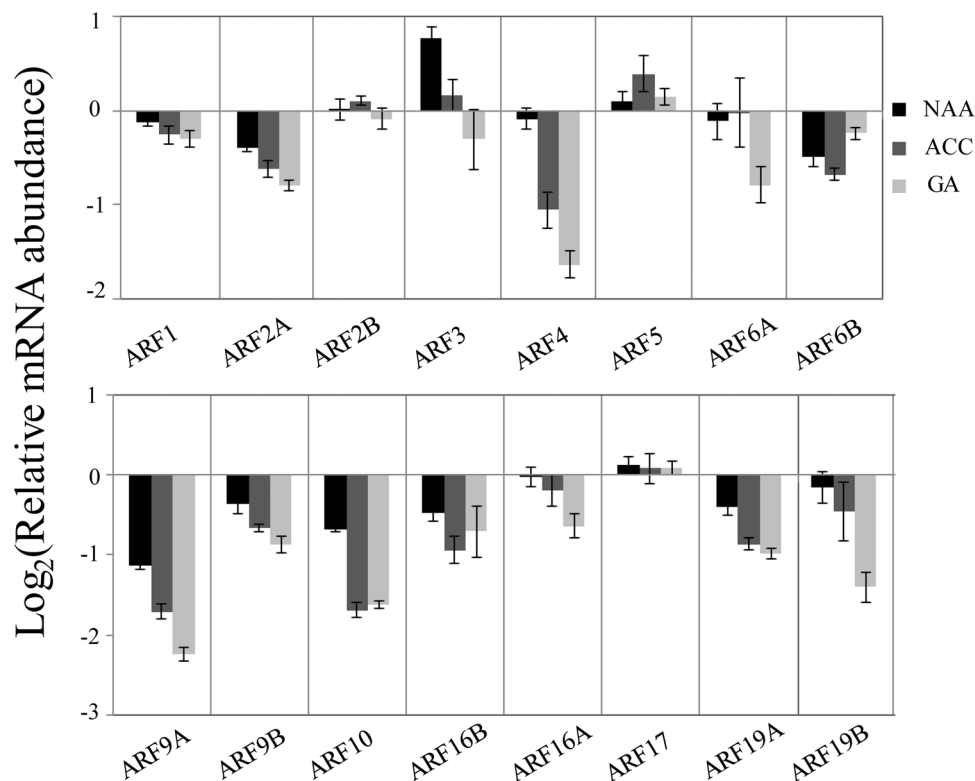


**Figure 4. Effect of environmental cues and developmental stages on *EgrARF* expression.** The heat map was constructed by using the relative expression values determined by qRT-PCR of *EgrARF* genes (indicated on the right) in various tissues and conditions (indicated at the top) normalized with a control sample (*in vitro* plantlets). In the heat map, red and green indicates relatively higher expression and lower expression ( $\log_2$ ratios) than in the control, respectively. The heat map and the hierarchical clustering were generated by MultiExperiment Viewer (MEV). doi:10.1371/journal.pone.0108906.g004

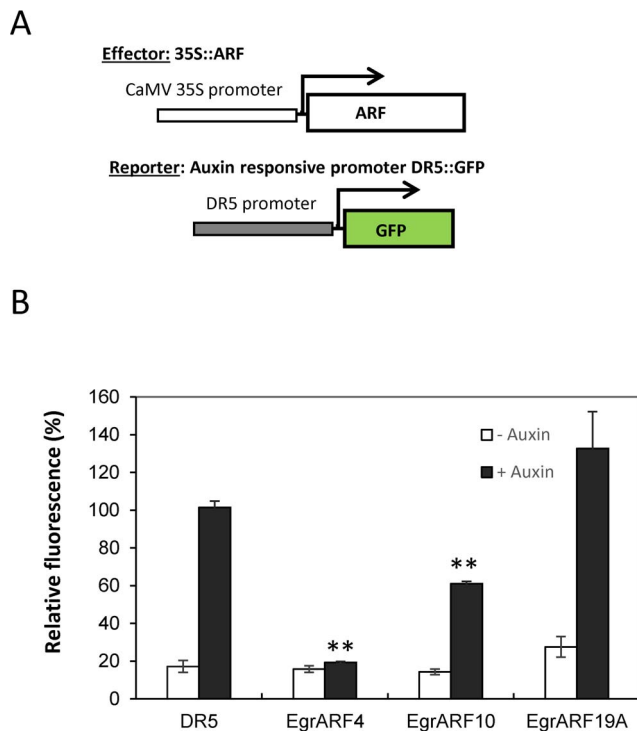
ortholog, *EgrARF19B* was highly expressed in root [24]. It should be noted that the activator *EgrARF5* is highly expressed in all samples analysed, with the highest expression in *in vitro* plantlets. Because *in vitro* plantlets were used to normalize the expression data in the heatmap, the expression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

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

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



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



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

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

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

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

Several hormones are known to regulate cambium activity and xylem formation [[31], [64] and references therein]. For instance,

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

#### Transcriptional activities of *EgrARF4*, *EgrARF 10* and *EgrARF19A*

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

#### Conclusions

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

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

## Supporting Information

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

(TIFF)

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

(TIFF)

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

(PDF)

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

(PDF)

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

(PDF)

**Figure S6 Comparative Phylogenetic relationships between ARF proteins from poplar, *Eucalyptus*, grapevine,**

***Arabidopsis*, tomato and rice.** Full-length protein sequences were aligned using the Clustal\_X program. The phylogenetic tree was constructed by using the MEGA5 program and the neighbour-joining method with predicted full-length ARF proteins. Bootstrap supports are indicated at each node.

(PDF)

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

(TIFF)

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

(TIFF)

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

(TIFF)

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

(TIFF)

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

(PDF)

**Table S2 Protein identity matrix between *EgrARF* and *AtARF*.**

(PDF)

**Table S3 Protein identity matrix among *EgrARF*.**

(PDF)

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

(PDF)

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

(PDF)

**Table S6 Potential small RNAs targeting *EgrARF* genes.**

(PDF)

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

(PDF)

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

(ZIP)

## Acknowledgments

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

Gion and E. Villar (CIRAD, FR), F. Melun and L. Harvenge (FCBA, France), C. Araujo and L. Neves (AltriFlorestal, Portugal) and C. Marques (RAIZ, Portugal) for kindly providing and/or allowing collection of *Eucalyptus* organ and tissue samples, C. Graça, and V. Carocha for help with sample collection and RNA extraction, N Ladouce for performing Fluidigm experiments. Thanks to the Bioinfo Genotoul Platform (<http://bioinfo.genotoul.fr>) for access to bioinformatic resources and to the Genome and transcriptome Platform (<http://get.genotoul.fr/>) for advice and technical assistance with high-throughput Biomark Fluidigm qRT-PCR amplifications. Thanks to Y. Martinez (FR3450) for assistance with

microscopy analysis to C. Pecher and A. Zakaroff-Girard for technical assistance and expertise in flow cytometry (Toulouse RIO imaging platform <http://tri.ups-tlse.fr>).

## Author Contributions

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

## References

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

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





# Comprehensive Genome-Wide Analysis of the *Aux/IAA* Gene Family in *Eucalyptus*: Evidence for the Role of *EgrIAA4* in Wood Formation

Hong Yu<sup>1</sup>, Marçal Soler<sup>1</sup>, Hélène San Clemente<sup>1</sup>, Isabelle Mila<sup>2,3</sup>, Jorge A.P. Paiva<sup>4,5</sup>, Alexander A. Myburg<sup>6</sup>, Mondher Bouzayen<sup>2,3</sup>, Jacqueline Grima-Pettenati<sup>1,\*</sup> and Hua Cassan-Wang<sup>1</sup>

<sup>1</sup>LRSV Laboratoire de Recherche en Sciences Végétales, UMR5546, Université Toulouse III, UPS, CNRS, BP 42617, Auzeville, F-31326 Castanet Tolosan, France

<sup>2</sup>Université de Toulouse, INP-ENSA Toulouse, Génomique et Biotechnologie des Fruits, Avenue de l'Agrobiopole BP 32607, F-31326 Castanet-Tolosan, France

<sup>3</sup>INRA, UMR990 Génomique et Biotechnologie des Fruits, Chemin de Borde Rouge, F-31326 Castanet-Tolosan, France

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

<sup>5</sup>IBET – Instituto de Biologia Experimental e Tecnológica, Apartado 12, 2781-901 Oeiras, Portugal

<sup>6</sup>Department of Genetics, Forestry and Agricultural Biotechnology Institute (FABI), Genomics Research Institute (GRI), University of Pretoria, Private Bag X20, Pretoria, 0028, South Africa

\*Corresponding author: Email, [grima@lrsv.ups-tlse.fr](mailto:grima@lrsv.ups-tlse.fr)

(Received October 31, 2014; Accepted December 23, 2014)

**Auxin plays a pivotal role in various plant growth and development processes, including vascular differentiation. The modulation of auxin responsiveness through the auxin perception and signaling machinery is believed to be a major regulatory mechanism controlling cambium activity and wood formation. To gain more insights into the roles of key *Aux/IAA* gene regulators of the auxin response in these processes, we identified and characterized members of the *Aux/IAA* family in the genome of *Eucalyptus grandis*, a tree of worldwide economic importance. We found that the gene family in *Eucalyptus* is slightly smaller than that in *Populus* and *Arabidopsis*, but all phylogenetic groups are represented. High-throughput expression profiling of different organs and tissues highlighted several *Aux/IAA* genes expressed in vascular cambium and/or developing xylem, some showing differential expression in response to developmental (juvenile vs. mature) and/or to environmental (tension stress) cues. Based on the expression profiles, we selected a promising candidate gene, *EgrIAA4*, for functional characterization. We showed that *EgrIAA4* protein is localized in the nucleus and functions as an auxin-responsive repressor. Overexpressing a stabilized version of *EgrIAA4* in *Arabidopsis* dramatically impeded plant growth and fertility and induced auxin-insensitive phenotypes such as inhibition of primary root elongation, lateral root emergence and agravitropism. Interestingly, the lignified secondary walls of the interfascicular fibers appeared very late, whereas those of the xylary fibers were virtually undetectable, suggesting that *EgrIAA4* may play crucial roles in fiber development and secondary cell wall deposition.**

**Keywords:** *Aux/IAA* • Auxin • *Eucalyptus* • Gene expression • Secondary cell walls • Wood formation.

**Abbreviations:** AFB, auxin signaling F-box; ARF, auxin-response factor; AuxRE, auxin-responsive *cis*-element; Aux/IAA, auxin/indole-3-acetic acid; CaMV, *Cauliflower mosaic*

*virus*; EAR, ethylene-responsive element-binding factor-associated amphiphilic repression; GFP, green fluorescent protein; MS, Murashige and Skoog; NLS, nuclear localization signal; qRT-PCR, quantitative reverse transcription-PCR; SCW, secondary cell wall; TIR1, transport inhibitor response1.

## Introduction

The plant hormone auxin plays an important role in regulating plant growth and developmental processes such as embryogenesis, apical dominance, lateral root formation, tropism, fruit development and vascular differentiation (Friml 2003). Auxin also plays a crucial role in specifying vascular stem cells (Miyashima et al. 2013), and regulating the activity of the vascular cambium (for a review, see Bhalerao and Fischer 2014)—a lateral meristem that contributes to secondary radial growth of wood in trees. Measurements of auxin levels across wood-forming tissues revealed a radial auxin concentration gradient where high auxin concentrations localize to the cambium, intermediate concentrations to the xylem elongation zone and low concentrations to the maturation zone (Uggla et al. 1996, Tuominen et al. 1997, Uggla et al. 1998). It was proposed that this gradient regulates cambial activity and differentiation of cambial derivatives by providing positional information to cells within the tissue (Uggla et al. 1998, Sundberg et al. 2000, Schrader et al. 2003); however, the hypothesis that auxin acts as a morphogen still lacks strong experimental support (Nilsson et al. 2008, Bhalerao and Fischer 2014). Moreover, the expression patterns of most of the auxin-responsive genes correlate only weakly with auxin concentration across the wood-forming gradient, arguing against a strong and direct impact of auxin levels on radial patterning (Nilsson et al. 2008). In addition, there is only subtle variation in the absolute auxin levels between active and dormant cambium in trees,

suggesting that there may be instead a seasonal fluctuation of auxin sensitivity (Uggla et al. 1996, Uggla et al. 2001, Schrader et al. 2003, Schrader et al. 2004). Indeed, the reduced auxin responsiveness of the dormant cambium correlates with reduced expression levels of components of the auxin perception machinery, implying that auxin signaling controls cambial activity by modulation of auxin responsiveness (Baba et al. 2011).

Wood is a highly variable material that is both developmentally and environmentally regulated (Plomion et al. 2001). For instance, in reaction to mechanical stress, a local increase in cambial cell division induces the formation of tension wood in the upper side of bent angiosperm tree stems. Tension wood has specific anatomical features such as the presence of a characteristic inner gelatinous cell wall layer (G layer) (Timell 1969, Pilate et al. 2004). Auxin is proposed to be implicated in the tension response, and application of either exogenous auxin or auxin transport inhibitors is known to induce the formation of G fibers (Morey and Cronshaw 1968, Mellerowicz et al. 2001). For many years it was assumed that auxin distribution is involved in the regulation of tension wood but Hellgren et al. (2004) showed that auxin levels were homogeneously balanced under gravitational stress in bent *Populus* stems. Auxin may instead exert its influence by means of components of its signaling pathway, as suggested by changes in expression of a large set of auxin-related genes (Andersson-Gunneras et al. 2006) including some members of the aspen Aux/IAA gene family (Moyle et al. 2002, Paux et al. 2005).

The perception and signaling of auxin involves central regulators such as the transport inhibitor response 1 (TIR1)/auxin signaling F-box (ABFs) proteins, auxin/IAA (Aux/IAA) proteins and auxin response factor (ARF) proteins (Mockaitis and Estelle 2008, Calderon Villalobos et al. 2012). Aux/IAA proteins are direct targets of TIR1 and of its paralogous ABFs (Dharmasiri et al. 2005, Kepinski and Leyser 2005, Tan et al. 2007). At low intracellular auxin concentrations, Aux/IAA proteins act as transcriptional repressors of auxin-mediated gene expression by interacting and sequestering ARF proteins, thus preventing them from regulating the transcription of their target genes (Guilfoyle and Hagen 2007). In contrast, high intracellular auxin levels foster interactions between Aux/IAA proteins and TIR1 E3 ubiquitin–ligase complexes, resulting in the degradation of Aux/IAA proteins by the 26 S proteasome (Gray et al. 2001, Woodward and Bartel 2005, Tan et al. 2007). As a consequence, ARF proteins are released from their Aux/IAA interaction partners and can regulate the transcription of their auxin-responsive target genes.

The Aux/IAA genes were first identified in soybean and pea and were described as early auxin-responsive genes (Walker and Key 1982, Theologis et al. 1985, Ainley et al. 1988). This plant-specific transcription factor family has 29 members in Arabidopsis (Overvoorde et al. 2005), the vast majority of which encode short-lived nuclear proteins. Canonical Aux/IAA contains four highly conserved domains (called I, II, III and IV). Domain I contains a conserved leucine repeat motif (LxLxLx) similar to the EAR (ethylene-responsive element-binding factor-associated amphiphilic repression) motif,

which is responsible for the transcriptional repressor activity of this protein family (Tiwari et al. 2004) and can also interact with the co-repressor TOPLESS (Szemenyei et al. 2008). Domain II has a motif of five highly conserved amino acids (VGWPP) that leads to the rapid degradation of Aux/IAA proteins through interaction with a component of the TIR1/ABFs ubiquitin–proteasome protein degradation pathway (Gray et al. 2001, Dharmasiri et al. 2005, Kepinski and Leyser 2005). This interaction is abolished by mutations within the core VGWPP motif of domain II, resulting in accumulation of the mutated protein and leading to defects in auxin responses (Gray et al. 2001, Ouellet et al. 2001, Reed 2001, Tian et al. 2003). Domains III and IV can mediate homo-dimerization and hetero-dimerization with other Aux/IAA family members, as well as dimerization with ARFs which also contain these two similar domains (Kim et al. 1997, Ulmasov et al. 1997, Ouellet et al. 2001). A high-throughput protein–protein interaction study of 29 Aux/IAAs and 23 ARFs in Arabidopsis showed that the majority of Aux/IAA proteins interact with themselves and with ARF activators (Vernoux et al. 2011).

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

Although, several studies have found that the auxin perception and signaling machinery are important in regulating cambial activity, cambial dormancy, secondary cell wall (SCW) deposition and tension wood formation (Bhalerao and Fischer 2014), the involvement of Aux/IAA genes in these processes remains largely underinvestigated. Work by Nilsson et al. (2008) showed that changes in endogenous auxin levels in wood-forming tissues modulate expression of a few key regulators such as Aux/IAA genes that control global gene expression patterns essential for normal secondary xylem development. Moreover, these authors found that overexpressing a stabilized version of *PttIAA3m* (mutation in the degron domain II) in *Populus* led to a reduction in cambium cell division and a decrease in secondary xylem width (Nilsson et al. 2008). To the best of our knowledge, this is the only example demonstrating the role of an Aux/IAA gene in xylem development in a woody species.

The sequence of the *Eucalyptus grandis* genome (Myburg et al. 2014) was published recently. This is the second forest tree genome to be sequenced, and it offers unique opportunities to analyze the characteristics of the Aux/IAA family in the most planted hardwood worldwide, which, in contrast to *Populus*, does not present cambium dormancy. In this study, we performed a comprehensive genome-wide identification and characterization of the Aux/IAA gene family in *E. grandis*. In addition to analyses of comparative phylogenetics, genomic organization and prediction of protein structural motifs, we investigated by quantitative reverse transcription–PCR (qRT–PCR) the expression profiles of the 24 Aux/IAA members

among various organs and tissues, at different developmental stages and in response to environmental cues, with a special focus on wood-forming tissues. Based on these phylogenetic and expression results, we identified *EgrIAA4* as the best candidate gene potentially involved in the regulation of wood formation in *Eucalyptus*. Overexpression of a stabilized version of *EgrIAA4* (*EgrIAA4m*) in transgenic Arabidopsis led to auxin-related aberrant phenotypes and strongly affected interfascicular and xylary fiber formation, thereby confirming the hypothesized role of *EgrIAA4* in the regulation of cambium and wood-forming tissues.

## Results

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

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

The coding sequences of the vast majority (83%) of the *EgrIAA* genes are disrupted by three or four introns, and the intron positions and phases are well conserved (**Fig. 1**). Variations were observed for some members, however, involving mainly loss of one or more introns (*EgrIAA3A*, 19 and 33A) and, in one case, gain of one additional intron (*EgrIAA9B*). The *EgrIAA* exon–intron patterns are similar to those of their orthologs in Arabidopsis (**Supplementary Fig. S2**).

The predicted *EgrIAA* proteins range in size from 154 (*EgrIAA20*) to 370 amino acids (*EgrIAA26B*) with corresponding molecular masses ranging from 17 to 41 kDa (**Table 1**). The theoretical isoelectric points also vary widely from 4.7 (*EgrIAA33B*) to 9.5 (*EgrIAA31*), indicating that they may function in different microenvironments. Pair-wise analysis of *EgrIAA* protein sequences showed that the identity levels vary greatly from 83.7% (*EgrIAA3A* and 3B) to 26.4% (*EgrIAA9A* and 28) (**Supplementary Table S1**). A similar wide variation of identity values was reported in Arabidopsis (Overvoorde et al. 2005) and tomato (Audran-Delalande et al. 2012). Sequence alignment of the predicted proteins and MEME protein motif analyses showed that 17 of the 24 *EgrIAA* proteins have the typical four highly conserved domains of canonical *Aux/IAA* proteins (**Fig. 1**; **Supplementary Fig. S4**), comprising (i) an N-terminal repressor domain I with a conserved leucine repeat motif (LxLxLx), which can recruit the co-repressor TOPLESS; (ii) a degron domain II, which is responsible for the stability of *Aux/IAA* proteins; and (iii) a C-terminal dimerization (CTD) domain, consisting of two highly conserved dimerization subdomains

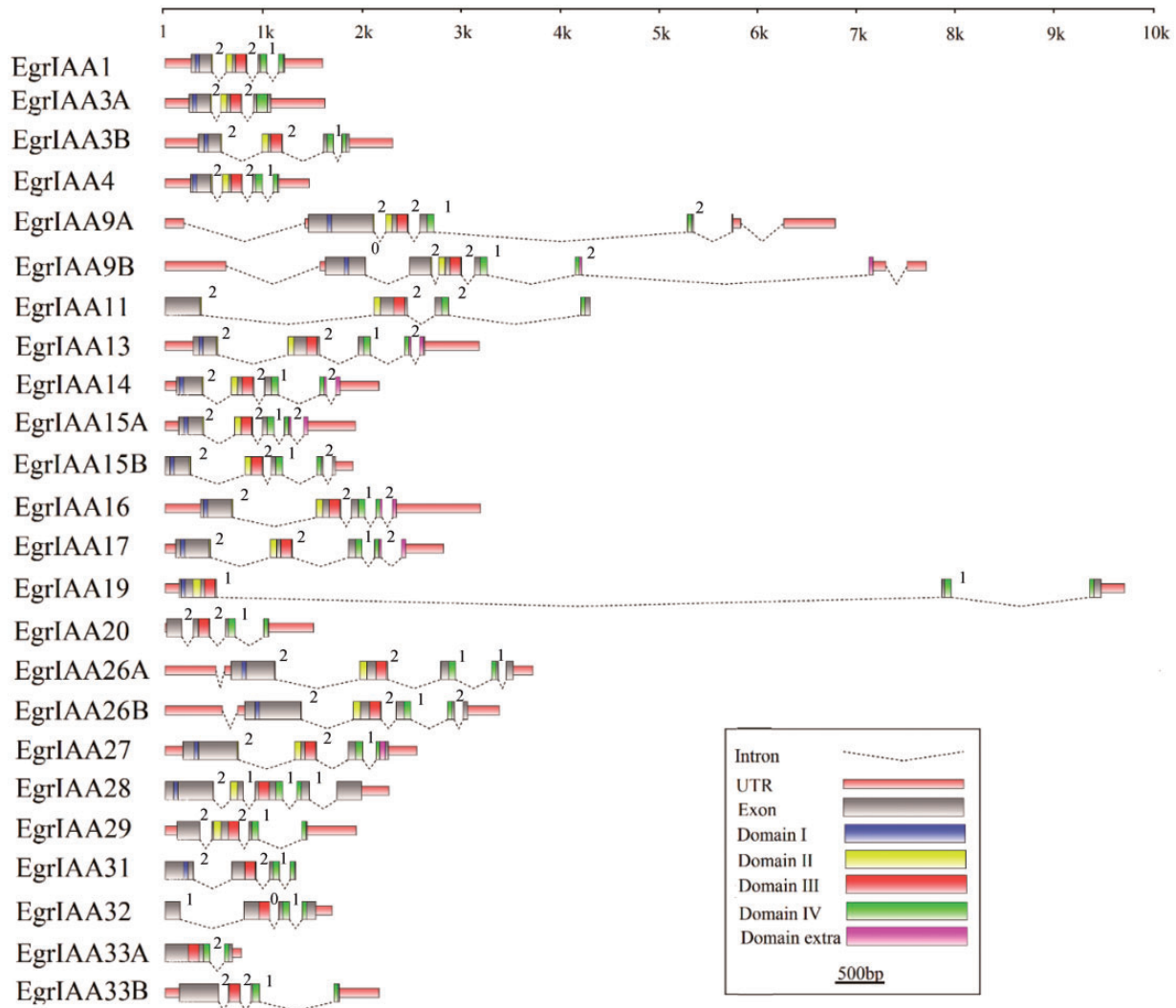
III and IV, similar to those found in ARFs (Reed 2001). Some members, however, had poorly conserved or missing domains I and/or II. For instance, the consensus sequence T/LELR/LGLPG in domain I is not well conserved in *EgrIAA11*, 20 and 33B, and is absent in *EgrIAA 29*, 32 and 33A (**Fig. 1**; **Supplementary Fig. S4**), and the conserved degron sequence VGWPP in domain II is not well conserved in *EgrIAA31* (QDWPP) and is missing in *EgrIAA20*, 32, 33A and 33B (**Fig. 1**; **Supplementary Fig. S4**). Rapid degradation of *Aux/IAA* proteins is essential for auxin signaling, and an amino acid substitution in the degron sequence was shown to cause *Aux/IAA* protein accumulation leading to auxin response defects in Arabidopsis (Woodward and Bartel 2005). Therefore, the *Aux/IAA* proteins with either no degron sequence or a degenerate sequence are likely to be more stable, and their molecular properties may be distinct from those of canonical *Aux/IAA* proteins.

Two types of putative nuclear localization signal (NLS) were detected in most of the *E. grandis* *Aux/IAA* proteins. The first has a bipartite structure comprising a conserved basic doublet, KR, between domains I and II and is associated with the presence of basic amino acids in domain II (**Supplementary Fig. S4**). The second one is a basic residue-rich region located in domain IV that resembles SV40-type NLSs (**Supplementary Fig. S4**). Most of the *EgrIAA* proteins contain both types of NLS and are therefore most probably targeted to the nucleus, consistent with their transcriptional activity. Family members such as *EgrIAA29* and 33B, however, contain a degenerate SV40-type NLS, whereas *EgrIAA20*, *EgrIAA32*, *EgrIAA33A* and *EgrIAA33B* lack the bipartite NLS. Using transient expression assays with tomato *Aux/IAA*, Audran-Delalande et al. (2012) have shown that despite having a degenerate NLS, *Sl-IAA29* protein was specifically targeted to the nucleus. In contrast, in the absence of the bipartite structure, for instance in *Sl-IAA32*, the accumulation of the protein was not restricted to the nuclear compartment (Audran-Delalande et al. 2012).

### Comparative phylogenetics and chromosomal distribution of *EgrIAA* genes

The *Aux/IAA* family in *E. grandis* with only 24 members is slightly contracted compared with the 29 members in Arabidopsis and 35 in *Populus*. Its size is quite similar to that of tomato, which contains only 25 members (**Table 2**). To investigate the phylogenetic relationships of *Aux/IAA* family members in various species, we constructed a phylogenetic tree by using the predicted full-length amino acid sequences of *Aux/IAA* from *Eucalyptus*, Arabidopsis and *Populus*. The *Aux/IAA* members fell into 11 groups named A–K (**Fig. 2**); and *EgrIAAs* were distributed among all groups and had representatives even within the four groups of non-canonical *Aux/IAAs* (groups H, I, J and K) lacking domain I and/or II (Overvoorde et al. 2005). When compared with Arabidopsis and *Populus*, *Eucalyptus* had fewer members in each group except in group E, which notably contains two *Eucalyptus* members (*EgrIAA15A* and 15B).





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

In silico mapping of the gene loci showed that the 24 *EgrIAA* genes and the two truncated genes were unevenly distributed among six of the 11 *E. grandis* chromosomes, with 1–8 genes per chromosome; chromosomes 1, 4, 5, 7 and 9 were devoid of *Aux/IAA* genes (Table 1; Supplementary Fig. S5). Similarly, in *Populus*, the *PtrIAA* genes were present on only 10 of the 19 chromosomes, with 1–5 genes per chromosome (Kalluri et al. 2007), whereas *Arabidopsis Aux/IAA* genes were scattered on all five chromosomes (Overvoorde et al. 2005). Notably, one cluster of tandem duplicated *EgrIAA* genes was detected on chromosome 3: *EgrIAA29*, *29A* and *29B* were located within a 45 kb fragment on chromosome 3 and resulted from a recent tandem duplication event, which led to two truncated versions and probably inactive versions of *EgrIAA29* (Supplementary Fig. S3). Three pairs of *EgrIAA* genes were found to result from segmental duplications (Myburg et al. 2014) although they are located very close to each other (within a distance of <25 kb) on chromosomes 8 (*EgrIAA4/16* and *EgrIAA3A/14*) and

11 (*EgrIAA1/17*) (Supplementary Fig. S5). These six genes are phylogenetically related, the more distant genes encoding protein sequences with 61% identical residues. Members of protein pairs have identity levels higher than 63% (Supplementary Table S1), although each member of a pair belongs to a distinct phylogenetic group. *EgrIAA1*, *3A* and *4* belong to group A, whereas *EgrIAA17*, *14* and *16* belong to group C (Supplementary Figs. S2, S5). Altogether, these data suggest that one ancestor gene underwent tandem duplication to give a first pair of *EgrIAA* genes that subsequently underwent segmental duplication events to generate the other two pairs. It seems that tandem duplication events occurred prior to the chromosomal block duplication in the *Eucalyptus Aux/IAA* family, in a similar way to the events proposed for *Arabidopsis* (Remington et al. 2004).

As duplication and alternative splicing are the two main mechanisms involved in diversification of function within gene families, we performed an in silico survey of the alternative transcripts predicted by the *E. grandis* genome JGI assembly

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

Gene name <sup>a</sup>	No. of predicted alternative transcripts	Accession No. <sup>b</sup>	Chromosome <sup>c</sup>	Genome location <sup>d</sup>	ORF <sup>e</sup> (bp)	Deduced polypeptide <sup>f</sup>			Exon No.
						Length (aa)	Mol. wt (kDa)	pI	
EgrIAA1	2	Eucgr.K03314.1	11	42,167,198–42,168,788	570	189	21	5.39	4
EgrIAA3A	2	Eucgr.H03171.1	8	46,743,317–46,744,932	612	203	23	6.02	3
EgrIAA3B	1	Eucgr.H00216.1	8	2,425,940–2,428,241	615	204	23	5.58	4
EgrIAA4	1	Eucgr.H04336.1	8	62,350,663–62,352,118	576	191	22	5.63	4
EgrIAA9A	2	Eucgr.H02407.1	8	32,905,397–32,912,178	1107	368	39	6.05	5
EgrIAA9B	4	Eucgr.F02172.1	6	29,479,509–29,487,212	1080	359	39	6.18	6
EgrIAA11 <sup>g</sup>	1	Eucgr.K01426.1	11	17,332,089–17,336,565	930	309	33	5.88	4
EgrIAA13	1	Eucgr.H02914.1	8	42,398,802–42,401,976	801	266	29	8.8	5
EgrIAA14	1	Eucgr.H03170.1	8	46,720,521–46,722,685	732	243	26	6.45	5
EgrIAA15A	1	Eucgr.J03016.1	10	37,360,412–37,362,334	648	215	23	5.76	5
EgrIAA15B <sup>g</sup>	1	Eucgr.C01083.1	3	17,083,069–17,086,237	645	214	23	7.54	5
EgrIAA16	1	Eucgr.H04335.1	8	62,338,061–62,341,249	813	270	29	8.28	5
EgrIAA17 <sup>g</sup>	1	Eucgr.K03313.1	11	42,152,995–42,155,812	807	268	29	6.06	5
EgrIAA19	1	Eucgr.F02578.1	6	35,331,299–35,341,010	585	194	22	8.33	3
EgrIAA20	1	Eucgr.K00561.1	11	6,262,537–6,264,036	465	154	17	4.99	4
EgrIAA26A	3	Eucgr.F03080.1	6	40,229,065–40,232,796	990	329	36	8.38	5
EgrIAA26B	2	Eucgr.J02934.1	10	36,614,769–36,618,150	1113	370	41	8.88	5
EgrIAA27	3	Eucgr.F03050.1	6	39,988,867–39,991,412	1050	349	37	6.27	4
EgrIAA28 <sup>g</sup>	2	Eucgr.C02984.1	3	56,397,825–56,400,087	1044	347	38	6.72	5
EgrIAA29	1	Eucgr.C01734.1	3	29,257,063–29,258,993	651	216	25	6.6	4
EgrIAA[T]29A <sup>h</sup>	1	Eucgr.C01731.1	3	29,214,163–29,215,275	390	129	14	9.87	2
EgrIAA[T]29B <sup>h</sup>	1	Eucgr.C01732.1	3	29,230,438–29,231,106	369	122	14	9.45	2
EgrIAA31 <sup>g</sup>	1	Eucgr.H04141.1	8	59,329,525–59,330,844	687	228	25	9.51	4
EgrIAA32	1	Eucgr.B02853.1	2	52,269,360–52,271,045	654	217	25	4.83	4
EgrIAA33A	1	Eucgr.C02329.1	3	43,423,640–43,424,409	531	176	19	8.71	2
EgrIAA33B	1	Eucgr.C01373.1	3	22,002,554–22,005,153	651	216	24	4.75	4

<sup>a</sup> Designation given to *E. grandis* in this work.<sup>b</sup> Accession number in phytozome.<sup>c, d</sup> Location of the *EgrIAA* genes on the chromosome.<sup>e</sup> Length of the open reading frame in base pairs.<sup>f</sup> The number of amino acids, molecular weight (kDa) and isoelectric point (pI) of the deduced polypeptides.<sup>g</sup> Using FGENESH+ to complete the complete sequence<sup>h</sup> [T] Truncated gene.**Table 2** Number of Aux/IAA family gene members in angiosperm species

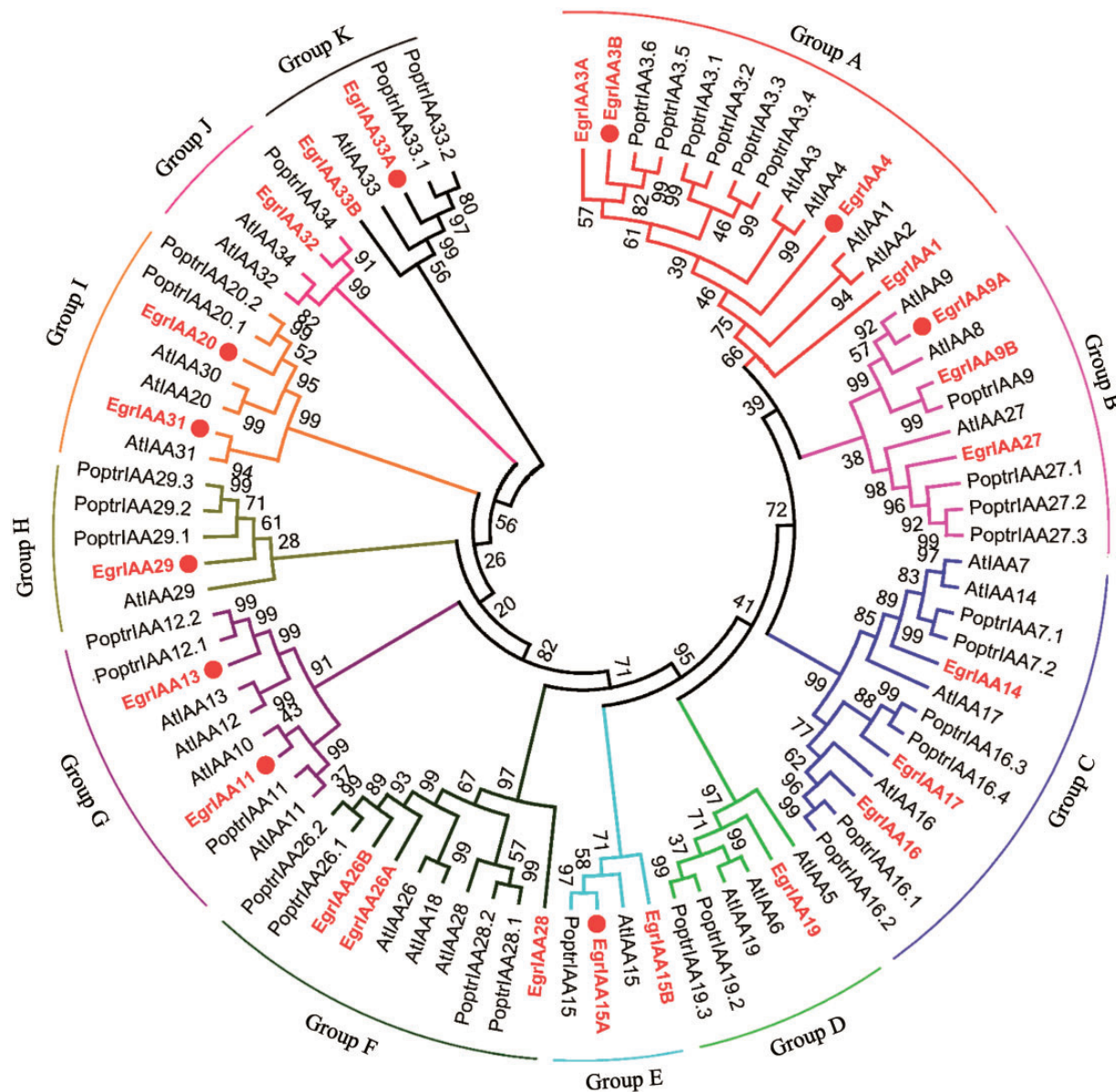
	Aux/IAA	Reference
<i>Eucalyptus grandis</i>	24	This study
<i>Solanum lycopersicon</i>	25	Audran-Delalande et al. (2012)
<i>Arabidopsis thaliana</i>	29	Overvoorde et al. (2005)
<i>Cucumis sativus</i>	29	Gan et al. (2013)
<i>Populus trichocarpa</i>	35	Kalluri et al. (2007)
<i>Zea mays</i>	34	Ludwig et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

v1.0, annotation v1.1 (<http://www.phytozome.net/eucalyptus>). Among the 24 *EgrIAA* genes, eight are predicted to have more than one splice variant (Table 1). This proportion of alternative transcripts is similar to that found in *Arabidopsis*. In *Eucalyptus*,

most of the alternative transcripts arise from members of groups B and F, whereas in *Arabidopsis* the majority of them belong to groups B and G (Fig. 2; Supplementary Table S2). Some alternative transcripts are lacking domains III and/or IV (Fig. 1; Supplementary Fig. S6), probably contributing to shape the functional diversity of the family.

### Expression profiling of *EgrIAA* genes in different tissues/organs and in response to environmental cues

To gain insights into the developmental patterns of expression of *EgrIAA* genes, we assessed their transcript levels by qRT-PCR in 13 different organs and tissues. The 24 *EgrIAA* genes could be detected in all the tissues tested, and most of them were preferentially expressed in certain tissues (Fig. 3). The relative transcript accumulation of all the *EgrIAA* genes is presented as a heat



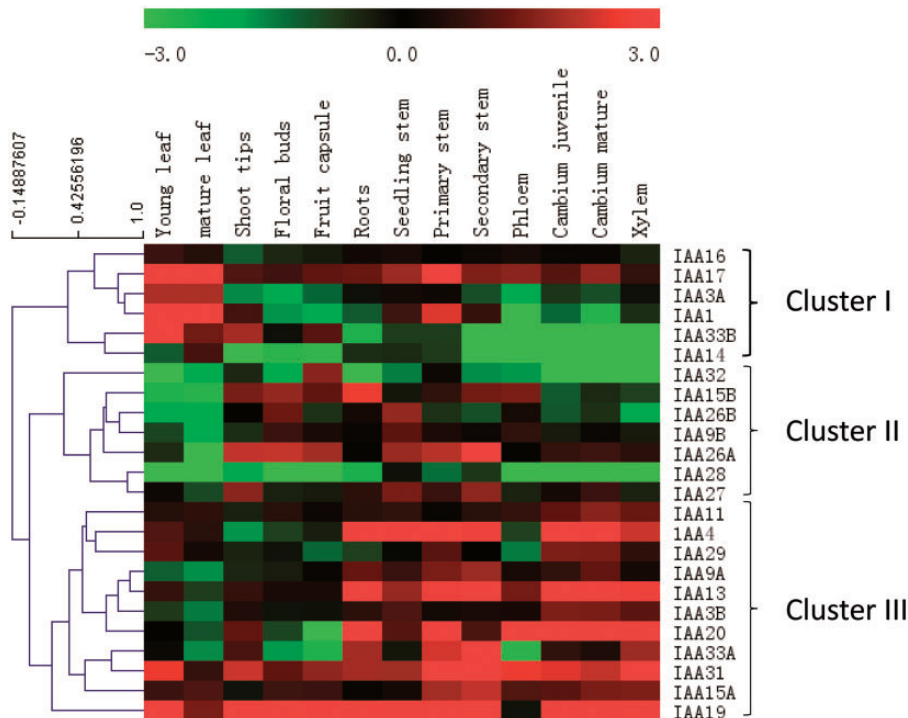
**Fig. 2** Phylogenetic analysis of *Eucalyptus*, *Arabidopsis* and *Populus* Aux/IAA proteins. Full-length protein sequences were aligned by using the Clustal\_X program. The phylogenetic tree was constructed by using the MEGA5 program and the Neighbor–Joining method with predicted Aux/IAA proteins. Bootstrap support is indicated at each node. Each Aux/IAA group (A–K) is indicated by a specific color. *EgrIAA*s are noted in red and bold. *EgrIAA* genes that are preferentially expressed in vascular tissue (cluster III, Fig. 3) are marked with a red solid circle.

map, and hierarchical clustering allowed us to group all the expression patterns into three distinct clusters (Fig. 3). Many members of the three clusters were differentially expressed in young and mature leaves. Most of them were more highly expressed in young leaves; only *EgrIAA14* was more highly expressed in mature leaves. Members of both cluster I and II were most highly expressed in non-vascular tissues, but they diverged regarding their differential expression in leaves: members of cluster I were highly expressed whereas those of cluster II were expressed at very low levels in leaves. Cluster III was the largest cluster, containing 11 members that were preferentially expressed in vascular tissues (phloem, cambium and xylem).

Within cluster III, *EgrIAA13*, 20 and 31 were relatively highly expressed in all three vascular tissues, whereas *EgrIAA4* and

33A were expressed at higher levels in cambium and xylem than in phloem. *EgrIAA29* was preferentially expressed in cambium as compared with phloem and xylem. Most of the members of cluster III were expressed at higher levels in juvenile xylem than in mature xylem (Fig. 4A), whereas no obvious differences in their expression patterns were detected between juvenile and mature cambium (Fig. 3). Of particular note, *EgrIAA4* was the only gene that was more highly expressed in mature xylem than in juvenile xylem (Fig. 4A). Most *EgrIAA* genes from cluster III were responsive to mechanical stress due to bending (Fig. 4B). Four *EgrIAA* genes were down-regulated (*EgrIAA4*, 11, 13 and 20) and two (*EgrIAA31* and 33A) were up-regulated in tension wood as compared with straight wood (Fig. 4B). Interestingly, *EgrIAA11* and 20





**Fig. 3** Expression profiles of *EgrIAA* genes in various organs and tissues. The hierarchically clustered heat map was constructed by using the log<sub>2</sub> ratios of relative expression values normalized to a control sample (in vitro plantlets) for 24 *EgrIAA* genes (indicated on the right) in 13 tissues and organs (indicated at the top). The genes belonging to cluster III are preferentially expressed in vascular tissues (phloem, cambium and/or xylem) and in stems and roots.

were down-regulated in both tension and opposite woods, whereas *EgrIAA33A* was up-regulated in these tissues (Fig. 4B).

### Candidate *EgrIAA* genes potentially involved in wood formation

To identify the best candidate gene(s) potentially involved in wood formation for further functional characterization in planta, we defined several criteria: transcript abundance in vascular tissues; high expression in cambium and/or xylem and low expression in phloem; responsiveness to bending (given that the response to bending involves dramatic changes in both xylem development and SCW structure and composition); and differential expression in mature and juvenile xylem, which are known to exhibit different biochemical and physical properties. The combination of all these criteria when presented as a Venn diagram (Supplementary Fig. S7) identified *EgrIAA4* and *EgrIAA33A* as the most suitable candidates matching the first three and main criteria. Ultimately, *EgrIAA4* was selected for further functional characterization because it was the only gene that was more highly expressed in mature than in juvenile xylem (Supplementary Fig. S7).

### Nuclear localization and transcriptional activity of *EgrIAA4*

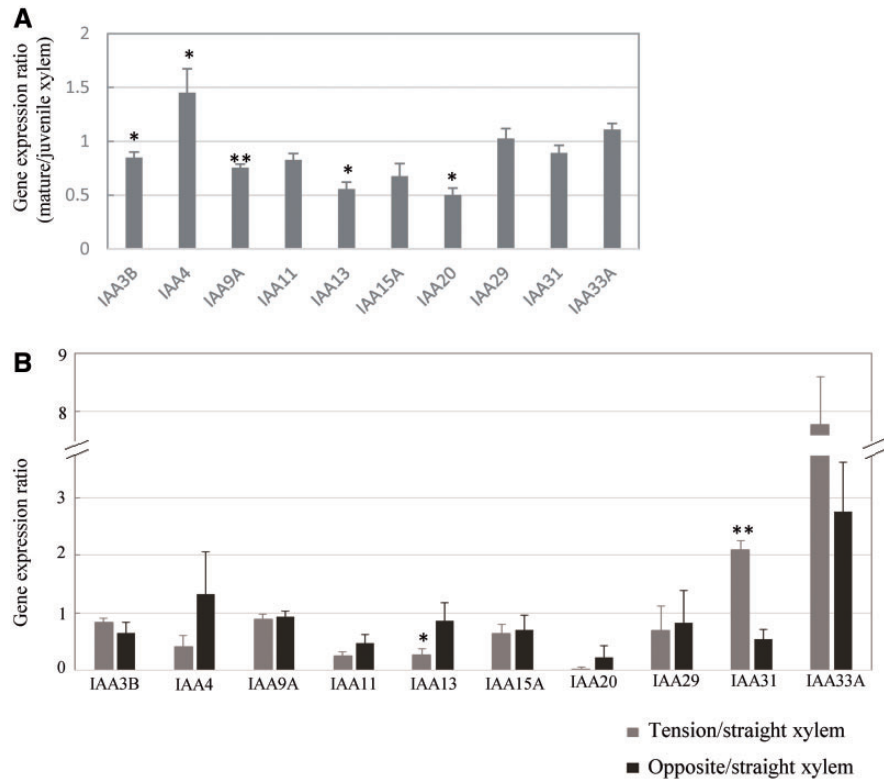
When transiently expressed as a green fluorescent protein (GFP) fusion protein in BY-2 tobacco protoplasts, *EgrIAA4*

protein was located exclusively in the nucleus (Fig. 5A) as predicted by the presence of two well-conserved NLSs (Supplementary Fig. S4). The impact of *EgrIAA4* on the transcriptional activity of target genes was assessed in tobacco protoplasts co-transfected with an effector construct expressing the full-length coding sequence of *EgrIAA4* under the 35S *Cauliflower mosaic virus* (CaMV) promoter and a reporter construct carrying DR5::GFP, an auxin-responsive promoter fused to the GFP coding sequence. The DR5::GFP reporter construct is commonly used to assess auxin-dependent transcriptional activity in planta (Ulmasov et al. 1999, Audran-Delalande et al. 2012). Basal DR5-driven GFP expression was low, but was induced up to 9-fold by exogenous auxin treatment (Fig. 5B). Co-transfection with the effector plasmid *EgrIAA4* resulted in strong repression (87%) of this auxin-induced expression of the reporter gene (Fig. 5B), indicating that *EgrIAA4* is able to mediate an auxin response in vivo and that it functions as a strong transcriptional repressor.

### Overexpression of *EgrIAA4m* affects plant growth and development

To gain insights into the role of *EgrIAA4*, we overexpressed in Arabidopsis a mutated version of the gene, referred to as *EgrIAA4m* (see the Materials and Methods; Supplementary Fig. S8A), encoding a stabilized form of the protein. We hypothesized that the mutation introduced in the degron domain would prevent auxin-mediated *EgrIAA4* protein degradation. Thus *EgrIAA4m* was cloned under the control





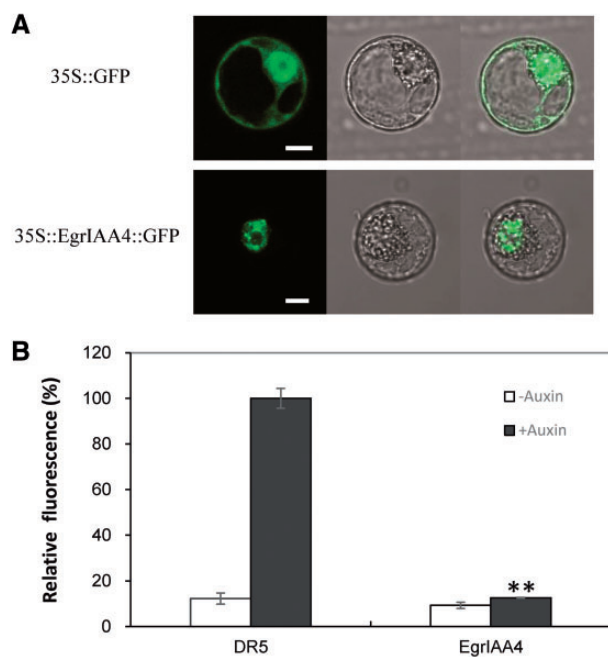
**Fig. 4** Effects of developmental and environmental cues on the transcript levels of 10 *EgrIAA* genes preferentially expressed in vascular tissues. Ratios of relative mRNA abundance in (A) mature and juvenile xylem, (B) tension vs. straight xylem, and opposite vs. straight xylem. Each relative mRNA abundance was normalized to a control sample (in vitro plantlets). Error bars indicate the SE of mean expression values from three independent experiments. Asterisks indicate values found to be significantly different (Student's *t*-test, *P*-value) \**P* < 0.05; \*\**P* < 0.01.

of the 35SCaMV promoter into the Gateway *pFAST-G02* expression vector and transformed into *Arabidopsis thaliana* Col-0. This *pFAST-G02* vector contains an *OLE1-GFP* selection marker that allows direct screening of transformed seeds that are fluorescent under UV light (Shimada et al. 2010). Of 10 seeds sown, six germinated giving independent *EgrIAA4m* transformants, all of similar and dramatically reduced size as compared with controls. We selected two phenotypically representative transformants (*EgrIAA4m\_1.3* and *EgrIAA4m\_2.3*) overexpressing *EgrIAA4m* (Supplementary Fig. S8B) for further characterization. *T*<sub>2</sub> seeds from both lines germinated and half of the seedlings had the drastic phenotype, with leaf rosettes of only about 1 cm diameter and very small, curled-down leaves that turned yellow or brown just before the plants died without flowering. (Supplementary Fig. S8C). The surviving transgenic plants (*T*<sub>2</sub>) had several distinctive phenotypes in their aerial parts (Fig. 6A, B). During the seedling stage, the cotyledons were curled up, whereas the leaves were curled down and had very short petioles (Fig. 6A). During the vegetative stage, the *EgrIAA4m* transgenic plants from both lines were significantly shorter and had smaller leaf rosettes (Fig. 6B), and the inflorescence stems were thinner than those of wild-type plants. In addition, *EgrIAA4m* plants had smaller siliques and much lower fertility (i.e. no seeds or <20 seeds per plant) than wild-type plants.

### *EgrIAA4m* affects root development and gravitropism

To examine the root phenotype of the transgenic plants, seeds collected from *T*<sub>1</sub> plants were grown in vitro on 1/2 Murashige and Skoog (MS) medium and the primary root length and the number of lateral roots were compared 10 d after germination with those of wild-type plants. The primary roots of *EgrIAA4m* seedlings were significantly shorter (around 15 mm) than those of wild-type seedlings (around 50 mm long; Fig. 6C). In addition, no lateral roots were observed in *EgrIAA4m* transgenic seedlings, whereas wild-type seedlings had well-developed lateral roots of 6.9 mm (Fig. 6D). Seventeen days after germination, there was still no lateral root emergence in *EgrIAA4m* transgenic plants. These findings indicate that *EgrIAA4m* overexpression inhibits both primary root elongation and lateral root emergence.

The gravitropic response is a typical auxin-related phenotype, and several *Aux/IAA* *Arabidopsis* mutants, such as *axr5/IAA1*, *axr2/IAA7*, *slr/IAA14*, *axr3/IAA17* and *msg2/IAA19*, showed agravitropism in roots and/or hypocotyls (Watahiki and Yamamoto 1997, Rouse et al. 1998, Nagpal et al. 2000, Fukaki et al. 2002, Yang et al. 2004). To assess the gravitropic response of *EgrIAA4m* transgenic plants, we grew them on vertically oriented 1/2 MS plates. Ten days after germination, the hypocotyls and roots of the *EgrIAA4m* transgenic plants



**Fig. 5** Subcellular localization and repressor activity of EgrIAA4 protein on a synthetic DR5 promoter. (A) Subcellular localization of EgrIAA4–GFP fusion protein in BY-2 tobacco protoplasts. The merged images of green fluorescence (left panels) and the corresponding bright-field (middle panels) are shown in the right panels. Scale bar = 10  $\mu$ m. (B) Repressor activity of EgrIAA4 protein on a synthetic DR5 promoter. Effector and reporter constructs were co-expressed in tobacco protoplasts in the presence or absence of a synthetic auxin (50  $\mu$ M 2,4-D). A mock effector construct (empty vector) was used as control. In each experiment, protoplast transformations were performed in independent biological triplicates. Three independent experiments were performed and similar results were obtained; the figure indicates the data from one experiment. Error bars represent SEs of mean fluorescence. Significant statistical differences (Student's *t*-test,  $P < 0.001$ ) from the control are marked with \*\*.

grew in random directions (agravitropically), whereas the wild-type seedlings all grew vertically (Fig. 6E). We then returned the plates to the horizontal and checked the roots' response. After 48 h, the wild-type Arabidopsis roots had changed their growth direction by 90°, whereas the *EgrIAA4m* transgenic plants did not respond to the change in orientation (data not shown), indicating that the transgenic roots had lost their ability to perceive gravity.

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

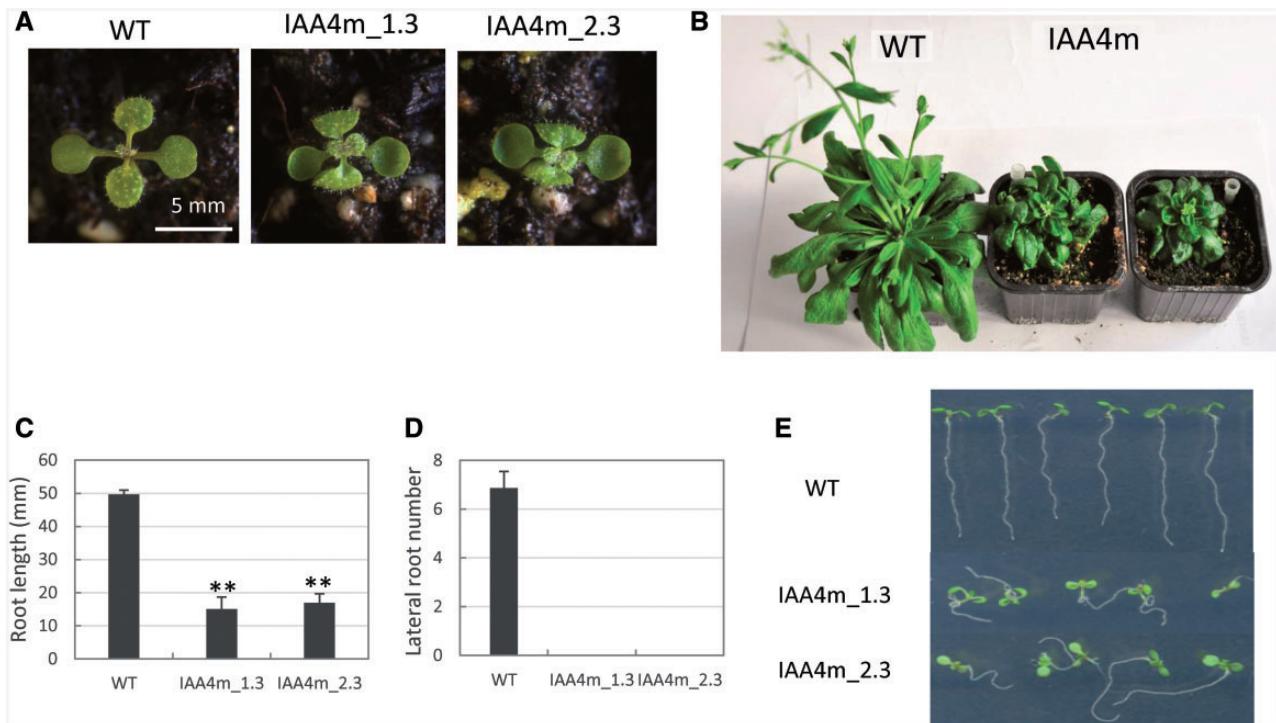
Next, we investigated the impact of *EgrIAA4m* overexpression on xylem and interfascicular fiber development in Arabidopsis inflorescence stems (Fig. 7). The lignification patterns were obtained by staining stem cross-sections from plants grown for 37 d (Fig. 7A–D) or 47 d (Fig. 7E, F) with phloroglucinol-HCl. When compared with 37-day-old wild-type plants (Fig. 7A, C), the proportion of lignified tissues stained red by phloroglucinol was dramatically reduced in transgenic lines (Fig. 7B,

D) suggesting that, at the same age, the activity of both fascicular and interfascicular cambium was greatly reduced and/or delayed in the transgenic plants (Fig. 7A–F). Closer examination revealed the virtual absence of lignified interfascicular fibers in 37-day-old *EgrIAA4m* transgenic Arabidopsis (Fig. 7D) as compared with control (Fig. 7C). A very weak and discontinuous light reddish staining appeared in the interfascicular zone of 47-day-old transgenic plants (Fig. 7F), however, suggesting that very limited lignification had occurred. In addition, the xylem fibers in the vascular bundle regions were completely absent and the xylem vessels were much smaller than in control plants (Fig. 7C–F). These results indicate that overexpression of *EgrIAA4m* strongly and negatively affected xylem bundle and interfascicular fiber development.

## Discussion

The *Aux/IAA* family in *E. grandis* contains 24 members and therefore is slightly contracted as compared with most angiosperm *Aux/IAA* families studied hitherto (Table 2). With 25 members, the gene family in tomato is the closest to *Eucalyptus* but it lacks non-canonical members in phylogenetic group I (Audran-Delalande et al. 2012), which we show here are present in *Eucalyptus*. All 11 phylogenetic groups defined in Arabidopsis are present in *Eucalyptus*, but *Eucalyptus* has fewer members in each group, except in group E, which has two members whereas Arabidopsis and *Populus* have only one member each. In Arabidopsis, transcripts of *AtIAA15*, the unique member of group E, have never been detected, suggesting that it may be a pseudogene (Remington et al. 2004). In contrast, in *E. grandis*, the transcripts of both *EgrIAA15A* and *15B* were detected by qRT–PCR, indicating they are probably functional genes. Whereas in Arabidopsis and *Populus* members of the *Aux/IAA* family arose predominantly through large-scale genomic duplication events (Remington et al. 2004, Kalluri et al. 2007), *Aux/IAA* family members in the *E. grandis* genome show evidence of only very few segmental duplication events. One cluster of recent tandem duplications was found on chromosome 3, but duplication of *EgrIAA29* led to two truncated genes lacking both domains III and IV. To the best of our knowledge, such a structure had never been reported in other species and these truncated genes are most probably not functional.

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



**Fig. 6** Phenotypic characterization of *EgrIAA4m* transgenic plants. (A) Nine-day-old plantlets of wild-type and two representative *EgrIAA4m* lines plantlets. (B) 45-day-old wild-type and *EgrIAA4m* transgenic plants. (C, D) Primary root length and lateral root numbers in 10-day-old wild-type and *EgrIAA4m* transgenic seedlings grown on 1/2 MS medium. Error bars represent SEs of the means of primary root length ( $n > 10$ ). Significant statistical differences (Student's *t*-test,  $P < 0.001$ ) from the control are marked with \*\*. (E) Three-day-old seedlings grown on vertical plates with 1/2 MS medium.

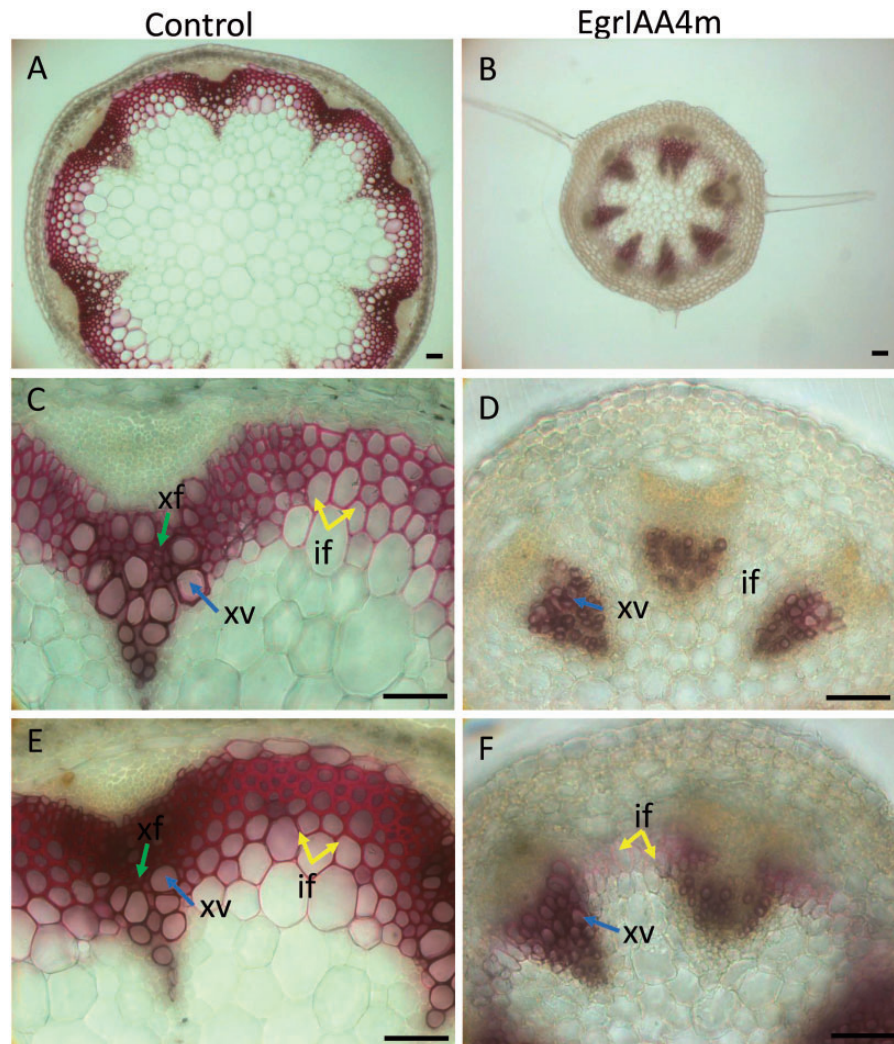
interaction with ARF proteins, a corresponding dosage relationship is probably needed (Remington et al. 2004).

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

Most of the *EgrIAA* genes in cluster III responded to bending stress, but their behavior seems quite different from that of their potential orthologs in *Populus* (Moyle et al. 2002). For example, *EgrIAA20* was down-regulated in tension wood, whereas its *Populus* orthologs *PttIAA3* and 4 were not affected (Moyle et al. 2002). In contrast, *EgrIAA9A* was not affected by long-term bending, whereas its *Populus* ortholog *PttIAA2* was down-regulated in tension wood (Moyle et al. 2002). In a previous study, we found that the transcript level of *EgrIAA9A* increased sharply between 6 and 24 h and then decreased between 24 h and 1 week to reach a level similar to that of the control (Paux et al. 2005). Thus Aux/IAA genes respond to bending in a time-dependent way, which may explain the apparent discrepancies between the responses of *Populus* and *Eucalyptus* Aux/IAA orthologs. Indeed, in our study, the plants were exposed to bending stress for 3 weeks, whereas Moyle et al. (2002) analyzed an early response to bending (from 30 min to 11 h). Indeed, our results are more consistent with those of Anderson-Gunneras et al. (2006) who analyzed *Populus* gene expression after 3 weeks bending. We found that *EgrIAA4*, 11 and 20 were all down-regulated in tension wood, as were their *Populus* orthologs (Anderson-Gunneras et al. 2006).

*EgrIAA4*, which belongs to subgroup A, was considered to be the best candidate to regulate cambium activity and wood formation based on its expression profile, and was chosen for functional characterization in planta. *EgrIAA4* is the ortholog of the *Arabidopsis* gene pair *AtIAA3* and *AtIAA4*.





**Fig. 7** Histochemical analysis of basal stem cross-sections from transgenic *Arabidopsis* plants expressing *Egr1AA4m*. Sections of wild-type and *Egr1AA4m* transgenic plants were stained with phloroglucinol-HCl. if, interfascicular fiber; xf, xylary fiber; xv, xylem vessel. Scale bar = 50  $\mu\text{m}$ .

Only the function of *AtIAA3* has been studied through gain-of-function experiments. The *AtIAA3* mutation *shy2* produces plants with shorter hypocotyls, fewer lateral roots and a slower gravitropic response than the wild type (Reed et al. 1998, Tian and Reed 1999). *AtIAA3* was shown recently to be part of the auxin signaling module (SLR/IAA14–ARF7–ARF19 and SHY2/IAA3–ARFs) regulating lateral root formation (Goh et al. 2012). Likewise the down-regulation of *Sl-IAA3*, the ortholog of *AtIAA3* in tomato, leads to reduced growth of primary roots and inhibition of lateral roots (Chaabouni et al. 2009). Overexpression of an auxin-insensitive version of *Egr1AA4* in *Arabidopsis* led to very similar but more severe aberrant auxin response phenotypes. The gravitropic response was lost completely, and growth and development were severely affected, with much reduced height and diameter of the inflorescence stem, reduced leaf size and fertility, and absence of lateral roots. Notably, *Egr1AA4m* overexpression also led to phenotypic alterations of the vascular system not reported yet for its *Arabidopsis* ortholog. The development of both xylem and interfascicular fibers was dramatically

delayed and reduced, suggesting reduced cambium activity in response to the altered auxin responsiveness induced by *Egr1AA4m* expression. The lignified secondary walls of the interfascicular fibers appeared very late, whereas that of the xylary fibers was virtually undetectable. This phenotype, together with the expression pattern of *Egr1AA4*, strongly suggests that auxin signaling and *Egr1AA4* are important in fiber development and SCW deposition.

## Materials and Methods

### Identification of Aux/IAA genes in *Eucalyptus*

We used 29 *Arabidopsis* Aux/IAA protein sequences and BLASTP to search for related proteins predicted by the *E. grandis* genome (JGI assembly v1.0, annotation v1.1, <http://www.phytozome.net/eucalyptus>), which identified 55 potential *E. grandis* Aux/IAA proteins. Then, we used the Pfam database (<http://pfam.sanger.ac.uk/search>) and NCBI conserved domain database (<http://www.ncbi.nlm.nih.gov/cdd>) web server to examine the conserved domains (Finn et al. 2010, Marchler-Bauer et al. 2011). The incomplete gene models were completed by FGENESH+ (<http://linux1.softberry.com>), and redundant

and invalid gene models were removed. The corrected *E. grandis* Aux/IAA protein sequences were used as query sequences in two subsequent searches: first, a BLASTP search of the *E. grandis* proteome for exhaustive identification of divergent *E. grandis* gene family members, and, secondly, TBLASTN searches of the *E. grandis* genome for any possible unpredicted genes. In total, 26 *EgrIAA* genes were identified in the *E. grandis* genome (*E. grandis* genome V1.1, May 2012). For validation, we also used *Populus* Aux/IAA proteins as queries in the complete search procedure described above, and we obtained exactly the same final 26 *EgrAux/IAA* genes. Gene information on chromosomal location was retrieved from the *E. grandis* genome browser (<http://www.phytozome.net/eucalyptus>) with manual curation, and we mapped their loci using MapChart 2.2 (Voorrips 2002). Basic physical and chemical parameters of Aux/IAA proteins were calculated by means of the online ProtParam tool (<http://web.expasy.org/protparam/>).

### Sequence, gene structure and phylogenetic analysis

Conserved protein motifs of *EgrIAA* were determined by use of MEME-MAST programs (<http://meme.sdsc.edu/meme>) (Bailey et al. 2009). The exon–intron structures were extracted from Phytozome with manual curation and visualized by Fancy Gene v1.4 (<http://bio.ieu.eu/fancygene/>). Multiple protein sequence alignments was performed by using the Clustal\_X2 program (Version 2.0.11). All predicted protein sequences were used for phylogenetic analysis, and the phylogenetic trees were constructed with the MEGA5 program by using the Neighbor–Joining method with 1,000 bootstrap.

### Plant materials and growth conditions

The provenance and preparation of all *Eucalyptus* organs and tissues were as described in Cassan-Wang et al. (2012). *Arabidopsis thaliana* ecotype Col-0 plants were grown in a growth chamber under the conditions: 16 h day : 8 h night for long days, 22:20°C day:night temperature, 70% relative humidity, 200  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$  light intensity (intense luminosity). The plants were watered every 2 d and fertilized weekly. Seeds were surface-sterilized for 1 min in 70% ethanol, 10 min in 25% bleach, rinsed five times in sterile water and plated on MS medium containing 1.0% sucrose solidified with 1% agar.

### RNA isolation, cDNA synthesis and qRT–PCR

Total RNAs were extracted from 100–200 mg of frozen material as described previously (Southerton et al. 1998), and genomic DNA contamination was removed by using the Turbo DNA-free™ kit (Ambion). RNA concentrations and purity were determined by using a NanoDrop spectrophotometer ND-1000 (Thermo Scientific), and integrity was assessed by using the Agilent 2100 Bioanalyzer. Only samples with RNA integrity number (RIN) >7 were used for reverse transcription. cDNA was reverse transcribed from 2  $\mu\text{g}$  of total RNA using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems).

Primers were designed by using the software QuantPrime (qPCR primer design tool: <http://www.quantprime.de/>; Arvidsson et al. 2008); the sequences are shown in Supplementary Table S4. Oligonucleotides were synthesized by Sigma Life Science. qRT–PCR was performed by the Genotoul service in Toulouse (<http://genomique.genotoul.fr/>) using BioMark® 96:96 Dynamic Array integrated fluidic circuits (Fluidigm Corporation) as described previously (Cassan-Wang et al. 2012). Only primers that produced linear amplification and qPCR products with single-peak melting curves were used for further analysis. The efficiency of each pair of primers was determined from the data of the amplification Ct value plot with a serial dilution of mixture cDNA and the equation  $E = 10^{(-1/\text{slope})} - 1$ . The  $E^{-\Delta\Delta\text{Ct}}$  method was used to calculate relative mRNA fold change compared with a control sample by using the formula  $(E_{\text{target}})^{\Delta\text{Ct}_{\text{target}}(\text{control-sample})} / (E_{\text{reference}})^{\Delta\text{Ct}_{\text{reference}}(\text{control-sample})}$  (Pfaffl 2001), and five reference genes (IDH, PP2A1, PP2A3, EF-1a and SAND, Supplementary Table S4) were used for data normalization. In vitro plantlets were used as control samples.

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

The *EgrIAA4* gene was amplified by PCR using Phusion Taq (Thermo) and a gene-specific primer pair: 5'-CACCATGGCAGCTCAAGGAGAGGAT-3' and 5'-AACCTCTGATGACCCCTTTCATGATT-3' based on sequence prediction from the *E. grandis* genome v.1.1 (Eucgr.H04336.1). To investigate the function of this gene, we created a mutation in the auxin degradation domain by changing the codon for amino acid residue 80 from one that encodes proline to one that encodes serine (P-to-S) in domain II by using overlap PCR. The overlap primers were 5'-ATCGGACCGGACTCCACCCAC GACTTGTGCCTTA-3' and 5'-CGTGGGGTGGAGTCCGGTCCGATCCTACCG AAA-3'. The underlined sequences indicate the overlap region and the bold bases are the mutated nucleotides. The *EgrIAA4m* fragment was cloned into the *pENTR/D-TOPO* vector (Invitrogen) for sequencing. After sequencing, we recombined the *EgrIAA4m* fragment into the destination vector *pFAST-G02* by using LR Clonase (Invitrogen). The vector *pFAST-EgrIAA4m* was transformed into *Agrobacterium tumefaciens* strain GV3101, and then transformed into *A. thaliana* ecotype Col-0 by using the floral dip method (Clough and Bent 1998).

### Transient gene expression in protoplasts

Protoplasts for transfection were obtained from suspension-cultured tobacco (*Nicotiana tabacum*) BY-2 cells according to the method of Leclercq et al. (2005). They were transfected by a modified polyethylene glycol method as described by Abel and Theologis (1994). To study the nuclear localization of selected Aux/IAAs, 'effector' constructs were created by fusing full-length cDNAs in-frame with the GFP sequence to produce a fusion protein at the C-terminus of GFP and under the control of the 35S<sub>CaMV</sub> promoter in the vector *pK7FWG2.0* (Karimi et al. 2002). Transfected protoplasts were incubated for 16 h at 25°C and examined for GFP fluorescence by using a Leica TCS SP2 laser scanning confocal microscope. Images were obtained with a  $\times 40$  water-immersion objective. For co-transfection assays, the full-length cDNAs encoding selected Aux/IAA proteins were cloned under the control of the 35S<sub>CaMV</sub> promoter in the vector *pGreen*. The 'reporter' constructs used a synthetic auxin-responsive promoter DR5 fused to the GFP reporter gene. Tobacco BY-2 protoplasts were co-transfected with the reporter and effector constructs as described in Audran-Delalande et al. (2012). After 16 h incubation, GFP expression was quantified by flow cytometry (LSR Fortessa, BD Bioscience), and the data were analyzed using BD CellQuest and BD FACS Diva software. Transfection assays were performed with three independent replicates, and 3,000–4,000 protoplasts were gated for each sample. GFP fluorescence corresponds to the average fluorescence intensity of the protoplast population after subtraction of autofluorescence determined with non-transformed protoplasts. 2,4-D (50  $\mu\text{M}$ ) was used for auxin treatment.

### Microscopy analysis

*Arabidopsis* hypocotyls and the basal  $\sim 1$  cm of inflorescence stems were harvested from 37- and 47-day-old plants, and stored in 70% ethanol. Sections were prepared by using a vibratome Leica VT1000 S. Lignin polymers are the characteristic components of SCW and are normally absent from primary cell walls; therefore, we used lignin deposition detection techniques to screen for SCW phenotype. Cross-sections of inflorescence stem and hypocotyl ( $\sim 80$   $\mu\text{m}$  thick) were stained with phloroglucinol-HCl, which stains violet-red specifically the lignin polymer precursors coniferaldehyde and *p*-coumaraldehyde in the SCW. Phloroglucinol-HCl was directly applied on the slide and images were recorded with a CCD camera (Photonic Science; <http://www.photonic-science.co.uk>).

### Supplementary data

Supplementary data are available at PCP online.



## Funding

This work was supported by the Centre National pour la Recherche Scientifique (CNRS); the University Paul Sabatier Toulouse III (UPS); the Agence Nationale de la Recherche (ANR); the French Laboratory of Excellence [project 'TULIP' (ANR-10-LABX-41; ANR-11-IDEX-0002-02)]; the Plant KBBE TreeForJoules [project ANR-2010-KBBE-007-01 (FR) and P-KBBE/AGR\_GPL/0001/2010 (FCT, PT) and the project microEgo (PTDC/AGR-GPL/098179/2008; FCT, PT)]; the China Scholarship Council [a PhD grant to H.Y.]; the Departament d'Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya [a post-doctoral fellowship 'Beatriu de Pinós' to M.S.]; the Fundação para a Ciência e a Tecnologia (FCT) [a research contract from the Ciência 2008 program and a post-doctoral fellowship SFRH/BPD/92207/2013 to J.A.P.P.].

## Acknowledgments

The authors are grateful to C. Dunand and Q. Li for their help with the bioinformatics analysis, to E. Camargo, J.M. Gion and E. Villar (CIRAD, FR), F. Melun and L. Harvengt (FCBA, France), C. Araujo and L. Neves (AltriFlorestal, Portugal) and C. Marques (RAIZ, Portugal) for kindly providing and/or allowing collection of *Eucalyptus* organ and tissue samples, to C. Graça and N. Saidi for help with sample collection and RNA extraction, and to N. Ladouce and V. Carocha for performing qRT-PCR. Thanks to the Bioinfo Genotoul Platform (<http://bioinfo.genotoul.fr>) for access to resources and to the Genome and Transcriptome Platform (<http://get.genotoul.fr/>) for advice and technical assistance with high-throughput Biomark Fluidigm, to Y. Martinez (FR3450) for assistance with microscopy (Plateforme Imagerie TRI; <http://tri.ups-tlse.fr/>), and to C. Pecher and A. Zakaroff-Girard for their technical assistance and expertise in flow cytometry (Cytometry and Cell Sorting Platform, INSERM UPS UMR 1048, Toulouse RIO imaging platform).

## Disclosures

The authors have no conflicts of interest to declare.

## References

- Abel, S. and Theologis, A. (1994) Transient transformation of *Arabidopsis* leaf protoplasts: a versatile experimental system to study gene expression. *Plant J.* 5: 421–427.
- Ainley, W.M., Walker, J.C., Nagao, R.T. and Key, J.L. (1988) Sequence and characterization of two auxin-regulated genes from soybean. *J. Biol. Chem.* 263: 10658–10666.
- Andersson-Gunneras, S., Mellerowicz, E.J., Love, J., Segerman, B., Ohmiya, Y., Coutinho, P.M. et al. (2006) Biosynthesis of cellulose-enriched tension wood in *Populus*: global analysis of transcripts and metabolites identifies biochemical and developmental regulators in secondary wall biosynthesis. *Plant J.* 45: 144–165.
- Arvidsson, S., Kwasniewski, M., Riano-Pachon, D.M. and Mueller-Roeber, B. (2008) QuantPrime—a flexible tool for reliable high-throughput primer design for quantitative PCR. *BMC Bioinformatics* 9: 465.
- Audran-Delalande, C., Bassa, C., Mila, I., Regad, F., Zouine, M. and Bouzayen, M. (2012) Genome-wide identification, functional analysis and expression profiling of the *Aux/IAA* gene family in tomato. *Plant Cell Physiol.* 53: 659–672.
- Baba, K., Karlberg, A., Schmidt, J., Schrader, J., Hvidsten, T.R., Bako, L. et al. (2011) Activity–dormancy transition in the cambial meristem involves stage-specific modulation of auxin response in hybrid aspen. *Proc. Natl Acad. Sci. USA* 108: 3418–3423.
- Bailey, T.L., Boden, M., Buske, F.A., Fritch, M., Grant, C.E., Clementi, L. et al. (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* 37: W202–W208.
- Bassa, C., Mila, I., Bouzayen, M. and Audran-Delalande, C. (2012) Phenotypes associated with down-regulation of *Sl-IAA27* support functional diversity among *Aux/IAA* family members in the tomato. *Plant Cell Physiol.* 53: 1583–1595.
- Bhalerao, R.P. and Fischer, U. (2014) Auxin gradients across wood— instructive or incidental? *Physiol. Plant.* 151: 43–51.
- Calderon Villalobos, L.I., Lee, S., De Oliveira, C., Ivetac, A., Brandt, W., Armitage, L. et al. (2012) A combinatorial TIR1/AFB–Aux/IAA co-receptor system for differential sensing of auxin. *Nat. Chem. Biol.* 8: 477–485.
- Cassan-Wang, H., Soler, M., Yu, H., Camargo, E.L., Carocha, V., Ladouce, N. et al. (2012) Reference genes for high-throughput quantitative reverse transcription–PCR analysis of gene expression in organs and tissues of *Eucalyptus* grown in various environmental conditions. *Plant Cell Physiol.* 53: 2101–2116.
- Chaabouni, S., Jones, B., Delalande, C., Wang, H., Li, Z., Mila, I. et al. (2009) *Sl-IAA3*, a tomato *Aux/IAA* at the crossroads of auxin and ethylene signalling involved in differential growth. *J. Exp. Bot.* 60: 1349–1362.
- Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* 16: 735–743.
- Deng, W., Yan, F., Liu, M., Wang, X. and Li, Z. (2012a) Down-regulation of *SlIAA15* in tomato altered stem xylem development and production of volatile compounds in leaf exudates. *Plant Signal. Behav.* 7: 911–913.
- Deng, W., Yang, Y., Ren, Z., Audran-Delalande, C., Mila, I., Wang, X. et al. (2012b) The tomato *SlIAA15* is involved in trichome formation and axillary shoot development. *New Phytol.* 194: 379–390.
- Dharmasiri, N., Dharmasiri, S. and Estelle, M. (2005) The F-box protein TIR1 is an auxin receptor. *Nature* 435: 441–445.
- Finn, R.D., Mistry, J., Tate, J., Coghill, P., Heger, A., Pollington, J.E. et al. (2010) The Pfam protein families database. *Nucleic Acids Res.* 38: D211–D222.
- Friml, J. (2003) Auxin transport—shaping the plant. *Curr. Opin. Plant Biol.* 6: 7–12.
- Fukaki, H., Tameda, S., Masuda, H. and Tasaka, M. (2002) Lateral root formation is blocked by a gain-of-function mutation in the *SOLITARY-ROOT/IAA14* gene of *Arabidopsis*. *Plant J.* 29: 153–168.
- Gan, D., Zhuang, D., Ding, F., Yu, Z. and Zhao, Y. (2013) Identification and expression analysis of primary auxin-responsive *Aux/IAA* gene family in cucumber (*Cucumis sativus*). *J. Genet.* 92: 513–21.
- Goh, T., Kasahara, H., Mimura, T., Kamiya, Y. and Fukaki, H. (2012) Multiple AUX/IAA–ARF modules regulate lateral root formation: the role of *Arabidopsis* SHY2/IAA3-mediated auxin signalling. *Philos. Trans. R. Soc. B: Biol. Sci.* 367: 1461–1468.
- Gray, W.M., Kepinski, S., Rouse, D., Leyser, O. and Estelle, M. (2001) Auxin regulates SCF(TIR1)-dependent degradation of AUX/IAA proteins. *Nature* 414: 271–276.
- Guilfoyle, T.J. and Hagen, G. (2007) Auxin response factors. *Curr. Opin. Plant Biol.* 10: 453–460.
- Hellgren, J.M., Olofsson, K. and Sundberg, B. (2004) Patterns of auxin distribution during gravitational induction of reaction wood in poplar and pine. *Plant Physiol.* 135: 212–220.
- Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L. et al. (2008) Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. *Adv. Bioinformatics*. 2008: 420747.

- Hussey, S.G., Saidi, M.N., Hefer, C.A., Myburg, A.A. and Grima-Pettenati, J. (2014) Structural, evolutionary and functional analysis of the NAC domain protein family in *Eucalyptus*. *New Phytol.* (in press).
- Jain, M., Kaur, N., Garg, R., Thakur, J.K., Tyagi, A.K. and Khurana, J.P. (2006) Structure and expression analysis of early auxin-responsive Aux/IAA gene family in rice (*Oryza sativa*). *Funct. Integr. Genomics* 6: 47–59.
- Kalluri, U.C., Difazio, S.P., Brunner, A.M. and Tuskan, G.A. (2007) Genome-wide analysis of Aux/IAA and ARF gene families in *Populus trichocarpa*. *BMC Plant Biol.* 7: 59.
- Karimi, M., Inze, D. and Depicker, A. (2002) GATEWAY vectors for *Agrobacterium*-mediated plant transformation. *Trends Plant Sci.* 7: 193–195.
- Kepinski, S. and Leyser, O. (2005) The *Arabidopsis* F-box protein TIR1 is an auxin receptor. *Nature* 435: 446–451.
- Kim, J., Harter, K. and Theologis, A. (1997) Protein–protein interactions among the Aux/IAA proteins. *Proc. Natl Acad. Sci. USA* 94: 11786–11791.
- Leclercq, J., Ranty, B., Sanchez-Ballesta, M.T., Li, Z., Jones, B., Jauneau, A. et al. (2005) Molecular and biochemical characterization of LeCRK1, a ripening-associated tomato CDPK-related kinase. *J. Exp. Bot.* 56: 25–35.
- Ludwig, Y., Zhang, Y. and Hochholdinger, F. (2013) The maize (*Zea mays* L.) AUXIN/INDOLE-3-ACETIC ACID gene family: phylogeny, synteny, and unique root-type and tissue-specific expression patterns during development. *PLoS One* 8: e78859.
- Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C. et al. (2011) CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res.* 39: D225–D229.
- Mellerowicz, E.J., Baucher, M., Sundberg, B. and Boerjan, W. (2001) Unravelling cell wall formation in the woody dicot stem. *Plant Mol. Biol.* 47: 239–274.
- Miyashima, S., Sebastian, J., Lee, J.Y. and Helariutta, Y. (2013) Stem cell function during plant vascular development. *EMBO J.* 32: 178–193.
- Mockaitis, K. and Estelle, M. (2008) Auxin receptors and plant development: a new signaling paradigm. *Annu. Rev. Cell Dev. Biol.* 24: 55–80.
- Morey, P.R. and Cronshaw, J. (1968) Induction of tension wood by 2,4-dinitrophenol and auxins. *Protoplasma* 65: 393–405.
- Moyle, R., Schrader, J., Stenberg, A., Olsson, O., Saxena, S., Sandberg, G. et al. (2002) Environmental and auxin regulation of wood formation involves members of the Aux/IAA gene family in hybrid aspen. *Plant J.* 31: 675–685.
- Muto, H., Watahiki, M.K., Nakamoto, D., Kinjo, M. and Yamamoto, K.T. (2007) Specificity and similarity of functions of the Aux/IAA genes in auxin signaling of *Arabidopsis* revealed by promoter-exchange experiments among MSG2/IAA19, AXR2/IAA7, and SLR/IAA14. *Plant Physiol.* 144: 187–196.
- Myburg, A.A., Grattapaglia, D., Tuskan, G.A., Hellsten, U., Hayes, R.D., Grimwood, J. et al. (2014) The genome of *Eucalyptus grandis*. *Nature* 509: 356–362.
- Nagpal, P., Walker, L.M., Young, J.C., Sonawala, A., Timpte, C., Estelle, M. et al. (2000) AXR2 encodes a member of the Aux/IAA protein family. *Plant Physiol.* 123: 563–574.
- Nilsson, J., Karlberg, A., Antti, H., Lopez-Vernaza, M., Mellerowicz, E., Perrot-Rechenmann, C. et al. (2008) Dissecting the molecular basis of the regulation of wood formation by auxin in hybrid aspen. *Plant Cell* 20: 843–855.
- Ouellet, F., Overvoorde, P.J. and Theologis, A. (2001) IAA17/AXR3: biochemical insight into an auxin mutant phenotype. *Plant Cell* 13: 829–841.
- Overvoorde, P.J., Okushima, Y., Alonso, J.M., Chan, A., Chang, C., Ecker, J.R. et al. (2005) Functional genomic analysis of the AUXIN/INDOLE-3-ACETIC ACID gene family members in *Arabidopsis thaliana*. *Plant Cell* 17: 3282–3300.
- Paux, E., Carocha, V., Marques, C., Mendes de Sousa, A., Borralho, N., Sivadon, P. et al. (2005) Transcript profiling of *Eucalyptus* xylem genes during tension wood formation. *New Phytol.* 167: 89–100.
- Pfaffl, M.W. (2001) A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Res.* 29: e45.
- Pilate, G., Déjardin, A., Laurans, F. and Leplé, J.C. (2004) Tension wood as a model for functional genomics of wood formation. *New Phytol.* 164: 63–72.
- Plomion, C., Leprovost, G. and Stokes, A. (2001) Wood formation in trees. *Plant Physiol.* 127: 1513–1523.
- Reed, J.W. (2001) Roles and activities of Aux/IAA proteins in *Arabidopsis*. *Trends Plant Sci.* 6: 420–425.
- Reed, J.W., Elumalai, R.P. and Chory, J. (1998) Suppressors of an *Arabidopsis thaliana* *phyB* mutation identify genes that control light signaling and hypocotyl elongation. *Genetics* 148: 1295–1310.
- Remington, D.L., Vision, T.J., Guilfoyle, T.J. and Reed, J.W. (2004) Contrasting modes of diversification in the Aux/IAA and ARF gene families. *Plant Physiol.* 135: 1738–1752.
- Rouse, D., Mackay, P., Stirnberg, P., Estelle, M. and Leyser, O. (1998) Changes in auxin response from mutations in an AUX/IAA gene. *Science* 279: 1371–1373.
- Sato, A. and Yamamoto, K.T. (2008) Overexpression of the non-canonical Aux/IAA genes causes auxin-related aberrant phenotypes in *Arabidopsis*. *Physiol. Plant.* 133: 397–405.
- Schrader, J., Baba, K., May, S.T., Palme, K., Bennett, M., Bhalerao, R.P. et al. (2003) Polar auxin transport in the wood-forming tissues of hybrid aspen is under simultaneous control of developmental and environmental signals. *Proc. Natl Acad. Sci. USA* 100: 10096–10101.
- Schrader, J., Moyle, R., Bhalerao, R., Hertzberg, M., Lundeberg, J., Nilsson, P. et al. (2004) Cambial meristem dormancy in trees involves extensive remodelling of the transcriptome. *Plant J.* 40: 173–187.
- Shimada, T.L., Shimada, T. and Hara-Nishimura, I. (2010) A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of *Arabidopsis thaliana*. *Plant J.* 61: 519–528.
- Soler, M., Camargo, E.L.O., Carocha, V., Cassan-Wang, H., San Clemente, H., Savelli, B. et al. (2014) The *Eucalyptus grandis* R2R3-MYB transcription factor family: evidence for woody growth related evolution and function. *New Phytol.* (in press).
- Southerton, S.G., Marshall, H., Mouradov, A. and Teasdale, R.D. (1998) Eucalypt MADS-box genes expressed in developing flowers. *Plant Physiol.* 118: 365–372.
- Su, L., Bassa, C., Audran, C., Cheniclet, C., Chevalier, C., Bouzayen, M. et al. (2014) The auxin *Sl-IAA17* transcriptional repressor controls fruit size via the regulation of endoreduplication-related cell expansion. *Plant Cell Physiol.* 55: 1969–1976.
- Sundberg, B., Uggla, C. and Tuominen, H. (2000) Cambial growth and auxin gradients. In *Cell and Molecular Biology of Wood Formation*. Edited by Savidge, R.A., Barnett, J.R. and Napier, R. pp. 169–188. BIOS Scientific Publishers, Oxford.
- Szemenyei, H., Hannon, M. and Long, J.A. (2008) TOPLESS mediates auxin-dependent transcriptional repression during *Arabidopsis* embryogenesis. *Science* 319: 1384–1386.
- Tan, X., Calderon-Villalobos, L.I., Sharon, M., Zheng, C., Robinson, C.V., Estelle, M. et al. (2007) Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature* 446: 640–645.
- Theologis, A., Huynh, T.V. and Davis, R.W. (1985) Rapid induction of specific mRNAs by auxin in pea epicotyl tissue. *J. Mol. Biol.* 183: 53–68.
- Tian, Q., Nagpal, P. and Reed, J.W. (2003) Regulation of *Arabidopsis* SHY2/IAA3 protein turnover. *Plant J.* 36: 643–651.
- Tian, Q. and Reed, J.W. (1999) Control of auxin-regulated root development by the *Arabidopsis thaliana* SHY2/IAA3 gene. *Development* 126: 711–721.



- Timell, T. (1969) The chemical composition of tension wood. *Svensk Papperstidning* 72: 173–181.
- Tiwari, S.B., Hagen, G. and Guilfoyle, T.J. (2004) Aux/IAA proteins contain a potent transcriptional repression domain. *Plant Cell* 16: 533–543.
- Tuominen, H., Puech, L., Fink, S. and Sundberg, B. (1997) A radial concentration gradient of indole-3-acetic acid is related to secondary xylem development in Hybrid Aspen. *Plant Physiol.* 115: 577–585.
- Uggla, C., Magel, E., Moritz, T. and Sundberg, B. (2001) Function and dynamics of auxin and carbohydrates during earlywood/latewood transition in scots pine. *Plant Physiol.* 125: 2029–2039.
- Uggla, C., Mellerowicz, E.J. and Sundberg, B. (1998) Indole-3-acetic acid controls cambial growth in scots pine by positional signaling. *Plant Physiol.* 117: 113–121.
- Uggla, C., Moritz, T., Sandberg, G. and Sundberg, B. (1996) Auxin as a positional signal in pattern formation in plants. *Proc. Natl Acad. Sci. USA* 93: 9282–9286.
- Ulmasov, T., Hagen, G. and Guilfoyle, T.J. (1999) Activation and repression of transcription by auxin-response factors. *Proc. Natl Acad. Sci. USA* 96: 5844–5849.
- Ulmasov, T., Murfett, J., Hagen, G. and Guilfoyle, T.J. (1997) Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. *Plant Cell* 9: 1963–1971.
- Vernoux, T., Brunoud, G., Farcot, E., Morin, V., Van den Daele, H., Legrand, J. et al. (2011) The auxin signalling network translates dynamic input into robust patterning at the shoot apex. *Mol. Syst. Biol.* 7: 508.
- Voorrips, R.E. (2002) MapChart: software for the graphical presentation of linkage maps and QTLs. *J. Hered.* 93: 77–78.
- Walker, J.C. and Key, J.L. (1982) Isolation of cloned cDNAs to auxin-responsive poly(A)RNAs of elongating soybean hypocotyl. *Proc. Natl Acad. Sci. USA* 79: 7185–7189.
- Wang, H., Jones, B., Li, Z., Frasse, P., Delalande, C., Regad, F. et al. (2005) The tomato *Aux/IAA* transcription factor *IAA9* is involved in fruit development and leaf morphogenesis. *Plant Cell* 17: 2676–2692.
- Watahiki, M.K. and Yamamoto, K.T. (1997) The *massugu1* mutation of *Arabidopsis* identified with failure of auxin-induced growth curvature of hypocotyl confers auxin insensitivity to hypocotyl and leaf. *Plant Physiol.* 115: 419–426.
- Woodward, A.W. and Bartel, B. (2005) Auxin: regulation, action, and interaction. *Ann. Bot.* 95: 707–735.
- Yang, X., Lee, S., So, J.H., Dharmasiri, S., Dharmasiri, N., Ge, L. et al. (2004) The *IAA1* protein is encoded by *AXR5* and is a substrate of SCF(TIR1). *Plant J.* 40: 772–782.
- Yu, H., Soler, M., Mila, I., San Clemente, H., Savelli, B., Dunand, C. et al. (2014) Genome-wide characterization and expression profiling of the *AUXIN RESPONSE FACTOR* (*ARF*) gene family in *Eucalyptus grandis*. *PLoS One* 9: e108906.





Article

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

Ying Dai <sup>1</sup>, Guojian Hu <sup>2</sup>, Annabelle Dupas <sup>1</sup>, Luciano Medina <sup>1</sup>, Nils Blandels <sup>1</sup>,  
Hélène San Clemente <sup>1</sup>, Nathalie Ladouce <sup>1</sup>, Myriam Badawi <sup>3,4</sup>,  
Guillermina Hernandez-Raquet <sup>3</sup>, Fabien Mounet <sup>1</sup>, Jacqueline Grima-Pettenati <sup>1</sup>  
and Hua Cassan-Wang <sup>1,\*</sup>

- <sup>1</sup> Laboratoire de Recherche en Sciences Végétales, Université 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.)
- <sup>2</sup> UMR 990, Génomique et Biotechnologie des Fruits, Université de Toulouse, INP-ENSA Toulouse, Avenue de l'Agrobiopole, 31326 Castanet-Tolosan, France; guojian.hu@etu.ensat.fr
- <sup>3</sup> TBI, Université de Toulouse, CNRS, INRAE, INSA, 31400 Toulouse, France; Myriam.Badawi@univ-lemans.fr (M.B.); hernandg@insa-toulouse.fr (G.H.-R.)
- <sup>4</sup> Laboratoire Mer Molécules Santé, MMS EA2160 Le Mans Université, 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 (DSDDecode, <http://dsdecode.scgene.com>). For *CCR1*-lines the majority of the direct sequencing of PCR amplicons could not be read by the DSDDecode 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 Tables S1 and 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). Noteworthily 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). Noteworthily, 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 DSDecode online tool failed to run, thereby confirming that this analytic tool is not suitable for chimeric or multiple mutations.

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

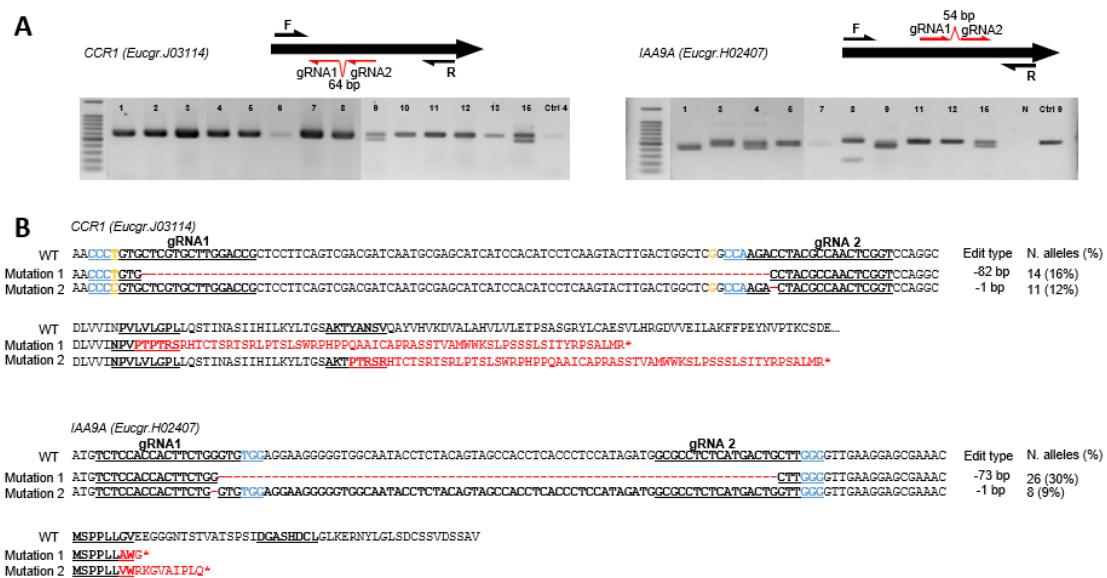


without adding the stress of transferring them to hydroponic culture. For the second batch and third batch of transformation CRISPR *CCR1*, we collected the samples at 73 days and 54 days, respectively. The genotyping results clearly showed that the younger the transgenic roots, the higher the edition 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).

**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 <sup>1</sup>	≥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%)		

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



**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.

### 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 DSDcode online tool on the 13 PCR amplicon sequences (13 transgenic hairy roots). Nine were successfully treated by DSDcode (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 DSDcode 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 DSDcode. 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 DSDcode 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 DSDcode, 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 15 bp) and large (larger than 15 bp) 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 (≤15 bp)			Large Deletion (>15 bp)			Small Insertion (≤15 bp)			Large Insertion (>15 bp)			Small Substitution (≤15 bp)			Large Substitution (>15 bp)				
			sg1	sg2	sg1&2	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	0	19
			52 (58.4%)			21 (23.6%)			4 (4.5%)			0			29 (32.6%)			0			(21.3%)	
IAA9A	88	20	29	39	19	38	27	27	5	11	0	12	0	0	12	2	0	0	0	0	0	27
			49 (55.7%)			38 (43.2%)			16 (18.2%)			12 (13.6%)			14 (15.9%)			0			(30.7%)	

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.

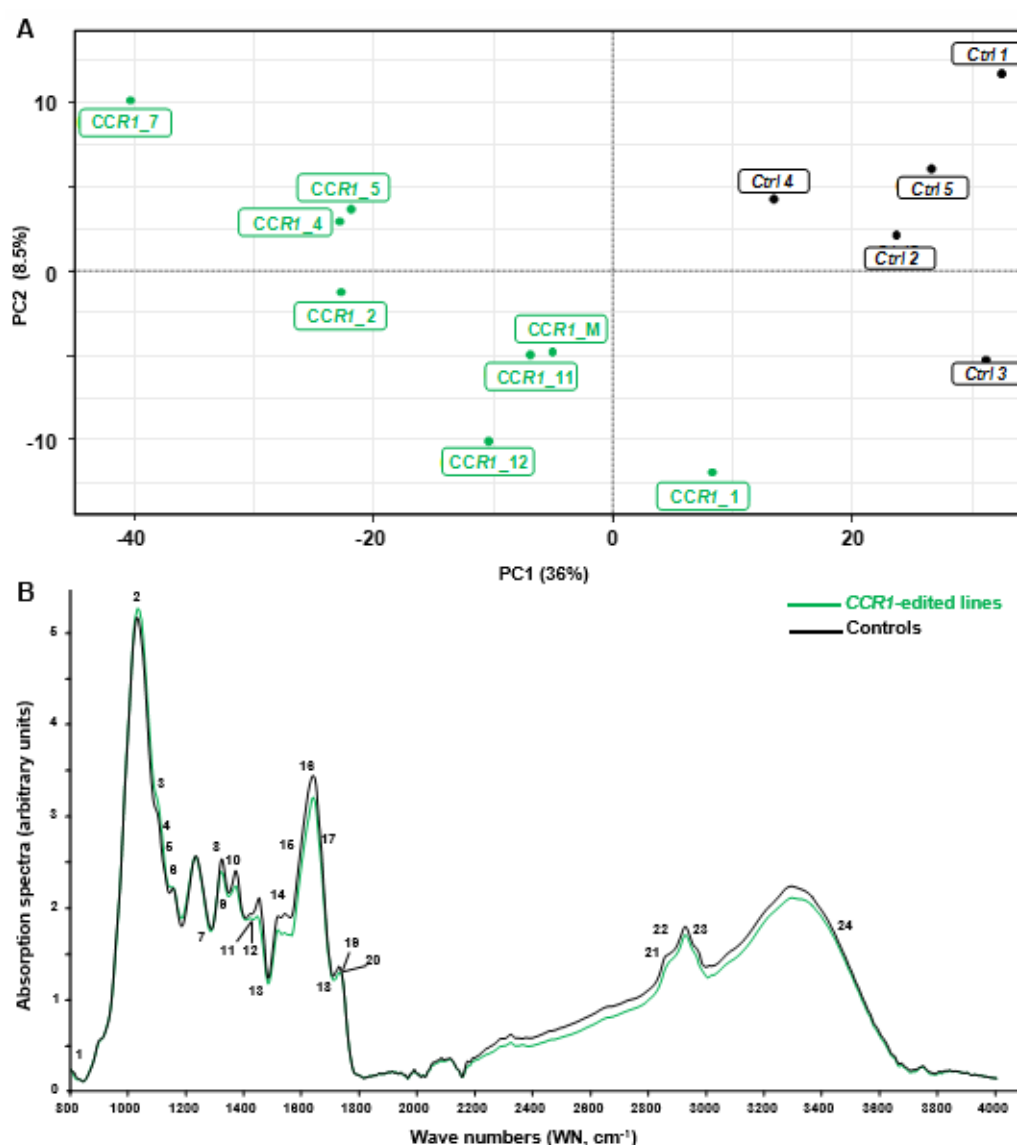
## 2.2. Phenotyping Revealed Expected Alterations of Lignification in *CCR1*-Edited Lines

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.

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: 1000–1200  $\text{cm}^{-1}$ , 1300–1800  $\text{cm}^{-1}$  and 2200–3600  $\text{cm}^{-1}$  (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 bands 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.

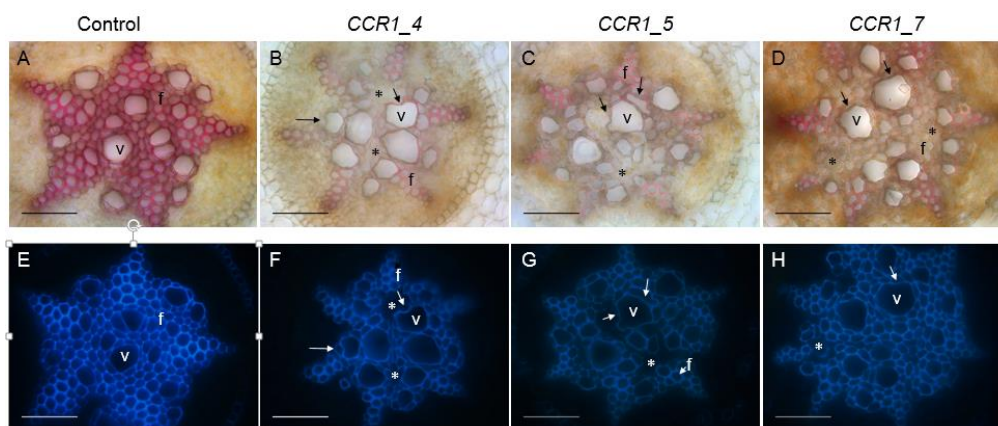


**Figure 2.** Comparison between FT-IR spectra obtained from controls and *CCR1* edited lines. (A) partial least square analysis (PLS-DA) analysis was performed using the normalized values of Fourier-transformed infra-red (FTIR) absorption spectra ( $800\text{--}4000\text{ cm}^{-1}$ ), acquired from 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].

### 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  $\mu$ 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 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.

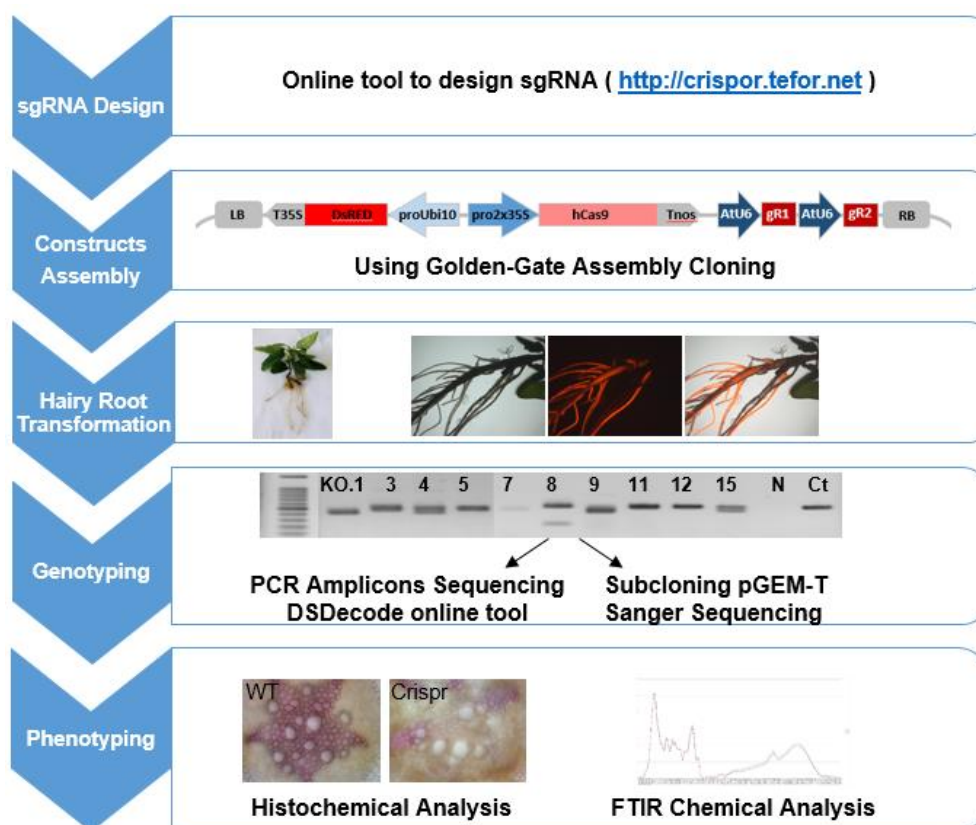
## 4. Materials and Methods

### 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 1 cm 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.

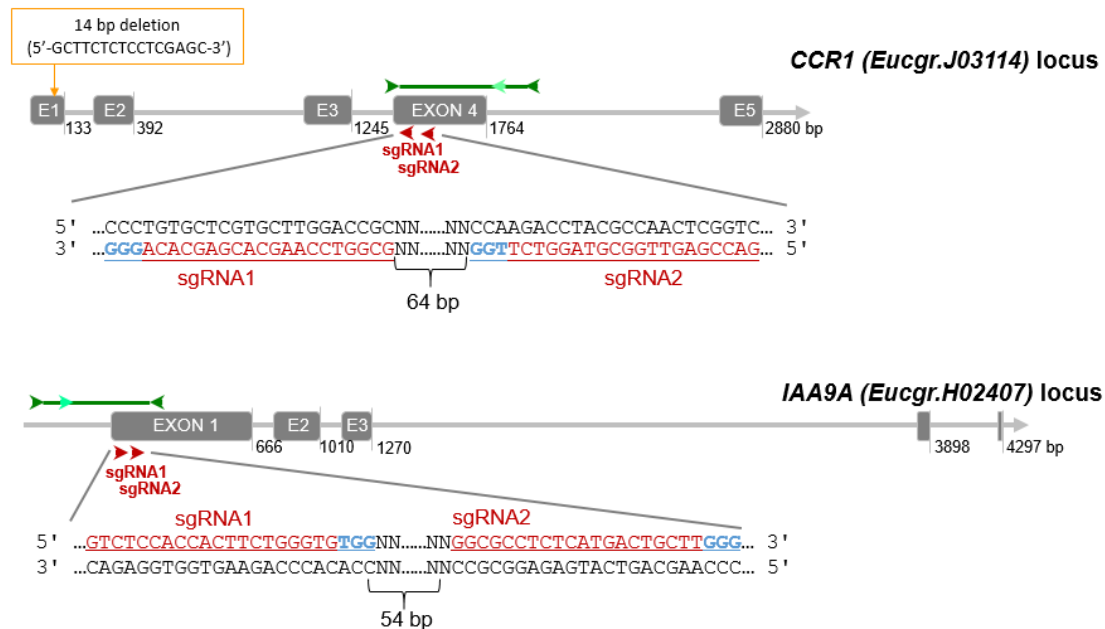
#### 4.3. CRISPR/Cas9 Target Site Selection and sgRNAs Design

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′-gcttctctctcgagc-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].

Finally, for *CCR1* (*Eucgr.J03114*), we chose two sgRNAs (sgRNA1\_ *CCR1* (5′-GCGGTCCAAGCAGCAGCACA-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\_IIA9A (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.

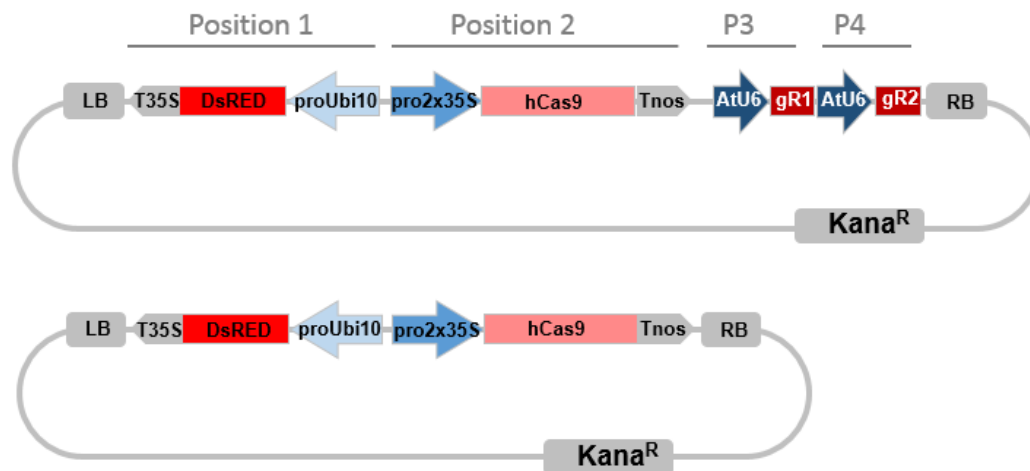


**Figure 5.** CRISPR/Cas9 sgRNA design and mutation detection in *CCR1* and *IIA9A*. Schematic representation of the target sites and the PCR assay for Sanger sequencing. Exons and introns are represented by gray boxes and gray lines, respectively. The target sites for each CRISPR/Cas9 nuclease are indicated 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.

#### 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<sup>R</sup>*, kanamycin resistance gene sequence; DsRed, DsRed fluorescent marker gene sequence, *ProUbi10* *Arabidopsis thaliana* Ubiquitin 10 gene promoter; *T35S*, terminator region of *CaMV*35S 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  $\mu$ L) was prepared to contain 2  $\mu$ L BsaI 10 $\times$  reaction buffer (NEB), 20 U of BsaI (NEB), 20 U of T4 DNA ligase (using high concentration ligase, 20 U/ $\mu$ 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 $\alpha$  *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  $\mu$ 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<sup>TM</sup> 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  $\frac{1}{2}$  strength of macro elements) for a period of three to ten weeks (7–12  $\mu\text{mol}/\text{m}^2/\text{s}$ , 8–16-h photoperiod, 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  $\frac{1}{2}$  strength of macro elements solution in a phytotron (130  $\mu\text{mol}/\text{m}^2/\text{s}$ , 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 (DSDecode, <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 DSDecode 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  $\text{cm}^{-1}$  with a 4  $\text{cm}^{-1}$  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 <http://www.mdpi.com/1422-0067/21/10/3408/s1>.

**Author Contributions:** Conceptualization, J.G.-P. and H.C.-W.; Funding acquisition, J.G.-P.; Investigation, Y.D., F.M. and H.C.-W.; Methodology, Y.D., G.H., A.D., L.M., N.B., H.S.C., N.L., M.B., G.H.-R., F.M. and H.C.-W.; Supervision, J.G.-P. and H.C.-W.; Validation, J.G.-P.; Writing—original draft, Y.D., F.M., J.G.-P. and H.C.-W.; Writing—review and editing, J.G.-P. and H.C.-W. 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
DSDecode	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. [[CrossRef](#)]
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*; U.S. Department of Energy: Richland, WA, USA, 2007; pp. 3552–3559.
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. [[CrossRef](#)]
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. [[CrossRef](#)]

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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
10. Tournier, V.; Grat, S.; Marque, C.; El Kayal, W.; Penchel, R.; de Andrade, G.; Boudet, A.-M.; Teulières, C. An Efficient Procedure to Stably Introduce Genes into an Economically Important Pulp Tree (*Eucalyptus grandis* × *Eucalyptus urophylla*). *Transgenic Res.* **2003**, *12*, 403–411. [[CrossRef](#)]
11. Girijashankar, V. Genetic transformation of *Eucalyptus*. *Physiol. Mol. Biol. Plants* **2011**, *17*, 9–23. [[CrossRef](#)]
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. [[CrossRef](#)]
13. Cao, P.B.; Ployet, R.; Nguyen, C.; Dupas, A.; Ladouce, N.; Martinez, Y.; Grima-Pettenati, J.; Marque, C.; Mounet, F.; Teulières, C. Wood architecture and composition are deeply remodeled in frost sensitive *Eucalyptus* overexpressing CBF/DREB1 transcription factors. *Int. J. Mol. Sci.* **2020**, *21*, 3019. [[CrossRef](#)]
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. [[CrossRef](#)]
15. Travis, J. Making the cut. *Science* **2015**, *350*, 1456–1457. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
21. Tsai, C.-J.; Xue, L.-J. CRISPRing into the woods. *GM Crops Food* **2015**, *6*, 206–215. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]

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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
44. Jia, H.; Wang, N. Targeted Genome Editing of Sweet Orange Using Cas9/sgRNA. *PLoS ONE* **2014**, *9*, e93806. [[CrossRef](#)]
45. Jia, H.; Xu, J.; Orbović, V.; Zhang, Y.; Wang, N. Editing Citrus Genome via SaCas9/sgRNA System. *Front. Plant Sci.* **2017**, *8*, 2135. [[CrossRef](#)] [[PubMed](#)]
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. [[CrossRef](#)] [[PubMed](#)]
47. Breitler, J.-C.; Dechamp, E.; Campa, C.; Zebral 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. [[CrossRef](#)]

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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
51. Bruegmann, T.; Deecke, K.; Fladung, M. Evaluating the Efficiency of gRNAs in CRISPR/Cas9 Mediated Genome Editing in Poplars. *Int. J. Mol. Sci.* **2019**, *20*, 3623. [[CrossRef](#)]
52. Piquemal, J.; Lapiere, 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. [[CrossRef](#)]
53. Leplé, J.-C.; Dauwe, R.; Morreel, K.; Storme, V.; Lapiere, 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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. *Int. J. Mol. Sci.* **2019**, *20*, 1155. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]
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. [[CrossRef](#)]

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. [[CrossRef](#)]
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. HyperSpec: A Package to Handle Hyperspectral Data Sets in R. R Package Version 0.99-20200213.1. Available online: <https://github.com/cbeleites/hyperSpec> (accessed on 10 February 2020).
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. [[CrossRef](#)]
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. [[CrossRef](#)]



© 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/>).





Article

# Overexpression of *EgrIAA20* from *Eucalyptus grandis*, a Non-Canonical *Aux/IAA* Gene, Specifically Decouples Lignification of the Different Cell-Types in *Arabidopsis* Secondary Xylem

Hong Yu <sup>1,†</sup>, Mingjun Liu <sup>1</sup>, Zhangsheng Zhu <sup>1,‡</sup>, Aiming Wu <sup>2</sup>, Fabien Mounet <sup>1</sup>, Edouard Pesquet <sup>3</sup>, Jacqueline Grima-Pettenati <sup>1</sup> and Hua Cassan-Wang <sup>1,\*</sup>

- <sup>1</sup> Laboratoire de Recherche en Sciences Végétales, Université de Toulouse, CNRS, UPS, Toulouse INP, 31320 Auzeville-Tolosane, France; hongyu@swmu.edu.cn (H.Y.); mingjunliu@pitt.edu (M.L.); zhuzs@scau.edu.cn (Z.Z.); fabien.mounet@univ-tlse3.fr (F.M.); grima@lrsv.ups-tlse.fr (J.G.-P.)
- <sup>2</sup> State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, Guangdong Key Laboratory for Innovative Development and Utilization of Forest Plant Germplasm, College of Forestry and Landscape Architectures, South China Agricultural University, Guangzhou 510642, China; wuaimin@scau.edu.cn
- <sup>3</sup> Arrhenius Laboratories, Department of Ecology, Environment and Plant Sciences (DEEP), Stockholm University, Svante Arrhenius väg 20A, 106 91 Stockholm, Sweden; edouard.pesquet@su.se
- \* Correspondence: hua.cassan@univ-tlse3.fr; Tel.: +33-534-323-851
- † Current Address: Public Centre for Experimental Technology, The School of Basic Medical Science, Southwest Medical University, Luzhou 64600, China.
- ‡ Current Address: College of Horticulture, South China Agricultural University, Guangzhou 510642, China.



**Citation:** Yu, H.; Liu, M.; Zhu, Z.; Wu, A.; Mounet, F.; Pesquet, E.; Grima-Pettenati, J.; Cassan-Wang, H. Overexpression of *EgrIAA20* from *Eucalyptus grandis*, a Non-Canonical *Aux/IAA* Gene, Specifically Decouples Lignification of the Different Cell-Types in *Arabidopsis* Secondary Xylem. *Int. J. Mol. Sci.* **2022**, *23*, 5068. <https://doi.org/10.3390/ijms23095068>

Academic Editor: Setsuko Komatsu

Received: 8 April 2022

Accepted: 26 April 2022

Published: 3 May 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 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 (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** Wood (secondary xylem) formation is regulated by auxin, which plays a pivotal role as an integrator of developmental and environmental cues. However, our current knowledge of auxin-signaling during wood formation is incomplete. Our previous genome-wide analysis of *Aux/IAAs* in *Eucalyptus grandis* showed the presence of the non-canonical paralog member *EgrIAA20* that is preferentially expressed in cambium. We analyzed its cellular localization using a GFP fusion protein and its transcriptional activity using transactivation assays, and demonstrated its nuclear localization and strong auxin response repressor activity. In addition, we functionally tested the role of *EgrIAA20* by constitutive overexpression in *Arabidopsis* to investigate for phenotypic changes in secondary xylem formation. Transgenic *Arabidopsis* plants overexpressing *EgrIAA20* were smaller and displayed impaired development of secondary fibers, but not of other wood cell types. The inhibition in fiber development specifically affected their cell wall lignification. We performed yeast-two-hybrid assays to identify *EgrIAA20* protein partners during wood formation in *Eucalyptus*, and identified *EgrIAA9A*, whose ortholog *PtoIAA9* in poplar is also known to be involved in wood formation. Altogether, we showed that *EgrIAA20* is an important auxin signaling component specifically involved in controlling the lignification of wood fibers.

**Keywords:** secondary xylem; auxin; non-canonical *Aux/IAA*; secondary fiber; cambium differentiation; *Eucalyptus*; syringyl lignin; *IAA20*; wood; *Arabidopsis*

## 1. Introduction

Wood, or secondary xylem, is one of the most abundant natural and renewable sources of energy, as well as an important bio-resource for global industry. It provides raw material for construction and building, pulp/paper making, energy production as well as a source of fine chemicals ranging from plasticizers to flavoring agents. Wood is one of the most environmental-friendly resources that enables both to replace our dependency on fossil fuels and to capture atmospheric CO<sub>2</sub> associated with global warming actively [1]. In trees, wood complex organization functions as a vascular/skeletal tissue due to the presence

of several specific cell types characterized by the presence of secondary cell walls (SCW). These include the sap-conducting cells called tracheary elements, the skeletal support cells called fibers, and ray cells allowing to load and unload the sap content. These wood cell types all derive from the differentiation of a lateral secondary meristem called the vascular cambium that undergoes orientated cell division, cell expansion, and massive deposition of SCW polysaccharides in patterns guided by microtubules, programmed cell death preceding or not cell wall lignification depending on the xylem cell type [2]. Albeit following the same general blueprint, each xylem cell type has specific cytological features (i.e., pattern and composition of SCWs) to fulfill their roles such as in the case of lignin. This cell wall polymer is composed of phenylpropanoids that vary in both their terminal aliphatic functions (alcohol, aldehyde, acid) and their level of aromatic substitutions (meta hydroxy/methoxylation). Canonical lignin residues include *p*-hydroxyphenyl (H) with no meta substitution, guaiacyl (G) with one meta substitution, and syringyl (S) with two meta substitutions [3]. Each distinct xylem cell type and the different cell wall layers are enriched with specific lignin residues. The SCWs in tracheary elements are enriched in G residues; the fiber SCWs are enriched in S residues, whereas the primary cell wall separating these cell types are enriched in H residue [4,5]. The succession of the differentiation stage of xylem cell types from their initiations and their coordination in time of specific cytological features to their full maturation depend on growth factors that have either overlapping or distinct effects and that depend on the xylem cell types, their differentiation stage, and/or the surrounding environmental constraints [6–8]. Among these growth factors, an increase in auxin levels plays an essential role for tracheary element initiation [9,10], but also uncouples xylem cell types—not affecting TE formation but fibers to give rise to either in parenchyma cells or fibers with a reduced SCW formation [11,12]. Yet, the molecular mechanisms allowing auxin to uncouple the formation of the different xylem cell types remain unclear.

Auxin signaling is thus capable of directing cellular behaviors from cell division, expansion, and differentiation [13,14]. However, our current knowledge of the auxin signaling pathway components involved in secondary xylem differentiation remains incomplete. Auxin signaling is mediated by two classes of transcription factors: AUXIN/INDOLE-3-ACETIC ACID (*Aux/IAA*) and AUXIN RESPONSE FACTOR (ARF) [15]. In the absence of auxin, the *Aux/IAA* repressors bind to the ARFs to stop transcription. In contrast, when auxin is present, *Aux/IAAs* associate with S-PHASE KINASE-ASSOCIATED PROTEIN1 SKP-Cullin-F BOX (SCF)<sup>TRANSPORT INHIBITOR RESISTANT1/AUXIN SIGNALING F-BOX</sup> (SCF<sup>TIR1/AFB</sup>), which triggers their ubiquitination and proteolysis. The interacting ARF transcription factors are thus freed and able to modulate the expression of the downstream auxin-responsive genes [16,17]. Despite numerous studies showing the role of ARF5/MONOPTEROS (ARF5/MP) and its interacting *Aux/IAA* members during primary xylem formation (i.e., IAA12 in embryogenesis, IAA20/30 in root and leaf) [18–22], only a handful of studies have investigated the role of *Aux/IAA* members and their interacting ARFs during secondary xylem formation [23]. These studies include the role of *PttIAA3* in hybrid aspen, which is highly expressed in wood forming tissue. The overexpression of a stabilized *PttIAA3* version reduced the width and the length for both tracheary elements and fibers in transgenic lines in addition to a reduction in cambial cell division [24]. *EgrIAA13* in *Eucalyptus* and its orthologs in *Populus* are preferentially expressed in the xylogenetic tissues and downregulated in tension wood. The downregulation of *EgrIAA13* using the RNAi knockdown construct impeded *Eucalyptus* xylem fiber and vessel development in Induced Somatic Sector Analysis experiments [25]. The *Eucalyptus EgrIAA4*, which is highly expressed in cambium, was shown to uncouple xylem cell types, maintaining the formation of tracheary elements but altering fibers when overexpressed in *Arabidopsis* [26]. It is worth noting that *Arabidopsis* undergoes true secondary growth wood formation (i.e., originating from cambial activity), although devoid of ray cells, in hypocotyls, roots, and at the basal part of inflorescence stems when stimulated by short-day growth conditions [27,28]. The formed secondary xylem in hypocotyls comprises two distinct phases: (i) phase I with

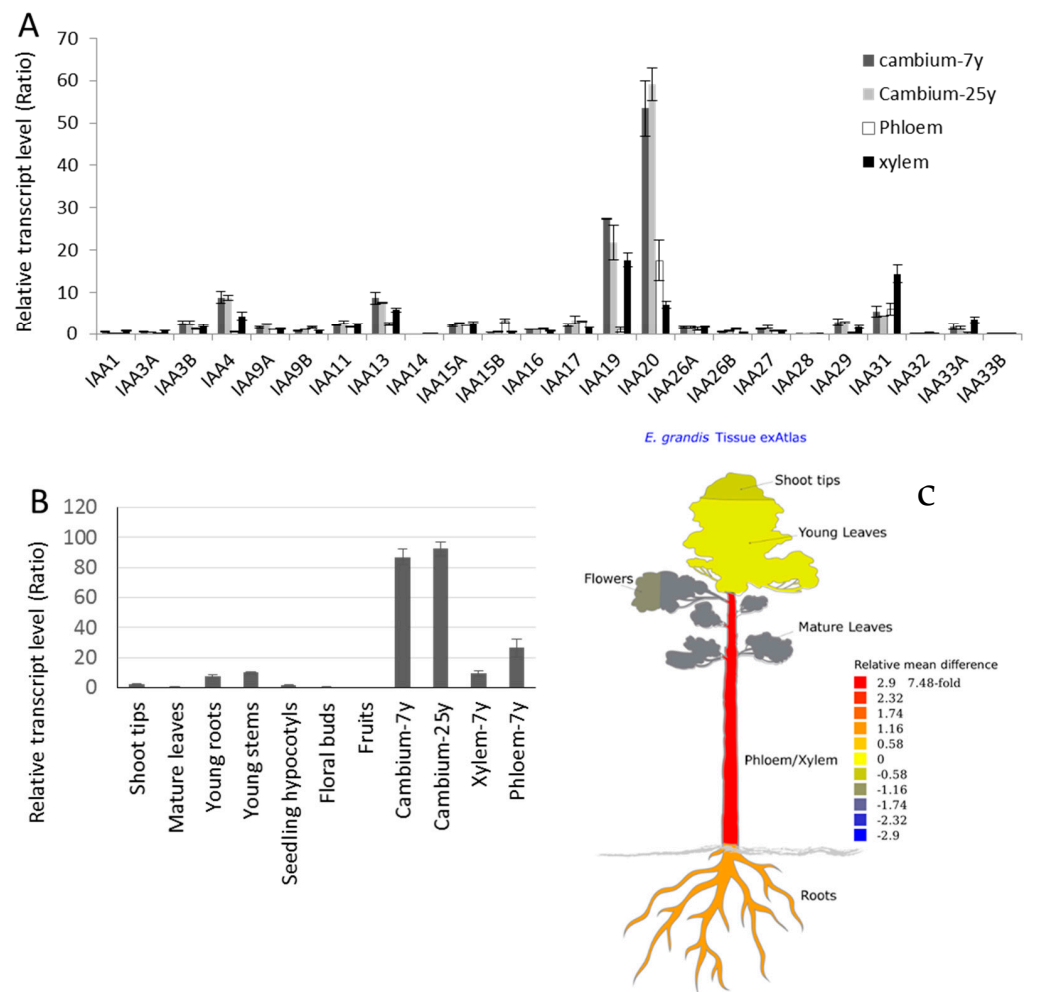
xylem composed mainly of tracheary elements and parenchyma cells, and (ii) phase II with both tracheary elements and fibers [28]. *Arabidopsis* can thus be used as a reliable simplified model of wood formation to investigate the function of key genes in trees, from *Eucalyptus* to poplar, using heterologous over-expression [26,29,30].

Within the genome of *Eucalyptus grandis* [31], several candidate *ARF* and *Aux/IAA* genes are associated with wood formation. Among them, *EgrIAA20*, a non-canonical *Aux/IAA* member, is preferentially expressed in cambium and developing secondary xylem of *Eucalyptus*. Herein, we performed the functional analysis of the atypical *EgrIAA20* that lacks the highly conserved domain II required for the rapid degradation of *Aux/IAA* protein in the presence of auxin. Using transient expression in tobacco protoplast, we showed that *EgrIAA20* functions as a strong transcriptional repressor of auxin responses. The overexpression of *EgrIAA20* not only impaired vascular patterning in cotyledons during primary growth, but also altered secondary xylem development by altering secondary fiber development but not tracheary elements. Screening for the protein partners of *EgrIAA20* during wood formation identified *EgrIAA9A* as the main interacting protein. This suggests that a complex *EgrIAA9A-EgrIAA20* is likely to be involved in specific wood cell type differentiation. Our study reveals that the ubiquitous and substantial over-expression of the auxin-dependent transcription factor *EgrIAA20* specifically uncoupled the differentiation and/or lignification of tracheary elements from fibers. This novel discovery not only opens new avenues to deepen our understanding about the regulation of wood cell type differentiation, but also provides a proof-of-concept strategy to control wood cell types and thus lignin composition without affecting sap conduction.

## 2. Results

### 2.1. *EgrIAA20* Transcripts Accumulate Preferentially in Cambium in *Eucalyptus*

We previously identified 26 *Aux/IAA* genes in the *E. grandis* genome and numbered them according to their phylogenetic relationship with their *Arabidopsis* orthologs [26]. Large scale RT-qPCR expression analyses of these genes across various tissues and organs, including cambium and developing xylem, enabled us to identify 11 members with significantly high expression levels in vascular tissues [26]. Among these *Aux/IAA* paralogs, *EgrIAA20* presented the highest expression level in vascular tissues, especially in cambium and xylem (Figure 1A,B), followed by *EgrIAA19*, *EgrIAA31*, *EgrIAA4*, and *EgrIAA13*. Complementary analyses using the EucGenIE database, an online resource for *Eucalyptus* genomics and RNAseq transcriptomics, further confirmed the preferential expression of *EgrIAA20* in tree trunks (Figure 1C). We therefore selected *EgrIAA20* as a promising candidate likely involved in the auxin regulatory mechanisms controlling the differentiation of wood cell types.



**Figure 1.** *EgrIAA20* transcript (*Eucgr.K00561*) shows preferential accumulation in the cambium of *Eucalyptus*. **(A)** Ratio of relative mRNA abundance of all *Eucalyptus* *Aux/IAA* members in vascular tissues including: juvenile vascular cambium (Cambium-7-year-old tree), mature vascular cambium (Cambium-25-year-old tree), phloem (7-year-old tree), and xylem (7-year-old tree). **(B)** Real-time PCR expression levels of *EgrIAA20* in various *Eucalyptus* tissues and organs. Each relative mRNA abundance was normalized to a control sample (in vitro *Eucalyptus* plantlets). Error bars indicate mean expression values  $\pm$ SE from three independent experiments. **(C)** Pictographic view of *EgrIAA20* expression across a diverse range of RNAseq expression datasets by exImage (<https://eucgenie.org/exImage>, accessed on 25 April 2022).

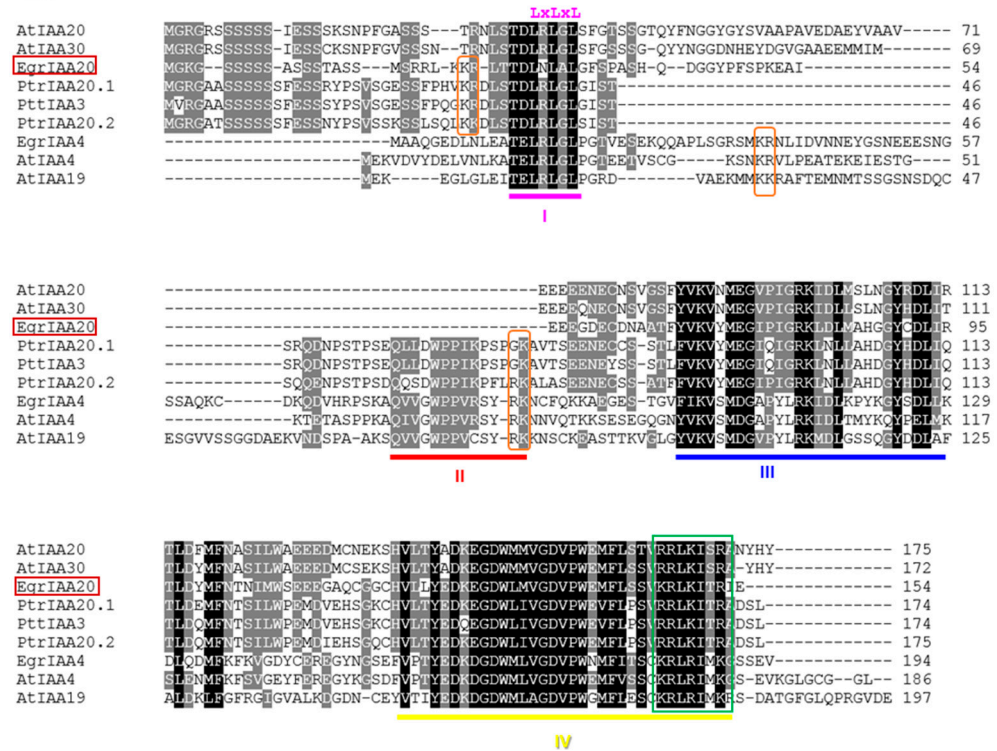
## 2.2. *EgrIAA20* Defines a Non-Canonical *Aux/IAA* Protein Which Lacks Highly Conserved *Aux/IAA* Characteristic DOMAIN II

Protein sequence analyses and alignments between *Aux/IAA* para/orthologs revealed that *EgrIAA20* lacks the highly conserved Domain II (Figure 2A,C). Domain II is conserved in the majority of the *Aux/IAA* para/orthologs, and single mutations in this domain increase protein stability thus hampering normal auxin responses that rely on the rapid degradation of *Aux/IAA* [32]. The closest *Arabidopsis* orthologs of *EgrIAA20* are AtIAA20 and AtIAA30, which also lack Domain II. These non-canonical *Aux/IAA* proteins such as AtIAA20 were shown to be long-lived protein in *Arabidopsis* in contrast to the other paralogs [33]. It is worth noting that non-canonical *Aux/IAA* devoid of Domain II are not present in grapevine [34], nor in poplar species such as black cottonwood (*PtrIAA20.1*, *PtrIAA20.2*) or aspen (*PttIAA3*) [24] (Figure 2A–C) in this phylogenetic clade. Additional analyses also revealed the presence of the highly conserved Domain I with the repression motif TDLRLGL. The core motif “LxLxL” is also called EAR repression motif (Ethylene Response Factor (ERF)-Associated amphiphilic Repression), which can recruit a TOPLESS

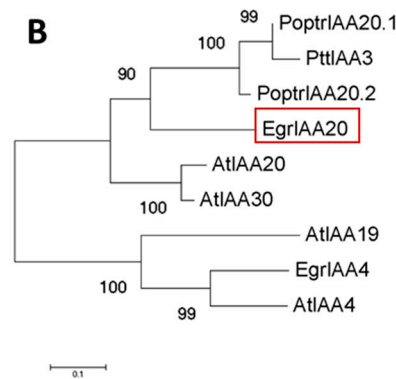


(TPL) transcriptional co-repressor [35]. We thus identified unique *Aux/IAA* proteins, lacking Domain II controlling their turnover rate and with a potential EAR repression motif, present in both *Arabidopsis* and *Eucalyptus*.

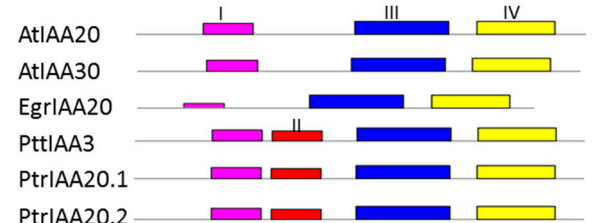
**A**



**B**



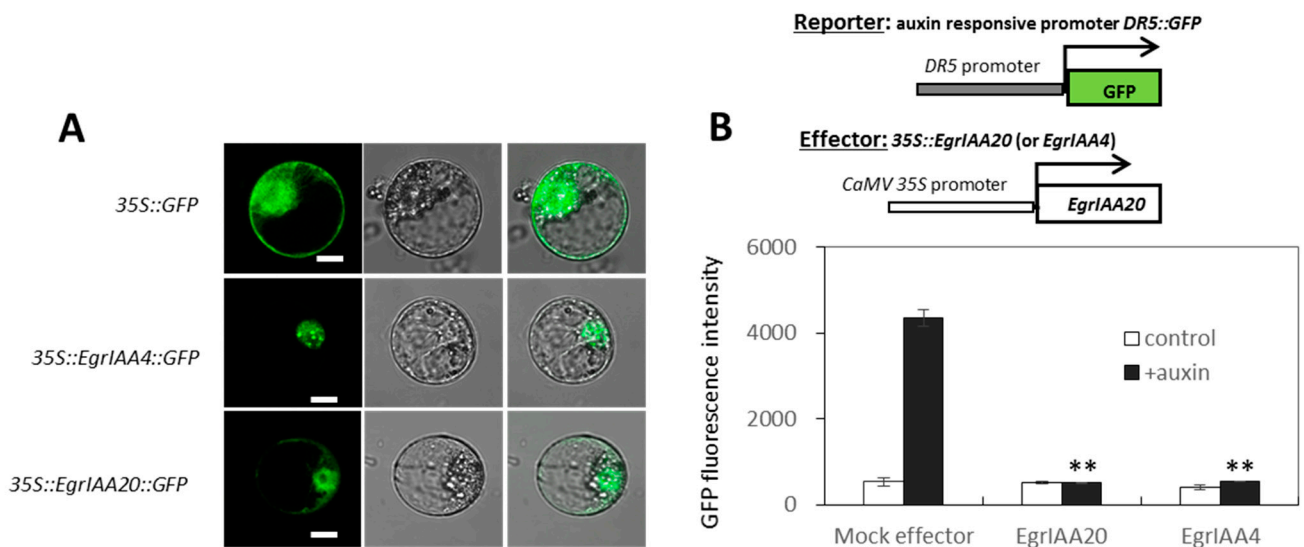
**C**



**Figure 2.** *EgrIAA20* defines a non-canonical *Aux/IAA* protein which lacks highly conserved *Aux/IAA* characteristic Domain II. (A) Sequence Analysis of *EgrIAA20*. Sequence comparison of *EgrIAA20*, putative orthologs from other plants, and paralog of *Aux/IAA* protein family. Conserved residues are shaded in black; gray shading indicates similar residues in at least 5 out of 9 of the sequences. Four conserved domains I, II, III, IV are underlined. Conserved basic residues that putatively function as NLS are indicated by orange frames (Type I) and green frame (Type II). The core repression EAR motif in Domain I is indicated by "LxLxL" on the top of the alignment. (B) The phylogenetic relationship of *EgrIAA20* with its related putative orthologs from *Arabidopsis* and poplar defines a distinct clade with *Arabidopsis* AtIAA20 and AtIAA30, and previously characterized wood-related aspen member PttIAA3. (C) Schemata of protein structures for *EgrIAA20* and its closely related putative orthologs from poplar and Arabidopsis. The domains of *Aux/IAA* proteins were predicted by Pfam (<http://pfam.xfam.org>, accessed on 23 April 2020) and are indicated by different colors; Domain II (red box) was absent for *EgrIAA20*, AtIAA20, and AtIAA30, but present in PtrIAA20.1, PtrIAA20.2, and PttIAA3.

### 2.3. *EgrIAA20* Is Mainly but Not Exclusively Nuclear Localized

In general, *Aux/IAA* proteins have two types of nuclear localization signals (NLSs). The type I NLS is a bipartite structure comprising a conserved KR basic doublet located between Domain I and II and associated with the basic amino acids from Domain II. The type II NLS corresponds to a basic residue-rich region located in Domain IV that resembles the SV40-type NLS (Figure 2A). Protein sequence analysis showed that *EgrIAA20* contains a type II NLS in Domain IV but lacks type I NLS (Figure 2A). We therefore investigated the subcellular localization of *EgrIAA20* using transient expression in tobacco protoplasts. To ensure the exclusive nuclear targeting, we used the canonical *Aux/IAA* *EgrIAA4* that harbors two conserved NLSs as a reference [26]. As expected, fluorescence microscopy analysis confirmed that *EgrIAA4* fused to GFP was exclusively targeted to nuclei (Figure 3A). In cells transformed with GFP alone, a fluorescence signal was found throughout the cell, in both nuclei and cytoplasm, thus confirming the functionality of NLS in *EgrIAA4* (Figure 3A). For *EgrIAA20* fused to GFP, the fluorescence signal was found mostly in nuclei but also in the cytoplasm with a less intense signal than with the GFP alone (Figure 3A). This extended localization of *EgrIAA20* in the extra-nuclear compartment is probably due to the lack of bipartite NLS, as similar wide localization profiles were also reported for tomato non-canonical *Aux/IAA* member *SlIAA32* also lacking the type I NLS [36]. To our knowledge, the function ensured by *IAA20* and its orthologs in the cytosol remains unknown; further experiments are needed to explore its function based on this cytosolic localization. We thus confirmed that *EgrIAA20* is mostly localized in nuclei to mediate auxin response similar to other *Aux/IAA* paralogs.



**Figure 3.** Subcellular localization and repressor activity of auxin response of *EgrIAA20* protein on a synthetic *DR5* promoter. **(A)** Subcellular localization of *EgrIAA20*-GFP fusion protein in BY-2 tobacco protoplasts. The merged images of green fluorescence (left panel) and the corresponding bright-field image (middle panels) are shown in the right panels. Canonical *Aux/IAA* member *EgrIAA4*-GFP serves as a positive control. Scale bar = 10  $\mu$ m. **(B)** Transcriptional repression of auxin response by *EgrIAA20* protein on a synthetic *DR5* promoter. Effector and reporter constructs were co-expressed in tobacco protoplasts in the presence or absence of a synthetic auxin (50  $\mu$ M 2,4-D). A mock effector construct (empty vector) was used as negative control, and *EgrIAA4* construct was used as positive control. Three independent experiments were performed, and similar results were obtained. In each experiment, protoplast transformations were performed in independent biological triplicates. The figure indicates the data from one experiment. Error bars represent SE of mean fluorescence. Significant statistical differences (Student's *t*-test,  $n > 400$ ,  $p < 0.01$ ) from the control are marked with \*\*. The schemes of the reporter and effector constructs are illustrated on the top of the figure.



#### 2.4. *EgrIAA20 Acts as a Potent Transcriptional Repressor of Auxin Response In Vivo*

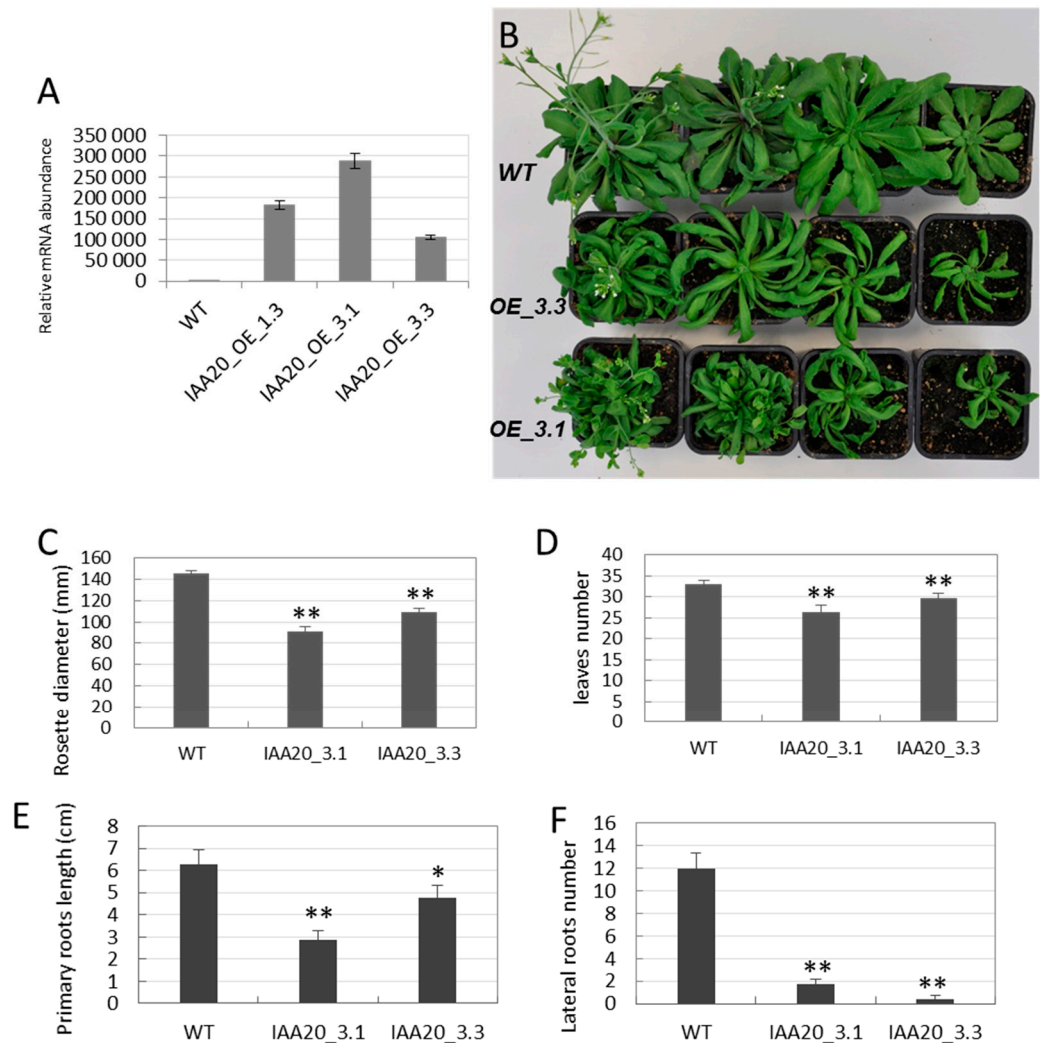
To ensure that *EgrIAA20* effectively affected auxin responses, we evaluated its trans-activation activity using transient expression experiments in tobacco BY-2 protoplasts co-transfected with a *DR5* auxin responsive-promoter controlling *GFP* as a reporter gene (Figure 3B). This *DR5::GFP* construct is a synthetic auxin responsive-promoter composed of several copies of the TGTCTC core motif of auxin response *cis* element [37]. The *GFP* activity of the *DR5::GFP* reporter was highly induced by auxin treatment (50  $\mu$ M 2,4-D) and not affected by the co-transfection with a mock effector (empty plasmid containing the *CaMV* 35S promoter without the effector gene) (Figure 3B). Auxin treatment combined with co-transfecting the canonical *Aux/IAA* protein *EgrIAA4* showed a strong repression of *GFP* fluorescence intensity, as previously demonstrated [26]. Co-transfection with *EgrIAA20* together with auxin treatment also reduced *GFP* fluorescence (12% of residual activity when compared to the mock effector) (Figure 3B). This result revealed that *EgrIAA20* harboring an EAR repression motif was effectively able to mediate auxin response in vivo and acted as a strong transcriptional repressor.

#### 2.5. *Over-Expression of EgrIAA20 in Transgenic Arabidopsis Results in Smaller and Bushier Plants with Floppy Stems*

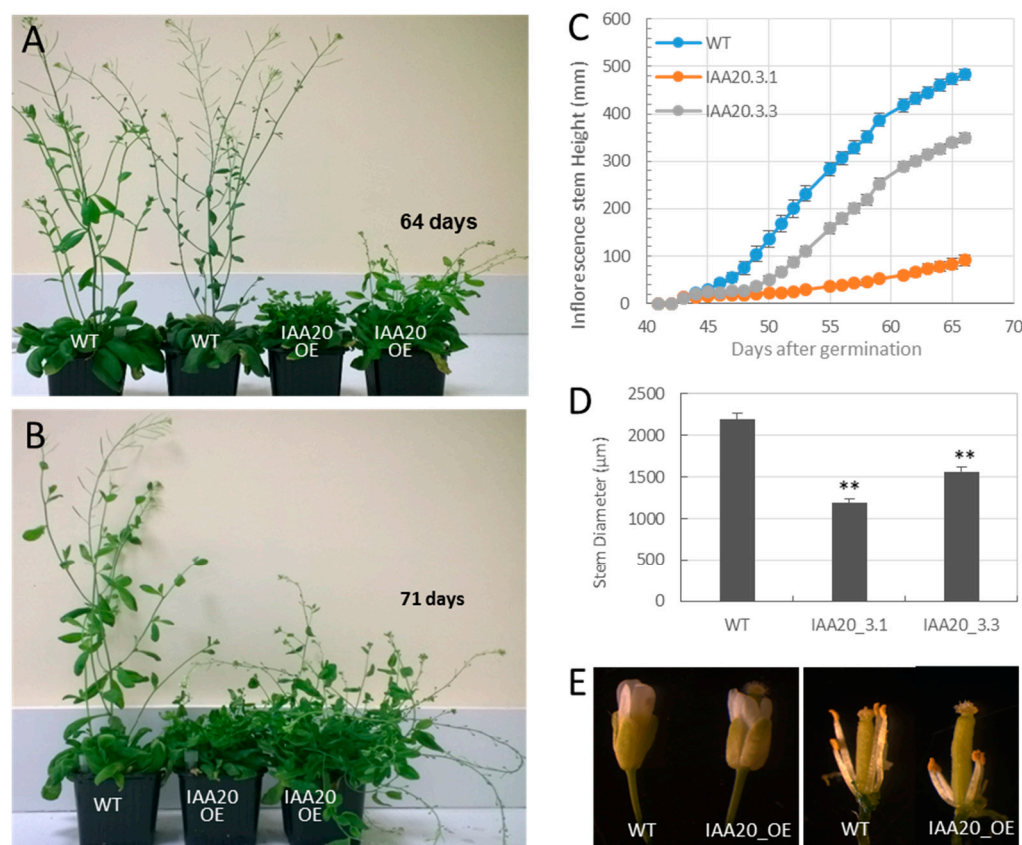
We then investigated the functional role of *EgrIAA20* in secondary xylem differentiation in planta by constitutively and ubiquitously overexpressing *EgrIAA20* in *Arabidopsis*. This approach is similar to previously successful functional analyses of *Eucalyptus* genes involved in the regulation of wood formation [26,30]. In general, the overexpression of the wild-type *Aux/IAA* genes results in no significant phenotypic changes, which were explained by the intrinsic rapid turnover [38]. The effect of *CaMV* 35S promoter-driven *EgrIAA20* was compared to plants transformed with empty vectors. More than 10 independent lines were obtained, all exhibiting reduced plant size compared to mock transformed plants. We chose three phenotypically representative transgenic lines to assess *EgrIAA20* mRNA accumulation. A higher expression level of *EgrIAA20* was detected in all the transgenic lines, with line 3.1 exhibiting the highest expression level compared to lines 1.3 and 3.3 (Figure 4A). We thus selected the strongest (Line 3.1) and the weakest (Line 3.3) expressing lines to characterize the effects of ubiquitous overexpression of *EgrIAA20* in *Arabidopsis*.

Strong phenotypic alterations were triggered by *EgrIAA20* overexpression (Figure 4B), some dependent and other independent from transgene expression levels, each affecting differently specific organs. Leaves and rosette diameter were significantly reduced in transgenic plants compared to mock controls (Figure 4C,D). Rosette leaves also displayed helical twisting and backward rolling in early stages of development (Figure 4B). Using vertically grown seedlings in vitro, root growth was also impacted with *EgrIAA20* overexpressing plants showing significant reduction in primary roots' elongation as well as a limited formation of lateral roots compared to controls (Figure 4E,F). For stems, although bolting time was not affected, their growth rates were significantly reduced in overexpressing plants compared to control plants, especially in the strong line 3.1 (Figure 5A–C). As a consequence of reduced inflorescence stem growth, the *EgrIAA20* overexpressing lines displayed a shorter and bushier morphology (Figure 5A,B), accentuated accordingly with the overexpression levels to 28% and 81% reduction in lines 3.3 and 3.1, respectively (Figure 5C). The inflorescence stems of the transgenic plants were also thinner than the WT. The diameter of the basal part of the inflorescence was significantly reduced in transgenic lines (46% in line 3.1 and 29% in line 3.3; Figure 5D), although the ratio of stem height to width did not significantly differ between genotypes. Remarkably, the *EgrIAA20* overexpressing lines displayed loss of stem rigidity in contrast to the upright stems of controls (Figure 4A,B), suggesting a likely alteration of skeletal support similarly to mutants such as *ifl1* with impaired fiber formation in secondary xylem [39]. Last, the flowering time in *EgrIAA20* overexpressing lines was delayed. The flowers from transgenic plants remained as closed buds for longer (>10 days) and generally failed to reach the anthesis stage. Closer observations showed that the stamens were dramatically shorter than in the control plants

(Figure 5E). The fertility of *EgrIAA20* overexpressing lines was also severely decreased often leading to barren fruits (0 to 20 seeds per transgenic plants compared to more than 200 seeds per WT plants). However, pollination of the transgenic lines using the WT pollen enabled seed formation, indicating that the ovules of *EgrIAA20* overexpressing lines were fertile. Altogether, *EgrIAA20* overexpressing plants showed phenotypic defects similar to *ref3/c4h* mutants that are affected in the lignin biosynthetic *CINNAMATE-4-HYDROXYLASE* gene with reduced stem size, loss of stem rigidity, reduced stamen filament growth, and barren fruits [40]. Thus, our results suggest that secondary xylem cell type and/or their lignin formation could be altered by the overexpression of *EgrIAA20*.



**Figure 4.** *EgrIAA20* overexpressing lines displayed reduced rosette development with helical twisting leaves and reduced root development. (A) *EgrIAA20* mRNA accumulation in three phenotypically representative independent transgenic lines. The *EgrIAA20* overexpressing lines (*IAA20\_OE\_3.1* and *IAA20\_OE\_3.3*) presented reduced rosette development with helices twisting and backward rolling leaf blades and reduced leaf size (B), significantly reduced diameter of rosette (C), significantly reduced leaves number (D), significantly reduced primary root length (E), and greatly reduced lateral roots number (F). Error bars represent standard error. Asterisks indicate values found to be significantly different from the wild-type control (Student's *t*-test,  $n > 10$ ); \*\* indicates  $p < 0.01$ , and \* indicates  $p < 0.05$ .



**Figure 5.** Phenotypes of *EgrIAA20* overexpressing lines. Comparison of plant architecture of aerial parts at 64 days old (A) and 71 days old (B); *EgrIAA20* overexpressing lines displayed a bushier plants architecture compared to the normal plant, with floppy inflorescence stems, in contrast to wild-type control up-right growth inflorescence stems. (C) Growth curve of the *Arabidopsis* primary inflorescence stem. The elongation was reduced in *EgrIAA20* overexpressing lines with significantly reduced inflorescence diameter (D); error bars represent standard error,  $n > 12$ ; stem diameter was measured at the base (<1 cm to the rosette level) when the first silique fully developed. Asterisks indicate values found to be significantly different (student's *t*-test,  $n > 12$ ) from the wild-type control. \*\* indicates  $p < 0.01$ . (E) The stamens of *EgrIAA20* overexpressing lines were dramatically shorter than the wild-type control.

### 2.6. *EgrIAA20* Over-Expression Reduces Fiber-Specific Lignin in *Arabidopsis* Hypocotyls

To investigate the causes of the reduced growth and loss of rigidity due to *EgrIAA20* overexpression, we analyzed the chemical composition of cell walls in ball-milled hypocotyls with pyrolysis gas chromatography/mass spectrometry (Py-GC/MS). The *EgrIAA20* overexpressing lines showed no changes in cell wall polysaccharides compared to WT plants (Table 1). In contrast, a significant reduction in lignin S residue content as well as of other phenolic residues was observed, together with an increase in lignin H residues but no changes in lignin G residue content compared to WT plants (Table 1). Consequently, changes in lignin composition shown by the S/G ratio were largely reduced by the overexpression of *EgrIAA20* (Table 1). Altogether, our results show that the constitutive and ubiquitous overexpression of *EgrIAA20* altered specifically the accumulation of S residues in lignin and not the other cell wall components.

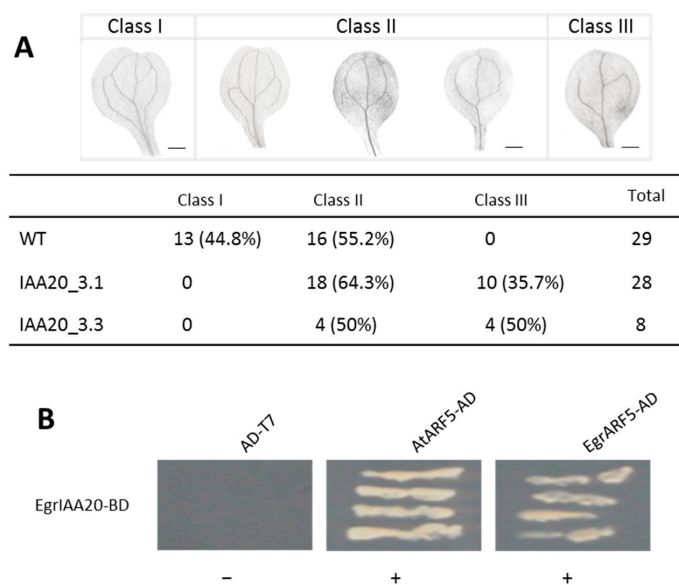
**Table 1.** Chemical composition of wild type and *EgrIAA20\_OE* transgenic *Arabidopsis* hypocotyls by Py-GC/MS analysis (n > 12).

Genotype	Carbohydrate	Guaiacyl	Syringyl	<i>p</i> -Hydroxy-Phenol	Phenolic	Lignin	S/G	C/L
WT	77.5 ± 2.0	8.5 ± 1.2	2.9 ± 0.7	1.7 ± 0.2	1.1 ± 0.1	14.2 ± 2.0	0.34 ± 0.04	5.55 ± 0.82
<i>EgrIAA20</i> <sup>1</sup>	76.9 ± 3.0	8.3 ± 0.8	2.0 ± 0.4 *	1.9 ± 0.2 *	1.0 ± 0.1 *	13.2 ± 2.4	0.25 ± 0.03 **	6.03 ± 1.4

Asterisks indicate values found to be significantly (Student's *t*-test) different from the wild type. \*  $p < 0.05$ , \*\*  $p < 0.01$ . <sup>1</sup>: results of pooled samples from three independent batches of *EgrIAA20\_OE* lines 3.1 and 3.3 (5 plants for each line each batch).

### 2.7. *EgrIAA20* Overexpression Alters Primary Xylem Development in Cotyledons

To assess the role of *EgrIAA20* overexpression on primary xylem development, we investigate the vascular patterning of cotyledons by measuring the number of secondary vein loops from the mid-vein as previously reported [41]. In control seedlings, cotyledons' vein patterns belonged to classes I and II, and 45% of them exhibited a more complex pattern (Class I with four complete loops, Class II with at least two complete loops) (Figure 6A). In contrast, the cotyledons' vein patterns of all *EgrIAA20* overexpressing lines exclusively belonged to classes II and III (class III showing no complete vascular loop), and none of them belonged to class I. This result indicated that the overexpression of *EgrIAA20* prevented the completion of vascular patterning during primary growth.



**Figure 6.** Over-expressing *EgrIAA20* in *Arabidopsis* regulated vascular patterning in cotyledons through the interaction with AtARF5. **(A)** *EgrIAA20* overexpression impaired vascular patterning in cotyledons. The vascular patterning was assessed by the number of secondary vein loops originating from the mid-vein. Class I presents four complete vascular loops; class II presents partially incomplete venation with at least two loops; class III presents entirely incompletely venation with no entire venation loop. In wild type seedlings, all the cotyledon venation patterns belonged to classes I and II, and 45% of them exhibited a more complex pattern (class I with four complete loops). In contrast, cotyledon venation patterns of all the transgenic plants belonged to classes II and III; none of them showed the complete vascular patterning with four loops (class I). Values in bracket indicate the percentage contribution of each class. The scale bars represent 0.5 mm. **(B)** *EgrIAA20* interacts with *Arabidopsis* ARF5/MP (AtARF5) and its potential ortholog in *Eucalyptus* EgrARF5 in yeast-2-hybrid assay. The *EgrIAA20* and AtARF/EgrARF5 proteins were fused with GAL4 DNA-binding domain (BD) and a GAL4 activation domain (AD), respectively. Yeast of co-transformed *EgrIAA20-BD* and *AtARF5-AD* or *EgrARF5-AD* grew on quadruple dropout medium lacking leucine, tryptophan, histidine, and adenine (THLA), and then scratched again on a TLHA plate; AD-T7 were used as negative controls.



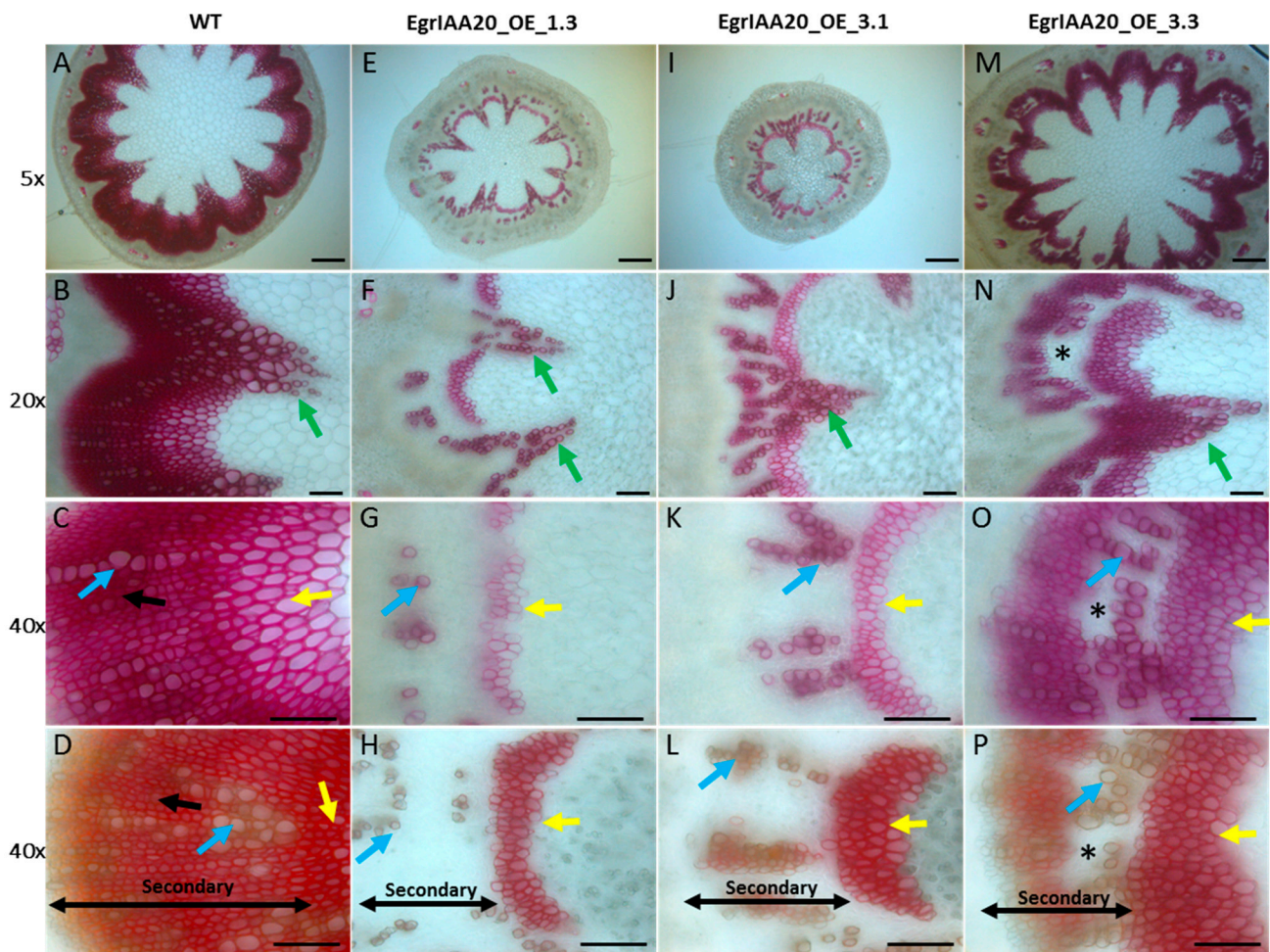
This incomplete vascular patterning in cotyledon phenotype mimics that of *Arabidopsis* ARF5 loss-of-function mutant *monopteros/arf5*. Considering that *AtARF5* was reported to control this development process in *Arabidopsis* [19,42,43], we tested if *EgrIAA20* interacts with *AtARF5* and its potential *Eucalyptus* ortholog *EgrARF5*, using yeast two-hybrid. The results confirmed that there was a strong interaction between *EgrIAA20* and *AtARF5*, as well as between *EgrIAA20* and *EgrARF5* (Figure 6B), which further supported the orthologous relationship between *Arabidopsis* and *Eucalyptus* defined by our phylogenetic analysis [26,44]. We thus also confirmed that the interaction between *Aux/IAA* and ARF is conserved between *Arabidopsis* and *Eucalyptus*.

### 2.8. *EgrIAA20* Overexpression Specifically Represses Xylary Fibers of Secondary Xylem in *Arabidopsis* Stem and Hypocotyl

The secondary growth producing secondary xylem cells occurs naturally in adult plants in the base part of the inflorescence stem, root and hypocotyl [27]. To investigate if the overexpression of *EgrIAA20* also impacted secondary xylem cells' development, we particularly examined the basal part of inflorescence stems (<1 cm to the rosette level) as well as hypocotyls of adult *Arabidopsis* plants using a short-day condition to stimulate secondary growth [28]. To detect primary and secondary xylem (all containing lignified SCW) in cross-sections easily, we used phloroglucinol/HCl or the Wiesner test that specifically stains the lignified SCW into red-purple by reacting with lignin coniferaldehyde residues [45], or the Mäule test that allows the staining of the SCW of the tracheary element in brown and fiber in red [46]. Cell types present in the primary xylem cells include tracheary elements, xylary fibers, and interfascicular fibers (Fiber I) [47], whereas secondary xylem cells mainly include secondary tracheary elements (Vessel II) and secondary fibers (Fiber II) (Figure 7C and 7D). In WT plants, cell walls of all primary and secondary xylem cells were all stained positively by the Wiesner test (Figure 7A), with some tracheary elements exhibiting more intense staining (Figure 7C, indicated by the blue arrow). In contrast, the intensity of the Wiesner staining was dramatically reduced in *EgrIAA20* overexpressing lines (Figure 7E–L), indicating a modified lignification with a reduced coniferaldehyde incorporation [45]. Additionally, the strongest overexpressing lines 1.3 and 3.1 also showed a complete absence of staining in secondary fibers but not in the primary fibers (Figure 7E–L). Only a few isolated stained cells were observed in the secondary growth region (both interfascicular and vascular bundle regions) (Figure 7E–L). In order to confirm the cell type identity, complementary staining was performed using the Mäule test to discriminate fiber cells stained in red due to SCW enriched lignin S residue, from tracheary element cells stained in brown due to SCW enriched in lignin G residue (Figure 7D). The isolated cells in the secondary region of transgenic plants were all stained in brown, confirming their tracheary element identity (Figure 7H,L). Moreover, cells' in-between tracheary elements in the secondary xylem zone of *EgrIAA20* overexpressing lines (Figure 7H) did not exhibit the red staining observed in WT plants (Figure D), thus confirming that S residue incorporation was affected in transgenic plants. The moderate level of overexpression of line 3.3 exhibited an intermediate phenotype between WT and strongly overexpressing lines (Figure 7M–P).

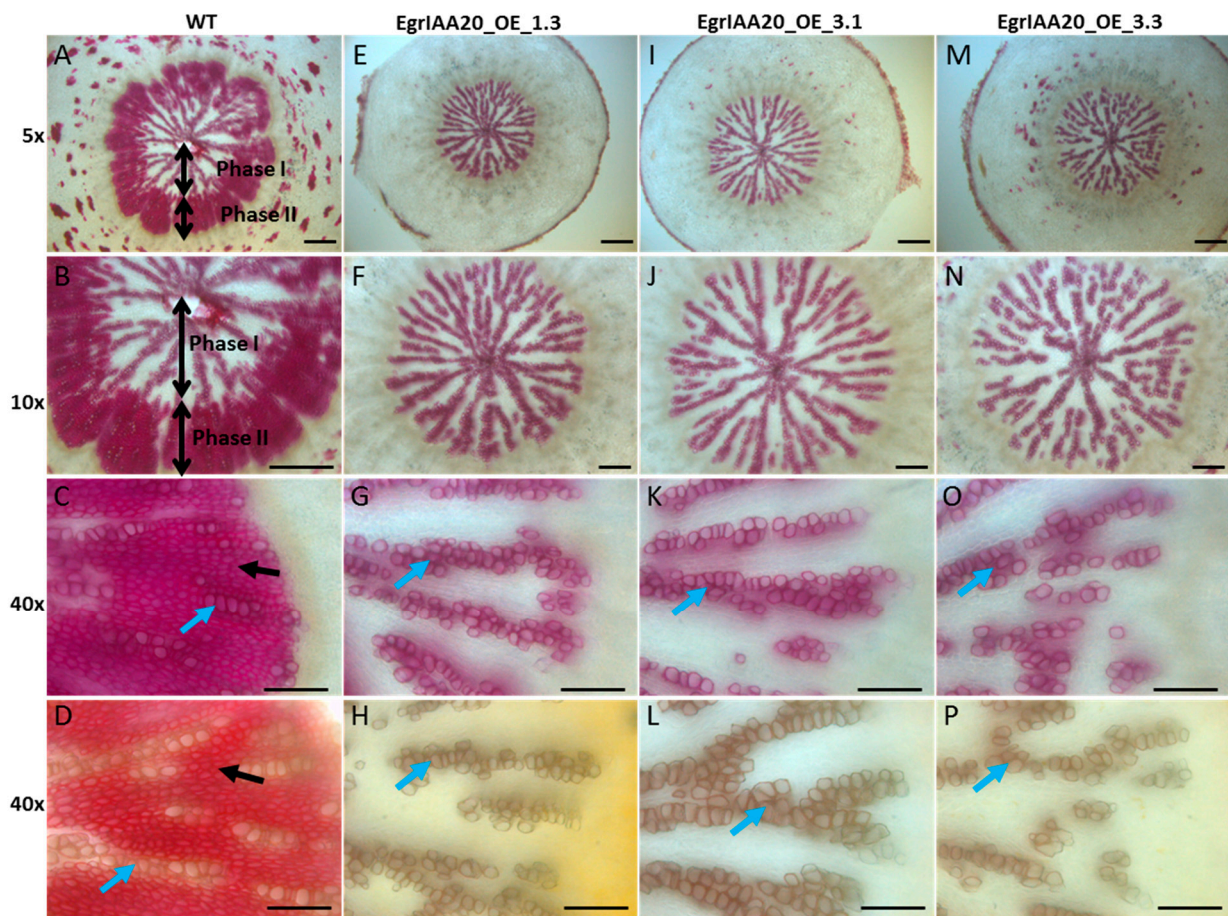
Further analyses on adult plant hypocotyls were performed due to its strong similarity with wood in perennial species such as poplar [28]. In WT plants, we confirmed the normal biphasic development of secondary xylem in short-day growth conditions with an early phase (Phase I), in which lignified tracheary elements are formed but not fibers (Figure 8A,B), followed by a late phase (Phase II), in which both lignified fibers and tracheary elements are formed [28]. However, secondary xylem lignification was greatly altered in *EgrIAA20* overexpressing lines: the Wiesner test staining showed normal staining for tracheary elements in Phase I and II in WT, but an absence of staining of fibers in Phase II of transgenic lines, independently of the *EgrIAA20*-overexpression level (Figure 8E–P). Complementary staining using the Mäule test confirmed the presence of tracheary elements enriched in G lignin, but also the complete absence of S lignin enriched secondary fibers in *EgrIAA20* overexpressing lines (Figure 8H,L,P). Overall, our results revealed that the lignifi-

cation of secondary fibers is specifically reduced by the dose-dependent overexpression of *EgrIAA20* as well as uncoupled from the lignification of tracheary elements.



**Figure 7.** *EgrIAA20* overexpression specifically repressed secondary fibers but not primary fibers nor secondary vessels in *Arabidopsis* inflorescence stems. Left panel (A–D) shows cross section of inflorescence stems at the basal part of wild-type control, <1 cm to the rosette level, 62-day-old plants; 5 $\times$ , 20 $\times$ , and 40 $\times$  objective image for the first row, second and third row, respectively, using Phloroglucinol-HCl staining, which stains the lignin polymers in the SCW of xylem cells into red-purple; the last row was 40 $\times$  objective observation image stained with Maüle method, which stains the fiber cells into bright red color due to the syringyl unit (S unit) of lignin, indicated by yellow and black arrows, and stains the vessels in to brown due to the G unit of lignin, indicated by blue arrow. The corresponding cross sections from three independent *EgrIAA20* overexpressing lines are shown in the middle and right panels: *EgrIAA20\_OE\_1.3* (strong line, middle left panel (E–H)), *EgrIAA20\_OE\_3.1* (strong line, middle right panel (I–L)), *EgrIAA20\_OE\_3.3* (weak line, right panel, (M–P)). Green arrows indicate primary xylem cells in fascicular bundles; blue arrows indicate secondary vessel cells (vessel II); yellow arrows indicate primary fiber cells (fiber I); black arrows indicate secondary fiber cells (Fiber II). \* indicates staining gaps dispersed in the secondary growth region. The first to third row used Phloroglucinol-HCl staining; the bottom row used Maüle staining methods. The secondary growth regions are indicated by double-arrow lines in the images stained by Maüle. For 5 $\times$  objective observation images, scale bar = 200  $\mu$ m, and for 20 $\times$  and 40 $\times$  objective observation images, scale bar = 50  $\mu$ m.



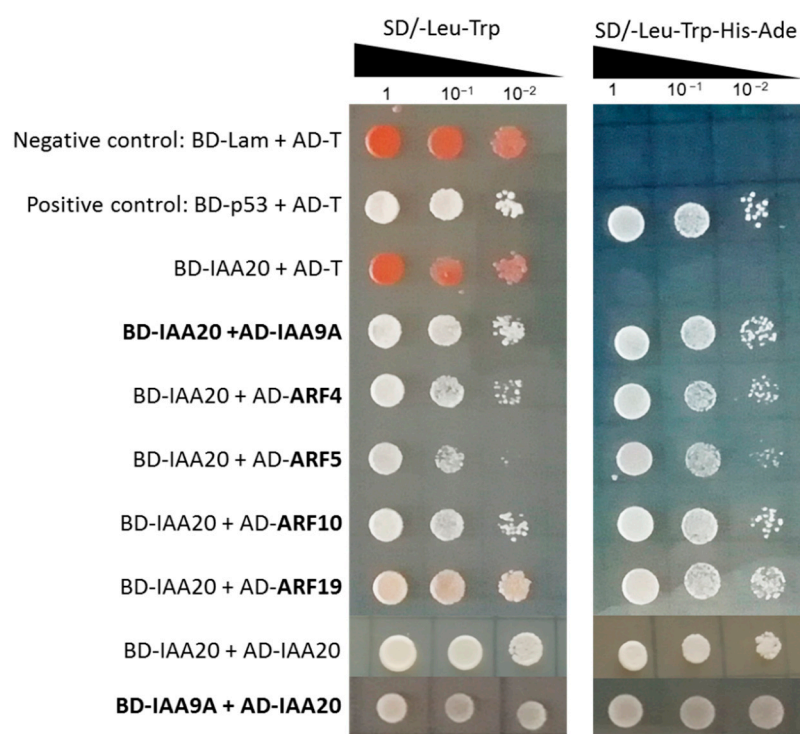


**Figure 8.** *EgrIAA20* overexpression specifically inhibits the lignification of fibers in secondary but not primary fibers nor any tracheary elements in *Arabidopsis* hypocotyl. Left panel (A–D) shows cross section of hypocotyl of wild-type control, 5×, 10×, and 40× objective image for the first row, second and third row, respectively, using Phloroglucinol-HCl staining, which stains the lignin polymers in the SCW of xylem cells into red-purple; the last row was 40× objective observation image stained with Maile method, which stains the fiber cells into bright red color, indicated by black arrow, and stains the vessels in to brown, indicated by blue arrow. The corresponding hypocotyl cross sections from three independent *EgrIAA20* overexpressing lines are shown in the middle and right panels: *EgrIAA20\_OE\_1.3* (strong line, middle left panel, (E–H)), *EgrIAA20\_OE\_3.1* (strong line, middle right panel, (I–L)), *EgrIAA20\_OE\_3.3* (weak line, right panel, (M–P)). Phase I growth region and phase II growth region are indicated by double-arrow lines in wild type cross section. Blue arrows indicate secondary vessel cells (vessel II); black arrows indicate secondary fiber cells (Fiber II). The first to third row used Phloroglucinol-HCl staining, the bottom row used Maile staining methods. For 5× and 10× images, scale bar = 200 μm, and for 40× images, scale bar = 50 μm.

### 2.9. Identification of *EgrIAA9A* as the Main Interacting Protein of *EgrIAA20*

To deeper our understanding of the molecular modality of the *EgrIAA20* regulation mechanism during wood formation in trees, we aimed to identify its interacting protein partners. To identify proteins that interact with *EgrIAA20* during wood formation in *Eucalyptus*, we screened a *Eucalyptus* wood yeast two hybrid (Y2H) library [30] with *EgrIAA20* as a bait. We successfully sequenced 35 clones among the 47 positive clones of Y2H (Supplementary Table S1), showing that the most likely interacting candidate was *EgrIAA9A*, which represented 10 out of 35 total sequenced colonies (29%). This interaction was confirmed by targeted Y2H between *EgrIAA20* and *EgrIAA9A*, independently of *EgrIAA20* being used as bait or prey (Figure 9). Three other *Aux/IAA* genes were also identified: *IAA1*, *IAA11*, and *IAA16* (two clones) with lower representation. Two out

of 35 clones corresponded to a *Eucalyptus* ortholog of *SGT1* (*SUPPRESSOR OF THE G2 ALLELE OF SKP1*), previously reported to be a positive regulator of auxin signaling and disease resistance in *Arabidopsis* [48]. Strikingly, no ARF proteins were identified from our Y2H screening although they are well-known interacting proteins of *Aux/IAA* proteins and were identified above as interacting with *EgrIAA20* for EgrARF5 (Figure 6B). We therefore performed a targeted Y2H to determine any interaction between *Eucalyptus* wood-expressing-ARF proteins [44] and *EgrIAA20*. The results showed that *EgrIAA20* interacted with all four wood-related EgrARFs: EgrARF4, EgrARF5, EgrARF10, and EgrARF19 (Figure 9). Our targeted Y2H results also showed that *EgrIAA20* was able to form homodimers with *EgrIAA20*, and confirmed the strong interaction with *EgrIAA9A* independently of its position as bait or prey (Figure 9). Overall, our results suggest that *EgrIAA20* may form a transcription regulator complex preferentially with *EgrIAA9A* as well as other *Aux/IAA* but also with ARFs.



**Figure 9.** Yeast-two-hybrid analysis of protein interactions between *EgrIAA20* and xylem expressing EgrARFs and EgrIAA proteins. The *EgrIAA20* and EgrARF/EgrIAA proteins were fused with Gal4 DNA-binding domain (BD) and a GAL4 activation domain (AD), respectively. The interaction between BD-Lam and AD-T was used as negative control, while the interaction between BD-p53 and AD-T was used as positive control. Yeast cells were inoculated on selective medium in a 10-fold gradient dilution. SD/-Leu-Trp: double dropout medium lacking leucine and tryptophan; SD/-Leu-Trp-His-Ade: quadruple dropout medium lacking leucine, tryptophan, histidine, and adenine.

### 3. Discussion

In this study, we performed the functional characterization of a novel non-canonical *Aux/IAA* member, *EgrIAA20*, which is preferentially expressed in wood cambium. This non-canonical *Aux/IAA* was characterized by the absence of Domain II, pointing to a lower turnover compared to the canonical short-lived paralogs, as well as an EAR transcriptional-repression motif on its Domain I (Figure 2). We experimentally confirmed these predictions by showing that *EgrIAA20* was mostly targeted to nuclei and presented a repressor activity on auxin-responsive transcription (Figure 3). As genetic engineering is strongly hampered in eucalypts species, we performed gain-of-function studies of *Eucalyptus* genes directly in *Arabidopsis* as previously achieved [26,30]. We ensured the biological relevance of the

heterologous *EgrIAA20* overexpression in *Arabidopsis* by confirming the conservation of both its sequence structure, both lacking Domain II and presenting the EAR motif (Figure 2), and its interactions with *ARF5* independently of the species used (Figure 6B). This result thus confirmed that true orthologs exist between *Arabidopsis* and *Eucalyptus*, unlike other angiosperms which have lost these specific paralogs (Figure 3). Homologous gain-of-function experiments in *Arabidopsis* plants by overexpressing *AtIAA20* [49], the ortholog of *EgrIAA20* [26], showed defects in vascular patterning in cotyledon [49] similar to the severe defects observed in *EgrIAA20* overexpressors (Figure 6). However, previous studies did not evaluate the impact of the genetic modulation of *AtIAA20* on secondary xylem formation or lignification as we performed in the herein study. Although the sweeping and pendulous inflorescence stem phenotype was also reported in the *AtIAA20* overexpression line, it was however interpreted as an impacted gravitropic response [49]. However, in light of our results, it is likely that these plants also lacked proper secondary fiber development to ensure the skeletal support of the inflorescence stem. The ectopic overexpression of *EgrIAA20* in *Arabidopsis* specifically uncoupled the development of the different secondary xylem cell types by inhibiting the lignification of secondary fibers but not of tracheary elements. This was only observed during secondary growth of both hypocotyls and basal parts of inflorescence stems, thereby showing that the development of fibers and tracheary elements formed in secondary xylem not only differ from primary xylem but also depend on distinct auxin-dependent regulations.

Furthermore, it would be interesting to investigate the mechanism of how non-canonical *Aux/IAA*, lacking the conserved degron Domain II, regulates auxin signaling. To this end, it is worth noting that another lacking Domain II non-canonical *Aux/IAA* member, *IAA33*, was shown to compete with canonical *Aux/IAA* member *IAA5* for binding *ARF10/16* to protect the latter from *IAA5*-mediated inhibition. Thus, *IAA33* maintains root distal stem cell identity and negatively regulates auxin signaling [50]. Whether *EgrIAA20* acts through a similar mechanism deserves further attention.

Wood and fiber constitute an extremely important renewable biological resource. The efficient exploitation of this lignocellulosic biomass requires a full understanding of the wood formation process. Our results represent a timely contribution to the rapidly expanding knowledge of the wood formation regulation, especially the regulation of wood cells' specification and maturation. The genetic engineering of wood biomass by uncoupling the development of the different cell types has been a target of recent studies, specifically to reduce lignin in fibers without altering tracheary elements. Many of these strategies use lignin monomer loss-of-function mutants that are complemented using a tracheary element specific promoter such as *ProVND6* [51] or *ProSNBE* [52,53]. We herein present another strategy to uncouple the lignification of fibers from tracheary elements by directly overexpressing *EgrIAA20*. Although our strategy enabled us to uncouple cell type specific lignification using one gain-of-function construct instead of combining loss-of-function with cell specific complementation, our strategy needs further refinement to control the level of transgene expression. Overall, our novel discovery not only opens new avenues to deepen our understanding of the mechanisms transducing auxin for cell type specific regulation in wood but also provides a proof-of-concept of a novel technological strategy to uncouple wood cell types genetically to modify lignin composition in fibers without affecting water and solutes' conduction, thus ultimately to improve wood composition for its better end-uses.

## 4. Materials and Methods

### 4.1. Plant Material and Growth Conditions

The provenance and preparation of all *Eucalyptus* organs and tissues were as described in [54]. For phenotype characterization, *Arabidopsis thaliana* ecotype Col-0 plants (wildtype and transgenic plants) were sown on soil, vernalized for 5 days at 4 °C, then were grown in a growth chamber, under 8h-day/16h-night short days condition to promote secondary growth. The growth chamber was set to 22 °C-day/20 °C-night temperature, 70% relative



humidity, 200  $\mu\text{mol photons m}^{-1} \text{s}^{-1}$  light intensity (intense luminosity). The plants were watered every two days and fertilized weekly since 4-weeks-old till harvest. Seeds for in vitro culture were surface-sterilized for 1 min in 70% ethanol, 10 min in 25% bleach, rinsed five times in sterile water, and planted on MS medium containing 1.0% sucrose solidified with 1% agar.

#### 4.2. Microfluidic Quantitative PCR Expression Analysis

Mature leaves, young roots, young stems, seedling, floral bud, fruits (capsules) were collected from *Eucalyptus globulus*. Shoot tips, secondary xylem, secondary phloem, and a cambium-enriched fraction were collected from 7-year-old trees of *Eucalyptus Gundal* hybrids (*Eucalyptus gunnii*  $\times$  *Eucalyptus dalrympleana*, genotype 850645) grown in south-west France (Longage) by the FCBA. Cambium fractions were also collected from a 25-year-old *Gundal* hybrid (genotype 821290). Vascular tissues were sampled as previously described [55]. Briefly, a piece of bark was removed from the trunk; cambium-enriched fractions which include cambium and some xylem and phloem mother cells were obtained by gently sweeping the inner side surface of the bark and outside surface of xylem (debarked trunk). After collecting the cambium, the procedure involved scarping the inner side of bark to obtain phloem, and scarping the outer side of debarked trunk to obtain developing secondary xylem. Gene-specific primers for quantitative PCR were designed using QuantPrime [56] with default parameters (Supplementary Table S2). Transcript abundance was assessed by microfluidic qPCR using the BioMark. The reference genes were selected as described in [54].

#### 4.3. Gene Cloning, Vector Construction, and Genetic Transformation

The coding sequence of *EgrIAA20* (*Eucgr.K00561*), *EgrARF5* (*Eucgr.F02090*) was amplified from *Eucalyptus grandis* cDNA using high fidelity PhusionTaq (Thermo Fisher Scientific, Waltham, MA, USA), and the amplicons were inserted into the pDONR207 vector using the BP clonase II (Invitrogen, Carlsbad, CA, USA) or into the pENTR/D-TOPO vector (Invitrogen, Carlsbad, CA, USA). LR clonase II was then used to generate the different destination vectors, including pFAST-G02 [57] for over-expression of *pro35S::EgrIAA20* in *Arabidopsis thaliana*, and pGBG-BD-GTW and pGAD-AD-GTW (kindly provided by Laurent Deslandes, LIPM, Auzeville Tolosane, France) for yeast two-hybrid (Y2H). For genetic transformation, we transformed the destination vector *pFAST-EgrIAA20* into *Agrobacterium tumefaciens* strain GV3101. Then, the *A. thaliana* ecotype Col-0 was transformed using *A. tumefaciens* and the floral dip method [58]. Primary transformants were selected using a fluorescent stereomicroscope with GFP filters which is able to detect the GFP fluorescent marker present in the pFAST-G02 binary vector that is expressed specifically in transformed seeds, as described by Shimada et al. [57]. Ten independent *EgrIAA20* over-expressing T1 *Arabidopsis* lines were generated; all presented similar phenotypes showing reduced plant size as compared to the control. For detailed characterization of their phenotypes, three independent T2 lines (two strong lines and one weak line) were selected according to their T1 phenotypes and the transgene transcript abundance (determined in leaves by RT-qPCR) (Figure 4A).

#### 4.4. Transient Expression in Protoplasts for Subcellular Localization and Transactivation Assay

Protoplasts for transfection were obtained from suspension-cultured tobacco (*Nicotiana tabacum*) BY-2 cells according to the method described by Yu et al. [26]. Protoplasts were transfected by a modified polyethylene glycol method as described by Abel and Theologis [59]. For nuclear localization of *EgrIAA20* and *EgrIAA4*, the full-length cDNAs were fused in frame at the C-terminus with GFP in the pK7FWG2.0 vector (Karimi et al. 2002) under the control of the *CaMV 35S* promoter. Transfected protoplasts were incubated for 16 h at 25 °C and examined for GFP fluorescence signals using a Leica TCS SP2 laser scanning confocal microscope. Images were obtained with a x40 water immersion objective. For co-transfection assays, the full-length cDNAs of the selected *Aux/IAA* were cloned

into pGreen vector under the *CaMV 35S* promoter to create the effector constructs. The reporter constructs used a synthetic auxin-responsive promoter *DR5* followed by the *GFP* reporter gene. For co-transfection assays, aliquots of protoplasts ( $0.5 \times 10^6$ ) were transformed with 10  $\mu\text{g}$  of the reporter vector, containing the *DR5* synthetic promoter fused to the *GFP* reporter gene, and with either the effector vector of *EgrIAA20* under the *CaMV 35S* promoter or the empty vector as mock treatment. After 16 h incubation, *GFP* expression was quantified by flow cytometry (FACS calibur II, BD Bioscience), and the data were analyzed using Cell Quest software (BD Biosciences, San Jose, CA, USA). Transfection assays were performed in three independent replicates, and 400–1000 protoplasts were gated for each sample. *GFP* fluorescence corresponds to the average fluorescence intensity of the protoplasts' population after subtraction of auto-fluorescence determined with non-transformed protoplasts. A total of 50  $\mu\text{M}$  2,4-D were used for auxin treatment.

#### 4.5. Microscopy Analysis

The basal end (<1 cm) of *Arabidopsis* inflorescence stems and hypocotyls were harvested at indicated dates, and then stored in 70% ethanol. The 62-day-old plants (2 months) grown on a short day condition present fully developed inflorescence stems, and the first siliques were clearly observed. The cross sections (100  $\mu\text{m}$  thick) were prepared using vibratome Leica VT1000 S (Leica, Paris, France). Lignin polymers are the characteristic components of the secondary cell wall (SCW) and are normally absent from the primary cell wall. Therefore, we used lignin deposition detection techniques to screen for the SCW phenotype. Cross sections of the inflorescence stem and hypocotyl were stained with phloroglucinol-HCl, which stains specifically the lignin polymer unit coniferaldehyde and p-coumaraldehyde in the SCW giving a violet-red color. Phloroglucinol-HCl was directly applied on the slide, and the sections were observed under a bright-field inverted microscope (DM IRBE; Leica). Images were recorded with a CCD camera (DFC300 FX; Leica) directly connected to microscope.

Mäule staining was used to distinguish vessel cells from fiber cells. Firstly, the sections were incubated in 1% sodium citrate solution of  $\text{KMnO}_4$  for 3 min. The stained sections were washed in abundant water for 5 min. Later, bleaching was conducted using HCl 6N for a few seconds. As soon as the sections became transparent, sections were taken out and washed in water for another 5 min. At last, 5%  $\text{NaHCO}_3$  was directly applied on the slices for observation under a bright-field inverted microscope (DM IRBE; Leica). Images were recorded with a CCD camera (DFC300 FX; Leica) directly connected to microscope.

#### 4.6. Pyrolysis Analysis

*EgrIAA20* overexpression transgenic (*EgrIAA20\_OE* lines 3.1 & 3.3) and wild type *Arabidopsis* were planted in 3 different batches in short days (5 plants for each line in one batch), and 10 days different between each batch. The hypocotyls were harvested in liquid nitrogen when the first silique was fully developed. Then, the hypocotyls were freeze-dried and ball-milled (MM400; Retsch) at 30 Hz in stainless steel jars (1.5 mL) for 2 min with one ball (diameter of 7 mm). A total of 50–70  $\mu\text{g}$  (XP6, Mettler-Toledo, Switzerland) of powder were transferred to auto sampler containers (Eco-cup SF, Frontier Laboratories, Japan) for the Py-GC/MS. The sample was carried to the oven pyrolyzer by an auto sampler (PY-2020iD and AS-1020E, FrontierLabs, Nagoya, Japan) and analyzed by a GC/MS system (Agilent, 7890A/5975C, Agilent Technologies AB, Sweden). The pyrolysis oven was set to 450  $^\circ\text{C}$ , the interface to 340  $^\circ\text{C}$ , and the injector to 320  $^\circ\text{C}$ . The pyrolysate was separated on a capillary column with a length of 30 m, diameter of 250  $\mu\text{m}$ , and film thickness of 25  $\mu\text{m}$  (JandW DB-5; Agilent Technologies, Stockholm, Sweden). The gas chromatography oven temperature program started at 40  $^\circ\text{C}$ , followed by a temperature ramp of 32  $^\circ\text{C}/\text{min}$  to 100  $^\circ\text{C}$ , 6  $^\circ\text{C}/\text{min}$  to 118.75  $^\circ\text{C}$ , 15  $^\circ\text{C}/\text{min}$  to 250  $^\circ\text{C}$ , and 32  $^\circ\text{C}/\text{min}$  to 320  $^\circ\text{C}$ . Total run time was 19 min, and full-scan spectra were recorded in the range of a 35 to 250 mass-to-charge ratio. Data processing, including peak detection, integration, normalization, and identification, was conducted as described by Gerber et al. [60].

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

#### 4.7. Yeast Two-Hybrid (Y2H) Screening

The Y2H library of cambium and developing xylem from field-grown *Eucalyptus globulus* (Labill.) was constructed as described by Soler et al. [30] using the Make Your Own Mate-&-Plate Library System (Clontech, Mountain View, CA, USA). RNAs were isolated from a mixture of scraping cambium and xylem from juvenile to mature trees. The complexity of the eucalypt xylem library obtained was c.  $1 \times 10^6$  clones  $\text{mL}^{-1}$ , with a density of c.  $5 \times 10^8$  cells.

For the library screening, the full length CDS of *EgrIAA20* was inserted into the pGBG-BD-GTW vector, transformed in Y2HGold yeast cells, and used as bait to screen the cDNA library. Y2H screening was performed using the mating protocol described for the Matchmaker Gold Yeast Two-Hybrid System (Clontech) with QDO medium (Quadruple synthetic drop-out autotrophic medium lacking tryptophan, leucine, adenine, and histidine). A total of 57 colonies were obtained and sequenced using primers designed for the pGADT7 vector (Clontech). Plasmids were isolated with the Easy Yeast Plasmid Isolation Kit (Clontech). Targeted Y2H assays to test the repeatability and the specificity of the interactions were performed using QDO/X/A medium (QDO medium supplemented with Aureobaisin A and 5-bromo-4-chloro-3-indolyl  $\alpha$ -Dgalactopyranoside (X-a-Gal)) and DDO medium (double synthetic drop-out auxotrophic medium lacking tryptophan and leucine), as specified by the manufacturer (Clontech, Mountain View, CA, USA).

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms23095068/s1>.

**Author Contributions:** Conceptualization, H.C.-W.; Funding acquisition, H.Y., J.G.-P. and H.C.-W.; Investigation, H.Y., M.L., Z.Z. and H.C.-W.; Methodology, H.Y., Z.Z., A.W., F.M., E.P. and H.C.-W.; Supervision, J.G.-P. and H.C.-W.; Validation, H.Y., M.L. and H.C.-W.; Writing—original draft, H.C.-W.; Writing—review and editing, H.Y., M.L., Z.Z., A.W., F.M., E.P., J.G.-P. and H.C.-W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the Centre National pour la Recherche Scientifique (CNRS), the University Paul Sabatier Toulouse III (UPS), the French Laboratory of Excellence project “TULIP” (ANR-10-LABX-41; ANR-11-IDEX-0002-02). H.Y. was funded by Ph.D. grants from the China Scholarship Council; M.L. was funded by an exchange student grant from the China Scholarship Council.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** Thanks to the Bioinfo Genotoul Platform (<http://bioinfo.genotoul.fr>, accessed from September 2011 to April 2021 for access to resources and to the Genome and transcriptome Platform (<http://get.genotoul.fr/>, accessed from September 2012 to April 2019) for advice and technical assistance with high-throughput Biomark Fluidigm qRT-PCR amplifications. Thanks to Y. Martinez (FR3450) for assistance with microscopy analysis (Plateforme Imagerie TRI [tri.ups-tlse.fr/](http://tri.ups-tlse.fr/), accessed from September 2014 to September 2020).

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

## References

1. Boudet, A.M.; Kajita, S.; Grima-Pettenati, J.; Goffner, D. Lignins and lignocellulosics: A better control of synthesis for new and improved uses. *Trends Plant Sci.* **2003**, *8*, 576–581. [[CrossRef](#)] [[PubMed](#)]
2. Ye, Z.H.; Zhong, R. Molecular control of wood formation in trees. *J. Exp. Bot.* **2015**, *66*, 4119–4131. [[CrossRef](#)] [[PubMed](#)]



3. Barros, J.; Serk, H.; Granlund, I.; Pesquet, E. The cell biology of lignification in higher plants. *Ann. Bot. London* **2015**, *115*, 1053–1074. [[CrossRef](#)]
4. Boerjan, W.; Ralph, J.; Baucher, M. Lignin biosynthesis. *Annu. Rev. Plant Biol.* **2003**, *54*, 519–546. [[CrossRef](#)] [[PubMed](#)]
5. Pesquet, E.; Wagner, A.; Grabber, J.H. Cell culture systems: Invaluable tools to investigate lignin formation and cell wall properties. *Curr. Opin. Biotech.* **2019**, *56*, 215–222. [[CrossRef](#)]
6. Brackmann, K.; Qi, J.; Gebert, M.; Jouannet, V.; Schlamp, T.; Grunwald, K.; Wallner, E.S.; Novikova, D.D.; Levitsky, V.G.; Agusti, J.; et al. Spatial specificity of auxin responses coordinates wood formation. *Nat. Commun.* **2018**, *9*, 875. [[CrossRef](#)]
7. De Zio, E.; Montagnoli, A.; Karady, M.; Terzaghi, M.; Sferra, G.; Antoniadis, I.; Scippa, G.S.; Ljung, K.; Chiatante, D.; Trupiano, D. Reaction Wood Anatomical Traits and Hormonal Profiles in Poplar Bent Stem and Root. *Front. Plant Sci.* **2020**, *11*, 590985. [[CrossRef](#)]
8. Fischer, U.; Kucukoglu, M.; Helariutta, Y.; Bhalerao, R.P. The Dynamics of Cambial Stem Cell Activity. *Annu. Rev. Plant Biol.* **2019**, *70*, 293–319. [[CrossRef](#)]
9. Fukuda, H.; Komamine, A. Establishment of an Experimental System for the Study of Tracheary Element Differentiation from Single Cells Isolated from the Mesophyll of *Zinnia elegans*. *Plant Physiol.* **1980**, *65*, 57–60. [[CrossRef](#)]
10. Pesquet, E.; Ranocha, P.; Legay, S.; Digonnet, C.; Barbier, O.; Pichon, M.; Goffner, D. Novel markers of xylogenesis in zinnia are differentially regulated by auxin and cytokinin. *Plant Physiol.* **2005**, *139*, 1821–1839. [[CrossRef](#)]
11. Johnsson, C.; Jin, X.; Xue, W.; Dubreuil, C.; Lezhneva, L.; Fischer, U. 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. *Physiol. Plant.* **2018**, *165*, 673–689. [[CrossRef](#)] [[PubMed](#)]
12. Morey, P.R.; Dahl, B.E. Histological and Morphological Effects of Auxin Transport Inhibitors on Honey Mesquite. *Bot. Gaz.* **1975**, *136*, 274–280. [[CrossRef](#)]
13. Perrot-Rechenmann, C. Cellular responses to auxin: Division versus expansion. *Cold Spring Harb. Perspect. Biol.* **2010**, *2*, a001446. [[CrossRef](#)] [[PubMed](#)]
14. Vanneste, S.; Friml, J. Auxin: A Trigger for Change in Plant Development. *Cell* **2009**, *136*, 1005–1016. [[CrossRef](#)]
15. Bhalerao, R.P.; Fischer, U. Auxin gradients across wood-instructive or incidental? *Physiol. Plant.* **2014**, *151*, 43–51. [[CrossRef](#)] [[PubMed](#)]
16. Leyser, O. Auxin Signaling. *Plant Physiol.* **2018**, *176*, 465–479. [[CrossRef](#)]
17. Wang, R.H.; Estelle, M. Diversity and specificity: Auxin perception and signaling through the TIR1/AFB pathway. *Curr. Opin. Plant Biol.* **2014**, *21*, 51–58. [[CrossRef](#)]
18. Berleth, T.; Jurgens, G. The Role of the *MONOPTEROS* Gene in Organizing the Basal Body Region of the Arabidopsis Embryo. *Development* **1993**, *118*, 575–587. [[CrossRef](#)]
19. Hamann, T.; Benkova, E.; Baurle, I.; Kientz, M.; Jurgens, G. The Arabidopsis *BODENLOS* gene encodes an auxin response protein inhibiting *MONOPTEROS*-mediated embryo patterning. *Genes Dev.* **2002**, *16*, 1610–1615. [[CrossRef](#)]
20. Hamann, T.; Mayer, U.; Jurgens, G. The auxin-insensitive *bodenlos* mutation affects primary root formation and apical-basal patterning in the Arabidopsis embryo. *Development* **1999**, *126*, 1387–1395. [[CrossRef](#)]
21. Muller, C.J.; Valdes, A.E.; Wang, G.; Ramachandran, P.; Beste, L.; Uddenberg, D.; Carlsbecker, A. *PHABULOSA* Mediates an Auxin Signaling Loop to Regulate Vascular Patterning in Arabidopsis. *Plant Physiol.* **2016**, *170*, 956–970. [[CrossRef](#)] [[PubMed](#)]
22. Wenzel, C.L.; Schuetz, M.; Yu, Q.; Mattsson, J. Dynamics of *MONOPTEROS* and *PIN-FORMED1* expression during leaf vein pattern formation in Arabidopsis thaliana. *Plant J.* **2007**, *49*, 387–398. [[CrossRef](#)] [[PubMed](#)]
23. Ruonala, R.; Ko, D.; Helariutta, Y. Genetic Networks in Plant Vascular Development. *Annu. Rev. Genet.* **2017**, *51*, 335–359. [[CrossRef](#)] [[PubMed](#)]
24. Nilsson, J.; Karlberg, A.; Antti, H.; Lopez-Vernaza, M.; Mellerowicz, E.; Perrot-Rechenmann, C.; Sandberg, G.; Bhalerao, R.P. Dissecting the molecular basis of the regulation of wood formation by auxin in hybrid aspen. *Plant Cell* **2008**, *20*, 843–855. [[CrossRef](#)]
25. Karannagoda, N.; Spokevicius, A.; Hussey, S.; Cassan-Wang, H.; Grima-Pettenati, J.; Bossinger, G. Eucalyptus grandis AUX/INDOLE-3-ACETIC ACID 13 (*EgrIAA13*) is a novel transcriptional regulator of xylogenesis. *Plant Mol. Biol.* **2022**. [[CrossRef](#)]
26. Yu, H.; Soler, M.; San Clemente, H.; Mila, I.; Paiva, J.A.; 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. [[CrossRef](#)]
27. Barra-Jimenez, A.; Ragni, L. Secondary development in the stem: When Arabidopsis and trees are closer than it seems. *Curr. Opin. Plant Biol.* **2017**, *35*, 145–151. [[CrossRef](#)]
28. Chaffey, N.; Cholewa, E.; Regan, S.; Sundberg, B. Secondary xylem development in Arabidopsis: A model for wood formation. *Physiol. Plant.* **2002**, *114*, 594–600. [[CrossRef](#)]
29. Legay, S.; Sivadon, P.; Blervacq, A.S.; Pavy, N.; Baghdady, A.; Tremblay, L.; Levasseur, C.; Ladouce, N.; Lapierre, C.; Seguin, A.; et al. *EgMYB1*, an R2R3 MYB transcription factor from eucalyptus negatively regulates secondary cell wall formation in Arabidopsis and poplar. *New Phytol.* **2010**, *188*, 774–786. [[CrossRef](#)]

30. Soler, M.; Plasencia, A.; Larbat, R.; Pouzet, C.; Jauneau, A.; Rivas, S.; Pesquet, E.; Lapierre, C.; Truchet, I.; Grima-Pettenati, J. The Eucalyptus linker histone variant EgH1.3 cooperates with the transcription factor EgMYB1 to control lignin biosynthesis during wood formation. *New Phytol.* **2017**, *213*, 287–299. [[CrossRef](#)]
31. 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. [[CrossRef](#)]
32. Worley, C.K.; Zenser, N.; Ramos, J.; Rouse, D.; Leyser, O.; Theologis, A.; Callis, J. Degradation of Aux/IAA proteins is essential for normal auxin signalling. *Plant J.* **2000**, *21*, 553–562. [[CrossRef](#)] [[PubMed](#)]
33. Dreher, K.A.; Brown, J.; Saw, R.E.; Callis, J. The Arabidopsis Aux/IAA protein family has diversified in degradation and auxin responsiveness. *Plant Cell* **2006**, *18*, 699–714. [[CrossRef](#)] [[PubMed](#)]
34. Sato, A.; Yamamoto, K.T. What's the physiological role of domain II-less Aux/IAA proteins? *Plant Signal. Behav.* **2008**, *3*, 496–497. [[CrossRef](#)] [[PubMed](#)]
35. Szemenyei, H.; Hannon, M.; Long, J.A. TOPLESS mediates auxin-dependent transcriptional repression during Arabidopsis embryogenesis. *Science* **2008**, *319*, 1384–1386. [[CrossRef](#)]
36. Audran-Delalande, C.; Bassa, C.; Mila, I.; Regad, F.; Zouine, M.; Bouzayen, M. Genome-wide identification, functional analysis and expression profiling of the Aux/IAA gene family in tomato. *Plant Cell Physiol.* **2012**, *53*, 659–672. [[CrossRef](#)]
37. Ulmasov, T.; Murfett, J.; Hagen, G.; Guilfoyle, T.J. Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. *Plant Cell* **1997**, *9*, 1963–1971. [[CrossRef](#)]
38. Reed, J.W. Roles and activities of Aux/IAA proteins in Arabidopsis. *Trends Plant Sci.* **2001**, *6*, 420–425. [[CrossRef](#)]
39. Zhong, R.; Taylor, J.J.; Ye, Z.H. Disruption of interfascicular fiber differentiation in an Arabidopsis mutant. *Plant Cell* **1997**, *9*, 2159–2170. [[CrossRef](#)]
40. Schillmiller, A.L.; Stout, J.; Weng, J.K.; Humphreys, J.; Ruegger, M.O.; Chapple, C. Mutations in the cinnamate 4-hydroxylase gene impact metabolism, growth and development in Arabidopsis. *Plant J.* **2009**, *60*, 771–782. [[CrossRef](#)]
41. Roschztardt, H.; Paez-Valencia, J.; Dittakavi, T.; Jali, S.; Reyes, F.C.; Baisa, G.; Anne, P.; Gissot, L.; Palauqui, J.C.; Masson, P.H.; et al. The VASCULATURE COMPLEXITY AND CONNECTIVITY gene encodes a plant-specific protein required for embryo provascular development. *Plant Physiol.* **2014**, *166*, 889–902. [[CrossRef](#)] [[PubMed](#)]
42. Piya, S.; Shrestha, S.K.; Binder, B.; Stewart, C.N., Jr.; Hewezi, T. Protein-protein interaction and gene co-expression maps of ARFs and Aux/IAAs in Arabidopsis. *Front. Plant Sci.* **2014**, *5*, 744. [[CrossRef](#)]
43. Vernoux, T.; Brunoud, G.; Farcot, E.; Morin, V.; Van den Daele, H.; Legrand, J.; Oliva, M.; Das, P.; Larrieu, A.; Wells, D.; et al. The auxin signalling network translates dynamic input into robust patterning at the shoot apex. *Mol. Syst. Biol.* **2011**, *7*, 508. [[CrossRef](#)] [[PubMed](#)]
44. Yu, H.; Soler, M.; Mila, I.; San Clemente, H.; Savelli, B.; Dunand, C.; Paiva, J.A.; 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. [[CrossRef](#)] [[PubMed](#)]
45. Blaschek, L.; Champagne, A.; Dimotakis, C.; Nuoendagula, Decou, R.; Hishiyama, S.; Kratzer, S.; Kajita, S.; Pesquet, E. Cellular and Genetic Regulation of Coniferaldehyde Incorporation in Lignin of Herbaceous and Woody Plants by Quantitative Wiesner Staining. *Front. Plant Sci.* **2020**, *11*, 109. [[CrossRef](#)]
46. Mitra, P.P.; Loque, D. Histochemical Staining of Arabidopsis thaliana Secondary Cell Wall Elements. *Jove-J. Vis. Exp.* **2014**, *87*, e51381. [[CrossRef](#)]
47. Sehr, E.M.; Agusti, J.; Lehner, R.; Farmer, E.E.; Schwarz, M.; Greb, T. Analysis of secondary growth in the Arabidopsis shoot reveals a positive role of jasmonate signalling in cambium formation. *Plant J.* **2010**, *63*, 811–822. [[CrossRef](#)]
48. Azevedo, C.; Betsuyaku, S.; Peart, J.; Takahashi, A.; Noel, L.; Sadanandom, A.; Casais, C.; Parker, J.; Shirasu, K. Role of SGT1 in resistance protein accumulation in plant immunity. *EMBO J.* **2006**, *25*, 2007–2016. [[CrossRef](#)]
49. Sato, A.; Yamamoto, K.T. Overexpression of the non-canonical Aux/IAA genes causes auxin-related aberrant phenotypes in Arabidopsis. *Physiol. Plant.* **2008**, *133*, 397–405. [[CrossRef](#)]
50. Lv, B.; Yu, Q.; Liu, J.; Wen, X.; Yan, Z.; Hu, K.; Li, H.; Kong, X.; Li, C.; Tian, H.; et al. Non-canonical AUX/IAA protein IAA33 competes with canonical AUX/IAA repressor IAA5 to negatively regulate auxin signaling. *EMBO J.* **2020**, *39*, e101515. [[CrossRef](#)]
51. Yang, F.; Mitra, P.; Zhang, L.; Prak, L.; Verhertbruggen, Y.; Kim, J.S.; Sun, L.; Zheng, K.; Tang, K.; Auer, M.; et al. Engineering secondary cell wall deposition in plants. *Plant Biotechnol. J.* **2013**, *11*, 325–335. [[CrossRef](#)] [[PubMed](#)]
52. De Meester, B.; de Vries, L.; Ozparpucu, M.; Gierlinger, N.; Corneillie, S.; Pallidis, A.; Goeminne, G.; Morreel, K.; De Bruyne, M.; De Rycke, R.; et al. Vessel-Specific Reintroduction of CINNAMOYL-COA REDUCTASE1 (CCR1) in Dwarfed ccr1 Mutants Restores Vessel and Xylary Fiber Integrity and Increases Biomass. *Plant Physiol.* **2018**, *176*, 611–633. [[CrossRef](#)] [[PubMed](#)]
53. De Meester, B.; Vanholme, R.; de Vries, L.; Wouters, M.; Van Doorslaere, J.; Boerjan, W. Vessel- and ray-specific monolignol biosynthesis as an approach to engineer fiber-hypolignification and enhanced saccharification in poplar. *Plant J.* **2021**, *108*, 752–765. [[CrossRef](#)] [[PubMed](#)]
54. Cassan-Wang, H.; Soler, M.; Yu, H.; Camargo, E.L.O.; Carocha, V.; Ladouce, N.; Savelli, B.; Paiva, J.A.P.; Leple, J.C.; Grima-Pettenati, J. Reference Genes for High-Throughput Quantitative Reverse Transcription-PCR Analysis of Gene Expression in Organs and Tissues of *Eucalyptus* Grown in Various Environmental Conditions. *Plant Cell Physiol.* **2012**, *53*, 2101–2116. [[CrossRef](#)]

55. Paux, E.; Tamasloukht, M.; Ladouce, N.; Sivadon, P.; Grima-Pettenati, J. Identification of genes preferentially expressed during wood formation in Eucalyptus. *Plant Mol. Biol.* **2004**, *55*, 263–280. [[CrossRef](#)]
56. Arvidsson, S.; Kwasniewski, M.; Riano-Pachon, D.M.; Mueller-Roeber, B. QuantPrime—a flexible tool for reliable high-throughput primer design for quantitative PCR. *BMC Bioinform.* **2008**, *9*, 465. [[CrossRef](#)]
57. Shimada, T.L.; Shimada, T.; Hara-Nishimura, I. A rapid and non-destructive screenable marker, FAST, for identifying transformed seeds of Arabidopsis thaliana. *Plant J.* **2010**, *61*, 519–528. [[CrossRef](#)]
58. Clough, S.J.; Bent, A.F. Floral dip: A simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J.* **1998**, *16*, 735–743. [[CrossRef](#)]
59. Abel, S.; Theologis, A. Transient transformation of Arabidopsis leaf protoplasts: A versatile experimental system to study gene expression. *Plant J.* **1994**, *5*, 421–427. [[CrossRef](#)]
60. Gerber, L.; Eliasson, M.; Trygg, J.; Moritz, T.; Sundberg, B. Multivariate curve resolution provides a high-throughput data processing pipeline for pyrolysis-gas chromatography/mass spectrometry. *J. Anal. Appl. Pyrol.* **2012**, *95*, 95–100. [[CrossRef](#)]



## C. Collaborations nationales et internationales

Ces travaux de recherche sur la caractérisation des gènes impliqués dans la formation du bois et auxine dépendants m'ont permis aussi établir des collaborations internationales sur les sujets de régulation de la formation du bois chez les arbres, avec SouthWest University of China (Prof. Keming LUO) sur la caractérisation fonctionnelle du gène *IAA9* chez peuplier (Xu et al., 2019), avec South China Agricultural University (Prof. Aimin WU) sur les ARNs non codants longs (Lin et al., 2019), avec University of Melbourne (Dr. Nadeeshani Karannagoda) sur la caractérisation fonctionnelle du gène *IAA13* chez *Eucalyptus* (Karannagoda et al., 2022), et avec University of Pretoria (Dr. Steven Hussey) une thèse en co-tutelle est en cours sur les facteurs de transcription MYB et ARF chez *Eucalyptus* (Mme Ipeleng Randome, thèse en cours). Ces travaux ont conduit à des publications dont un en co-corresponding author (Annexe 9).

### Résumés des publications issues des collaborations nationales et internationales

1. Karannagoda, N., et al. (2022).

**N Karannagoda, A Spokevicius, S Hussey, H Cassan-Wang, J Grima-Pettenati, G Bossinger. 2022 *Eucalyptus grandis* AUX/INDOLE-3-ACETIC ACID 13 (*EgrIAA13*) is a novel transcriptional regulator of xylogenesis. *Plant Mol Biol* 2022, doi:10.1007/s11103-022-01255-y. (IF=4.1)**

**KEY MESSAGE:** Our Induced Somatic Sector Analysis and protein-protein interaction experiments demonstrate that *Eucalyptus grandis* IAA13 regulates xylem fibre and vessel development, potentially via *EgrIAA13* modules involving ARF2, ARF5, ARF6 and ARF19. Auxin is a crucial phytohormone regulating multiple aspects of plant growth and differentiation, including regulation of vascular cambium activity, xylogenesis and its responsiveness towards gravitropic stress. Although the regulation of these biological processes greatly depends on auxin and regulators of the auxin signalling pathway, many of their specific functions remain unclear. Therefore, the present study aims to functionally characterise *Eucalyptus grandis* AUX/INDOLE-3-ACETIC ACID 13 (*EgrIAA13*), a member of the auxin signalling pathway. In *Eucalyptus* and *Populus*, *EgrIAA13* and its orthologs are preferentially expressed in the xylogenic tissues and downregulated in tension wood. Therefore, to further investigate *EgrIAA13* and its function during xylogenesis, we conducted subcellular localisation and Induced Somatic Sector Analysis experiments using overexpression and RNAi knockdown constructs of *EgrIAA13* to create transgenic tissue sectors on growing stems of *Eucalyptus* and *Populus*. Since Aux/IAAs interact with Auxin Responsive Factors (ARFs), in silico predictions of IAA13-ARF interactions were explored and experimentally validated via yeast-2-hybrid experiments. Our results demonstrate that *EgrIAA13* localises to the nucleus and that downregulation of *EgrIAA13* impedes *Eucalyptus* xylem fibre and vessel development. We also observed that *EgrIAA13* interacts with *Eucalyptus* ARF2, ARF5, ARF6 and ARF19A. Based on these results, we conclude that *EgrIAA13* is a regulator of *Eucalyptus* xylogenesis and postulate that the observed phenotypes are likely to result from alterations in the auxin-responsive transcriptome via IAA13-ARF modules such as *EgrIAA13-EgrARF5*. Our results provide the first insights into the regulatory role of *EgrIAA13* during xylogenesis.

2. Lin, Z., et al. (2019).

**Lin Z, Long JM, Yin Q, Wang B, Li HL, Luo JZ, Wang HC, Wu AM (2019) Identification of novel lncRNAs in *Eucalyptus grandis*. *Industrial Crops and Products* 129: 309-317 (IF=3.85)**

*Eucalyptus grandis* is the world's most widely planted hardwood tree. However, despite the availability of a sequenced genome and easily accessible functional genetic tools, the quantities and roles of long non-coding RNA (lncRNA) in *E. grandis* developmental processes remain largely unknown. In this study, we constructed RNA libraries for high-throughput sequencing of *Eucalyptus grandis* samples, and 551 novel lncRNAs were identified by selections. By annotation analysis of lncRNA targets, the results showed that lncRNAs participated in many apparent physiological processes, such as photosynthesis, ubiquitin-mediated proteolysis, protein export and processing in the endoplasmic reticulum. Finally, ten lncRNAs were confirmed to express in *Eucalyptus grandis* by quantitative real-time reverse transcription PCR, indicating that our reported lncRNAs existed. The identification of lncRNAs will lead to a greater understanding of the role and function of lncRNAs in the physiology, growth and development of *Eucalyptus grandis* trees.

3. Xu, C., et al. (2019).

**C Xu, Y Shen, F He, X Fu, H Yu, W Lu, Y Li, C Li, D Fan, HCassan-Wang\*, K Luo\* 2019 Auxin-mediated Aux/IAA-ARF-HB signaling cascade regulates secondary xylem development in Populus. New Phytol 222: 752-767. co-corresponding author (IF=7.4)**

Wood development is strictly regulated by various phytohormones and auxin plays a central regulatory role in this process. However, how the auxin signaling is transduced in developing secondary xylem during wood formation in tree species remains unclear. Here, we identified an Aux/INDOLE-3-ACETIC ACID 9 (IAA9)-AUXIN RESPONSE FACTOR 5 (ARF5) module in *Populus tomentosa* as a key mediator of auxin signaling to control early developing xylem development. PtoIAA9, a canonical Aux/IAA gene, is predominantly expressed in vascular cambium and developing secondary xylem and induced by exogenous auxin. Overexpression of PtoIAA9m encoding a stabilized IAA9 protein significantly represses secondary xylem development in transgenic poplar. We further showed that PtoIAA9 interacts with PtoARF5 homologs via the C-terminal III/IV domains. The truncated PtoARF5.1 protein without the III/IV domains rescued defective phenotypes caused by PtoIAA9m. Expression analysis showed that the PtoIAA9-PtoARF5 module regulated the expression of genes associated with secondary vascular development in PtoIAA9m- and PtoARF5.1-overexpressing plants. Furthermore, PtoARF5.1 could bind to the promoters of two Class III homeodomain-leucine zipper (HD-ZIP III) genes, PtoHB7 and PtoHB8, to modulate secondary xylem formation. Taken together, our results suggest that the Aux/IAA9-ARF5 module is required for auxin signaling to regulate wood formation via orchestrating the expression of HD-ZIP III transcription factors in poplar.

4. Zouine, M., et al. (2014).

**Zouine M, Fu Y, Chateigner-Boutin AL, Mila I, Frasse P, Wang H, Audran C, Roustan JP, Bouzayen M (2014) Characterization of the tomato ARF gene family uncovers a multi-levels post-transcriptional regulation including alternative splicing. PLoS One 9: e84203 (IF=4.2)**

**BACKGROUND:** The phytohormone auxin is involved in a wide range of developmental processes and auxin signaling is known to modulate the expression of target genes via two types of transcriptional regulators, namely, Aux/IAA and Auxin Response Factors (ARF). ARFs play a major role in transcriptional activation or repression through direct binding to the promoter of auxin-responsive genes. The present study aims at gaining better insight on distinctive structural and functional features among ARF proteins. **RESULTS:** Building on the most updated tomato (*Solanum lycopersicon*) reference genome sequence, a comprehensive set of ARF genes was identified, extending the total number of family members to 22. Upon correction of structural annotation inconsistencies, renaming the tomato ARF family members provided a consensus nomenclature for all ARF genes across plant species. In silico search predicted the presence of putative target site for small interfering RNAs within twelve SI-




ARFs while sequence analysis of the 5'-leader sequences revealed the presence of potential small uORF regulatory elements. Functional characterization carried out by transactivation assay partitioned tomato ARFs into repressors and activators of auxin-dependent gene transcription. Expression studies identified tomato ARFs potentially involved in the fruit set process. Genome-wide expression profiling using RNA-seq revealed that at least one third of the gene family members display alternative splicing mode of regulation during the flower to fruit transition. Moreover, the regulation of several tomato ARF genes by both ethylene and auxin, suggests their potential contribution to the convergence mechanism between the signaling pathways of these two hormones. CONCLUSION: All together, the data bring new insight on the complexity of the expression control of *SlARF* genes at the transcriptional and post-transcriptional levels supporting the hypothesis that these transcriptional mediators might represent one of the main components that enable auxin to regulate a wide range of physiological processes in a highly specific and coordinated manner.



Annexe 9. Article IAA9 chez Peuplier



# Auxin-mediated Aux/IAA-ARF-HB signaling cascade regulates secondary xylem development in *Populus*

Changzheng Xu<sup>1\*</sup>, Yun Shen<sup>1\*</sup>, Fu He<sup>1\*</sup>, Xiaokang Fu<sup>1\*</sup>, Hong Yu<sup>1,2\*</sup>, Wanxiang Lu<sup>1</sup>, Yongli Li<sup>1</sup>, Chaofeng Li<sup>1,3</sup>, Di Fan<sup>1</sup>, Hua Cassan Wang<sup>4</sup> and Keming Luo<sup>1</sup> 

<sup>1</sup>Chongqing Key Laboratory of Plant Resource Conservation and Germplasm Innovation, School of Life Sciences, Southwest University, Chongqing 400715, China; <sup>2</sup>School of Basic Medical Sciences, Southwest Medical University, Luzhou, Sichuan 646000, China; <sup>3</sup>Key Laboratory of Adaptation and Evolution of Plateau Biota, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining 810008, China; <sup>4</sup>UMR5546, Laboratoire de Recherche en Sciences Végétales, Université de Toulouse III Paul Sabatier, CNRS, UPS, 31326, Castanet-Tolosan, France

## Summary

Authors for correspondence:  
Hua Cassan Wang  
Tel: +33 5 34323851  
Email: huawang76@yahoo.com

Keming Luo  
Tel: +86 23 68253021  
Email: kemingl@swu.edu.cn

Received: 13 November 2018  
Accepted: 14 December 2018

New Phytologist (2019)  
doi: 10.1111/nph.15658

**Key words:** ARF5, Aux/IAA9, auxin, HD-ZIP III transcription factors, *Populus*, xylem development.

- Wood development is strictly regulated by various phytohormones and auxin plays a central regulatory role in this process. However, how the auxin signaling is transduced in developing secondary xylem during wood formation in tree species remains unclear.
- Here, we identified an Aux/INDOLE-3-ACETIC ACID 9 (IAA9)-AUXIN RESPONSE FACTOR 5 (ARF5) module in *Populus tomentosa* as a key mediator of auxin signaling to control early developing xylem development.
- *PtoIAA9*, a canonical *Aux/IAA* gene, is predominantly expressed in vascular cambium and developing secondary xylem and induced by exogenous auxin. Overexpression of *PtoIAA9m* encoding a stabilized IAA9 protein significantly represses secondary xylem development in transgenic poplar. We further showed that *PtoIAA9* interacts with *PtoARF5* homologs via the C-terminal III/IV domains. The truncated *PtoARF5.1* protein without the III/IV domains rescued defective phenotypes caused by *PtoIAA9m*. Expression analysis showed that the *PtoIAA9*-*PtoARF5* module regulated the expression of genes associated with secondary vascular development in *PtoIAA9m*- and *PtoARF5.1*-overexpressing plants. Furthermore, *PtoARF5.1* could bind to the promoters of two Class III homeodomain-leucine zipper (HD-ZIP III) genes, *PtoHB7* and *PtoHB8*, to modulate secondary xylem formation.
- Taken together, our results suggest that the *Aux/IAA9*-*ARF5* module is required for auxin signaling to regulate wood formation via orchestrating the expression of HD-ZIP III transcription factors in poplar.

## Introduction

Meristems, specialized structures of reservoirs of stem cells undergoing proliferation, drive postembryonic developmental programs in plants (Nakajima & Benfey, 2002). In comparison with primary/longitudinal growth of shoots and roots mediated by apex-localized meristems, secondary/radial growth is dominated by vascular cambium (Campbell & Turner, 2017). Cambial cells sequentially undergo proliferation and differentiation into secondary vascular tissues (Matte Risopatron *et al.*, 2010). Different from herbaceous plants, trees exhibit perennial stem thickening during their life cycle due to continuous competence of the vascular cambium (Sanchez *et al.*, 2012). The secondary xylem, commonly called wood, overwhelmingly contributes to stem thickening of trees (Sanchez *et al.*, 2012). Wood production is a predominant proportion of biomass accumulation in terrestrial

ecosystems and is also of outstanding economic value (Ragauskas *et al.*, 2006; Bonan, 2008).

Cambium-originated wood formation starts with specification of secondary xylem cell precursors, which in turn undergo a series of cellular events for maturation, including expansion, secondary cell-wall deposition and programmed cell death (Ye & Zhong, 2015). Cambial homeostasis between stem cell identity and xylem specification affects the deposition of xylem cell layers (Campbell & Turner, 2017). Expansion of newly specified xylem cells contributes to the size of woody elements, while secondary wall deposition is important for functional maturation of xylem. Although vessels and fibers are lignified components of secondary xylem, they differ in cell size, secondary cell-wall deposition and cell viability (Fukuda, 2004). These differences result from the cellular changes in cytoskeletal arrangement, secondary wall patterning and autolysis that occur during differentiation of tracheary elements, the unit of xylem vessels (Fukuda, 2004; Turner *et al.*, 2007). These coordinated behaviors of cambial cells and

\*These authors contributed equally to this work.

their differentiating descendants are intricately and dynamically orchestrated by various genetic and environmental factors (Dejardin *et al.*, 2010; Ye & Zhong, 2015). Recent high-spatial-resolution profiling analyses in *Populus* have revealed three transcriptome reprogramming events coinciding with transitions of distinct cellular behaviors in the context of wood formation (Sundell *et al.*, 2017). Although a comprehensive picture of the gene regulatory machinery involved in secondary growth has emerged in recent decades, precise regulation of spatially organized and temporally coordinated cellular events for wood differentiation remains largely unknown.

Auxin is a crucial phytohormone for cell–cell communication in meristems (Leyser, 2005). Auxin signaling is able to direct cellular behaviors, including cell division, expansion and differentiation, for various developmental programs (Vanneste & Friml, 2009; Perrot-Rechenmann, 2010). Previous studies have shown that auxin accumulates in a radial gradient pattern across wood-forming tissues with a peak in concentration in cambium and bilateral decay towards differentiating secondary xylem and phloem (Tuominen *et al.*, 1997; Uggla *et al.*, 1998; Immanen *et al.*, 2016). Auxin is thus considered as a primary regulator of secondary xylem formation in a morphogen-like manner (Uggla *et al.*, 1996; Sundberg *et al.*, 2000). In *Pinus*, a lack of auxin supply from shoot apex leads to a loss of fusiform shape of cambial derivatives (Savidge, 1983). Introduction of a stabilized version of PttIAA3 (homolog of IAA20s in *Arabidopsis* and *Populus*; Supporting Information Fig. S1) that represses auxin response in hybrid aspen attenuates periclinal division of cambial cells, but enhances stem cell characteristics as indicated by an enlarged zone of anticlinal cell division in cambium (Nilsson *et al.*, 2008). These findings reveal a dual role of auxin in coordinating cambial identity and proliferation activity. Recently, detection of auxin signaling via a high-affinity sensor indicates its moderate level in cambial stem cells, while increased level in differentiating cambial descendants (Brackmann *et al.*, 2018). These results implicate the key role of auxin signaling in cambial differentiation. Consistently, exogenous auxin is able to induce differentiation of intact and functionally normal secondary xylem cells (Björklund *et al.*, 2007). Local enhancement of auxin signaling is required for wood formation, also supporting the action of auxin in recruiting cells for differentiation (Bargmann *et al.*, 2013; Muller *et al.*, 2016). However, the precise regulation of auxin signaling on these coordinated cellular events for wood differentiation remains obscure in trees.

Auxin perception starts with auxin binding to TIR1 (TRANSPORT INHIBITOR RESPONSE1)/AFB (AUXIN SIGNALING F-BOX) receptors, and leads to subsequent degradation of the Aux/IAA proteins that repress auxin signaling via physical interactions with auxin response factor (ARF) proteins (Mockaitis & Estelle, 2008). The auxin-stimulated protein turnover of Aux/IAAs releases the transcriptional activity of their partner ARFs to activate downstream auxin responsive gene expression (Vanneste & Friml, 2009). Different Aux/IAA-ARF modules are known to regulate corresponding auxin-responsive genes and developmental processes (Vanneste & Friml, 2009). In *Arabidopsis*, the IAA12/BODENLOS (BDL)-ARF5/MONOPTEROS (MP) module was identified to

control provascular specification and patterning during embryogenesis (Hardtke & Berleth, 1998; Hamann *et al.*, 2002). ARF5/MP is also crucial for leaf (pre)procambial specification during leaf vein patterning (Przemeck *et al.*, 1996). In hybrid aspen, eight *Aux/IAA* members were assayed for expression of auxin responsive genes in wood tissues and during transitions of cambial activity (Moyle *et al.*, 2002), and *PttIAA3* was identified as an important mediator of auxin-dependent regulation of cambial proliferation activity (Nilsson *et al.*, 2008). When a *Eucalyptus* Aux/IAA member *IAA4* (*EgrIAA4*) was ectopically expressed in *Arabidopsis*, secondary xylem formation of the transgenic plants was dramatically reduced (Yu *et al.*, 2015).

Various plant model systems have been used to study the mechanisms governing plant vascular development. The construction of genetic networks underlying vascular tissue specification has been mainly achieved based on research on four organs/systems: provascular specification during embryogenesis, procambial cell specification during root development (primary growth), (pre)procambial cell specification during leaf vein patterning (primary growth), and cambium specification during secondary growth in stem for wood formation (Campbell & Turner, 2017). For embryogenesis and procambial specification during primary growth in roots and leaf veins, the key components of auxin signaling have been identified: the IAA12/BDL-ARF5/MP module for embryogenesis (Hardtke & Berleth, 1998; Hamann *et al.*, 2002), the ARF5/MP-mediated pathway for leaf vein patterning (Przemeck *et al.*, 1996; Donner *et al.*, 2009), and the IAA20/IAA30-ARF5/MP combinations for root development (Muller *et al.*, 2016). To date, however, it remains unknown which Aux/IAA-ARF combination mediates auxin-dependent cambial specification of secondary growth for wood formation.

Here, we identified *PtoIAA9* as a novel key component in modulating wood formation in *Populus tomentosa*. Functional characterization of the PtoIAA9–PtoARF5 module revealed auxin-dependent differentiation of secondary xylem derived from the cambium, and demonstrated their roles in orchestrating xylem cell specification, woody cell size and vessel density. Moreover, we provided biochemical and genetic evidence that *PtoHB7* and *PtoHB8*, encoding HD-ZIP III transcription factors, are direct targets of the PtoIAA9–PtoARF5 module to coordinate cellular behaviors associated with woody cell differentiation. Our results provide novel insights into auxin-dependent regulation on the dynamic and coordinated cellular events for secondary xylem differentiation in woody species.

## Materials and Methods

### Gene cloning and plasmid construction

The full-length coding sequence of *PtoIAA9*, *PtoARF5.1* and *PtoARF5.2* was amplified from cDNA of *P. tomentosa* using gene-specific primer pairs (Table S1), and constructed into the pCXSN vector (Chen *et al.*, 2009), respectively. To substitute the first proline (P) with serine (S) within the Degron motif, overlap PCR was performed using pCXSN-*PtoIAA9* as templates. The resulting PCR product was ligated into pCXSN under control of



the *CaMV 35S* promoter to generate the plasmid for overexpressing *PtoIAA9m*. The full-length sequence of *PtoARF5.1* was subcloned into the modified pCAMBIA1305 with a kanamycin resistant gene. The sequence of *PtoARF5.1Δ* harboring the N-terminal 1995 nucleotides of *PtoARF5.1* was amplified from pCXS-*PtoARF5.1* and constructed into the modified pCAMBIA1305 vector. A 2-kb promoter fragment upstream of *PtoIAA9* was amplified from genomic DNA of *P. tomentosa* using the primer pair *Pro-PtoIAA9-fw/rv* (Table S1), and inserted into pXGUS-P to drive the *GUS* reporter gene (Chen *et al.*, 2009).

### Genetic transformation and growth conditions

*P. tomentosa* was stably transformed using the method of *Agrobacterium*-mediated infiltration of leaf disks as described previously (Jia *et al.*, 2010). PCR genotyping with the primers of hygromycin/kanamycin-resistant genes was performed for the identification of positive transgenic plants. Transgenic and WT poplar plants were propagated via *in vitro* microcutting. For clonal propagation, shoot segments of 3–4 cm with two or three young leaves were cut from sterilized seedlings and cultivated on woody plant medium solid medium at 25°C with 16 h light of 5000 lux and 8 h dark. For phenotype analyses, 4-wk-old microcutting-propagated seedlings were transferred to soil in pots, and cultivated in a glasshouse at 23–25°C with the light of 10 000 lux under a 16 h: 8 h, day : night cycle.

### Cross-sectioning and histological staining

The 7<sup>th</sup> internodes were sectioned with a razor blade, and then stained with 0.05% (w/v) toluidine blue for 5 min. The cross-sections were observed and captured using a microscope (Zeiss). The images were analyzed using IMAGEJ (<https://imagej.nih.gov/ij/>) for quantifying morphological parameters of xylem cells.

### RNA *in situ* hybridization

For probe preparation, 276- and 223-bp gene-specific cDNA fragments were amplified for *PtoIAA9* and *PtoARF5.1*, respectively (Table S1). The two probes were labeled using a DIG RNA Labeling Kit (Roche). Section pretreatment, hybridization and immunological detection were performed as previously described (Sang *et al.*, 2012).

### GUS staining

For GUS staining (Jefferson, 1987), cross-sections of the 7<sup>th</sup> internodes of 2-month-old plants were fixed in acetone for 1 h at –20°C, and then washed twice in double distilled H<sub>2</sub>O (ddH<sub>2</sub>O). The cross-sections were soaked in GUS staining solution (0.5 M Tris, pH 7.0, 10% Triton X-100 with 1 mM X-Gluc (5-bromo-4-chloro-3-indolyl-D-glucuronide)) for 15 min at 37°C in the dark. After reaction, Chl was removed by use of 75% ethanol three times at 65°C. The Chl-free stained stems were observed under an Olympus 566 SZX16 microscope (Tokyo, Japan) and documented using an Olympus DP73 camera.

### qRT-PCR

Total RNA was extracted from tissues of 2-month-old *P. tomentosa* plants using a Plant RNeasy Mini Kit (Qiagen). cDNA was synthesized using a PrimeScript<sup>TM</sup> RT reagent Kit with gDNA Eraser (Takara, Dalian, China). Quantitative real time polymerase chain reaction (qRT-PCR) was performed using SYBR Premix ExTaq<sup>TM</sup> (Takara) in a TP800 Real-Time PCR machine (Takara). The poplar *18S rRNA* gene was used as the reference gene as an internal standard. The primers used for qRT-PCR are listed in Table S1.

### Transactivation test in yeast

The recombinant plasmids were introduced into the yeast strain *Saccharomyces cerevisiae* Gold2 using the PEG/LiAc method. The transformed strains were screened on synthetic dropout (SD medium) lacking tryptophan (Trp; SD/-T) for selection of positive clones. Subsequently, positive clones were transferred to SD medium lacking Trp, histidine (His) and adenine (Ade; SD/-ATH) and cultivated at 28°C for 2 d. Positive clones were used for transactivation analysis on X- $\alpha$ -gal indicator plates. A digital camera (EOS 550D, Canon, Tokyo, Japan) was used to photograph yeast cells.

### Yeast two-hybrid (Y2H) assays

Y2H assays were performed based on the manufacturer's instructions (Clontech, Palo Alto, CA, USA). The AD and BD fusion constructs were co-transformed into yeast strain Gold2 as described above. The transformants were screened on SD medium lacking tryptophan (Trp) and leucine (Leu; SD/-TL) at 28°C. After 3 d, positive clones were cultured in YPD liquid medium for 1 d to an OD<sub>600</sub> of 1.0, and the yeast solution was diluted with ddH<sub>2</sub>O to 1–10<sup>3</sup> $\times$ , and then transferred to SD medium lacking Trp, Leu, histidine (His) and adenine (Ade; SD/-AHTL) and cultured for 2 d. A EOS 550D digital camera was used to photograph yeast cells.

### Bimolecular fluorescence complementation (BiFC)

For BiFC assays, full-length coding sequences of *PtoIAA9* and *PtoARF5.1* were cloned from pCXS-*PtoIAA9* and pCXS-*PtoARF5.1*, and ligated into pXY104 and pXY106 vectors, respectively. The vectors pXY104 and pXY106 carrying the N- and C-terminal halves of yellow fluorescent protein (YFP) were used. The open reading frame (ORF) of *PtoIAA9* was cloned into pXY106, while that of *PtoARF5.1* was constructed into pXY104. The constructs were co-transformed into tobacco as described above, and observed by confocal laser microscopy (Olympus 589 FV1200).

### Chromatin immunoprecipitation (ChIP) assay

ChIP analysis was performed as previously described (Yang *et al.*, 2012). One-month-old *PtoARF5.1* transgenic poplar plants

harboring the HA epitope were used as samples; the HA antibody and normal mouse IgG were used for immunoreaction. All primers used in ChIP assays are listed in Table S1.

### Effector-reporter test

Promoters of the *PtoHB7* and *PtoHB8* genes were amplified via specific primers (Table S1) and constructed into pCXGUS-P. The *35S:PtoARF5.1* and *35S:PtoARF5.1Δ* vectors were used as effectors. Tobacco leaves were transformed by *Agrobacteria* containing the effectors and reporters as described above. After 3 d of infiltration, GUS activity was measured by monitoring the cleavage of 4-methyl umbelliferyl β-D-glucuronide (MUG), the substrate of β-glucuronidase, which produces the fluorescent 4-methyl umbelliferone (4MU) upon hydrolysis. Protein concentrations were determined via the Bradford method.

### Gene accessions

GeneBank accession numbers of the *P. tomentosa* genes are: *PtoIAA9* (MH345700), *PtoARF5.1* (MH352401), *PtoARF5.2* (MH352402) and *PtoHB7* (MH345699).

## Results

### *IAA9* is highly expressed across wood-forming tissues in poplar

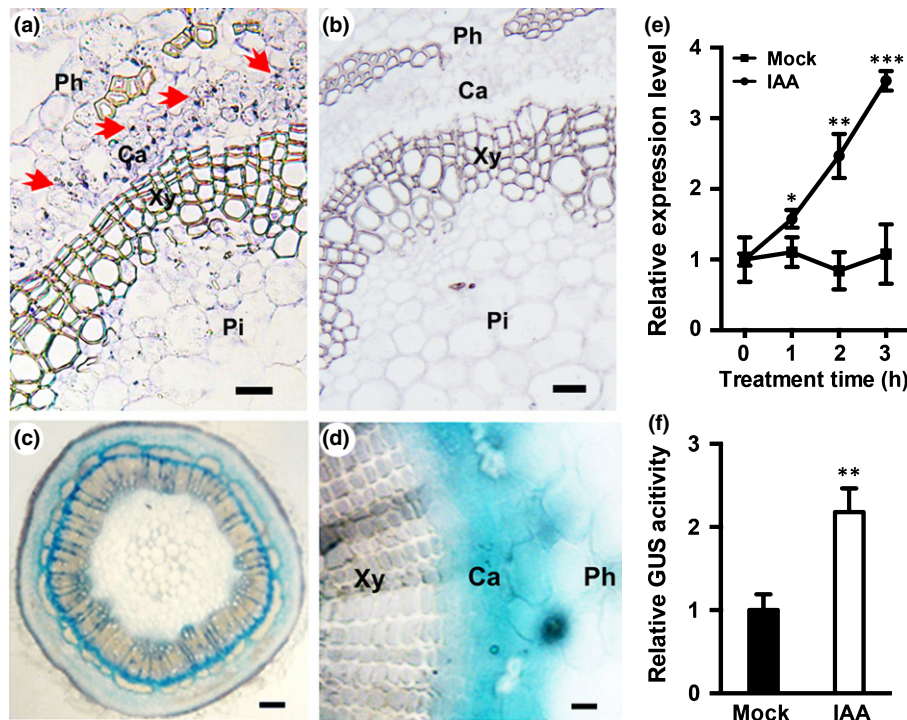
A total of 35 *Aux/IAA* genes were previously predicted in the poplar genome (Kalluri *et al.*, 2007). These *Aux/IAA* genes were identified in the more recent release of the *Populus trichocarpa* genome (v3.0; <https://phytozome.jgi.doe.gov>) and designated according to phylogenetic relationships with their Arabidopsis orthologues (Fig. S1a). To determine the members that play a key role during wood formation, expression levels of all *Aux/IAA* genes were comprehensively evaluated using the high-spatial-resolution RNA sequencing data of the wood-forming tissues in poplar (Sundell *et al.*, 2017). Among them, *IAA9*, *-11*, *-16.1*, *-16.2* and *-29.1* showed high expression (Fig. S1). Continuous assays across secondary xylem tissues in poplar stems (Sundell *et al.*, 2017) allowed us to evaluate spatial expression patterns of these highly expressed *Aux/IAAs* and some other members, such as *IAA12.1*, *-12.2*, *-20.1* and *-20.2* (Fig. S2b). A highly similar expression pattern was shared by *IAA9*, *-16.1* and *-16.2*, extending from the cambium to differentiated xylem (Fig. S2b), which was validated by correlation tests (Pearson correlation coefficient  $R^2 = 0.82$  or  $0.51$ ,  $P < 0.01$ ; Fig. S2c). For this pattern, the transcripts of these three *Aux/IAAs* accumulated in the cambial zone and developing xylem, and then decreased towards differentiated xylem (Fig. S2d), consistent with the previously reported auxin distribution across wood-forming tissues of poplar stems (Uggla *et al.*, 1998; Immanen *et al.*, 2016). By contrast, *IAA11* and *-29.1* displayed totally different expression patterns (Fig. S2b). Thus, *IAA9*, *-16.1* and *-16.2* were considered key *Aux/IAA* genes of auxin signaling involved in wood formation in poplar. Tissue-specific expression assays by qRT-PCR in *P. tomentosa* further

confirmed the enriched transcript abundance of *IAA9*, *-16.1* and *-16.2* in stems (Fig. S2d). We selected *PtoIAA9* from *P. tomentosa* as a candidate gene to further investigate auxin-dependent regulation of wood formation.

To determine the exact expression pattern of *PtoIAA9* in secondary vascular tissues, RNA *in situ* hybridization was performed using the fifth internode of 1.5-month-old poplar (Fig. 1a,b). Transcripts of *PtoIAA9* were preferentially accumulated in the cambial zone and neighboring cells (Fig. 1a). Furthermore, the GUS reporter driven by the *PtoIAA9* promoter in transgenic poplar confirmed its expression in the cambium zone and closely neighboring cell layers of the wood-developing stem (Fig. 1c,d). qRT-PCR assays revealed a more than 3-fold induced transcript abundance of *PtoIAA9* in stems within 3 h of exogenous auxin (Fig. 1e). The auxin-inducible expression was also examined by quantification of GUS activity driven by the *PtoIAA9* promoter in wood-forming tissues (Fig. 1f). These results indicated that *PtoIAA9* is an auxin-inducible gene predominantly expressed in cambium and adjacent cells towards xylem differentiation.

### *PtoIAA9* encodes a canonical Aux/IAA protein belonging to a distinct clade

*PtoIAA9* protein harbors the characteristic domains (Domain I, II, III and IV; Reed, 2001) of Aux/IAAs in high sequence similarity to its close homologs in other plant species (Fig. S3). Remarkably, *PtoIAA9* belongs to a distinct clade among all Aux/IAA members due to its 50% longer amino acid sequence (*c.* 300 aa) compared with the average length (*c.* 200 aa) of other Aux/IAA members (Wang *et al.*, 2005). Transient expression of a *PtoIAA9-GFP* fusion gene revealed that *PtoIAA9* is a nucleus-localized protein (Fig. S4a; Methods S1), consistent with the prediction of both bipartite and SV40-type nuclear localization signals (NLS) present within its sequences (Fig. S3). Aux/IAA proteins usually contain a short amino acid stretch (VGWPP) called Degron that confers auxin-stimulated protein turnover (Worley *et al.*, 2000). Due to the presence of a typical Degron motif in its Domain II, auxin-responsive protein stability of *PtoIAA9* fused with a green fluorescent protein (GFP) tag was monitored in transiently expressed epidermal cells of tobacco leaves (Fig. S4b; Methods S1). The fluorescent signals of *PtoIAA9-GFP* were significantly reduced in the leaf epidermal cells subject to IAA (Fig. S4b), as validated by quantifying the percentage of fluorescent nuclei and fluorescence intensity per nucleus (Fig. S4c,d). By contrast, this auxin-induced protein instability disappeared when the first proline (P) within the Degron motif of *PtoIAA9* was replaced by serine (S) (ns, not significant;  $P > 0.05$ ; Fig. S4b–e). Moreover, stronger fluorescence was detected in these leaf epidermal cells transiently transformed by the *PtoIAA9* mutant gene with impaired Degron motif, compared to the wild-type control (Fig. S4b–d), indicating that the auxin-inducible rapid turnover of *PtoIAA9* depends on its Degron motif. Self-activation assays in yeast showed that *PtoIAA9* has no transcriptional activating activity (Fig. S4f). Additionally, *PtoIAA9* compromised the activating capability of the VP16 domain, demonstrating that it is a transcriptional repressor. The truncated *PtoIAA9* protein without



**Fig. 1** Expression pattern and auxin induction of *PtoIAA9* transcripts in wood-forming tissues of *Populus tomentosa* stem. (a, b) RNA *in situ* hybridization of *PtoIAA9* in secondary vascular tissues of poplar. The 5<sup>th</sup> internodes of 1.5-month-old poplar plants cultivated in soil were cross-sectioned for hybridization with antisense (a) and sense (b) probes of *PtoIAA9*. Red triangles indicate *in situ* hybridization signals for *PtoIAA9* transcripts. Ca, cambium; Ph, phloem; Pi, pith; Xy, xylem. (c, d) Histological staining of the GUS reporter driven by the promoter of *PtoIAA9* in poplar stems. The 7<sup>th</sup> internodes of 1.5-month-old poplar plants cultivated in soil were cross-sectioned for GUS staining. Ca, cambium; Ph, phloem; Xy, xylem. (e) Time-course assays of auxin-induced transcript abundance of *PtoIAA9* in poplar stems. The microcutting-propagated poplar seedlings cultivated *in vitro* for 4 wk were subjected to 5  $\mu$ M IAA for 0, 1, 2 and 3 h, and stem tissues were collected for RNA extraction followed by qRT-PCR assays. 18S rRNA was used as a reference gene. Expression levels are indicated relative to values for 0 h (with 0 h set arbitrarily to 1). Error bars represent  $\pm$  SD. Asterisks indicate significant differences between mock and auxin treatment at each time point (Student's *t* test): \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ;  $n = 3$ . (f) Quantification of auxin-induced GUS activity driven by the promoter of *PtoIAA9* in poplar stem. IAA treatment (5  $\mu$ M) was performed for 6 h as indicated in (d). The values under mock treatment were normalized to 1. Error bars represent SD. Asterisks indicate significant differences with respect to mock (Student's *t* test): \*\*,  $P < 0.01$ ;  $n = 4$ . Bars: (a, b) 200  $\mu$ m; (c) 500  $\mu$ m; (d) 100  $\mu$ m.

Domain I at the C-terminal region (BD-VP16-PtoIAA9t) showed repressing activity (Fig. S4f). Taking this together, PtoIAA9 exhibits molecular features typical of canonical Aux/IAA repressors.

### Overexpression of an auxin-resistant *PtoIAA9m* represses wood formation

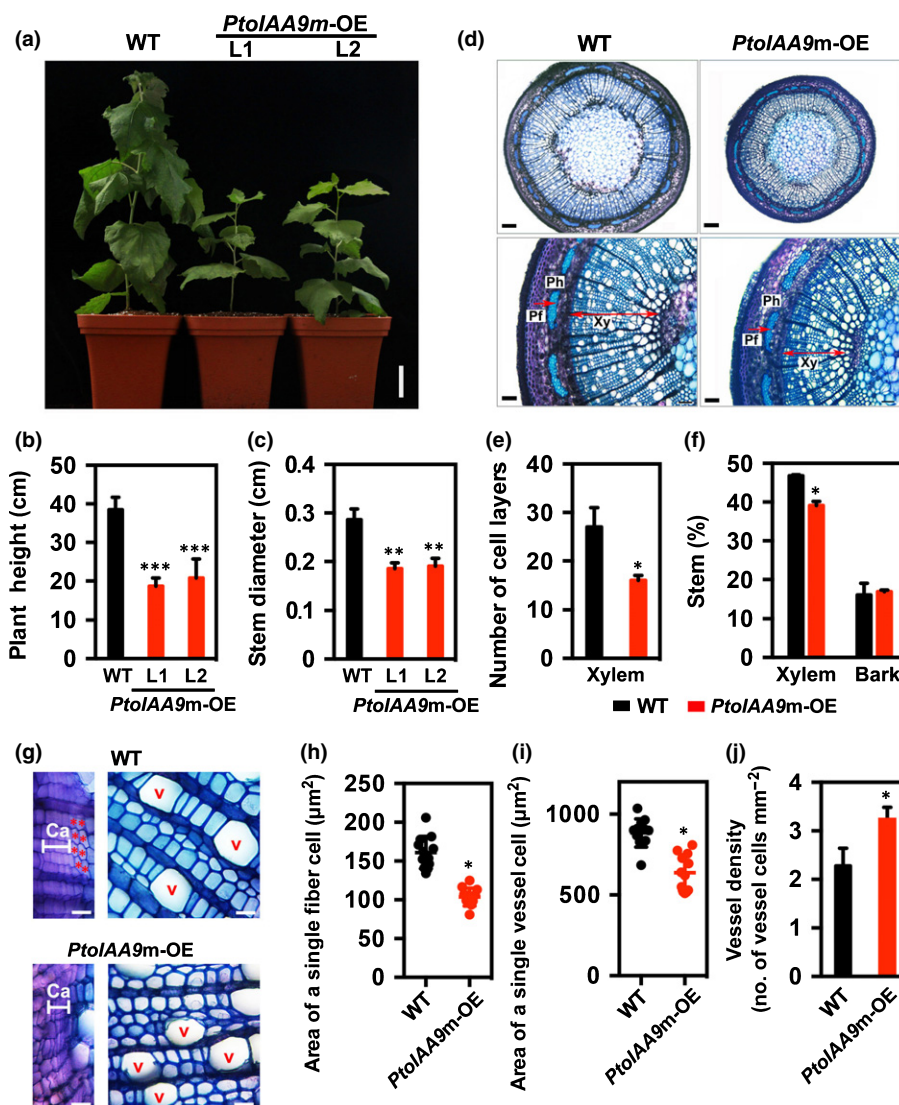
To identify its role in regulating secondary xylem development during wood formation, the auxin-resistant *PtoIAA9* containing the impaired Degron sequence (Fig. S4e; *PtoIAA9m*) was constitutively expressed in poplar (*PtoIAA9m*-OE; Fig. S5a). Representative overexpressing lines displayed significantly reduced plant growth compared to the wildtype (WT; Fig. 2a). Quantitative measurement revealed that overexpression of *PtoIAA9m* led to a 40–50% reduction in plant height and 35% reduction in stem diameter (Fig. 2b,c). Similar changes were indicated by a time-course analysis of these two parameters (Fig. S5b,c). The internode number of WT and transgenic plants was not changed (ns;  $P > 0.05$ ; Fig. S5d). Secondary vascular tissues were examined in *PtoIAA9m*-OE plants stained with toluidine O (Fig. 2d).

Secondary xylem development was significantly repressed by overexpressing *PtoIAA9m* (Fig. 2d), with a 40% decrease in the number of xylem cell layers, compared with WT (Fig. 2e). The percentage of secondary xylem occupying the whole stem was 46.8% in WT, but was attenuated to 39.1% in the *PtoIAA9m*-OE lines, whereas that of phloem was not affected (Fig. 2f). These results indicated that *PtoIAA9m* might inhibit wood formation.

### *PtoIAA9*-attenuated auxin signaling impairs secondary xylem formation

To determine PtoIAA9-dependent regulation of wood formation, we first examined phenotypes of wood-associated cell types, including cambium, xylem fibers and vessels in transgenic poplar. The results showed that cambial cells were arrayed more tightly in *PtoIAA9m*-OE plants than WT (Fig. 2g), but the number of cambial cell layers was not significantly affected by overexpression of *PtoIAA9m* (ns;  $P > 0.05$ ; Fig. S5e). Cross-sections of different internodes in WT revealed the presence of one or two particular cell layers on the side of the cambium zone towards the xylem





**Fig. 2** Wood phenotypes resulting from *PtoIAA9m* overexpression in *Populus tomentosa*. (a) Dwarf phenotypes of 2-month-old plants of independent *PtoIAA9m*-overexpressing (*PtoIAA9m*-OE) transgenic poplar lines (L1 and L2). (b, c) Measurement of plant height (b) and stem diameter (c) of *PtoIAA9m*-OE transgenic poplar lines corresponding to (a). (d) Cross-sectioning and staining with toluidine blue of the 7<sup>th</sup> internode of 2-month-old wild-type (WT) and *PtoIAA9m*-OE transgenic plants (L1). Ph, phloem; Pf, phloem fibers; Xy, xylem. (e) Quantification of secondary xylem cell layers in WT and *PtoIAA9m*-OE transgenic plants (L1). The number of secondary xylem cell layers was counted in toluidine blue-stained anatomical sections of the 7<sup>th</sup> internode of WT and *PtoIAA9m*-OE transgenic plants. (f) Percentage of secondary xylem and bark in the stem of WT and *PtoIAA9m*-OE transgenic plants (L1). The area of secondary xylem, bark and total stem was measured via IMAGEJ in toluidine blue-stained anatomical sections of the 7<sup>th</sup> internode of WT and *PtoIAA9m*-OE transgenic plants. (g) Detailed observation of the cambial zone and woody cells of secondary xylem in WT and *PtoIAA9m*-OE plants. The images were captured on toluidine blue-stained anatomical sections of the 7<sup>th</sup> internode of the corresponding lines. White lines indicate cambium (Ca), and red stars represent the cells of early developing xylem (EDX). V, vessel. (h, i) Quantification of the size of a single fiber (h) and vessel (i) cell in stem of WT and *PtoIAA9m*-OE plants. The area of fiber and vessel cells was measured and calculated via IMAGEJ based on the images of toluidine blue-stained anatomical sections as described in the Materials and Methods section. (j) Density of vessels in stem of WT and *PtoIAA9m*-OE plants. The number of vessels was counted based on the images of toluidine blue-stained anatomical sections. Error bars represent SD. Asterisks indicate significant differences with respect to values of WT (Student *t*-test): \*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ;  $n = 4$ . Bars: (a) 5 cm; (d) upper, 250  $\mu\text{m}$ , lower, 100  $\mu\text{m}$ ; (g) 50  $\mu\text{m}$ .

(Figs 2g, S5f). These cells were not stained by toluidine blue, suggesting that lignin deposition was not initiated for secondary cell wall formation. These cells were distinctly larger than cambial cells but smaller than mature xylem cells (Figs 2g, S5f). These features indicated that these cells were newly differentiated from cambium zone towards xylem and undergoing cell expansion, and thereby considered as early developing xylem (EDX).

However, these EDX cell layers were greatly reduced and almost absent in cross-sections of *PtoIAA9m*-OE stems (Fig. 2g), resembling EDX cells just before dormancy due to reduced rate of cambium cell periclinal division. This greatly reduced EDX cell layer suggests that overexpression of stabilized PtoIAA9 inhibited the periclinal division of cambium, thus leading to reduced wood formation.

Quantitative measurement showed that the cell size of xylem fibers and vessels was reduced by 36% and 28% in *PtoIAA9m*-OE plants relative to WT (Fig. 2g–i). By contrast, vessel density was significantly increased by 42% in *PtoIAA9m*-OE plants (Fig. 2g,j). Cell wall thickness was not affected by overexpressing *PtoIAA9m* (ns;  $P > 0.05$ ; Fig. S5g). Together, these data suggested that *PtoIAA9m* inhibited developing xylem cell (vessel and fiber) differentiation and expansion.

### PtoIAA9 interacts with PtoARF5s

PtoIAA9 contains intact Domains III and IV (also called PB1 domain) that mediate physical interaction with ARFs, a family of plant-specific B3-type transcription factors directly driving the expression of auxin responsive genes (Vanneste & Friml, 2009). We speculated that PtoIAA9 might interact with ARF proteins to regulate wood formation in poplar. A previous study has shown that ARF5/MP is a key regulator of auxin-dependent vascular patterning during embryogenesis and leaf vein patterning in Arabidopsis (Donner *et al.*, 2009; Brackmann *et al.*, 2018). The poplar genome harbors duplicated *ARF5* members, designated *ARF5.1* and *ARF5.2*, which are expressed in secondary vascular tissues of stems (Johnson & Douglas, 2007). Poplar ARF5s share conserved modular structure with their Arabidopsis homolog containing intact Domains III and IV (Fig. S6a), implying that they may interact with PtoIAA9. To determine the interactions between PtoIAA9 and PtoARF5s, we performed Y2H assay (Fig. 3a). Owing to the presence of the B3 domain with DNA-binding activity, the coding region of PtoARF5s was fused to AD, and the resulting fused protein was subsequently coexpressed in yeast cells with PtoIAA9-BD. Both PtoARF5 paralogs were shown to interact with PtoIAA9, but PtoARF5.1 displayed stronger interaction ability than PtoARF5.2 (Fig. 3a). The PtoIAA9–PtoARF5 interaction was subsequently confirmed by BiFC analyses (Fig. 3b).

Despite the physical interactions, the PtoIAA9–PtoARF5-dependent regulation on wood formation requires their coexpression in secondary vascular tissues. Analysis of expression patterns based on RNAseq datasets (Sundell *et al.*, 2017) revealed the highly correlated transcript accumulation between IAA9 and ARF5s in poplar wood-forming tissues (Fig. S6b,c). Similar to the expression patterns of *PtoIAA9*, RNA *in situ* hybridization results showed that *PtoARF5.1* was highly expressed in developing wood-associated tissues of poplar (Fig. 3c,d).

### *PtoARF5.1* rescues *PtoIAA9m*-affected secondary xylem differentiation

Physical interactions and overlapping expression of *PtoIAA9* and *PtoARF5s* suggested their cooperative involvement in the regulation of wood formation. To test this hypothesis, functional complementation of *PtoARF5.1* was performed on *PtoIAA9m*-OE-resulting phenotypes of secondary xylem development (Fig. 4). To avoid the undesirable effects of *PtoIAA9* overexpression and other endogenous *Aux/IAA* genes on its function, a full-length *PtoARF5.1* and a truncated form without C-terminal

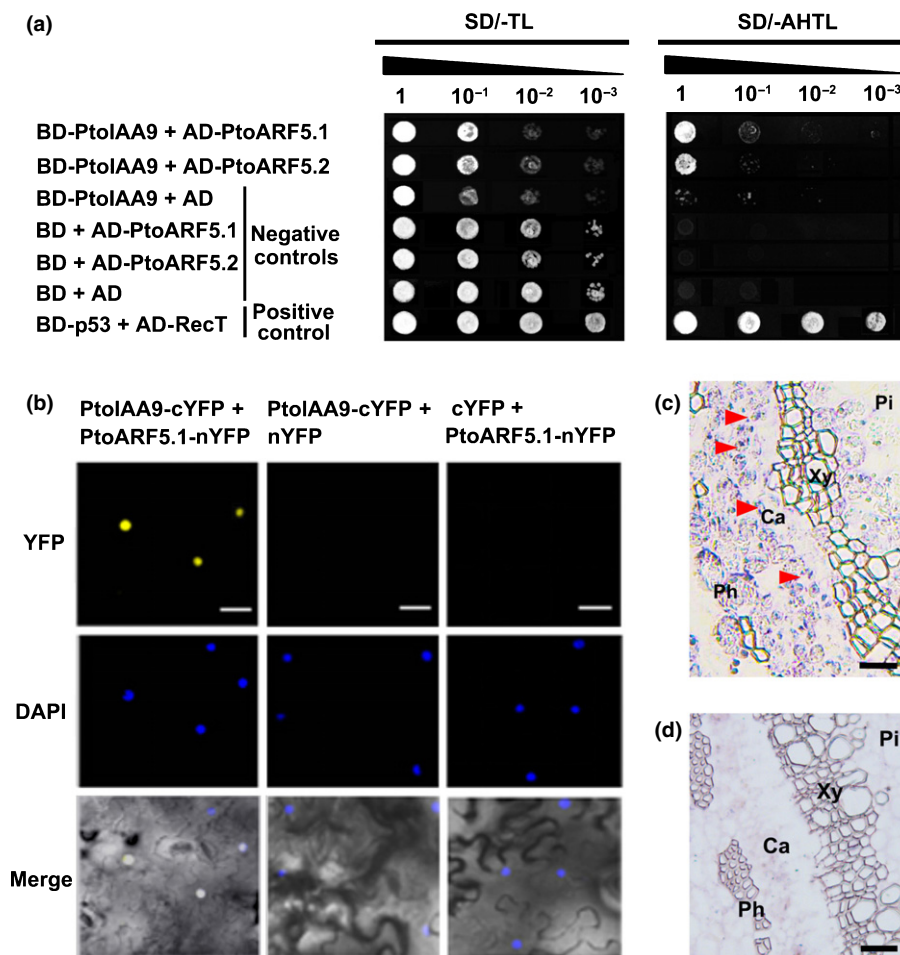
Domains III and IV responsible for Aux/IAA interactions (*PtoARF5.1Δ*) were introduced into the *PtoIAA9m*-OE transgenic lines, respectively (Figs 4, S6). Y2H tests showed that the C-terminal deletion abolished the interactions of PtoARF5.1 with the PtoIAA9 protein (Fig. S7a), but did not affect transcriptional activation (Fig. S7b).

Phenotypic analysis of transgenic plants showed that reduced growth of the *PtoIAA9m*-OE lines could be partially rescued by *PtoARF5.1*, whereas the truncated *PtoARF5.1* displayed stronger recovery than the full-length form (Fig. 4a). Quantitative measurement of plant height and stem diameter confirmed these changes in different transgenic lines (Fig. 4b,c). By contrast with the partial recovery (*c.* 50%) by the full-length *PtoARF5.1*, the *PtoARF5.1Δ* almost completely rescued the decreased stem diameter of the *PtoIAA9m*-OE plants relative to that of WT (Figs 4c, S7d). Stem cross-sectioning was conducted to establish if rescue of secondary xylem growth by constitutive expression of *PtoARF5.1* occurred in the *PtoIAA9m*-OE lines (Fig. 4d–f). Both the full-length and the truncated *PtoARF5.1* were able to rescue the inhibition of overexpressing *PtoIAA9m* on secondary xylem development, as evidenced by the increased number of xylem cell layers and xylem percentage (40% by *PtoARF5.1* and 94% by *PtoARF5.1Δ*) in stems (Fig. 4d–f). Noticeably, *PtoARF5.1Δ* conferred stronger recovery of the *PtoIAA9m*-OE-resulting defective wood formation than the full-length *PtoARF5.1* (Fig. 4d–f), leading to even more xylem cell layers than WT.

Subsequently, the cambium zone and its adjacent cell layers were characterized in detail to reveal the possible modulation of cell differentiation by the PtoIAA9–PtoARF5 module (Fig. 5). By contrast with the absence of EDX in *PtoIAA9m*-OE plants, these cell layers reoccurred in both *PtoARF5.1* and *PtoARF5.1Δ*-complementing lines (Fig. 5a). Similar recovery of defective xylem cell expansion resulting from overexpression of *PtoIAA9m* was detected for *PtoARF5.1* and *PtoARF5.1Δ*, as indicated by measurement of fiber and vessel cell size (Fig. 5b–d). Moreover, elevated vessel density was also rescued by introduction of *PtoARF5.1* and *PtoARF5.1Δ* into the *PtoIAA9m*-OE lines (Fig. 5b,e). The stronger recovery of *PtoARF5.1Δ* than its full-length form indicated that the PB1 domain mediating AUX/IAA-ARF protein interactions compromises the rescue of *PtoIAA9m*-resulting phenotypes of secondary xylem differentiation. Therefore, PtoARF5 is able to drive the PtoIAA9-dependent cellular behaviors for secondary xylem differentiation in poplar.

### The PtoIAA9–PtoARF5 module directly regulates expression of *PtoHB7/8*

Diverse auxin-triggered developmental phenotypes depend on expression of specific batteries of auxin responsive genes targeted by Aux/IAA-ARF pairs (Vanneste & Friml, 2009). To identify auxin responsive genes targeted by the PtoIAA9–PtoARF5 module during wood formation in poplar, comparative transcriptomic profiling via RNAseq was performed with the *PtoIAA9m*-OE lines and WT, revealing 3873 differentially expressed genes (fold change  $\geq 2$  and false discovery rate  $\leq 1\%$ ; Fig. S8a;



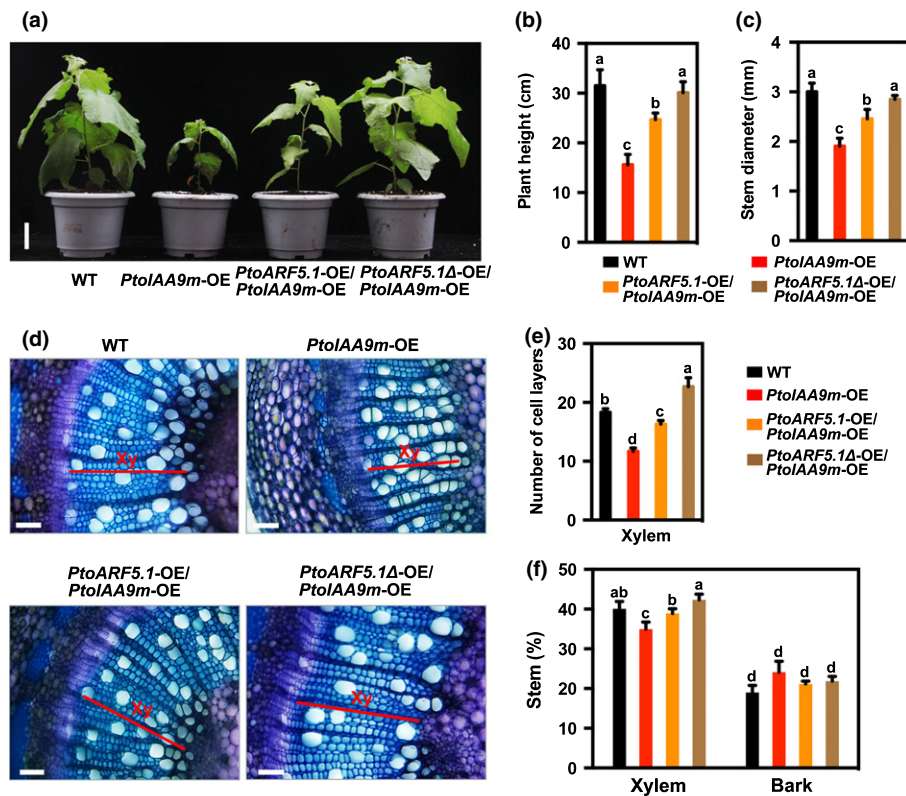
**Fig. 3** Protein interactions of PtoIAA9 with PtoARF5s and expression of *PtoARF5* in secondary vascular tissues of *Populus tomentosa* stem. (a) Yeast-two-hybrid analysis of protein interactions between PtoIAA9 and PtoARF5.1/5.2. PtoIAA9 and PtoARF5.1/5.2 were fused with a Gal4 DNA-binding domain (BD) and a GAL4 activation domain (AD), respectively. The interaction between BD-p53 and AD-RecT (SV40 large T-antigen) was used as a positive control, while those between blank constructs (BD or AD) with BD-PtoIAA9 or AD-PtoARF5.1/5.2 were used as negative controls. Yeast cells were inoculated on selective medium in a 10-fold gradient dilution. SD/-TL, double dropout medium lacking tryptophan and leucine; SD/-AHTL, quadruple dropout medium lacking adenine, histidine, tryptophan and leucine. (b) Bimolecular fluorescence complementation (BiFC) assays validating physical interactions between PtoIAA9 and PtoARF5 in nuclei. nYFP and cYFP represent the N- and C-terminal part of yellow fluorescent protein (YFP), respectively. The PtoIAA9-nYFP and PtoARF5.1-cYFP constructs were cotransfected into tobacco epidermal leaf cells via *Agrobacterium*-mediated infiltration. The blank constructs of cYFP or nYFP were cotransfected with PtoIAA9-nYFP or PtoARF5.1-cYFP, respectively, as negative controls. The fluorescence emitted by YFP was examined with a confocal microscope. Nuclei were identified by DAPI staining. (c, d) RNA *in situ* hybridization of *PtoARF5* in secondary vascular tissues of poplar stem. The 7<sup>th</sup> internodes of 6-wk-old poplar plants cultivated in soil were cross-sectioned for hybridization with sense (c) and anti-sense (d) probes. The probes were designed within the identical region of *PtoARF5.1* and *PtoARF5.2* for detection of both *ARF5* paralogs in poplar. Red triangles indicate the *in situ* hybridization signals. Ca, cambium; Ph, phloem; Pi, pith; Xy, xylem. Bars: (b) 50  $\mu$ m; (c, d) 200  $\mu$ m.

Methods S1; Dataset S1). A number of *Aux/IAAs*, *SAURs* and *GH3s*, which fall into classic families of early auxin responsive genes in plants (Guilfoyle, 1999), displayed significantly reduced transcript abundance (2.1- to 15.2-fold) in the *PtoIAA9m*-OE lines (Fig. S7b), suggesting that auxin signaling was weakened by *PtoIAA9m* overexpression in poplar. Gene ontology (GO) analysis of differentially expressed genes showed enrichment of the term of nucleic acid binding transcription factor activity (Fig. S7c). We thus investigated the differentially expressed genes encoding transcription factors that are known to regulate wood formation in poplar (Fig. S7d). All of these genes, including *WNDs*, *ANT* and *HBs*, displayed decreased expression in the *PtoIAA9m*-OE lines (Fig. S8d). It has previously been established

that class III HD-ZIP transcription factors are critical regulators of vascular formation and patterning in plants (Ramachandran *et al.*, 2017). This family, designated HB transcription factors, comprises eight members in poplar (Fig. S8e). RNAseq and qRT-PCR showed that five of them displayed *PtoIAA9m*-inhibited expression (Fig. S8d,f). Among them, PtoHB7 and PtoHB8 were identified as critical regulators of vascular cambium differentiation to secondary xylem in poplar (Zhu *et al.*, 2013). Therefore, we investigated whether the *PtoHB7/8* genes are downstream targets of PtoIAA9–PtoARF5 during wood formation.

To test this hypothesis, we determined the regulation of PtoIAA9–PtoARF5 on *PtoHB7* in the *PtoIAA9m*-OE lines. As



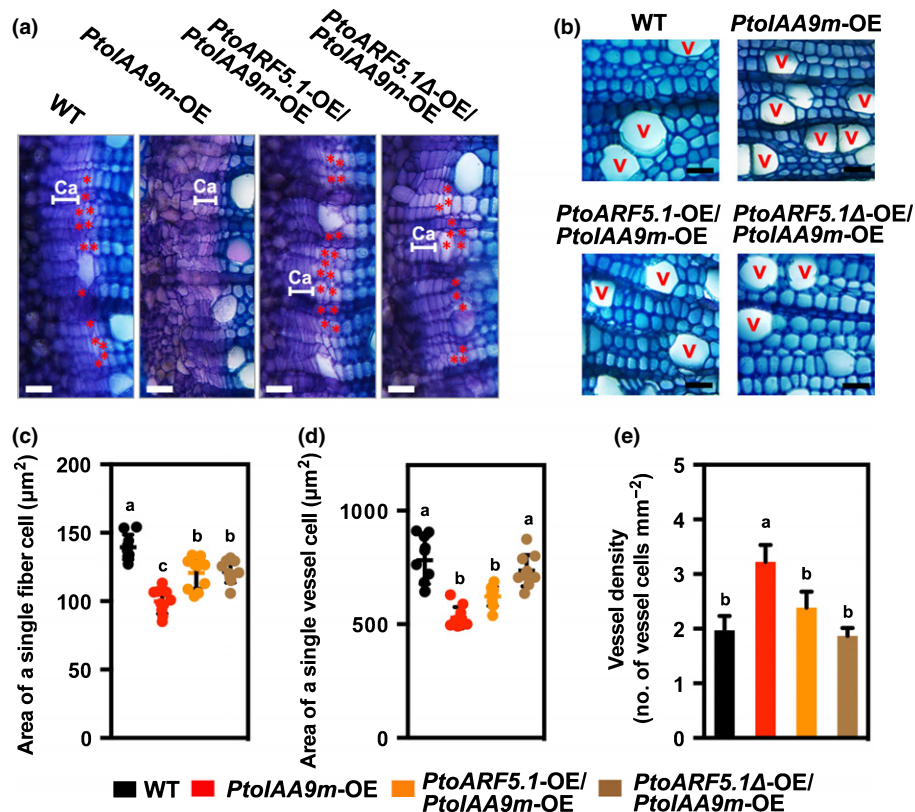


**Fig. 4** Rescue of *PtoIAA9m*-resulting wood-associated phenotypes in *Populus tomentosa* by *PtoARF5*. (a) Recovered morphological phenotypes of 1.5-month-old *PtoIAA9m*-overexpressing (OE) transgenic poplar plants by *PtoARF5*. The constructs of *35S:PtoARF5.1* and *35S:PtoARF5.1Δ* (encoding for 1–665 amino acids of *PtoARF5.1* without C-terminal III/IV domains) were transformed into poplar in the background of *PtoIAA9m-OE*. (b, c) Measurement of plant height and stem diameter of wild-type (WT), *PtoIAA9m-OE*, *PtoARF5.1-OE/PtoIAA9m-OE* and *PtoARF5.1Δ-OE/PtoIAA9m-OE* transgenic poplar lines. (d) Anatomical sections stained with toluidine blue of the 7<sup>th</sup> internode of 1.5-month-old WT, *PtoIAA9m-OE*, *PtoARF5.1-OE/PtoIAA9m-OE* and *PtoARF5.1Δ-OE/PtoIAA9m-OE* plants. Red lines indicate xylem. Xy, xylem. (e) Quantification of secondary xylem cell layers of WT, *PtoIAA9m-OE*, *PtoARF5.1-OE/PtoIAA9m-OE* and *PtoARF5.1Δ-OE/PtoIAA9m-OE* lines. The number of xylem cell layers was counted in toluidine blue-stained anatomical sections of the 7<sup>th</sup> internode of WT, *PtoIAA9m-OE*, *PtoARF5.1-OE/PtoIAA9m-OE* and *PtoARF5.1Δ-OE/PtoIAA9m-OE* plants. (f) Percentage of secondary xylem and bark in the stem of WT, *PtoIAA9m-OE*, *PtoARF5.1-OE/PtoIAA9m-OE* and *PtoARF5.1Δ-OE/PtoIAA9m-OE* plants. The area of secondary xylem, bark and total stem was quantified via IMAGEJ in toluidine blue-stained cross-sections of the 7<sup>th</sup> internode of WT, *PtoIAA9m-OE*, *PtoARF5.1-OE/PtoIAA9m-OE* and *PtoARF5.1Δ-OE/PtoIAA9m-OE* plants. Error bars represent SD. The letters above error bars indicate statistically significant differences (one-way ANOVA followed by Dunnett's test for pairwise comparisons;  $n = 4$ ). Bars: (a) 5 cm; (d) 100  $\mu$ m.

shown in Fig. 6(a), the expression of both *PtoHBs* was significantly decreased by 71% and 74% in the *PtoIAA9m-OE* lines, respectively, and could be rescued by the introduction of the truncated *PtoARF5.1* to even higher levels than that of WT. We also found that expression of *PtoHB7* and *PtoHB8* was increased 3.9- and 1.9-fold, respectively, in response to auxin treatment (Fig. 6b). Noticeably, overexpression of *PtoIAA9* almost completely blocked the auxin-induced expression of both *PtoHB7* and *PtoHB8* in the wood-forming stem (Fig. 6b), indicating that the auxin responsive expression of *PtoHB7* and *PtoHB8* is dependent on *PtoIAA9*. Promoter sequence analysis revealed that a series of core and canonical auxin response elements (*AuxREs*), potential DNA binding sites for ARFs, are predicted in the 1.5 kb upstream regions of the start codons of *PtoHB7* and *PtoHB8* (Fig. 6c,d). ChIP was performed using *PtoARF5.1-HA* transgenic plants and the enrichment of *AuxREs* harbored by the *PtoHB7* and *PtoHB8* promoters was quantified by qPCR (Fig. 6e,f). In comparison with negative controls, some *AuxRE*-containing promoter regions of *PtoHB7* (Regions II/III) and *PtoHB8* (Regions

III/IV) were significantly enriched after immunoprecipitation (Fig. 6e,f), indicating the direct binding of *PtoARF5.1* to their promoters. The *PtoIAA9/PtoARF5*-dependent regulation of *PtoHB7/8* was subsequently confirmed by effector–reporter assays in transiently expressed tobacco leaves. The reporter constructs carrying the *GUS* reporter gene driven by the 1.5 kb promoter regions of *PtoHB7* and *PtoHB8*, respectively, were co-transfected with different combinations of the effectors harboring *PtoIAA9* and *PtoARF5.1* (Fig. S9). Determination of *GUS* activity revealed that *PtoARF5.1* was able to independently activate expression of *PtoHB7*, whereas the activation was completely abolished by the addition of *PtoIAA9* ( $P < 0.05$ ; Fig. S9a). By contrast, *PtoIAA9* did not compromise the activation induced by *PtoARF5.1Δ* without the C-terminal III/IV domains, indicating that *PtoIAA9*-mediated repression of *PtoHB7* expression relies on its interactions with *PtoARF5.1* via the III/IV domains. Similar results were obtained from the promoter of *PtoHB8* (Fig. S9b).

To further validate direct regulation of *PtoARF5.1* on *PtoHBs* via binding to *AuxREs*, short promoter fragments harboring



**Fig. 5** *PtoARF5* restores *PtoIAA9*-affected phenotypes of secondary xylem cells in *Populus tomentosa* during wood formation. (a, b) Detailed observation of cambial zone and woody cells of secondary xylem in wild-type (WT), *PtoIAA9m-OE*, *PtoARF5.1-OE/PtoIAA9m-OE* and *PtoARF5.1Δ-OE/PtoIAA9m-OE* plants. The images were captured on toluidine blue-stained anatomical sections of the 7<sup>th</sup> internode of the corresponding lines. White lines indicate cambium (Ca), and red stars represent the cells of early developing xylem (EDX). V, vessel. Bars, 50  $\mu\text{m}$ . (c, d) Quantification of size of a single fiber and vessel cell in stems of WT, *PtoIAA9m-OE*, *PtoARF5.1-OE/PtoIAA9m-OE* and *PtoARF5.1Δ-OE/PtoIAA9m-OE* plants. The area of fiber and vessel cells was measured and calculated via IMAGEJ based on images of toluidine blue-stained anatomical sections as described in the Materials and Methods. (e) Density of vessels in stems of WT, *PtoIAA9m-OE*, *PtoARF5.1-OE/PtoIAA9m-OE* and *PtoARF5.1Δ-OE/PtoIAA9m-OE* plants. The number of xylem vessels was counted based on images of toluidine blue-stained anatomical sections. Error bars represent SD. The letters above error bars indicate significant differences (one-way ANOVA followed by Dunnett's test for pairwise comparisons;  $n = 10-12$ ).

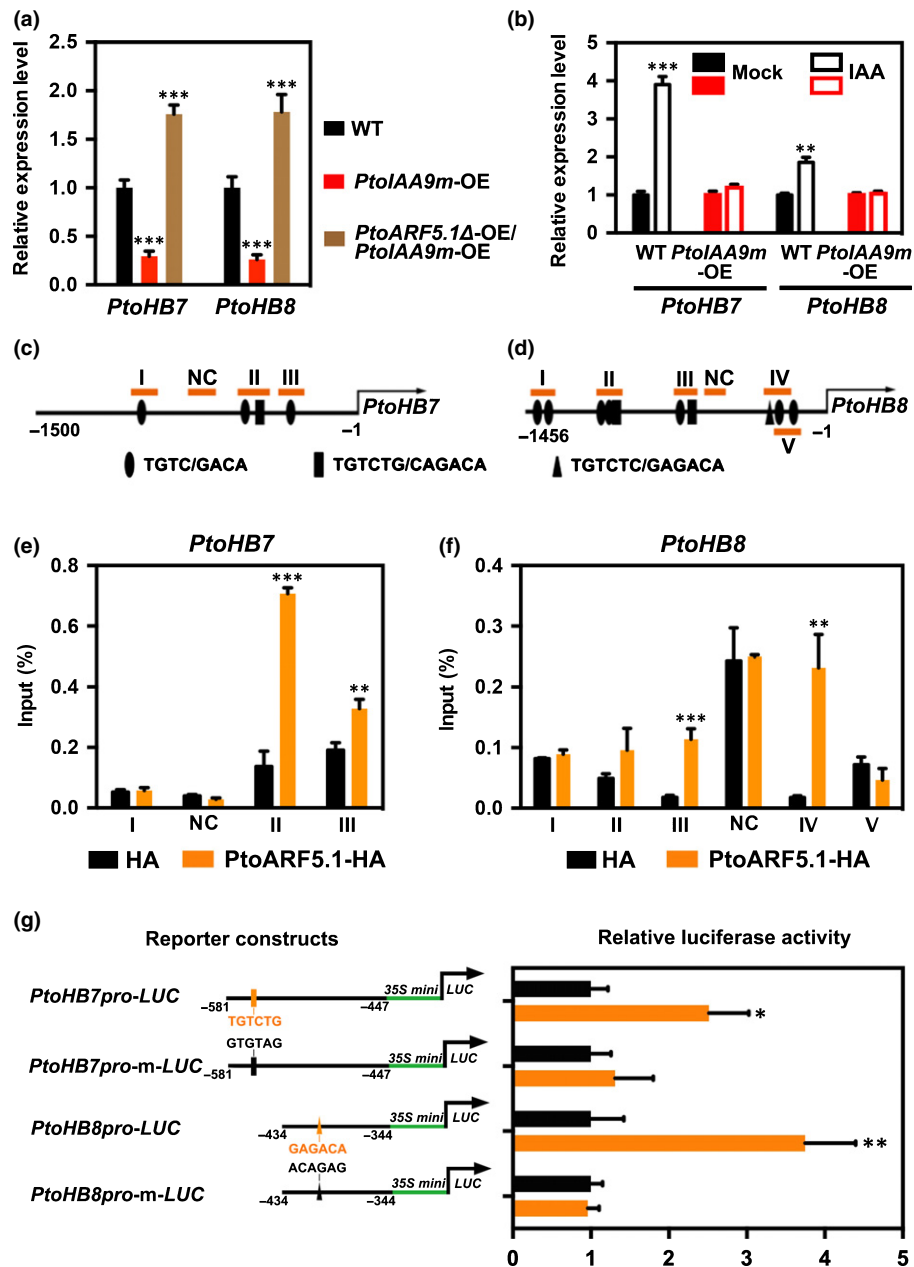
canonical *AuxREs* (−447 to −581 bp upstream of the start codon of *PtoHB7*, and −344 to −434 bp upstream of the start codon of *PtoHB8*), which were strongly enriched in ChIP assays, were constructed to drive the firefly luciferase reporter. Consistently, *PtoARF5.1* significantly activated the expression of luciferase reporter driven by the *AuxRE*-harbored promoter fragments of *PtoHB7/8* ( $P < 0.05$ ; Fig. 6g). By contrast, *PtoARF5.1*-driven activation disappeared when the *AuxREs* harbored in the promoter fragments of *PtoHB7* and *PtoHB8* were disrupted by site-directed mutagenesis (Fig. 6g). Therefore, the *PtoIAA9*–*PtoARF5* module directly regulates the expression of *PtoHB7* and *PtoHB8* via binding of *PtoARF5* to the *AuxREs* within their promoters.

#### *PtoHB7* partially rescues *PtoIAA9m*-resulting phenotypes of wood formation

To establish the link of biological functions between auxin signaling and the HB transcription factors during wood formation, we conducted complementation of the *PtoIAA9m*-resulting impaired secondary xylem phenotypes by *PtoHB7* via stable transformation

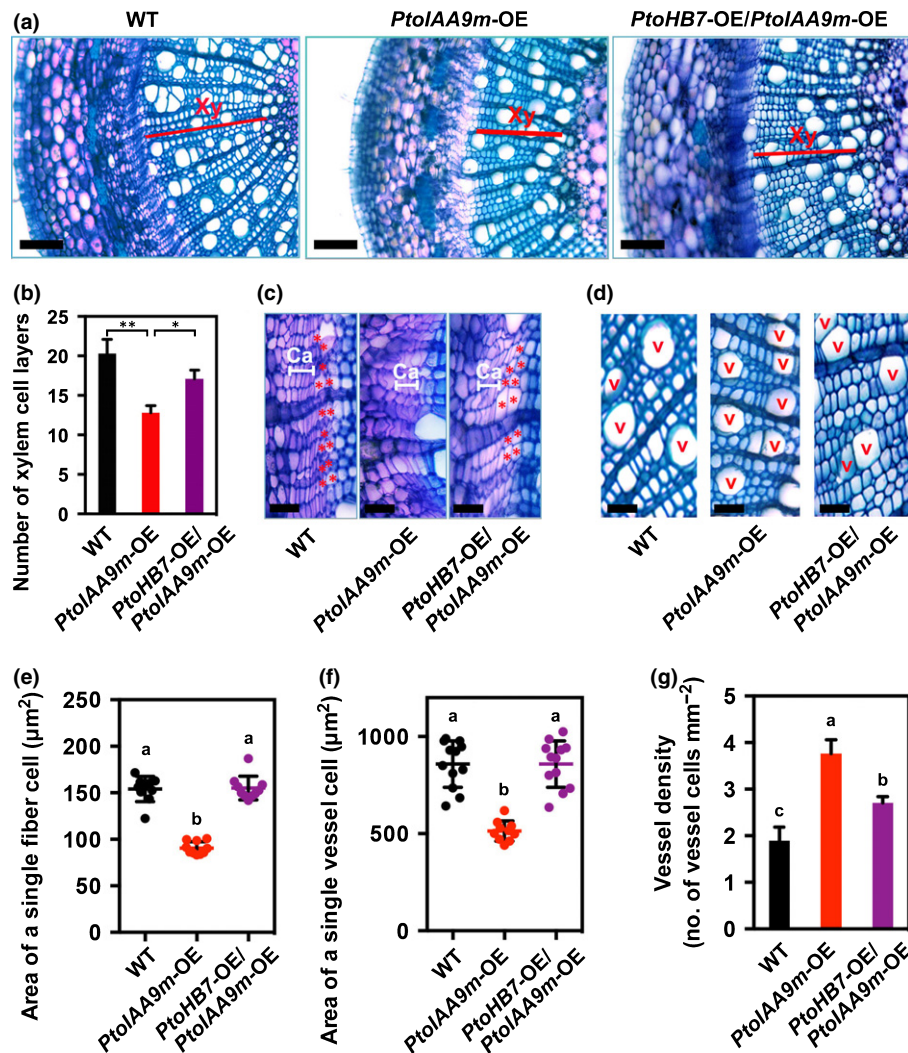
(Figs 7, S9). Phenotypic characterization revealed partial recovery of the *PtoIAA9m*-repressing plant growth and stem development by *PtoHB7* (Fig. S10b–d). Stem cross-sections showed that introduction of *PtoHB7* into the *PtoIAA9m-OE* background could partially rescue the impaired secondary xylem development, as validated by the quantification of xylem cell layers (Fig. 7a,b). Constitutive expression of *PtoHB7* led to the reoccurrence of the EDX cell layers, which were repressed in the *PtoIAA9m-OE* lines (Fig. 7c). Similarly, the decreased cell size of xylem fibers and vessels as well as the enhanced vessel density were partially rescued by the introduction of *PtoHB7* into the *PtoIAA9m-OE* plants (Fig. 7d–h).

Moreover, we determined the expression of several genes, including *ACL5*, *CesA7A*, *PAL4*, *GT43B* and *WND1B*, which were previously reported *HB7/8*-regulated genes involved in wood formation of poplar (Milhinhos *et al.*, 2013; Zhu *et al.*, 2013). The down-regulated expression of these genes in *PtoIAA9m-OE* plants was at least partially rescued in *PtoARF5.1*- and *PtoHB7*-complementing transgenic lines (Fig. S11). A similar result was obtained in *EXPA1*, a marker gene for cell expansion during wood formation (Sundell *et al.*, 2017). We also



**Fig. 6** Direct regulation of *PtoIAA9/ARF5* on *PtoHB7* and *PtoHB8* in *Populus tomentosa*. (a) Expression levels of *PtoHB7* and *PtoHB8* in wild-type (WT), *PtoIAA9m-OE* and *PtoARF5.1Δ-OE/PtoIAA9m-OE* lines determined by qRT-PCR. Stem tissues of 6-wk-old poplar plants cultivated in soil were collected for RNA extraction. 18S rRNA was used as a reference gene. WT values were normalized to 1. (b) Auxin-induced expression of *PtoHB7* and *PtoHB8* in WT and *PtoIAA9m-OE* lines. Microcutting-propagated poplar seedlings cultivated *in vitro* for 4 wk were subjected to 5 μM IAA for 6 h, and stem tissues were collected for RNA extraction followed by qRT-PCR assays. 18S rRNA was used as a reference gene. The values for mock treatment were normalized to 1. For (a, b), error bars represent SD, and one-way ANOVA followed by Dunnett's test for pairwise comparisons with respect to values of WT (a) or mock treatment (b) was performed to detect statistically significant differences (\*\*,  $P < 0.001$ ; \*\*\*,  $P < 0.001$ ;  $n = 4$ ). (c, d) Distribution of canonical and core auxin response elements (*AuxREs*) harbored in the promoter regions of *PtoHB7* (c) and *PtoHB8* (d). Orange lines with I–III indicate *AuxRE*-harbored fragments amplified by ChIP-qPCR, while NC represents *AuxRE*-free negative control. (e, f) ChIP-qPCR analysis of *PtoARF5* protein fused with HA tag with the promoter region of *PtoHB7* (e) and *PtoHB8* (f). Shoot tissues of 1-month-old poplar plants were used. Error bars represent SD. Student's *t*-test was performed to evaluate significant differences between values of WT and those of *PtoARF5.1-HA* for each region (\*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ;  $n = 3$ ). (g) *AuxRE*-dependent transactivation via transient cotransformation assay. For reporters, the 1-kb promoter fragments of *PtoHB7* and *PtoHB8* were constructed to drive the expression of firefly luciferase (LUC). The *AuxREs* found in these promoter fragments were disrupted via site-directed mutagenesis to generate *PtoHB7/8pro-m-LUC*. The effector encodes *PtoARF5.1* driven by the *CaMV 35S* promoter. Ratios of firefly luciferase to *Renilla* luciferase activity after cotransformation into tobacco leaf epidermal cells with different reporter and effector construct combinations were tested. Values for the blank effector were normalized to 1. Error bars represent SD. Student's *t*-test was performed to evaluate significant differences between values of blank effector and those of *PtoARF5.1* (\*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ;  $n = 3$ ).





**Fig. 7** *PtoHB7* partially rescues deficient phenotypes of wood formation resulting from constitutive expression of *PtoIAA9* in *Populus tomentosa*. (a) Cross-sections stained with toluidine blue of the 7<sup>th</sup> internode of 1-month-old WT, *PtoIAA9m-OE* and *PtoHB7-OE/PtoIAA9m-OE* lines. Red lines indicate xylem. Xy, xylem. (b) Number of xylem cell layers of wild-type (WT), *PtoIAA9m-OE* and *PtoHB7-OE/PtoIAA9m-OE* lines. (c, d) Detailed phenotypes of cambial zone and woody cells of secondary xylem in wood-forming stem of WT, *PtoIAA9m-OE* and *PtoHB7/PtoIAA9m-OE* plants. The images were captured on toluidine blue-stained anatomical sections of the 7<sup>th</sup> internode of the corresponding lines. White lines indicate cambium (Ca), and red stars represent the cells of early developing xylem (EDX). V, vessel. (e, f) Quantification of size of a single fiber and vessel cell in stems of WT, *PtoIAA9m-OE*, and *PtoHB7/PtoIAA9m-OE* plants. The area of fiber and vessel cells was measured and calculated via IMAGEJ based on images of toluidine blue-stained anatomical sections as described in the Materials and Methods. (g) Density of vessels in stem of WT, *PtoIAA9m-OE* and *PtoHB7/PtoIAA9m-OE* plants. The number of xylem vessels was counted based on the images of toluidine blue-stained anatomical sections. Error bars represent SD. The letters above error bars indicate significant differences (one-way ANOVA followed by Dunnett's test for pairwise comparisons;  $n = 10-12$ ). Bars: (a) 100  $\mu\text{m}$ ; (c, d) 50  $\mu\text{m}$ .

found that expression of *WND6A* and *B*, which are the closest homologs of Arabidopsis *VND6* and *VND7* (Zhong *et al.*, 2011), was enhanced by *PtoIAA9m* overexpression, but was rescued by *PtoARF5.1* and *PtoHB7* (Fig. S11), in accordance with elevated vessel formation by *PtoIAA9m-OE* and complemented by *PtoARF5* and *PtoHB7*. These results strongly suggest that auxin-induced secondary xylem development depends on the activation of the *PtoHB* genes by the *PtoIAA9/PtoARF5* complex in poplar.

## Discussion

Auxin is considered a positional signal that drives cambium-derived wood formation (Uggla *et al.*, 1996). Previous studies

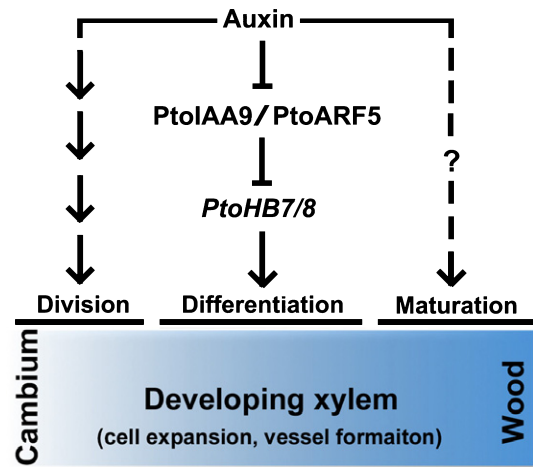
have demonstrated auxin-mediated regulation of cambial cell division in trees (Sundberg *et al.*, 2000; Nilsson *et al.*, 2008). However, a highly sensitive analysis of auxin signaling in Arabidopsis revealed its maximal levels in differentiating cambial descendants, spatially divergent from the maximum auxin concentration in cambial initials (Brackmann *et al.*, 2018). In this study, we report that auxin coordinates multiple aspects of the cellular changes occurring for cambium specification into secondary xylem through the IAA9-ARF5 pathway in *Populus*. We further demonstrate that the *PtoIAA9-PtoARF5* module directly targets *PtoHB7* and *PtoHB8* to mediate auxin-triggered early developing xylem (EDX) cell differentiation during wood formation.

Aux/IAA proteins function as molecular switches of auxin signaling conserved among flowering plants (Paponov *et al.*, 2009). Comprehensive analyses of *Populus Aux/IAAs* revealed variable levels and spatial patterns of expression across wood-forming tissues (Fig. S2), implying functional diversification in wood formation (Moyle *et al.*, 2002; Kalluri *et al.*, 2007). The transcripts of *IAA9*, *-16.1* and *-16.2* were enriched in cambium-neighboring cells undergoing differentiation into xylem but declined with xylem maturation (Fig. S2c). Highly correlated expression patterns among these *Aux/IAA* genes (Fig. S2c,d) and common protein interactions with ARF5 (Fig. S6d) suggested their functional redundancy in wood formation. Although compromised, wood differentiation could not be completely suspended by overexpressing *PtoIAA9*, implying regulatory complexity depending on some other pathways besides auxin signaling.

Compromised auxin signaling by expressing stable Aux/IAAs allows auxin-mediated coordination of cellular behaviors during wood formation to be explored. Overexpression of a hybrid aspen *IAA3* (*PttIAA3*) attenuates periclinal cell division in cambium but enlarges cell files harboring anticlinal cell division, suggesting dual regulation of auxin signaling on cambial proliferation (Nilsson *et al.*, 2008). Despite more tightly arranged cambial cells (Fig. 2g), however, overexpressing *PtoIAA9* did not lead to significant changes in cambial cell layers (Fig. S5e), implicating its weak role in cambial activity. *IAA20.1* and *-20.2*, the *PttIAA3* homologs in our nomenclature, displayed much lower expression levels than *IAA9*, and different expression patterns from *IAA9* during wood formation (Fig. S2). Aux/IAAs usually require functional cooperation of ARFs via protein interactions, and harbor variable interacting affinities with different ARFs (Vanneste & Friml, 2009). Thus, specific ARF partners of *PttIAA3* and *PtoIAA9* may cause their differential regulation on cambial activity.

Compared to WT, the proportion of woody tissues of transgenic plants overexpressing *PtoIAA9m* was significantly reduced (Fig. 2e, f), implying inhibition of secondary xylem development by *PtoIAA9*. Both reduced xylem cell layers and restricted xylem cell size led to *PtoIAA9*-inhibited wood formation that was rescued by overexpressing *PtoARF5* (Figs 2, 4, 5). Cambial proliferation and cambium specification to xylem cooperatively determine the number of xylem cell layers (Sanchez *et al.*, 2012). Given the excluded *PtoIAA9*-mediated effects on cambial proliferation, we propose that the reduced periclinal cell division rate may be responsible for *PtoIAA9*-repressed wood formation. This is supported by the disappearance and reappearance of the EDX cell layers in *PtoIAA9*-overexpressing and *PtoARF5.1*-complementing lines, respectively (Figs 2, 5). These cell layers are initially specified from cambial cells and at early stage of wood formation. Active cambium periclinal cell division leads to rapid xylem mother cell accumulation, while a low rate of cambium periclinal cell division slowly produces xylem mother cells, resulting in an absence of the EDX cell layers in the *PtoIAA9m*-OE lines. Similar phenotypes were also found in trees during dormancy (Gričar *et al.*, 2014). Therefore, specification and expansion of xylem cells during wood differentiation are coordinated by IAA9-ARF5-mediated auxin signaling in poplar.

Recent studies have shown that Arabidopsis MP/ARF5 directly represses *WOX4* expression for restrained stem cell quantity



**Fig 8** A model for Aux/IAA-ARF-HB-dependent wood formation in *Populus tomentosa*. Auxin displays complex regulation of various steps during wood formation in poplar. The auxin-dependent cambial cell division was demonstrated by Nilsson *et al.* (2008). The role of auxin in xylem maturation, mainly including secondary wall deposition, remained unclear. In this study, *PtoIAA9*-ARF5-HB7/8-mediated wood differentiation is proposed. The auxin-dependent *PtoIAA9*-ARF5 module directly regulates the expression of *PtoHB7/8*, which drives multiple cellular events including xylem cell specification, expansion and vessel formation during secondary xylem development. Arrows represent a positive regulatory action of one component on another. Lines ending with a trait represent a negative regulatory action. Dotted lines indicate indirect regulation.

(Brackmann *et al.*, 2018), and switches of MP/ARF5 phosphorylation mediated by BIL1 kinase integrate peptide and cytokinin signaling for cambial activity (Han *et al.*, 2018). Here we provide evidence in *Populus* that ARF5 coordinates multiple behaviors for woody cell differentiation via directly targeting HB7/8 paralogs. BDL/IAA12 is a canonical Aux/IAA partner of ARF5/MP in Arabidopsis during embryogenesis (Hamann *et al.*, 2002). Due to promiscuous protein interactions between Aux/IAAs and ARFs, cooperative functions of IAA12-ARF5 could not be excluded in poplar. The significantly lower transcript abundance of *IAA12* paralogs than *IAA9* and their different expression patterns in wood-forming stems (Fig. S2) suggest possible variable roles in wood differentiation. We expected that *IAA9* also regulates procambial specification during root development and leaf vein patterning. Indeed, the knockdown of *IAA9* in tomato led to abnormal leaf formation with different vascular tissue patterning (Wang *et al.*, 2005).

Arabidopsis HD-ZIP III transcription factors are key regulators in orchestrating vascular patterning in developmental contexts of root, leaf and stem (Ramachandran *et al.*, 2017). *AtHB8* displays specific expression in procambial cells and is involved in vascular differentiation (Baima *et al.*, 2001). *HB7*, the poplar ortholog of *AtHB8*, has been revealed to be a dose-dependent regulator that balances cambium activity and xylem differentiation during secondary growth (Zhu *et al.*, 2013). We found that a series of core and canonical auxin response elements (*AuxREs*), the potential DNA binding sites for ARFs, are present in the promoters of poplar *HB7* and *HB8*. ChIP and effector-reporter tests revealed transcriptional activation of *PtoARF5.1* via

direct binding to the *AuxRE* elements harbored in the promoters of *HB* genes (Fig. 6). Importantly, PtoHB7 is able to functionally rescue phenotypes of xylem cell specification, expansion and vessel density (Fig. 7). Therefore, our data suggested that the PtoIAA9–PtoARF5 module directly targets *PtoHB7* and *PtoHB8* for regulating secondary cell differentiation. In Arabidopsis, ARF5/MP has been identified to control leaf vein patterning by directly targeting *AtHB8* (Donner *et al.*, 2009), demonstrating ARF5-HB7/8 as a conserved pathway for vascular patterning of leaves and stems in herbaceous and woody species. Moreover, the *HB8* orthologs can directly regulate the expression of *ACL5* encoding thermospermine synthase in both Arabidopsis and poplar for xylem differentiation, and also form a feedback regulation on the auxin pathway (Milhinhos *et al.*, 2013; Baima *et al.*, 2014). We also found that the expression of *ACL5* was significantly decreased by *PtoIAA9* overexpression and rescued by *PtoARF5* complementation (Fig. S11).

Collectively, our data allow us to propose a model for IAA9-ARF5-mediated coordination for wood differentiation (Fig. 8). Auxin releases the repression of PtoIAA9, by triggering its protein degradation, on PtoARF5-activated auxin responsive gene expression during cambium-derived wood formation. In parallel, auxin-inducible transcript abundance of *PtoIAA9* switches-off auxin signaling in a self-controlled manner. The PtoIAA9–PtoARF5 module directly targets *PtoHB7* and *PtoHB8* to mediate auxin regulation on multiple cellular events, including xylem specification, expansion and vessel formation for secondary xylem development. In conclusion, our findings show that the auxin-Aux/IAA-ARF pathway plays a key role in controlling wood formation via HB-driven cambium differentiation in *Populus*.

## Acknowledgements

We thank Drs Guoqing Niu, Xinqiang He and Jianquan Liu for helpful comments. This work was supported by the National Natural Science Foundation of China (31500544, 31670669, 31870657, 31800505 and 31870175), National Key Project for Research on Transgenic Plant (2016ZX08010-003) and Fundamental Research Funds for the Central Universities (XDJK2018AA005, XDJK2014a005). The authors declare no conflicts of interest.

## Author contributions

KL, CX and HCW designed the work; CX, FH, YS, XF, WL, CL, HY, YL and DF performed experiments and data analyses; CX drafted the manuscript; KL and HCW revised the manuscript. All authors approved the final version of the manuscript for publication. CX, YS, FH, XF and HY contributed equally to this work.

## ORCID

Keming Luo  <https://orcid.org/0000-0003-4928-7578>

## References

- Baima S, Forte V, Possenti M, Penalosa A, Leoni G, Salvi S, Felici B, Ruberti I, Morelli G. 2014. Negative feedback regulation of auxin signaling by ATHB8/ACL5-BUD2 transcription module. *Molecular Plant* 7: 1006–1025.
- Baima S, Possenti M, Matteucci A, Wisman E, Altamura MM, Ruberti I. 2001. The Arabidopsis ATHB-8 HD-zip protein acts as a differentiation-promoting transcription factor of the vascular meristems. *Plant Physiology* 126: 643–655.
- Bargmann BO, Vanneste S, Krouk G, Nawy T, Efroni I, Shani E, Choe G, Friml J, Bergmann DC, Estelle M *et al.* 2013. A map of cell type-specific auxin responses. *Molecular Systems Biology* 9: 688.
- Björklund S, Antti H, Uddestrand I, Moritz T, 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. *The Plant Journal* 52: 499–511.
- Bonan GB. 2008. Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *Science* 320: 1444–1449.
- Brackmann K, Qi J, Gebert M, Jouannet V, Schlamp T, Grünwald K, Wallner ES, Novikova DD, Levitsky VG, Agustí J *et al.* 2018. Spatial specificity of auxin responses coordinates wood formation. *Nature Communications* 9: 875.
- Campbell L, Turner S. 2017. Regulation of vascular cell division. *Journal of Experimental Botany* 68: 27–43.
- Chen S, Songkumarn P, Liu J, Wang G. 2009. A versatile zero background T-vector system for gene cloning and functional genomics. *Plant Physiology* 150: 1111–1121.
- Dejardin A, Laurans F, Arnaud D, Breton C, Pilate G, Leple JC. 2010. Wood formation in Angiosperms. *Comptes Rendus Biologies* 333: 325–334.
- Donner TJ, Sherr I, Scarpella E. 2009. Regulation of preprocambial cell state acquisition by auxin signaling in Arabidopsis leaves. *Development* 136: 3235–3246.
- Fukuda H. 2004. Signals that control plant vascular cell differentiation. *Nature Reviews Molecular Cell Biology* 5: 379–391.
- Gričar J, Prislán P, Gryc V, Vavřík H, de Luis M, Čufar K. 2014. Plastic and locally adapted phenology in cambial seasonality and production of xylem and phloem cells in *Picea abies* from temperate environments. *Tree Physiology* 34: 869–881.
- Guilfoyle TJ. 1999. Auxin-regulated genes and promoters. In: Hooykaas PJJ, Hall M, Libbenga KL, eds. *Biochemistry and molecular biology of plant hormones*. Leiden, the Netherlands: Elsevier, 423–459.
- Hamann T, Benkova E, Baurle I, Kientz M, Jurgens G. 2002. The Arabidopsis *BODENLOS* gene encodes an auxin response protein inhibiting MONOPTEROS-mediated embryo patterning. *Genes & Development* 16: 1610–1615.
- Han S, Cho H, Noh J, Qi J, Jung HJ, Nam H, Lee S, Hwang D, Greb T, Hwang I *et al.* 2018. BIL1-mediated MP phosphorylation integrates PXY and cytokinin signalling in secondary growth. *Nature Plants* 4: 605–614.
- Hardtke CS, Berleth T. 1998. The Arabidopsis gene *MONOPTEROS* encodes a transcription factor mediating embryo axis formation and vascular development. *EMBO Journal* 17: 1405–1411.
- Immanen J, Nieminen K, Smolander OP, Kojima M, Alonso Serra J, Koskinen P, Zhang J, Elo A, Mähönen AP, Street N *et al.* 2016. Cytokinin and auxin display distinct but interconnected distribution and signaling profiles to stimulate cambial activity. *Current Biology* 26: 1990–1997.
- Jefferson RA. 1987. Assaying chimeric genes in plants: the GUS gene fusion system. *Plant Molecular Biology Reporter* 5: 387–405.
- Jia Z, Sun Y, Yuan L, Tian Q, Luo K. 2010. The chitinase gene (*Bbchit1*) from *Beauveria bassiana* enhances resistance to *Cytospora chrysosperma* in *Populus tomentosa* Carr. *Biotechnology Letters* 32: 1325–1332.
- Johnson LA, Douglas CJ. 2007. *Populus trichocarpa* MONOPTEROS/AUXIN RESPONSE FACTOR5 (*ARF5*) genes: comparative structure, subfunctionalization, and *Populus Arabidopsis* microsynteny. *Canadian Journal of Botany-Revue Canadienne De Botanique* 85: 1058–1070.
- Kalluri UC, Difazio SP, Brunner AM, Tuskan GA. 2007. Genome-wide analysis of *AuxI* IAA and *ARF* gene families in *Populus trichocarpa*. *BMC Plant Biology* 7: 59.



- Leyser O. 2005. Auxin distribution and plant pattern formation: how many angels can dance on the point of PIN? *Cell* 121: 819–822.
- Matte Risopatron J, Sun Y, Jones B. 2010. The vascular cambium: molecular control of cellular structure. *Protoplasma* 247: 145–161.
- Milhinhos A, Prestele J, Bollhoner B, Matos A, Vera-Sirera F, Rambla JL, Ljung K, Carbonell J, Blázquez MA, Tuominen H *et al.* 2013. Thermospermine levels are controlled by an auxin-dependent feedback loop mechanism in *Populus* xylem. *The Plant Journal* 75: 685–698.
- Mockaitis K, Estelle M. 2008. Auxin receptors and plant development: a new signaling paradigm. *Annual Review of Cell and Developmental Biology* 24: 55–80.
- Moyle R, Schrader J, Stenberg A, Olsson O, Saxena S, Sandberg G, Bhalerao RP. 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.
- Muller CJ, Valdes AE, Wang G, Ramachandran P, Beste L, Uddenberg D, Carlsbecker A. 2016. PHABULOSA mediates an auxin signaling loop to regulate vascular patterning in Arabidopsis. *Plant Physiology* 170: 956–970.
- Nakajima K, Benfey P. 2002. Signaling in and out: control of cell division and differentiation in the shoot and root. *Plant Cell* 14(Suppl): S265–S276.
- Nilsson J, Karlberg A, Antti H, Lopez-Vernaza M, Mellerowicz E, Perrot-Rechenmann C, Sandberg G, Bhalerao RP. 2008. Dissecting the molecular basis of the regulation of wood formation by auxin in hybrid aspen. *Plant Cell* 20: 843–855.
- Papouov IA, Teale W, Lang D, Paponov M, Reski R, Rensing SA, Palme K. 2009. The evolution of nuclear auxin signalling. *BMC Evolutionary Biology* 9: 126.
- Perrot-Rechenmann C. 2010. Cellular responses to auxin: division versus expansion. *Cold Spring Harbor Perspectives in Biology* 2: a001446.
- Przemeck GK, Mattsson J, Hardtke CS, Sung ZR, Berleth T. 1996. Studies on the role of the Arabidopsis gene *MONOPTEROS* in vascular development and plant cell axialization. *Planta* 200: 229–237.
- Ragauskas AJ, Williams CK, Davison BH, Britovsek G, Cairney J, Eckert CA, Frederick WJ Jr, Hallett JP, Leak DJ, Liotta CL *et al.* 2006. The path forward for biofuels and biomaterials. *Science* 311: 484–489.
- Ramachandran P, Carlsbecker A, Etschells J. 2017. Class III HD-ZIPs govern vascular cell fate: an HD view on patterning and differentiation. *Journal of Experimental Botany* 68: 55–69.
- Reed JW. 2001. Roles and activities of Aux/IAA proteins in Arabidopsis. *Trends in Plant Science* 6: 420–425.
- Sanchez P, Nehlin L, Greb T. 2012. From thin to thick: major transitions during stem development. *Trends in Plant Science* 17: 113–121.
- Sang XC, Li YF, Luo ZK, Ren D, Fang L, Wang N, Zhao F, Ling Y, Yang Z, Liu Y *et al.* 2012. CHIMERIC FLORAL ORGANS1, encoding a monocot-specific MADS box protein, regulates floral organ identity in rice. *Plant Physiology* 160: 788–807.
- Savidge RA. 1983. The role of plant hormones in higher plant cellular differentiation. II. Experiments with the vascular cambium, and sclereid and tracheid differentiation in the pine, *Pinus contorta*. *The Histochemical Journal* 15: 447–466.
- Sundberg B, Uggla C, Tuominen H. 2000. Cambial growth and auxin gradients. In: Savidge R, Barnett J, Napier R, eds. *Cell and molecular biology of wood formation*. Oxford, UK: BIOS Scientific Publishers, 169–188.
- Sundell D, Street NR, Kumar M, Mellerowicz EJ, Kucukoglu M, Johnsson C, Kumar V, Mannapperuma C, Delhomme N, Nilsson O *et al.* 2017. AspWood: high-spatial-resolution transcriptome profiles reveal uncharacterized modularity of wood formation in *Populus tremula*. *Plant Cell* 29: 1585–1604.
- Tuominen H, Puech L, Fink S, Sundberg B. 1997. A radial concentration gradient of indole-3-acetic acid is related to secondary xylem development in hybrid aspen. *Plant Physiology* 115: 577–585.
- Turner S, Gallois P, Brown D. 2007. Tracheary element differentiation. *Annual Review of Plant Biology* 58: 407–433.
- Uggla C, Mellerowicz EJ, Sundberg B. 1998. Indole-3-acetic acid controls cambial growth in scots pine by positional signaling. *Plant Physiology* 117: 113–121.
- Uggla C, Moritz T, Sandberg G, Sundberg B. 1996. Auxin as a positional signal in pattern formation in plants. *Proceedings of the National Academy of Sciences, USA* 93: 9282–9286.
- Vanneste S, Friml J. 2009. Auxin: a trigger for change in plant development. *Cell* 136: 1005–1016.
- Wang H, Jones B, Li Z, Frasse P, Delalande C, Regad F, Chaabouni S, Latche A, Pech JC, Bouzayen M. 2005. The tomato Aux/IAA transcription factor IAA9 is involved in fruit development and leaf morphogenesis. *Plant Cell* 17: 2676–2692.
- Worley CK, Zenser N, Ramos J, Rouse D, Leyser O, Theologis A, Callis J. 2000. Degradation of Aux/IAA proteins is essential for normal auxin signalling. *The Plant Journal* 21: 553–562.
- Yang H, Han Z, Cao Y, Fan D, Li H, Mo H, Feng Y, Liu L, Wang Z, Yue Y *et al.* 2012. A companion cell-dominant and developmentally regulated H3K4 demethylase controls flowering time in Arabidopsis via the repression of *FLC* expression. *PLoS Genetics* 8: e1002664.
- Ye ZH, Zhong R. 2015. Molecular control of wood formation in trees. *Journal of Experimental Botany* 66: 4119–4131.
- Yu H, Soler M, San Clemente H, Mila I, Paiva JA, Myburg AA, Bouzayen M, Grima-Pettenati J, 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.
- Zhong R, McCarthy RL, Lee C, Ye ZH. 2011. Dissection of the transcriptional program regulating secondary wall biosynthesis during wood formation in poplar. *Plant Physiology* 157: 1452–1468.
- Zhu Y, Song D, Sun J, Wang X, Li L. 2013. *PttHB7*, a class III HD-Zip gene, plays a critical role in regulation of vascular cambium differentiation in *Populus*. *Molecular Plant* 6: 1331–1343.

## Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

**Dataset S1** Differentially expressed genes in the RNAseq-based transcriptomic analysis of *PtoIAA9m*-OE vs WT.

**Fig. S1** Phylogenetic relationship of poplar and Arabidopsis *Aux/IAA* family members.

**Fig. S2** Expression levels and patterns of poplar *Aux/IAA* genes in wood-forming stem tissues.

**Fig. S3** Sequence alignment of *PtoIAA9*, Arabidopsis *IAA8/9* and tomato *IAA9*.

**Fig. S4** Attributes of a typical Aux/IAA protein harbored by the *IAA9* protein.

**Fig. S5** Generation and stem phenotypes of *PtoIAA9m*-OE poplar lines.

**Fig. S6** Sequence, expression and protein interactions of ARF5s in poplar.

**Fig. S7** Constitutive expression of *PtoARF5* in the *PtoIAA9m*-OE lines.

**Fig. S8** Comparative transcriptomic analysis of the *PtoIAA9m*-OE and WT lines via RNAseq.

**Fig. S9** Regulation of the *PtoIAA9/ARF5* module on *PtoHB7/8* promoter activities via effector–reporter tests using GUS as a reporter gene.

**Fig. S10** Constitutive expression of *PtoHB7* in the *PtoIAA9m*-OE lines.

**Fig. S11** Expression of some secondary xylem development-related genes in *PtoIAA9m*-OE, *PtoARF5.1*- and *PtoHB7*-complementing transgenic lines.

**Methods S1** Experimental methods for Figures S4 and S8, including subcellular localization, protein stability assay and mRNA sequencing.

**Table S1** Primer sequences used in this study.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.



## About *New Phytologist*

- *New Phytologist* is an electronic (online-only) journal owned by the New Phytologist Trust, a **not-for-profit organization** dedicated to the promotion of plant science, facilitating projects from symposia to free access for our Tansley reviews and Tansley insights.
- Regular papers, Letters, Research reviews, Rapid reports and both Modelling/Theory and Methods papers are encouraged. We are committed to rapid processing, from online submission through to publication 'as ready' via *Early View* – our average time to decision is <26 days. There are **no page or colour charges** and a PDF version will be provided for each article.
- The journal is available online at Wiley Online Library. Visit **www.newphytologist.com** to search the articles and register for table of contents email alerts.
- If you have any questions, do get in touch with Central Office (np-centraloffice@lancaster.ac.uk) or, if it is more convenient, our USA Office (np-usaoffice@lancaster.ac.uk)
- For submission instructions, subscription and all the latest information visit **www.newphytologist.com**

## CHAPITRE 4 - Projets de recherche en cours et à venir



## Introduction

Grâce à l'expertise et au savoir-faire acquis au cours de mes cursus professionnels, je construis, avec notre équipe ou en collaboration avec d'autres équipes et laboratoires, des projets répondant aux besoins de la Recherche. Les axes des projets que j'ai pu développer sont principalement centrés sur la régulation transcriptionnelle de la formation du bois.

### **A) Le cœur de mes activités de recherche en cours et à venir repose sur la régulation de la différenciation cellulaire au cours de la formation du bois et de sa réponse aux stress abiotiques :**

A1) La poursuite des travaux sur la caractérisation de la régulation auxine dépendante de la formation du bois.

A2) Trade-off entre croissance et adaptation aux stress abiotiques dans le bois chez les arbres : via la signalisation de l'auxine.

A3) La régulation et la dynamique de la différenciation cellulaire au cours de la formation du bois par approche transcriptomique type séquençage de l'ARN unicellulaire (scRNA-seq). Nous allons utiliser cet outil pointu pour caractériser en profondeur les lignées générées issues des parties A1 et A2, afin d'élucider les mécanismes moléculaires et modes de fonction pour les gènes à étudier.

### **B) Ensuite, je participe aux projets ANR de notre équipe :**

B1) « La régulation post transcriptionnelle de la formation du bois : Rôle des protéines Musashi » 2023-2026, dans ce projet je vais profiter de mes expériences réussies de l'édition du génome de *Eucalyptus*, et générer et caractériser des lignées Crispr/Cas9 des gènes Musashi chez *Eucalyptus*.

B2) PEPR Typex (Toward highly Predictable Editing of the plant genome leXicon, Vers une édition spécifique et précise du génome végétal, 2023-2028) ; dans ce projet nous allons profiter de nos résultats sur le gène *Egr1AA9A* (cf. A1 ci-dessus), et générer et caractériser des lignées hairy root ou transformation permanent d'*Eucalyptus* de gain-de-fonction et perte-de-fonction, grâce aux ressources partagées dans ce projet. Les résultats de gain-de-fonction seront comparés avec nos résultats obtenus dans la partie A.

### **C) Enfin, je m'investis dans les projets de la caractérisation des modifications et adaptations des parois :**

C1) En réponse aux stress biotiques (Projet FRAIB AWARE : Analysis of primary cell WALL in REsistance under combined stress) et abiotiques (sécheresse etc).

C2) Au cours de leur développement (Tomato locular gel tissue formation and regulation: the missing part of tomato fruit development)

## A : La différenciation cellulaire au cours de la formation du bois et lors d'une réponse aux stress abiotiques.

Le bois, ou xylème secondaire, est l'une des sources d'énergie naturelles et renouvelables les plus abondantes de notre terre ainsi qu'une bioressource importante pour l'industrie mondiale. Il fournit des matières premières pour la construction, la fabrication de pâte à papier, la production d'énergie ainsi qu'une source de produits chimiques fins allant des plastifiants aux agents aromatisants. Le bois est l'une des ressources les plus respectueuses de l'environnement qui permet à la fois de remplacer notre dépendance aux énergies fossiles et de capter activement le CO<sub>2</sub> atmosphérique pour réduire le réchauffement climatique (Boudet et al., 2003).

Chez les arbres, le bois fonctionne comme un tissu vasculaire et squelettique grâce aux cellules du bois, de plusieurs types cellulaires spécifiques, morphologiquement et fonctionnellement distincts, mais tous caractérisés par la présence de parois cellulaires secondaires (SCW). Celles-ci comprennent les cellules conductrices de sève avec une forme de tube appelées éléments trachéaires, les cellules longues et fines de support squelettiques appelées fibres et les cellules cubiques des rayons servant à charger et décharger le contenu de la sève. Ces types de cellules du bois, très diverses morphologiquement et fonctionnellement, sont toutes issues d'un méristème secondaire appelé cambium vasculaire, qui subit des étapes séquentielles telles que la division cellulaire orientée, l'expansion cellulaire, le dépôt massif de polysaccharides SCW dans des motifs guidés par des microtubules, et la mort cellulaire programmée précédant ou non la lignification des parois cellulaires en fonction du type de cellule du xylème (Ye & Zhong, 2015).

Bien que suivant le même schéma général, chaque type de cellule du xylème possède des caractéristiques cytologiques spécifiques (c'est-à-dire le type cellulaire et la composition des SCW) pour remplir leurs rôles, comme dans le cas des lignines. Ces polymères, déposés spécifiquement dans les parois cellulaires du xylème, sont composés de phénylpropanoïdes qui varient à la fois dans leur fonction aliphatique terminale (alcool, aldéhyde, acide) et dans leur niveau de substitution aromatique (méta hydroxy/méthoxylation). Les résidus canoniques de lignine comprennent le p-hydroxyphényl (H) sans méta-substitution, le gäiäcyle (G) avec une méta-substitution et le syringyle (S) avec deux méta-substitutions dans les polymères de lignine (Barros et al., 2015). Des types de cellules de xylème distincts et leurs différentes couches de parois cellulaires sont enrichis en résidus de lignine spécifiques : les SCW dans les éléments trachéaires sont enrichis en résidus G, les SCW en fibres sont enrichis en résidus S, tandis que la paroi cellulaire primaire séparant ces types de cellules est enrichie en résidu H (Boerjan et al., 2003; Pesquet et al., 2019). La succession de l'étape de différenciation de chaque type de cellule du xylème depuis leur initiation, la coordination dans le temps des caractéristiques cytologiques spécifiques, jusqu'à leur pleine maturation dépend de facteurs de croissance qui ont des effets superposés ou distincts selon les types de cellules du xylème, leur stade de différenciation et/ ou les contraintes environnementales (Brackmann et al., 2018; De Zio et al., 2020; Fischer et al., 2019). Donc la différenciation cellulaire et l'anatomie du bois varient en fonction du développement et des contraintes de l'environnement, au niveau des proportions des fibres *versus* vaisseaux et de la composition physicochimique des parois, pour assurer une meilleure adaptation de leur croissance à leur environnement. On sait maintenant qu'après la sécheresse, le re arrosage induit dans la formation du bois l'apparition de plus de vaisseaux que de fibres. Parmi les facteurs régulateurs de la différenciation cellulaire, l'augmentation des niveaux d'auxine joue un rôle essentiel pour l'initiation de l'élément trachéaire (Fukuda & Komamine, 1980; Pesquet et al., 2005). La suraccumulation de l'auxine dans les feuilles (par traitement NPA) provoque également l'hypertrophie



du xylème primaire (Berleth & Mattsson, 2000; Berleth et al., 2000; H. Wang et al., 2005). Les travaux pionniers ont démontré que l'application de l'auxine peut déclencher la différenciation des cellules mésophylles des feuilles en cellules du bois caractérisées par la présence de parois secondaires lignifiées (Fukuda & Komamine, 1980). Il a été démontré que la modification de la signalisation de l'auxine régule différenciellement les deux types principaux de cellules du xylème. Elle n'affecte pas la formation des vaisseaux mais celle des fibres : elle peut provoquer leur différenciation en cellules de parenchyme, ou la réduction de formation de SCW (Johnsson et al., 2018; Morey & Dahl, 1975; Yu et al., 2022). Pourtant, le mécanisme moléculaire permettant à l'auxine de découpler la formation des différents types de cellules du xylème reste mal connu.

La différenciation cellulaire au cours de la formation du bois, de la cellule méristématique de cambium aux cellules spécialisées et variées des xylèmes, et leurs réponses aux stress abiotiques pour mieux s'adapter ou survivre dans leur environnement de croissance, constituent un sujet fascinant pour moi à découvrir. Pour améliorer notre connaissance sur ce sujet, je propose une stratégie en 3 volets.

**Le premier volet** se focalise sur la poursuite des travaux « La régulation auxinique de la formation du bois », et approfondit le mécanisme moléculaire de la signalisation de l'auxine au cours de la formation du bois, la caractérisation fonctionnelle des gènes clés impliqués, la recherche sur des régulateurs en amont, les partenaires de ces facteurs de transcription et leurs cibles potentielles. Actuellement, la caractérisation fonctionnelle du gène *EgrIAA9A* est en cours. Nous avons obtenu des résultats prometteurs montrant que la surexpression de la protéine EgrIAA9A stabilisée chez *Arabidopsis* stimule la différenciation des vaisseaux au sein des cellules du bois et réprime la maturation et la lignification des fibres. Je voudrais continuer sa validation fonctionnelle dans un système homologue chez *Eucalyptus* par la technique « hairy root ». En plus, j'aimerais approfondir certains points inattendus et intéressants soulevés pendant nos recherches :

\* Comment un membre non-canonique tel qu'EgrIAA20 absent du domaine II (qui est responsable de la dégradation de cette protéine type « short life » ce qui censé être essentiel pour la transduction de signal) régule la signalisation de l'auxine ?

\* Les mécanismes de régulation de transcription de certains membres ont leurs protéines qui se localisent en dehors du noyau cellulaire. Serait-il possible que la délocalisation de ces facteurs se fasse en réponse à des stimuli pour réguler leur activité transcriptionnelle ?

\* En collaboration avec l'équipe miPEP de notre laboratoire (Dr. Jean-Philippe Combier), nous sommes en train de chercher les micro peptides (miPEP) qui potentiellement régulent la formation du bois en ciblant les facteurs de transcription de réponse à l'auxine.

**Le deuxième volet** se porte sur la recherche du trade-off entre la croissance et l'adaptation aux stress environnementaux. Sous les contraintes environnementales telles que la sécheresse ou une forte température, les arbres réduisent leur croissance pour mieux gérer leurs stress. Nous prenons pour hypothèse que les arbres utilisent la régulation des phytohormones telles que l'auxine pour réduire leur formation du bois sous les contraintes environnementales et adaptent leur stratégie de mode de croissance (sous environnement favorable) ou de mode survie (sous environnement défavorable). En particulier, quels sont les composants de la signalisation de l'auxine qui participent à ces processus et comment ? En cas de ré arrosage, il apparaît que plus de vaisseaux se forment pour mieux assurer leur rôle de transport de l'eau des racines vers les parties aériennes. Est-ce que l'auxine participe à réguler ce processus et comment ?

**Le troisième** et dernier volet étudie ce processus de différenciation cellulaire du bois en profitant des derniers avancements techniques de single cell RNA sequencing. L'analyse du transcriptome unicellulaire a révolutionné la recherche génomique chez l'homme et les modèles animaux en permettant de caractériser des programmes d'expression génique spécifiques aux cellules et de découvrir de nouveaux types de cellules. Le séquençage d'ARN unicellulaire (scRNA-seq) a conduit à la création d'atlas cellulaires d'organogenèse chez les animaux et les plantes (principalement pour les racines des plantes herbacées). Cette approche a permis de déterminer les modifications cellulaires qui se produisent au cours du développement d'organes. L'analyse de la trajectoire de développement des types cellulaires individuels a identifié le programme transcriptionnel qui détermine leur différenciation (Farrell et al., 2018). Il a été démontré sa capacité à découvrir des éléments de régulation clés qui déterminent le destin cellulaire au cours des programmes de développement (Conde et al., 2021). Nous voudrions profiter de cette avancée technique pour nous aider à découvrir le mécanisme moléculaire de la différenciation terminale de cambium (méristématique) aux cellules du bois (xylème secondaire) chez *Eucalyptus*, notamment pour identifier des nouveaux gènes marqueurs et des gènes régulateurs de la différenciation des cellules du bois. Nous souhaiterions reconstruire l'atlas cellulaire d'organogenèse du bois : la distribution spatiale et temporelle des cellules individuelles à différents stades, allant de la prolifération cellulaire, à la détermination de l'identité terminale cellulaire.

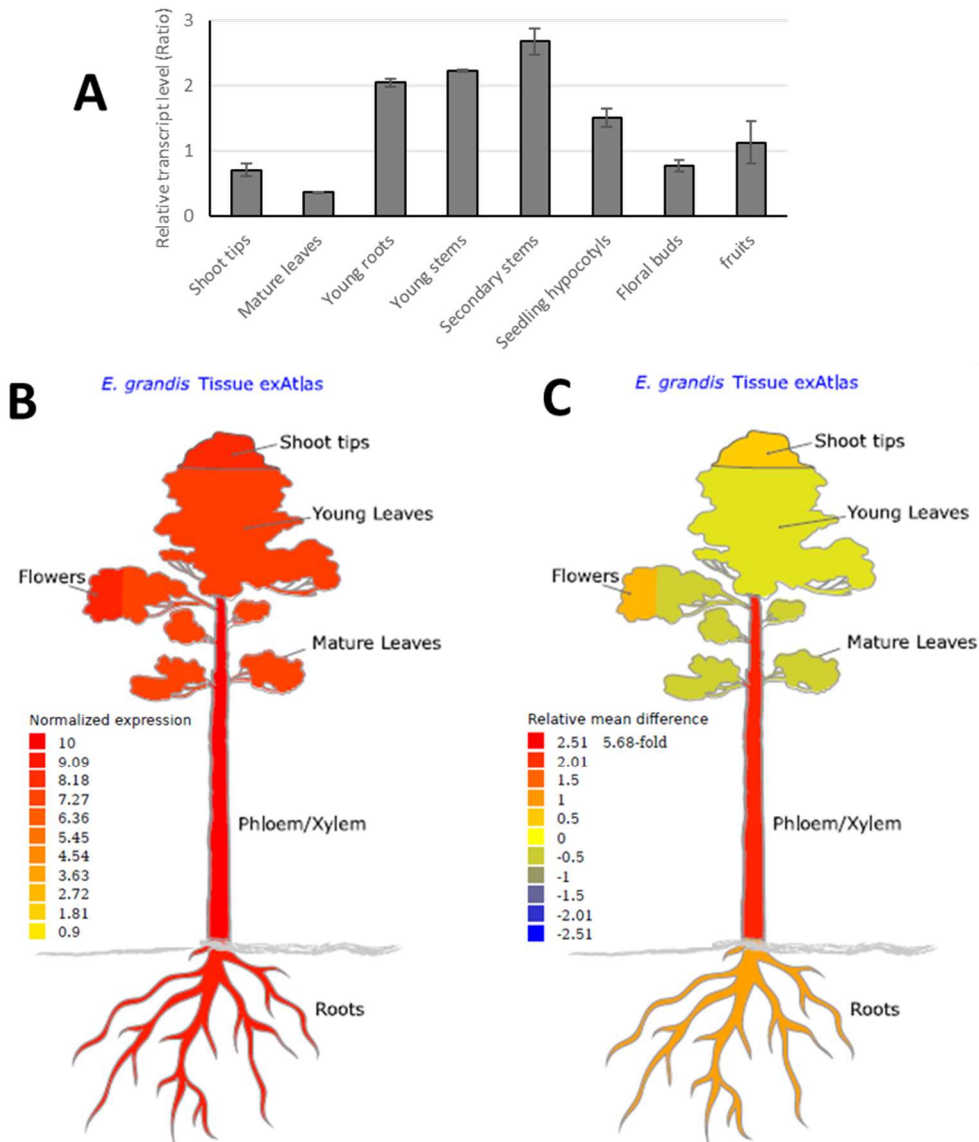
## A1 : la poursuite des travaux sur la caractérisation de la régulation auxine dépendante de la formation du bois.

Les travaux pionniers ont pu démontrer que l'application de l'auxine déclenche la transformation des cellules de mésophylle en cellules de bois, mais leur mécanisme moléculaire reste encore mal connu. Un an après mon intégration au sein de l'équipe GFE (Génomique Fonctionnelle d'*Eucalyptus*), grâce à la publication primaire du séquençage génome complet de *Eucalyptus*, et l'obtention d'un financement d'une thèse de 4 ans (Dr. YU Hong) en réponse à l'appel d'offre ministériel du gouvernement Chinois (CSC China Scholarship Council), j'ai proposé à l'équipe un nouvel axe de recherche visant à identifier des régulateurs de la voie de signalisation de l'auxine impliqués dans la formation du bois d'*Eucalyptus*. Nous avons identifié tous les membres des familles des gènes en relation avec la signalisation de l'auxine, y compris les familles de facteur de transcription de réponse à l'auxine *ARF*, *Aux/IAA* (Yu et al., 2014; Yu et al., 2015), et les familles de transporteurs de l'auxine, les transporteurs d'influx d'auxine *AUX1/LAX* et les transporteurs d'efflux *PIN-FORMED* (*PIN*) (Li et al., 2015). En analysant leur profil d'expression spatio-temporelle, nous avons pu identifier les membres *ARF* et *Aux/IAA* potentiellement impliqués dans la formation de bois. La caractérisation fonctionnelle des gènes candidats est partiellement publiée (*EgrIAA4* et *EgrIAA20*) (Yu et al., 2022; Yu et al., 2015) et se poursuit toujours, notamment sur le gène *EgrIAA9A*.

### *1. La surexpression de EgrIAA9A régule sélectivement la maturation et la lignification des vaisseaux et des fibres secondaires dans le xylème secondaire*

Notre étude d'expression à grande échelle des membres *Aux/IAA* a montré que *EgrIAA9A* a une expression très forte sur tous les organes/tissus testés. Elle est aussi forte que les gènes « house-keeping », mais elle l'est encore plus dans les tiges (Figure 1A). Des analyses complémentaires utilisant la base de données EucGenIE (<https://eucgenie.org/exImage>), une ressource en ligne pour la génomique de *Eucalyptus* et la transcriptomique RNAseq, ont en outre confirmé l'expression

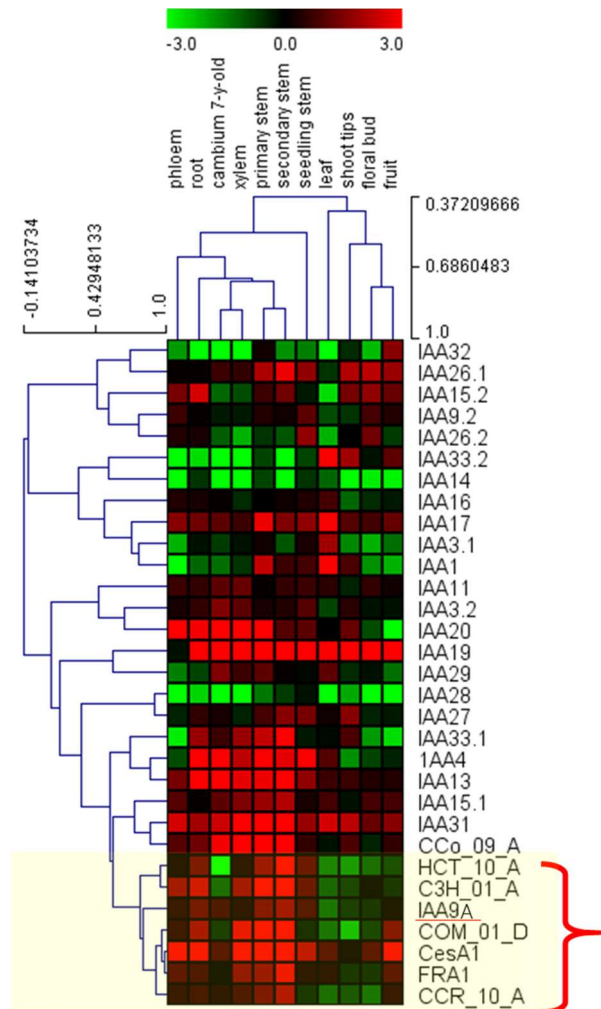
d'*EgrIAA9A* particulièrement forte partout (Figure 1B) et en particulier plus forte dans le tronc d'arbre (Figure 1C).



**Figure 1. La transcription *EgrIAA9A* (*Eucgr.H02407*) montre une accumulation préférentielle distincte dans les tiges et racines de l'eucalyptus. (A) Niveaux d'expression PCR en temps réel d'*EgrIAA9A* dans divers tissus et organes d'eucalyptus. Chaque abondance relative d'ARNm a été normalisée par rapport à un échantillon témoin (plantules d'*Eucalyptus in vitro*). Les barres d'erreur indiquent  $\pm$  SE des valeurs d'expression moyennes  $\pm$  SE de trois expériences indépendantes. (B) Vue pictographique de l'expression absolue d'*EgrIAA9A* sur une gamme diversifiée d'ensembles de données d'expression RNAseq par exlImage (<https://eucgenie.org/exlImage>). Niveaux d'accumulation d'ARNm (Normalized expression) sont indiqués par l'échelle des couleurs jaune/rouge. (C) Vue pictographique de l'expression relative d'*EgrIAA9A* sur des données d'expression RNAseq par exlImage.**

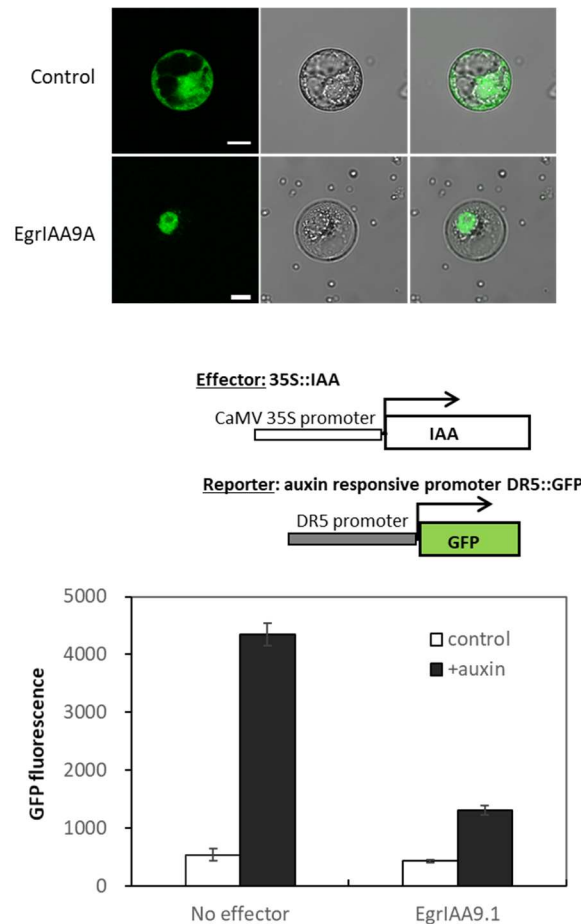
De plus, par qRT-PCR dans onze organes et tissus différents, nous avons évalué les niveaux de transcription des tous les membres *Aux/IAA* avec les gènes marqueurs de la formation de la paroi secondaire (Carocha et al., 2015), qui est caractéristique des cellules du bois. Les 24 gènes *EgrIAA* ont pu être détectés dans tous les tissus testés, et la plupart d'entre eux ont montré une expression préférentielle dans certains tissus distincts (Figure 2). Les accumulations relatives de transcription des

gènes *EgrIAAs* sont présentées sous forme de carte thermique. L'analyse clustering hiérarchique d'expression des gènes a permis de montrer que *EgrIAA9A* (souligné en rouge sur la figure ci-dessous) est regroupé dans le cluster des gènes de marqueur des cellules bois (boîte jaune sur la figure indiqué par l'accolade rouge). Cette analyse montre une co-expression du gène *EgrIAA9A* avec les gènes marqueurs de la formation de la paroi secondaire, et suggère également qu'il est probablement impliqué dans la formation du bois chez l'*Eucalyptus*.



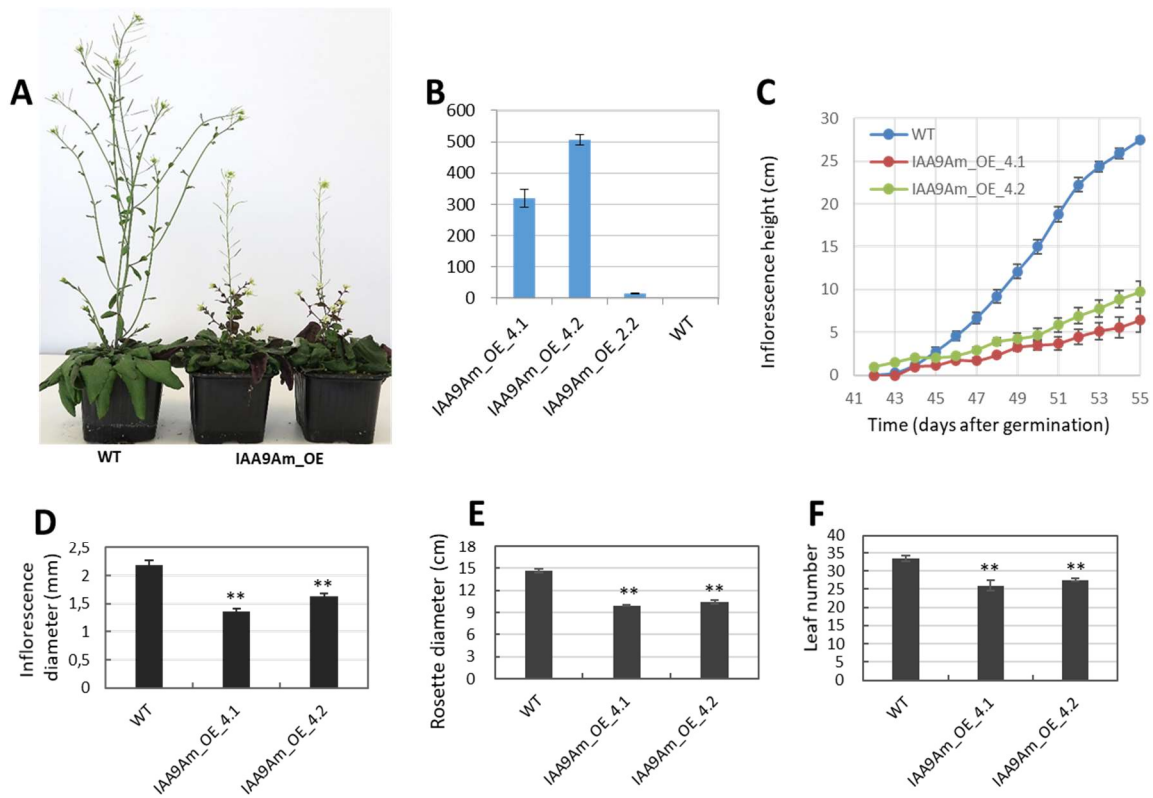
**Figure 2. *EgrIAA9A* montre une co-expression avec les gènes marqueurs des cellules bois.** Profils d'expression des gènes *EgrIAA* dans divers organes et tissus. La carte thermique regroupée hiérarchiquement des gènes a été construite en utilisant les ratios log<sub>2</sub> des valeurs d'expression relatives normalisées par rapport à un échantillon témoin (plantules *in vitro*) pour 24 gènes *EgrIAA* et sept gènes marqueurs de paroi secondaire (indiqués à droite) dans 11 tissus et organes (indiqués en haut de la carte thermique). Les gènes marqueurs de paroi secondaire forment un cluster distinct avec une profil d'expression préférentielle dans les tissus vasculaires (phloème, cambium et/ou xylème) et dans les tiges et racines, indiqué par la surface jaune et l'accolade rouge.

Nous avons montré que la protéine *EgrIAA9A* se dirige et s'exprime exclusivement dans le noyau à l'aide d'une protéine de fusion *EgrIAA9A::GFP* dans les protoplastes de Tabac (Figure 3A). En utilisant des essais de transactivation, nous avons dévoilé son rôle de répresseur très fort de la réponse à l'auxine (Figure 3B).



**Figure 3. Localisation subcellulaire et activité répressive de la réponse auxine de la protéine EgrIAA9A sur un promoteur DR5 synthétique. (A)** Localisation subcellulaire de la protéine de fusion EgrIAA9A-GFP dans les protoplastes du tabac BY-2. Les images fusionnées de fluorescence verte (panneau de gauche) et l'image en champ clair correspondante (panneaux du milieu) sont affichées dans les panneaux de droite. Barre d'échelle = 10 μm. **(B)** Répression transcriptionnelle de la réponse auxine par la protéine EgrIAA9A sur un promoteur DR5 synthétique. Les constructions effectrices et rapporteuses ont été coexprimées dans des protoplastes de tabac en présence ou en absence d'une auxine synthétique (50 μM 2,4-D). Une construction effectrice factice (vecteur vide) a été utilisée comme contrôle négatif. Trois expériences indépendantes ont été réalisées et des résultats similaires ont été obtenus. Dans chaque expérience, les transformations de protoplastes ont été réalisées en triples biologiques indépendants. La figure indique les données d'une expérience. Les barres d'erreur représentent SE de la fluorescence moyenne. Les différences statistiques significatives (test t de Student,  $n > 400$ ,  $P < 0,01$ ) par rapport au contrôle sont marquées de \*\*. Les schémas des constructions rapporteur et effecteur ont été illustrés en haut de la figure.

De plus, nous avons testé fonctionnellement le rôle de *EgrIAA9A* par surexpression constitutive (sous promoteur 35S) chez *Arabidopsis* pour rechercher des changements phénotypiques dans la formation secondaire du xylème. Les plantes transgéniques d'*Arabidopsis* surexprimant *EgrIAA9A* étaient légèrement plus petites avec un renforcement de la dominance apicale (Figure 4), et présentaient un développement altéré des cellules du xylème secondaire : une très forte inhibition de développement des fibres secondaires, et une stimulation du nombre de vaisseaux (Figure 5).



**Figure 4. Les lignées surexprimant *EgrIAA9A* présentaient un développement réduit de la rosette et de la hampe florale avec un renforcement de la dominance apicale. (A)** Comparaison de l'architecture végétale des parties aériennes, les lignées surexprimant *EgrIAA9A* présentaient des rosettes plus petites et des hampes florales plus petites mais une dominance apicale plus forte par rapport à la plante sauvage. **(B)** Accumulation d'ARNm *EgrIAA9A* dans trois lignées transgéniques indépendantes phénotypiquement représentatives. **(C)** Courbe de croissance de la tige de l'inflorescence primaire d'*Arabidopsis*. L'allongement de la hampe a été réduit dans les lignées surexprimant *EgrIAA9A* (IAA9A\_OE\_4.1 et IAA9A\_OE\_4.2) avec un diamètre d'inflorescence **(D)** et un diamètre des rosettes **(E)** considérablement réduit, plus un nombre réduit de feuilles significatif **(F)**. Les barres d'erreur représentent l'erreur standard,  $n > 12$  ; le diamètre de la tige a été mesuré à la base (<1 cm jusqu'au niveau de la rosette) lorsque la première silique était complètement développée. Les astérisques indiquent les valeurs jugées significativement différentes (test t de Student,  $n > 12$ ) par rapport au contrôle de type sauvage. \*\* indique  $p < 0,01$ .

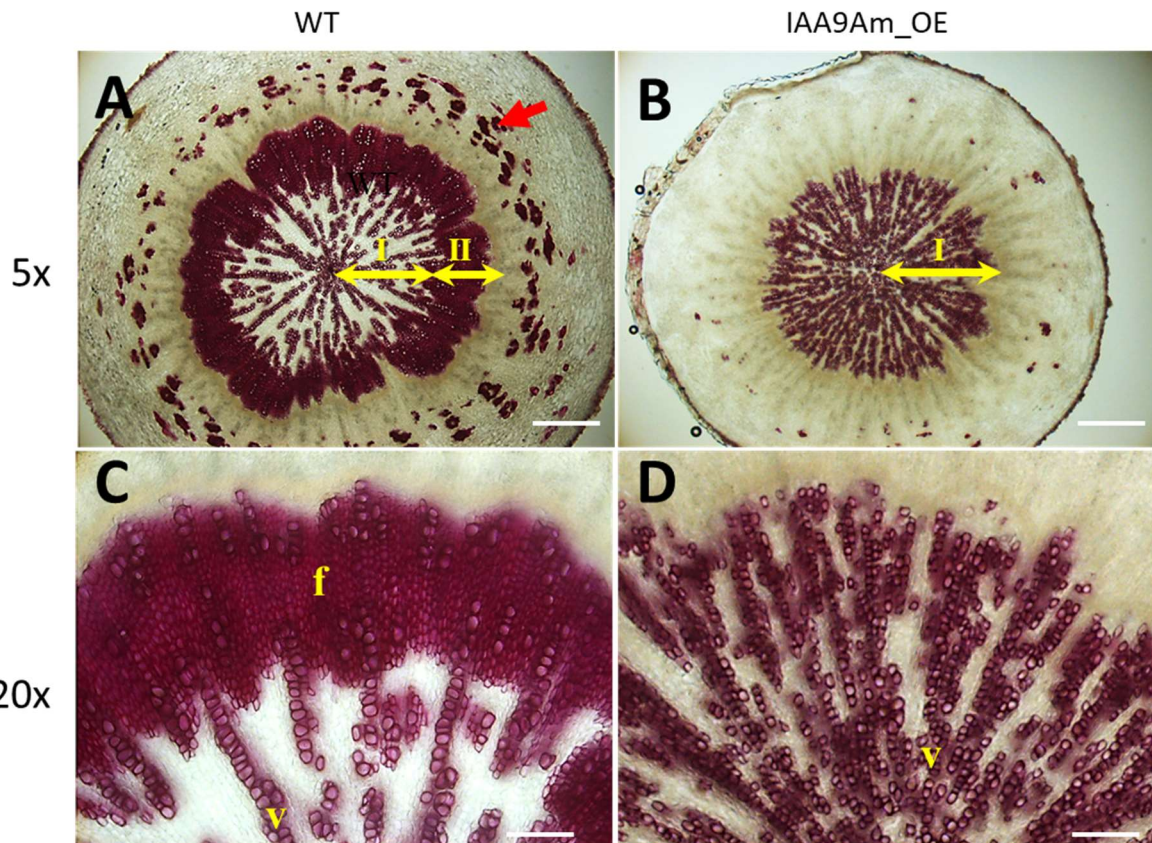
Nous avons analysé la composition chimique des parois cellulaires dans des hypocotyles broyés (broyeur à billes) en effectuant une pyrolyse couplée de chromatographie en phase gazeuse/spectrométrie de masse (Py-GC/MS). Les lignées surexprimant *EgrIAA9A* ont une réduction significative de la teneur totale en lignine ainsi une réduction significative en résidus de lignine S, de lignine G et d'autres résidus phénoliques (Table 1). L'ensemble des analyses biochimiques et histochimiques montre que l'inhibition du développement des fibres a spécifiquement affecté la lignification de leur paroi cellulaire.



**Tableau 1.** Composition chimique des hypocotyles d'Arabidopsis transgéniques de type sauvage et EgrIAA9A\_OE par analyse Py-GC/MS (n > 12).

Genotype	Carbohydate	Guaiacyl	Syringyl	p-hydroxyphenol	Phenolic	Lignin	S/G	C/L
WT	77.5±2.0	8.5±1.2	2.9±0.7	1.7±0.2	1.1±0.1	14.2±2.0	0.34±0.04	5.55±0.82
IAA9Am OE	78.0±1.2	<b>6.7±1.0*</b>	<b>1.9±0.4*</b>	1.5±0.4	<b>0.8±0.2**</b>	<b>11.0±1.9*</b>	<b>0.29±0.03*</b>	<b>7.30±1.51*</b>

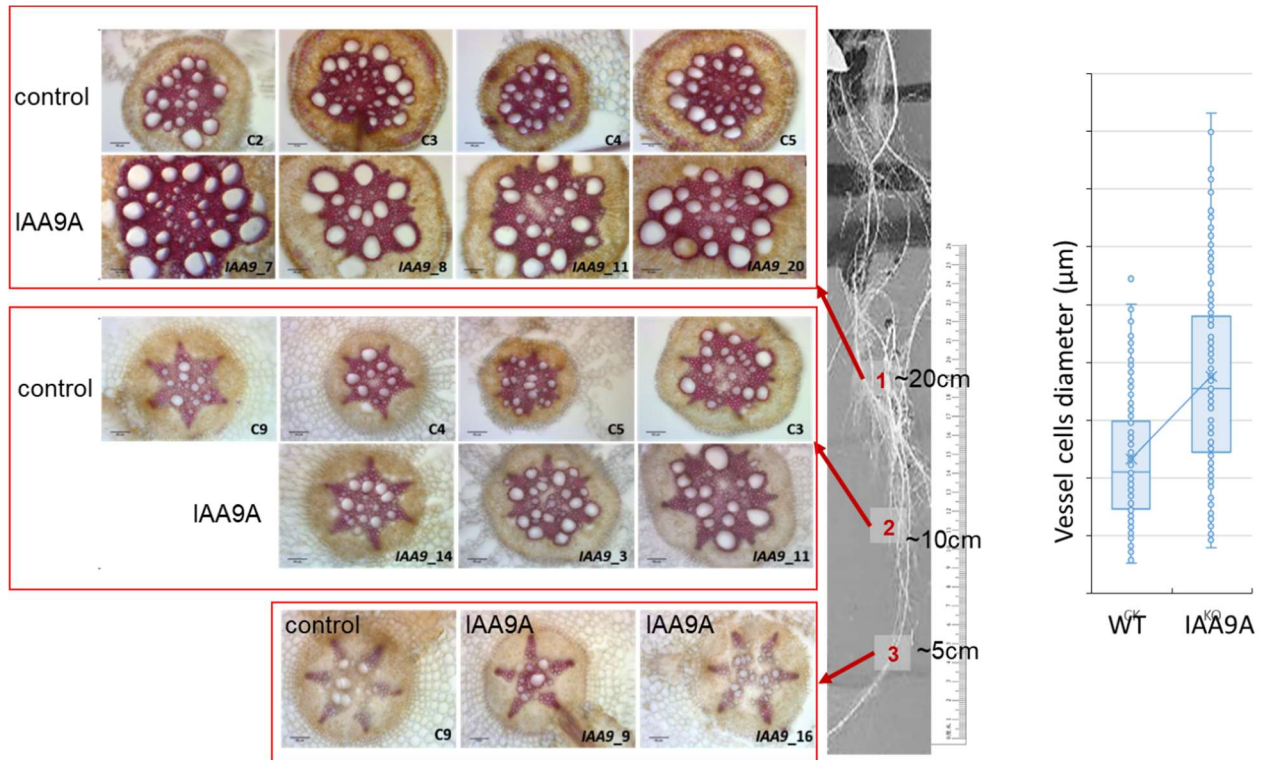
Les astérisques indiquent les valeurs jugées significativement (test t de Student) différentes de celles du type sauvage. \*p<0,05, \*\*p<0,01. Résultats d'échantillons groupés de trois lots indépendants de lignées EgrIAA9A\_OE 4.1 et 4.2 (5 plantes pour chaque lignée et chaque lot).



**Figure 5.** La surexpression de *EgrIAA9A* inhibe spécifiquement la lignification des fibres dans les fibres secondaires mais pas les éléments trachéaux, et stimule la formation des vaisseaux dans l'hypocotyle d'Arabidopsis. Le panneau de gauche (A & C) montre une coupe transversale de l'hypocotyle du contrôle de type sauvage (image objective 5x pour la première ligne, et 20x pour la deuxième ligne, respectivement, en utilisant la coloration au phloroglucinol-HCl, qui colore les polymères de lignine dans les parois secondaires des cellules de xylème en rouge-violet. La région de croissance de phase I et la région de croissance de phase II sont indiquées par des lignes à double flèche dans la coupe transversale de type sauvage. v indiquent les cellules de vaisseaux, f indiquent cellules de fibres secondaires. Pour les images 5x barre d'échelle = 200 µm et pour les images 20x, barre d'échelle = 50 µm.

Nous avons également effectué des tests de doubles hybrides de levure pour identifier les partenaires protéiques EgrIAA9A lors de la formation du bois dans l'*Eucalyptus* et identifié EgrIAA9A lui-même (suggèrent que EgrIAA9A pourrait former un homodimère dans le complexe de facteurs de transcription), et histone-linker qui a été démontré interagir avec EgrMYB1 pour réprimer la

différentiation du xylème secondaire (Soler et al., 2017). Ensuite, nous avons tenté de valider leur fonction dans le système homologue chez *Eucalyptus* par la technique « hairy root ». Les lignées de perte-de-fonction générées par la technique de CRISPR-Cas9 ont montré une stimulation du développement accéléré du xylème, avec une stèle plus grande et un diamètre accru des vaisseaux secondaires (Figure 6), suggérant que les lignées transgéniques IAA9A CRISPR présentaient un développement accéléré du xylème. La génération des lignées de gain-de-fonction sont actuellement en cours. Leur analyse phénotypique est prévue, en plus des caractérisations moléculaires pour connaître leur mode de fonctionnement et leurs régulateurs en amont et en aval.



**Figure 6. Les lignées de CRISPR-IAA9A ont montré une stimulation du développement accéléré du xylème, avec une stèle plus grande et un diamètre accru des vaisseaux secondaires.** Le panel gauche montre les sections transversales de racines de lignées transgéniques CRISPR-IAA9A et de contrôle dans trois positions : vieille (20 cm à partir de pointe racinaire), moyenne (10 cm) et jeune (5 cm). Les lignées transgéniques ont des diamètres de racines et des stèles plus grands, ainsi que des diamètres de vaisseaux accrus dans les positions 20 cm et 10 cm mais moins évident en position 5 cm. Le panel droit montre les mesures de diamètre des vaisseaux. La taille des vaisseaux des lignées CRISPR-IAA9A était clairement plus grande que celle du contrôle.

## 2. mi-PEP, la nouvelle clé pour découvrir un nouveau mécanisme de la régulation auxine-dépendant post-transcriptionnelle du développement des xylèmes secondaires ?

Les microARNs (miRs) sont des petits ARNs non codants, d'environ 21 nucléotides après maturation, qui contrôlent l'expression de gènes cibles au niveau post-transcriptionnel, en dégradant l'ARNm cible ou en inhibant sa traduction. Les gènes cibles sont souvent des gènes clés de processus développementaux. Ils codent par exemple des facteurs de transcription ou des protéines du protéasome. Plusieurs travaux ont montré que certains microARN (miRNA) régulent l'expression des

facteurs de transcription de l'auxine (Fahlgren et al., 2006; Gutierrez et al., 2009; Gutierrez et al., 2012; Liu et al., 2007; Mallory et al., 2005; J. W. Wang et al., 2005; Wu et al., 2006). Par exemple, certains miRs sont montrés pour stimuler le développement d'enracinement, qui est actuellement l'obstacle principal de la propagation massive des génotypes élités d'*Eucalyptus*, qui est l'espèce d'arbre de type bois dur la plus plantée dans le monde entier. Plusieurs études ont montré que les miRs peuvent réguler finement au niveau post-transcriptionnel l'abondance des facteurs de transcription de réponse à l'auxine et contrôler plusieurs processus développementaux y compris la formation des racines et le développement de tissus vasculaires et de bois (Baima et al., 2014; Gutierrez et al., 2009; Gutierrez et al., 2012; Hendelman et al., 2012; Itoh et al., 2008; Liu et al., 2007; Lu et al., 2008; Marin et al., 2010; Muller et al., 2016; Ohashi-Ito & Fukuda, 2010; Popko et al., 2010; Robischon et al., 2011; Tang et al., 2016; Yu et al., 2014). Par exemple, certains miRs peuvent contrôler le développement des racines latérales en régulant la division cellulaire et l'élongation cellulaire. Le développement de racines adventives, est indispensable à la propagation massive des génotypes élités. Il s'agit d'un trait génétique complexe doté d'une plasticité phénotypique élevée due à de multiples facteurs de régulation endogènes et environnementaux. Il a été démontré qu'un équilibre subtil entre les transcrits activateurs et répresseurs des facteurs de transcription de réponse à l'auxine contrôle l'initiation des racines adventives (Gutierrez et al., 2012). De plus, l'activité des miRs semble être nécessaire pour affiner ce processus. Plusieurs couples de miRs et facteurs de transcription de la réponse à l'auxine ont été démontrés être des régulateurs positifs de l'enracinement adventif. De plus, ces facteurs de transcription affichent des domaines d'expression qui se chevauchent, interagissent génétiquement et régulent mutuellement leur expression aux niveaux transcriptionnel et post-transcriptionnel en modulant la disponibilité des miRs. Ce réseau de régulation complexe comprend une régulation inattendue de l'homéostasie des miRs par des facteurs de transcription cibles directs et non directs (Gutierrez et al., 2012). Il a été démontré aussi chez les porte-greffes de peuplier et de pommier que certains miRs ciblent les facteurs de transcription de la réponse à l'auxine impliquées dans la formation des racines latérales en contrôlant la prolifération cellulaire (Gomes & Scortecci, 2021; Li et al., 2019).

La régulation de l'expression des miRs est très peu connue, mais des études récentes montrent que les micro-peptides (miPEPs, peptides codés par des miRs) stimulent la transcription de leurs miRs (Couzigou et al., 2015; Laressergues et al., 2015; Laressergues et al., 2022). Les micro-peptides définissent un peptide qui est codé par un cadre ouvert de lecture présent sur le transcrit primaire d'un miR (pri-MiR), c'est-à-dire le premier ORF du pri-miR dans la direction 5'-3' peut être traduit en miPEP. Ces miPEPs sont majoritairement formés d'une dizaine d'acides aminés et jouent un rôle de régulateur positif sur la transcription du gène miR dont ils sont issus. Il a été démontré que les miPEP conservés entre deux espèces (identité à 90 %) peuvent être actifs entre les différentes espèces végétales pour lesquelles ils sont conservés (Laressergues et al., 2022). L'équipe pionnière de cette découverte (Jean-Philippe Combier) partage le même laboratoire (LRSV). En profitant cette situation favorable, nous avons débuté une collaboration pour tester l'hypothèse : miPEP régule la formation du bois chez les plantes pérennes via la stimulation de la transcription de leur miRs régulant des facteurs de transcription de réponse à l'auxine. Nous avons réalisé des traitements exogènes d'*Eucalyptus grandis* avec des miPEP dérivés des gènes miRs régulant la signalisation de l'auxine. Ce test permet de détecter si le traitement avec miPEP peut affecter le développement du xylème chez les *Eucalyptus*. Pour des raisons de PI (Protection Intellectuelle), les noms des gènes miRs dont sont issus les miPEP et la cible des miRs ne seront pas divulgués.

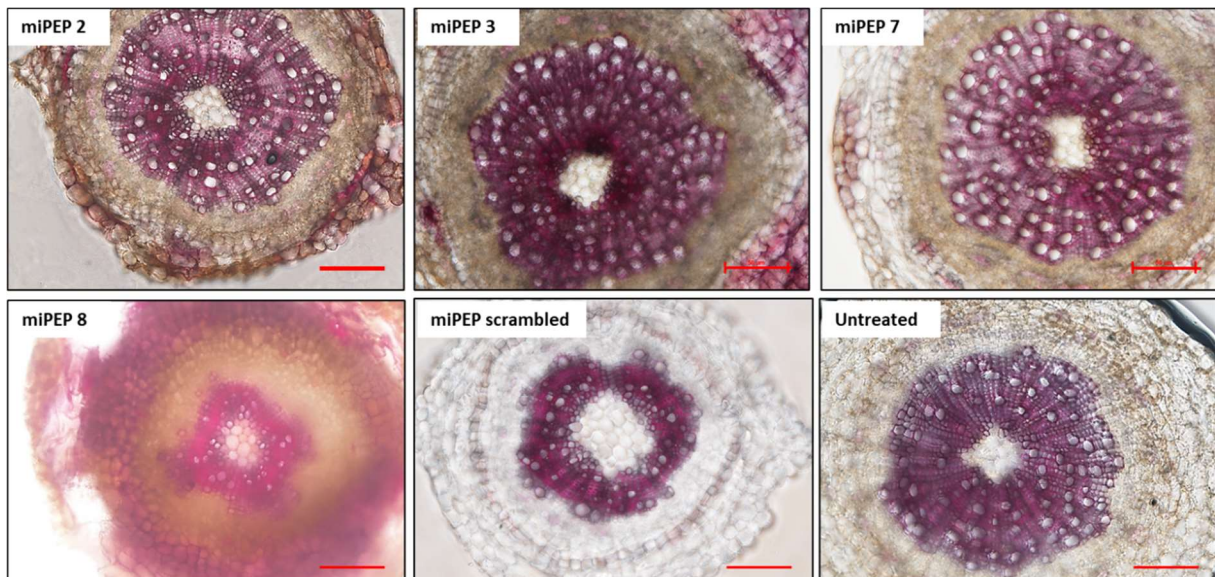
Des *Eucalyptus grandis* âgés de trois à cinq mois cultivés *in vitro* ont été traités avec 9 miPEP différents (dont 1 miPEP aléatoire comme contrôle négative). L'objectif est de voir si :



1) le développement et la croissance globale des plantules traitées au miPEP sont modifiés.

2) les cellules du xylème secondaire des plantes présentent des différences en termes de structure ou de composition par rapport à l'hypocotyle des *Eucalyptus* non traités. Pour le premier essai, des traitements miPEP ont été réalisés sur 4 séries différentes (76, 104, 125 et 153 jours après germination) pour maximiser la chance à détecter des effets du miPEP développement/âge-dépendant sur la formation du xylème.

Pour l'objectif 1, nous n'avons pas pu tirer de conclusion sur l'impact sur le développement ou la croissance à cause d'une grande variabilité naturelle entre les plantes testées (ce sont des graines d'*Eucalyptus grandis* à l'état sauvage, pas de lignée homozygote). Pour l'objectif 2, à l'aide des observations microscopiques, nous avons examiné des parois cellulaires secondaires lignifiées dans les tissus du xylème secondaire par coloration au phloroglucinol-HCL (test de Wiesner) sur des coupes d'hypocotyle des tiges d'*Eucalyptus*. Un phénotype d'ecto-lignification apparaît chez les plantes traitées avec quatre peptides différents (miPEP 2 ; miPEP 3, miPEP 7 et miPEP 8) (Figure 7).



**Figure 7. Un phénotype d'ecto-lignification apparaît suite à un traitement avec certains miPEP.** Coupe transversale de l'hypocotyle d'*Eucalyptus grandis*. Phloroglucinol-HCL marque spécifiquement les lignes en fuchsia/rose. La présence de lignines est normalement située dans les xylèmes secondaires à l'intérieur du cercle de cambium vasculaire, en revanche, l'ecto lignification se trouve dans l'épiderme et en zone de cortex, à l'extérieur du cercle de cambium vasculaire. La section est prélevée sur une zone de la tige située à 0,5 cm au-dessus du collet. L'observation des coupes est au grossissement 20x. Les coupes ont été prélevées sur des *Eucalyptus* traités ou non au 9 miPEP. Les répétitions biologiques varient selon les traitements. La barre d'échelle représente 100  $\mu$ m.

L'ecto-lignification est caractérisée par le dépôt de lignine dans les groupes cellulaires périphériques de la tige (cellules corticales et cellules épidermiques). Ce phénotype apparaît naturellement aussi chez les individus non traités suite à des stress mécaniques ou biotiques mais dans une proportion plus faible. Alors que ce phénotype d'ecto-lignification est présent dans les *Eucalyptus* traités avec miPEP 2,3,7 et 8, avec un pourcentage respectivement de 50 %, 50 %, 45 % et 72 % sur les 4 séries de tests, les plantes non traitées présentent un taux d'ecto-lignification de 20 % (Table 2). Une forte variabilité phénotypique concernant l'intensité de la coloration phloroglucinol-HCL est également observée au sein de ces ecto-lignification. Sur certaines coupes, des portions du xylème présentent

une très faible intensité de coloration rose fuchsia, apparaissant même en blanc (miPEP 2-3 ; miPEP 3-3 ; miPEP 5-1 ; miPEP brouillé 3). Cette ecto-lignification semble dépendre de l'âge (en effet bien que les plantes aient le même âge leur stade de développement varie entre elles). Par exemple, les plantes traitées avec miPEP 2 ne présentent une ecto-lignification que chez les plantes âgées de 153 jours, pour miPEP 7, toutes les plantes âgées de 104 jours présentent une ecto-lignification et le phénotype n'est pas trouvé chez les plantes plus âgées. Concernant miPEP 8, l'ecto-lignification est constatée chez les plantes à différents âges, miPEP8 est donc un candidat prometteur pour la régulation de la lignine ou la biosynthèse de la lignine.

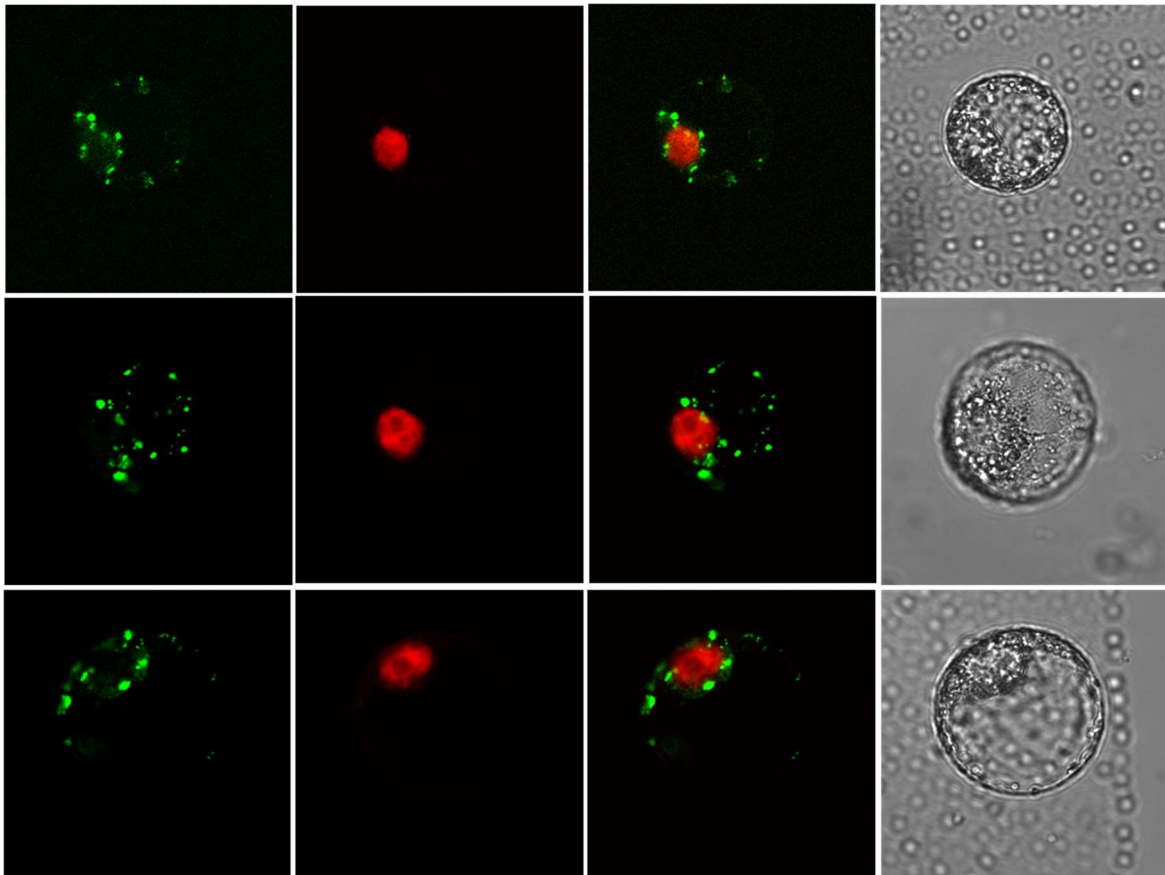
Table 2. Présence of ecto-lignification suite au traitement de différents miPEP

Eucalyptus	Untreated	miPEP scrambled	miPEP 2	miPEP 3	miPEP 7	miPEP 8
76-d-old	0/2	0/1	0/1	1/2	1/2	3/3
104-d-old	0/2	0/1	0/1	1/2	3/3	2/3
125-d-old	1/3	0/1	0/1	2/3	1/3	1/3
153-d-old	1/3	1/3	3/3	1/3	0/3	2/2
Total	2/10	1/6	3/6	5/10	5/11	8/11
%	20%	16%	50%	50%	45%	72%

Afin de limiter les variations naturelles et les variations liées aux conditions de culture *in vitro*, pour les prochaines expériences, nous allons tester ces miPEPs sur des jeunes plantes d'*Eucalyptus grandis*, cultivées dans du terreau (par rapport à la culture *in vitro*). Les traitements des 9 miPEPs seront appliqués sur la même plante ; chaque miPEP sur une branche, donc 9 branches différentes pour 9 différents miPEP mais sur la même plante. Les répétitions se réalisent entre les différentes plantes. En fait, la jeune plante d'*Eucalyptus* présente naturellement plein de ramifications assez homogènes, nous allons profiter de cette particularité pour surmonter, au moins partiellement, le problème de variation génétique des graines d'*Eucalyptus*.

### 3. Vers de nouveaux modes de régulation de la signalisation de l'auxine par les membres non-classiques des familles Aux/IAA et ARF, possible par trans-localisation subcellulaire de ces facteurs de transcription ?

Pendant les recherches de localisation subcellulaire de nos candidats de facteurs de transcription associés au bois, nous avons obtenus des résultats non attendus mais qui méritent d'être approfondis. Par exemple, les membres non-canoniques tel que EgrIAA20 s'expriment principalement dans le noyau mais on retrouve également une expression résiduelle dans le cytoplasme. Le facteur de transcription EgrARF19 est prédit activateur par analyse de séquence, nos essais de transactivation ont également suggéré son rôle d'activateur de transcription (Yu et al., 2014). Cependant, nous étions étonnés de voir que sa localisation subcellulaire n'était pas dans le noyau (Figure 8). Les signaux se trouvent dans des spots cytoplasmiques bien particuliers. Nous avons testé la première hypothèse : la présence d'auxine pourrait déclencher la trans localisation de cytoplasme au noyau de la protéine EgrARF19 pour ensuite pouvoir activer la transcription de réponse de l'auxine. Nous avons suivi la localisation subcellulaire d'EgrARF19 après traitement d'auxine 8H, 18H et 25H, mais nous n'avons pas trouvé de changement de localisation. Cependant, cette localisation subcellulaire bien particulière d'EgrARF19 et des membres non-canoniques m'intrigue fortement et je souhaiterais continuer pour peut-être mettre en évidence de nouveaux modes de régulation de la signalisation de l'auxine ?



**Figure 8. Localisation subcellulaire de la protéine de fusion EgrARF19A-GFP dans les protoplastes du tabac BY-2.** Les images sont prises 25 heures après un traitement de l'auxine (50  $\mu$ M 2,4-D). Les images fusionnées de fluorescence verte (EgrARF19A-GFP) (première colonne de gauche) et fluorescence rouge du marquage de noyau par DAPI (deuxième colonne de gauche) sont présentées dans la troisième colonne en partant de la gauche. L'image en champ clair correspondante est présentée dans la colonne de droite. Barre d'échelle = 10  $\mu$ m.

## A2. Trade-off entre croissance et adaptation aux stress dans le bois chez les arbres : via la signalisation de l'auxine ?

Quelle est la contribution de la régulation de croissance via des facteurs de transcription de la réponse à l'auxine, pour une meilleure tolérance de l'arbre vis-à-vis des stress abiotiques tels que le stress hydrique, la température (froid/gel et forte chaleur) et stress biotiques.

Les contraintes environnementales (stress hydrique forte chaleur etc.) de plus en plus répandues liées au réchauffement climatique de notre planète, affectent gravement la croissance des plantes et entraînent une réduction de la productivité des cultures. L'impact des contraintes environnementales est encore plus accentué pour les espèces pérennes comme les arbres. Pour mieux survivre aux stress, les plantes réduisent leur croissance et développent une série de régulateurs de transcription pour faire face aux changements environnementaux, la famille DESHYDRATION-RESPONSIVE ELEMENT BINDING PROTEIN (DREB), appartenant à la famille des facteurs de transcription AP2/ERF, est une sous-famille clé de facteurs de transcription régulant la réponse aux stress abiotiques (Yamaguchi-Shinozaki & Shinozaki, 1994). Par exemple, les gènes *DREB1/CBF* sont principalement responsables de la réponse au froid, tandis que les gènes *DREB2* sont principalement impliqués dans



le stress dû à la sécheresse, au sel et à la chaleur chez *Arabidopsis* (Nakano et al., 2006). Parmi eux, le gène *DREB2A* activé par la sécheresse régule l'expression de divers gènes sensibles au stress en se liant aux éléments DRE (ACCGAC), tels que RD29A et RD29B (Sakuma et al., 2006). La surexpression du *CBF* (*CBF-OE*) imite l'acclimatation au froid en plus des modifications du développement et de la croissance chez de nombreuses espèces de plantes (Akhtar et al., 2012; Cao et al., 2020; Lata & Prasad, 2011; Mizoi et al., 2012; Navarro et al., 2011), dont *Eucalyptus* (Navarro et al., 2011). Notamment, les gènes *CBF* sont surreprésentés dans le génome d'*Eucalyptus grandis* par rapport à la plupart des autres génomes végétaux, à l'exception de *Medicago* qui contient également 17 membres *CBF* (Cao et al., 2015). Cette expansion de sous-famille est encore plus importante chez *Eucalyptus gunnii*, qui est l'un des *Eucalyptus* les plus tolérants au gel (Nguyen et al., 2017). Certains de ces *CBF* sont fortement induits dans les tiges en réponse au traitement par le froid (Nguyen et al., 2017). La surexpression de *CBF* dans *Eucalyptus* transgénique entraîne une réduction de la croissance primaire et secondaire et a déclenché des changements dans l'architecture du xylème avec des vaisseaux et des fibres plus petits et plus fréquents présentant des lumières réduites, et confère une résistance au froid à la fois dans les feuilles et les tiges (Cao et al., 2020).

Nous sommes encore loin de comprendre l'ensemble des mécanismes qui permettent aux plantes de faire face à leur environnement. Mais récemment, la relation régulatrice entre la famille *DREB* et les hormones végétales a également été étudiée. *DREB1A* est membre de la sous-famille *DREB A-1*, également connue sous le nom de *CBF3*, qui régule négativement l'accumulation d'auxine dans les jeunes tissus d'*Arabidopsis* (Li et al., 2017). La surexpression de *ZmDREB4.1* dans le tabac inhibe l'accumulation d'auxine et de cytokinine et régule négativement la croissance et le développement des plantes (Li et al., 2018). Shani et al 2017 ont mis en évidence que de nombreux membres de la famille *DREB/CBF* se liaient aux promoteurs des gènes *Aux/IAA* par un crible simple hybride entre les promoteurs de 15 gènes *Aux/IAA* et environ 2000 facteurs de transcription. En plus ils ont confirmé par *ChIP* que *CBF1* et *DREB2A* se lient directement au promoteur de *IAA5* et *IAA19*, favorisant directement leur transcription en réponse au stress abiotique. Les mutations perte-de-fonction récessives de ces gènes *Aux/IAA* (*IAA5*, *IAA6* et *IAA19*) entraînent une diminution de la tolérance aux conditions de stress, démontrant le rôle de l'auxine dans le stress abiotique (Shani et al., 2017). Il a été alors proposé que les gènes *Aux/IAA* fonctionnent comme des centres intégrant les informations génétiques et environnementales pour obtenir des réponses physiologiques appropriées (Shani et al., 2017). De plus, les travaux récemment réalisés par notre équipe ont montré que le facteur de transcription *EguCBFQ* (impliqué dans la résistance au froid chez *Eucalyptus*) se lie au promoteur des gènes *EgrIAA11* et *EgrIAA29* (HADJ BACHIR, 2022), qui sont des membres *Aux/IAA* identifiés préférentiellement exprimés dans le bois chez *Eucalyptus* (Yu et al., 2015).

En plus, la formation du bois est un processus modulable et sensible aux changements climatiques/environnementaux. La grande plasticité de la formation du bois, qui s'exprime aux niveaux anatomiques, structuraux et physicochimiques, est une réponse remarquable aux stress environnementaux. Elle contribue certainement à leur adaptation, au sens fonctionnel du terme, à leur survie face aux conditions contraignantes (Zinkgraf et al., 2017). La formation du bois est un témoignage et une image du temps passé : la dendrochronologie (du grec *dendron* = arbre, *chronos* = temps, et *logos* = étude) consiste à analyser des cernes annuels de croissance du bois afin d'obtenir des informations sur des événements passés, par exemple des chutes de blocs rocheux, des incendies ou des avalanches, ainsi que sur des conditions climatiques passées. Chez les arbres en condition de stress abiotiques, on observe une réduction d'activité du cambium (réduction croissance) et de profondes modifications anatomiques et physicochimiques des cellules du bois, qui influencent probablement les propriétés hydrauliques de conduction de sève (Coleman et al., 2021; Polle et al., 2018). En réponse à la sécheresse, la réduction de la taille des vaisseaux, l'augmentation de leur

densité et la lignification de la paroi secondaire semblent être des modifications anatomiques fréquemment observées, et peuvent être interprétées comme une réponse aux risques d'embolies des vaisseaux (Polle et al., 2018). Cependant, les processus moléculaires associés aux caractéristiques du bois en condition de stress, qui devraient permettre de désigner des outils pour améliorer l'adaptation des arbres, restent très peu documentés à ce jour.

Dans ce contexte, nous proposons la formation du bois chez les arbres comme notre système de recherche sur le thème de « trade-off » croissance et résistance aux stress. Nous émettons l'hypothèse que le contrôle transcriptionnel des gènes *Aux/IAA* joue un rôle central dans l'établissement des voies de signalisation de l'auxine qui régulent la croissance, l'organogenèse, et la réponse aux stress environnementaux. Je souhaite mieux comprendre les mécanismes biologiques qui expliquent les réponses aux stress afin de découvrir les paramètres qui pourraient être utilisés à terme pour une meilleure sélection génétique des arbres faces aux contraintes climatiques.



**Figure 9. Phénotypes des plantes de Peuplier transgéniques surexprimant les gènes *EgrIAA4* et *EgrIAA13* après un arrêt d'arrosage.** Les lignées *EgrIAA13mOE* présentent un stress hydrique retardé. Photo de peupliers (*Populus alba x tremula*) à 5 mois, les peupliers ont été photographiés à 2 moments différents : Jour 14 et Jour 33 après un arrêt d'arrosage. Les principaux symptômes visibles sont la perte des feuilles, le brunissement des feuilles, le flétrissement des feuilles et le ramollissement de l'apex chez certains individus.

Au cours de ces dix dernières années, nous avons pu identifier plusieurs membres de *Aux/IAA* associés à la formation du bois à partir des données d'analyse d'expression spatiotemporelle et

caractérisation fonctionnelle (Hong, 2014; Karannagoda et al., 2022; Yu et al., 2022; Yu et al., 2015). Les expériences préliminaires à un arrêt de l'arrosage ont montré que les jeunes peupliers transgéniques sur-exprimant les gènes *Eucalyptus EgrIAA4m* et *EgrIAA13* ont une meilleure tolérance à la sécheresse (Figure 9), pour certains sans grandement modifier leur croissance. J'aimerais valoriser ces plantes transgéniques pour approfondir notre connaissance sur les mécanismes d'adaptation « trade-off » entre la croissance et la résistance, afin de finalement proposer certains gènes/marqueurs génétiques pour aider à la sélection génétique finalement. Pour cela, nous pouvons tout d'abord évaluer la résistance aux stress (sécheresse, froid etc.) des différentes lignées transgéniques des *Aux/IAA*. Ensuite, nous caractérisons la structure et la composition du bois et le profil d'expression des gènes (RNAseq) des lignées plus tolérantes ou plus sensibles aux stress. Après, nous établissons des corrélations multi-variables entre les phénotypes physicochimiques du bois, la tolérance aux stress et le réseau d'expression des gènes afin d'expliquer le remodelage de la paroi du bois en réponse aux stress et identifier de nouveaux gènes candidats agissant dans ce contexte. Enfin, nous proposons des paramètres du bois caractéristiques d'un arbre susceptible de mieux résister aux stress.

### A3 : Régulation et dynamique de la différenciation cellulaire au cours de la formation du bois par approche transcriptomique type séquençage de l'ARN unicellulaire (scRNA-seq)

De plus en plus d'évidences ont montré que la différenciation cellulaire au cours de la formation du bois est très plastique et modulable. Cela aide les arbres à mieux s'adapter à leur environnement. Au cours de ces dernières années, en analysant les plantes mutantes que j'ai développées affectant la formation du bois, en particulier *Arabidopsis* transgéniques surexprimant *Aux/IAAs*, leur phénotype de contrôle sélectif de la différenciation cellulaire du bois m'intrigue fortement. Plus les recherches avancent, plus je suis fascinée par cette thématique de la différenciation cellulaire de cambium aux xylèmes secondaires (vaisseaux, fibres secondaires, cellule de rayon, parenchyme du xylème, trachéides etc.).

La croissance latérale, qui épaissit le corps de la plante, est de nature bifaciale avec une seule couche de cellules souches / cellules initiales dans le cambium vasculaire se développant vers l'intérieur dans le xylème et vers l'extérieur dans le phloème (Smetana et al., 2019). Il y a deux types d'initiales cambiales, les initiales cubiques qui peuvent se développer en cellules de rayons et les initiales fusiformes qui peuvent se développer vers l'extérieur en phloème secondaire (y compris les tubes criblés, cellules compagnes, fibres du phloème et parenchyme du phloème) ou vers l'intérieur en xylème secondaire, y compris les vaisseaux (tube-like), les trachéides, les fibres secondaires et les parenchymes du xylème. Nous avons commencé à identifier certains facteurs hormonaux et génétiques impliqués dans ce processus de différenciation, mais le mécanisme précis contrôlant la spécification des cellules du cambium et la maturation des dérivés cambiaux reste encore un sujet fascinant à découvrir.

Les méthodes standard telles que les puces à ADN et le séquençage d'ARN en masse analysent l'expression de l'ARN à partir de grandes populations de cellules. Ces mesures peuvent masquer les différences critiques entre les cellules individuelles dans les populations de cellules mixtes. Le séquençage de l'ARN unicellulaire (scRNA-seq) fournit les profils d'expression de cellules individuelles et est considéré maintenant comme la référence en matière de définition des états et des phénotypes

cellulaires. Bien qu'il soit impossible d'obtenir des informations complètes sur chaque ARN exprimé par chaque cellule, en raison de la petite quantité de matériel disponible, les modèles d'expression génique peuvent être identifiés grâce à des analyses de regroupement de gènes. Cela peut révéler des types de cellules rares au sein d'une population cellulaire qui n'auraient peut-être jamais été observées auparavant. La transcriptomique unicellulaire permet de démêler des populations cellulaires hétérogènes, de reconstruire les voies de développement cellulaire (trajectoire développementale) et de modéliser la dynamique transcriptionnelle – le tout auparavant masqué par le séquençage massif d'ARN (RNA-seq) ou des recours des puces à ADN, les analyses globales ne parviennent pas à reconnaître si un changement dans le profil d'expression est dû à un changement de régulation ou de composition – par exemple si un type de cellule apparaît (début de la différenciation) pour dominer la population. Enfin, lorsque l'objectif est d'étudier la progression cellulaire par différenciation, les profils d'expression moyens ne peuvent classer les cellules que par temps plutôt que par stade de développement. Par conséquent, ils ne peuvent pas montrer de tendances dans les niveaux d'expression génique spécifiques à certaines étapes. Pour ces raisons, les études de la différenciation cellulaire ont beaucoup avancé ces derniers temps grâce à l'évolution rapide des méthodes et technologies, notamment le séquençage d'ARN unicellulaire, quel que soit le domaine biologique humain, animal ou végétal.

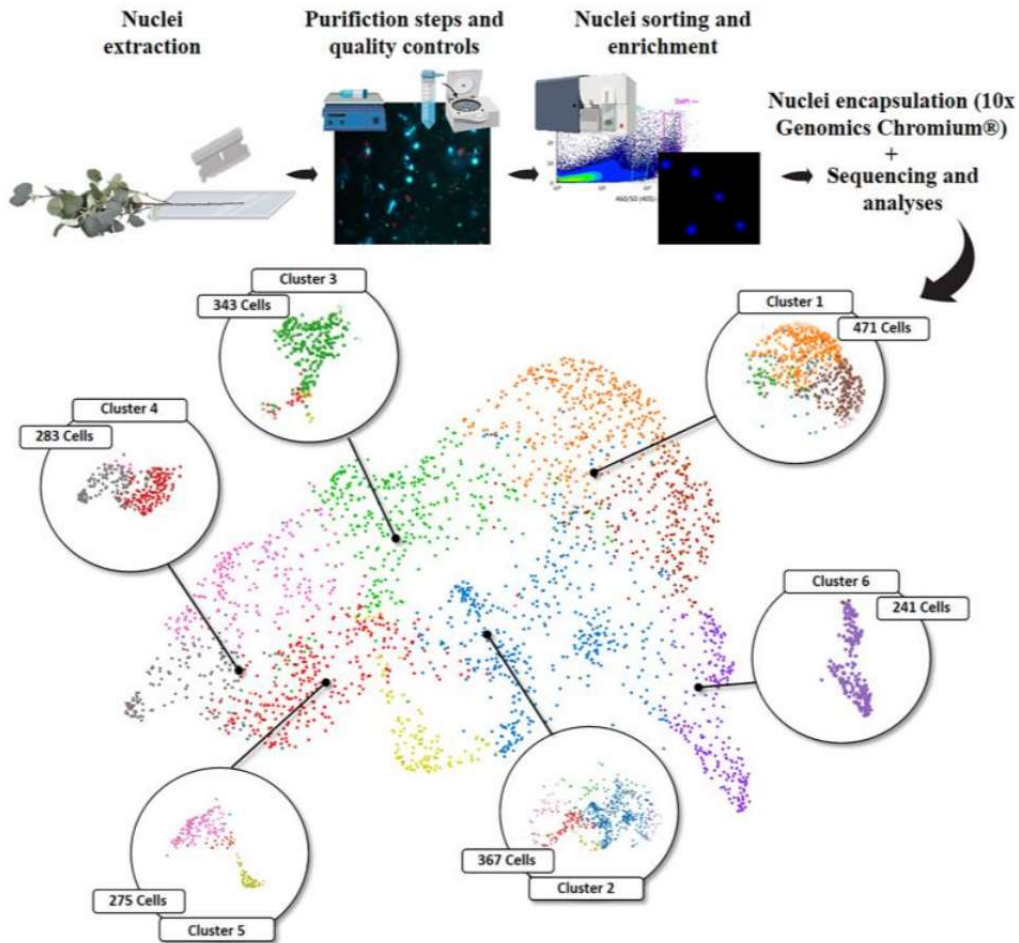
Les progrès récents en scRNA-seq chez les plantes ont montré des transcriptomes hautement hétérogènes de lignées cellulaires de la détermination de l'identité au cours de la différenciation cellulaires (Denyer et al., 2019; Jean-Baptiste et al., 2019; Kim et al., 2021; Neumann et al., 2022; Nobori et al., 2023; Rodriguez-Villalon & Brady, 2019; Shulze et al., 2019; Wendrich et al., 2020; Zhang et al., 2019). Les expériences scRNA-seq réalisées se sont concentrées sur la croissance apicale des plantes dans les feuilles (Kim et al., 2021; Lopez-Anido et al., 2021), les fleurs (Neumann et al., 2022; Xu et al., 2021), Shoot apex (méristème apical) (T. Q. Zhang et al., 2021) et en majorité pour l'instant sur le développement et croissance des racines (méristème racinaire) (Denyer et al., 2019; Jean-Baptiste et al., 2019; Ryu et al., 2019; Shulze et al., 2019; Turco et al., 2019; Wendrich et al., 2020; T.-Q. Zhang et al., 2021; Zhang et al., 2019), laissant les tiges et les troncs d'arbres (méristème latérale) comme l'organe le moins étudié avec des méthodes de profilage à haute résolution. Ces expériences de scRNA-seq ont été principalement réalisées à l'aide de plantes herbacées, principalement *Arabidopsis*, qui tend à produire un nombre limité de cellules de tissu vasculaire en raison de la proportion relativement faible de tissu vasculaire dans les quatre organes mentionnés ci-dessus. Les cellules de xylème identifiées dans les analyses précédentes de scRNA-seq de plantes herbacées ne se situaient que dans la plage de 72 à 1205 cellules, ce qui serait insuffisant pour mieux comprendre les lignées cellulaires dans le développement du xylème. De plus, *Arabidopsis* manque de cellules de parenchyme des rayons dans des conditions de croissance normales (Chaffey et al., 2002; Nieminen et al., 2004; Strabala & Macmillan, 2013) et a un développement de fibres immatures sans mort cellulaire programmée complète (Bollhöner et al., 2012; Sundell et al., 2017). Ces limitations des plantes herbacées soulignent le besoin de plantes ligneuses dans une étude approfondie du développement du xylème. L'année dernière, plusieurs études concomitantes ont montré qu'il était possible d'utiliser cette approche sur des échantillons ligneux en utilisant les méthodes de séquençage d'ARN à protoplaste unique ou à noyau unique chez peuplier (Chen et al., 2021; Conde et al., 2021; Li et al., 2021; Tung et al., 2023).

Il existe plusieurs façons d'isoler des cellules individuelles avant l'amplification et le séquençage du génome entier. Le tri cellulaire activé par fluorescence (FACS) est une approche largement utilisée. Les cellules individuelles peuvent également être collectées par micromanipulation, par exemple par dilution en série ou en utilisant une pipette patch ou un nanotube pour récolter une seule cellule. Les avantages de la micromanipulation sont sa facilité et son faible coût, mais elle est

laborieuse et susceptible de donner lieu à une mauvaise identification des types de cellules au microscope. La microdissection par capture laser (LCM) peut également être utilisée pour collecter des cellules individuelles. Bien que le LCM préserve la connaissance de l'emplacement spatial d'une cellule échantillonnée dans un tissu, il est difficile de capturer une cellule entière sans collecter également les matériaux des cellules voisines. Les méthodes à haut débit pour l'isolement d'une seule cellule incluent également la microfluidique, ces méthodes plus récentes encapsulent des cellules individuelles dans des gouttelettes dans un dispositif microfluidique, où se produit la réaction de transcription inverse, convertissant les ARN en ADNc. Chaque gouttelette porte un « code-barres » d'ADN qui marque de manière unique les ADNc dérivés d'une seule cellule. Une fois la transcription inverse terminée, les ADNc de nombreuses cellules peuvent être mélangés pour le séquençage, car les transcriptions d'une cellule particulière sont identifiées par le code-barres unique. Le FACS et la microfluidique sont précis, automatiques et capables d'isoler des échantillons impartiaux. Cependant, les deux méthodes nécessitent d'abord de détacher les cellules de leur microenvironnement, provoquant ainsi une perturbation des profils transcriptionnels dans l'analyse de l'expression de l'ARN.

Au cours de l'année dernière 2022, Nathalie Ladouce, Fabien Mounet et moi-même, en collaboration avec Frederic Martins (Plateforme GET-SANTE) et Delphine Labourdette (Plateforme GET-BioPuce) avons essayé d'adapter cette approche pour analyser l'effet des stress abiotiques sur le transcriptome des cellules du bois chez *Eucalyptus*. Le principe consiste à isoler des noyaux à partir de cellules de tige puis d'encapsuler ces noyaux grâce à la technologie 10X Genomics Single Cell Sequencing afin de séquencer les ARNm totaux de chacun d'entre eux. Nos premières expériences préliminaires, réalisées sur des jeunes tiges d'*Eucalyptus*, ont permis d'aboutir, avec succès, au transcriptome de plus de 2400 cellules, avec une moyenne de >170 000 lectures par cellules, ce qui a permis l'analyse de 1000 gènes marqueurs par cellule en moyenne. Sur la base de ces résultats de transcriptome, au moins 9 populations cellulaires ont pu être définies (Figure 10). Il reste cependant à identifier les programmes génétiques activés dans ces cellules, comparer leur transcriptome avec celui de peuplier et déterminer les voies de régulations au cours de leur développement.

A ces profils d'expression de gènes « cellules spécifiques », nous avons pour objectif de combiner des données de microscopie vibrationnelle : un ensemble de techniques non destructrices qui permettent d'obtenir des informations fonctionnelles sur la structure et la composition de la paroi des différents types de cellules du xylème. Couplée à des spectroscopes Infra Rouge ou Raman, la microscopie optique ou à fluorescence permet d'établir une « empreinte chimique » de la paroi à l'échelle micrométrique, de comparer des types cellulaires et de suivre leur composition en fonction de conditions variables (Bhagia et al., 2022; Cuello et al., 2020). Il est également possible de recueillir ces informations après l'utilisation de traitements visant à déconstruire de façon spécifique les différents polymères pariétaux (Gierlinger, 2018). Ces approches de microscopie fine ont été utilisées pour révéler des différences structurales au niveaux des couches de la paroi secondaire, pour comparer différentes espèces d'arbres, après l'application de contraintes physiques ou au niveau des anneaux de croissance saisonniers (Gierlinger, 2018). Je suis convaincue que l'intégration de données de transcriptomique et de microscopie fonctionnelle pourraient nous permettre de mieux décortiquer les régulations mis en œuvre dans les différentes cellules qui composent le xylème et de comprendre en quoi leur reprogrammation dans un contexte de stress peut modifier le développement de ce tissu particulier, et conséquent la physiologie de l'arbre.



**Figure 10. Analyse préliminaire des données de transcriptome de 2300 noyaux isolées à partir de jeunes tiges d'*Eucalyptus grandis*.** Le protocole, adapté de la référence (Conde et al., 2021), vise à extraire des noyaux de jeunes tiges. Après lavages, filtrations et tri, les noyaux sont encapsulés au sein de gouttelettes de gel par la technologie 10x Genomics Chromium®, qui assure une haute efficacité de capture des noyaux issus des différents types cellulaires présent dans les tissus (Plateforme GET-SANTE). Les ARNm de chaque noyau, « reverse transcrits » et barcodés, constituent autant de banques d'ADNc séquencées par la technologie Illumina NOVAseq 4000 (Plateforme GET-PLAGE). Un clustering K-means, réalisé à partir des données de séquençage, a permis d'identifier 6 groupes comprenant >200 à >400 noyaux partageant des transcriptomes similaires. Ces différents clusters contiennent plus de 50 gènes marqueurs du processus de xylogénèse (facteurs de transcription, synthèse de paroi secondaire, régulation hormonales).



## B : Participation aux projets de notre équipe

Par ailleurs, je participe activement aussi aux projets obtenus par notre équipe.

### B1. Régulation post transcriptionnelle de la formation du bois : Rôle des protéines Musashi

**Titre du projet :** Mechanisms and contributions of Musashi-mediated control in cell wall polymers synthesis in plants (MusaWall)

**Coordinateur du projet :** Thierry Lagrange, DR CNRS Laboratoire Génome et Développement des Plantes Perpignan

**Partner 2 :** Biopolymères Interactions Assemblages (BIA, INRAE, Nantes), PI : Richard Sibout

**Partner 3 :** LRSV, UT3/CNRS/INP ENSAT, PI : Fabien Mounet

#### **Résumé du projet :**

Chez les animaux comme chez les végétaux, la traduction est régulée par l'action de protéines capables de se fixer aux ARN messagers. Ces « RNA Binding Proteins » (RBPs) interviennent dans une multitude de processus cellulaires, en régulant la synthèse, le repliement, la translocation, l'assemblage et le tri des ARN messagers. À ce jour, un nombre très restreint d'études se sont intéressées au rôle des RBPs dans la formation du xylème. Chez *Arabidopsis thaliana*, deux protéines de type CCCH zinc-finger, peuvent se fixer aussi bien à l'ADN qu'aux ARNm, et semblent impliquées dans la formation de la paroi secondaire (Chai et al., 2022; Kim et al., 2014). De la même façon, le facteur d'initiation de la traduction eIF5a, qui favorise la synthèse de protéines en participant à l'exportation des ARNm du noyau, peut stimuler ou inhiber la formation du xylème selon s'il est surexprimé ou sous exprimé chez *Arabidopsis* (Liu et al., 2008). Au LGDP de Perpignan, l'équipe de Thierry Lagrange s'intéresse à une famille de gènes qui codent pour des RBPs contenant des motifs RRM de liaison à l'ARN. Ces gènes, nommés MLG-1 à 4 chez *Arabidopsis*, partagent de fortes similarités de séquence, de structure et de fonction avec certains régulateurs de la transcription appelés MUSASHI chez les animaux. Les MUSASHI peuvent se fixer dans la région 3'UTR de leur cible et moduler de façon positive ou négative la traduction des ARNm en recrutant des partenaires protéiques. Bien que les protéines MLG aient été isolées spécifiquement à partir de protéomes associés aux ARNm, nous ne savons presque rien de leur fonction *in planta*. La perte de fonction des gènes *MLG* entraîne une diminution drastique de la rigidité de la hampe florale, ce qui laisse supposer un défaut dans la formation des tissus vasculaires que les équipes de T. Lagrange et F. Mounet ont cherché à caractériser. Des analyses d'imagerie et d'histochimie à partir de coupes transversales de tiges ont révélé une réduction de l'épaisseur de la paroi secondaire des fibres interfasciculaires. La paroi de ces cellules est également moins lignifiée, comme le confirment les marquages au phloroglucinol et les dosages au bromure d'acétyle. L'utilisation d'un marquage au Calcofluor suggère un dépôt de polysaccharides plus important que dans les plantes sauvages. Cette hypothèse est renforcée par des résultats de saccharification, qui montrent une libération de sucres plus importante à partir des parois des plantes mutantes *mlg24* en comparaison des plantes Wild-type. Ces modifications sont complétées par la restauration de la fonction MLG dans des plantes *mlg24* + MLG2. L'utilisation de la spectroscopie infra-rouge à transformée de fourrier (FT-IR) montre que la composition des parois des plantes mutantes *mlg24* est radicalement différente de celle du WT, pointant des régions du spectre associées aussi bien à l'absorption de composés de la lignine, de la cellulose et des hémicelluloses dans la littérature. L'analyse fine des monosaccharides pariétaux par l'équipe de Richard Sibout (INRAE de Nantes) a montré que la perte de fonction des gènes *MLG2* et 4 n'entraîne en réalité aucune modification de la

teneur en sucres neutres présents dans la cellulose ou les hémicelluloses, mais une réduction importante des acides glucuroniques (GlucA) qui sont les constituants les plus abondants des xylan/glucuronoxylan chez *Arabidopsis*. Ces mutants *mlg* présentent en effet des phénotypes très similaires à ceux des mutants *gux*, au sein desquels l'inactivation du gène codant pour une glucuronosyl transférase entraîne une forte diminution des GlucA au niveau des xylanes et une perte de rigidité des tissus vasculaires de la tige (Lee et al., 2012; Mortimer et al., 2010). Bien que les cibles exactes des protéines MLG ne soient pas identifiées, nos observations mettent en évidence l'existence d'un nouveau mécanisme de régulation de la formation de la paroi secondaire qui implique des RBPs spécifiques. Étant donné la présence de gènes présentant une forte homologie avec les MLG chez d'autres plantes, ce mécanisme pourrait être conservé au sein du règne végétal. On retrouve par exemple 8 MLG putatifs dans le génome du peuplier, qui semblent s'exprimer dans tous les tissus. Chez *Eucalyptus*, seulement 4 gènes sont de potentiels orthologues des *MLG* et 3 d'entre eux s'expriment préférentiellement dans les tissus vasculaires. La caractérisation fonctionnelle des gènes *MLG* chez les arbres, en utilisant des lignées transgéniques sur/sous-expresses, nous permettra de mieux comprendre comment ces gènes sont impliqués dans la formation du bois chez les plantes pérennes.

Plus particulièrement, pour les premières étapes du projet, je me charge de la construction des vecteurs sur-expresses sous contrôle de promoteur 35S et knock-out par la technique de CRISPR Cas9 avec une Ingénieure dédiée à ce projet.

## B2 : Projet « TYPEX » (2023-2028) : Vers une édition spécifique et précise du génome végétal.

**Titre du projet :** Toward highly Predictable Editing of the plant genome leXicon.

**Responsable du projet :** Pierre-Marc Delaux, DR CNRS, Fabien Nogué, DR INRAE

**Type de projet :** ANR, PEPR (Programmes et équipements prioritaires de recherche).

### **Résumé du projet :**

Depuis sa découverte en 2012, la technologie CRISPR-Cas a rapidement supplanté toutes les autres méthodes d'édition du génome en raison de son efficacité et de sa facilité d'utilisation. Cette technologie s'est considérablement améliorée au cours de la dernière décennie pour répondre aux besoins des scientifiques et des sélectionneurs, élargissant leur boîte à outils pour créer des modifications dans le génome. Cependant, nous sommes encore loin de la situation idéale où n'importe quel type de modification peut être systématiquement introduit à n'importe quelle position dans le génome de n'importe quelle espèce végétale. L'objectif principal du projet TYPEX est de permettre à la communauté française des scientifiques et des sélectionneurs de maîtriser l'édition précise du génome végétal, et ce, chez un nombre important d'espèces modèles et cultivées. À l'heure actuelle, les limitations à une édition prédictible et précise du génome sont : (1) les caractéristiques inhérentes aux outils d'édition du génome limitant leur polyvalence et/ou leur efficacité et (2) la capacité de régénération des plantes après l'introduction et l'action du module CRISPR-Cas dans les cellules végétales. Récemment, une nouvelle stratégie a été développée, nommée prime-editing (PE), qui permet une très grande variété d'édits (substitution de base, insertions, suppressions, remplacements). Depuis qu'elle a été utilisée pour la première fois dans des cellules humaines, cette méthode a été appliquée à quelques espèces végétales. Les résultats de ces études indiquent qu'une

amélioration substantielle des méthodes de PE est nécessaire avant qu'elles puissent être utilisées en routine en sélection variétale.

Le premier objectif du projet TYPEX sera de développer des outils et des règles générales pour un PE efficace chez les plantes. Trois plantes modèles, *Marchantia*, *Physcomitrium* et *Arabidopsis*, chez lesquelles les stratégies basées sur le système CRISPR-Cas classique sont utilisées en routine, seront utilisées en parallèle pour améliorer le PE. L'amélioration du PE, basée sur une stratégie de type « cycle DBTL », portera sur plusieurs paramètres clés de l'efficacité du PE, tels que le type d'enzymes utilisées pour l'édition, les conditions expérimentales, ou la conception du pegRNA qui porte l'information de l'édition à réaliser. Parce que les stratégies de transformation utilisées pour ces modèles sont différentes, les améliorations apportées aux trois plantes seront très complémentaires et transférables à une majorité de plantes, y compris les plantes de grandes cultures. L'objectif final est de construire une capacité partagée d'outils pour l'édition précise et prédictible du génome végétale, avec la mise en place d'une plateforme en charge de la synthèse de gènes, des stratégies de clonage et de la collecte des modules CRISPR optimisés. Cette plateforme distribuera aux participants du projet les différents modules/protocoles à utiliser pour une édition du génome efficace. Ce sera le deuxième objectif du projet, qui est de déployer le prime editing dans une gamme large d'espèces de plantes cultivées. Pour cela, les règles générales établies sur les espèces modèles seront testées de manière exhaustive sur des plantes où la stratégie CRISPR-Cas classique est déjà maîtrisée et qui sont des cultures d'intérêt pour l'agriculture française. Le troisième objectif sera de permettre l'édition de gène chez des plantes pour lesquelles des techniques efficaces de transformations, et donc d'édition du génome, ne sont pas encore disponibles. À cet effet, de nouvelles stratégies, originales, de livraison au sein de la cellule des modules CRISPR-Cas optimisés, seront déployées. Ainsi, le projet TYPEX va renforcer fortement la capacité des acteurs français de la sélection et de la recherche pour faire de la création variétale précise et sur un très grand nombre d'espèces.

### **Les objectifs scientifiques et le programme de notre participation**

Le bois est crucial pour la conduction de la sève et le support mécanique des arbres. Il s'agit du réservoir de biomasse carbonée le plus abondant sur terre, utilisé pour la bioénergie, comme matière première et source de nouveaux produits chimiques. Le bois est principalement composé de SCW enrichis en polysaccharides et en lignine, un polymère issu du métabolisme des phénylpropanoïdes. La lignine représente une barrière hydrophobe, importante pour la conduction de la sève et la résistance au stress (a)biotique, mais elle est également un facteur de récalcitrance majeur nuisant au traitement de la biomasse. Une réduction drastique de la teneur en lignine nuit souvent à la croissance des arbres sur le terrain. Les progrès les plus récents en matière d'édition du génome ont ouvert la voie à de nouvelles stratégies de sélection visant à générer des génotypes hypo-lignifiés sans pénalité sur le terrain, en utilisant des mutations spécifiques des isoformes ou des allèles défectueux d'enzymes spécifiques de la lignine (Chanoca et al., 2019). À titre de preuve de concept, l'équipe ReDyWood a réussi à réduire la teneur en lignine dans le KO/KD édité par CRISPR/Cas9 du gène de la cinnamoyl-CoA réductase (Dai et al., 2020). De plus, en tant que régulateurs clés de la signalisation de l'auxine, Aux/IAA jouent un rôle clé dans la structuration des tissus du bois. ReDyWood a démontré que la surexpression des gènes *Aux/IAA* (*EgrIAA20*, *EgrIAA9A*) d'*Eucalyptus* réduisait spécifiquement la lignification des fibres sans affecter les vaisseaux SCW (Yu et al., 2022). Ceci est très recherché pour l'application de la réduction de la lignification sans impacter les vaisseaux qui assurent la conduction des sèves, réduisant ainsi le problème de goulot d'étranglement lié à la pénalité de rendement.

Les meilleures technologies d'édition, adaptées aux génotypes d'arbres élites, laissent espérer une avancée majeure dans les stratégies de sélection des arbres forestiers. Parmi les arbres forestiers industriels, l'*Eucalyptus* à croissance rapide, très productif, est l'arbre à larges feuilles le plus planté au

monde et une essence ligneuse prometteuse pour faire face au changement climatique dans les pays du Nord. Profitant de l'expertise de ReDyWood, nous visons à produire des lignées transgéniques d'*Eucalyptus* hypo-lignifiées à travers l'édition premium de *CCR1* (*Eucgr.J03114*) et *IAA9A* (*Eucgr.H02407*), afin d'améliorer la transformation de la biomasse ligneuse.

Le domaine II élevé de la famille de gènes *Aux/IAA* est responsable de la stabilité de la protéine, un changement d'acide aminé entraîne une stabilisation de la protéine affectée et une diminution de la réponse auxine (Ouellet et al., 2001; Overvoorde et al., 2005; Reed, 2001; Worley et al., 2000). Nous visons à muter le 229<sup>ème</sup> acide aminé de la proline à la sérine (P-to-S) de *EgrIAA9A*, sur la base des résultats de l'édition, nous pouvons obtenir à la fois les mutants à gain de fonction (portant exactement la mutation « P-to-S ») ou mutants avec perte de fonction (portant une mutation du cadre de lecture par décalage).

#### **Les objectifs intermédiaires et les livrables :**

- 1 : Régénération des Hairy Root fluorescentes transgéniques avec des constructions d'édition Prime ciblant les gènes *EgrCCR1* et *EgrIAA9A*
- 2 : Sélection de Hairy Root transgéniques éditées par KO-CCR et gain de fonction et perte de fonction *EgrIAA9A*
- 3 : Identification des lignées éditées présentant des tissus vasculaires hypo lignifiés dans les fibres et/ou dans les vaisseaux.

## **C : Caractérisation des modifications et adaptations des parois au cours du développement et en réponse aux stress biotiques et abiotiques**

Aux cours de mes activités de recherches de ces quinze dernières années, je suis en permanence confrontée à la problématique de la caractérisation des modifications et adaptations des parois. En fait, les cellules de bois sont des cellules spécialisées caractérisées par la présence des parois secondaires épaisses et lignifiées. Nous avons pu développer, à travers nombreux projets de recherches, des compétences et des méthodologies pour étudier la formation des parois secondaires et primaires. Naturellement, je souhaiterais développer davantage cet axe pour avancer nos recherches.

### **C1. Caractérisation parois en réponse aux stress biotiques**

Suite aux réponses d'appel d'offre projet inter-unités FRAIB 2021-2022 (The Agrobiosciences, Interactions and Biodiversity Research Federation), nous avons obtenu un financement pour le projet "Analysis of primary cell WALL in REsistance under combined stress" avec Dr. Marta MARCHETTI/Richard BERTHOME équipe REACH (Dynamique de la Réponse Immunitaire et Adaptation au Changement Climatique) du LIPMe (laboratoire des interactions plantes - microbes – environnement). Dans ce projet, nous avons caractérisé et comparé des modifications fines des parois des deux différents mutants de *cesa3* chez *Arabidopsis thaliana* qui ont une résistance différente au pathogène *Ralstonia solanacearum* lié à l'élévation de la température.

Les plantes terrestres sont soumises à de nombreux stress biotiques et abiotiques au cours de leur vie et ont développé en conséquence un large éventail de mécanismes pour se développer et survivre. En particulier, l'immunité des plantes repose à la fois sur des barrières constitutives préformées ainsi que sur des mécanismes induits lorsqu'elles entrent en contact avec des pathogènes envahisseurs conduisant soit à une résistance, soit au développement de maladies en fonction du génotype de la plante et du type d'interactions entre pathogènes (ou pathosystèmes). De plus, une majorité de résistances végétales sont impactées négativement par les stress abiotiques.

De nouvelles preuves indiquent que les cellules végétales exploitent des mécanismes sophistiqués de détection de l'altération de l'intégrité de la paroi cellulaire lors d'un stress biotique (Hamann, 2012; Pogorelko et al., 2013). En raison de sa localisation stratégique à l'interface plante-environnement, la paroi cellulaire contribue probablement aux résultats des interactions plante-pathogène (Huckelhoven, 2007). En fait, il a été démontré qu'un ensemble de mutants de la paroi cellulaire d'*Arabidopsis* ont un impact sur les phénotypes de résistance aux maladies et de condition physique, établissant un lien entre la composition de la paroi cellulaire végétale et les phénotypes de développement/immunité des plantes (Molina et al., 2021). En effet, plusieurs mutants de la paroi cellulaire d'*Arabidopsis* qui développent une résistance à *Agrobacterium tumefaciens* (*rat*), aux agents pathogènes de l'oïdium (*pmr*) et au champignon *Plectosphaerella cucumerina* ont été identifiés (Vogel & Somerville, 2000; Vogel et al., 2002; Vogel et al., 2004). De plus, l'inactivation d'*Arabidopsis* *WAT1* (*Walls Are Thin1*), un gène nécessaire au dépôt secondaire sur la paroi cellulaire, ou des gènes de la cellulose synthase spécifique à la paroi secondaire (*CesA*) (*ixr*) confère une résistance à large spectre aux agents pathogènes (Denance et al., 2013; Hernandez-Blanco et al., 2007).

L'élévation de la température, l'un des stress abiotiques majeurs auxquels les plantes doivent faire face actuellement lié au réchauffement climatique (Desaint et al., 2021), devrait favoriser l'émergence de nouveaux agents pathogènes et augmenter l'apparition d'épidémies (Garrett et al., 2006). En fait, Les recherches menées par nos collaborateurs de l'équipe REACH du LIPMe ont montré que l'élévation de la température a un impact drastique sur les réponses de défense des plantes et s'inhibe la principale source de résistance à *Ralstonia solanacearum* chez *A. thaliana*, médiée par la paire d'immuno-récepteurs RSP4/RRS1-R. Des conditions de températures plus élevées de 3°C (27-30 °C) suffisent pour inhiber les principaux mécanismes de résistance au pathogène *Ralstonia solanacearum* (Aoun et al., 2017; Desaint et al., 2021). *R. solanacearum* est une bactérie du sol, pathogène des végétaux, Gram-négative. Présente dans tous les continents, particulièrement dans les régions tropicales et subtropicales, elle colonise le xylème, causant une pourriture bactérienne ou bactériose vasculaire chez de nombreuses plantes-hôtes, touchant plus de 53 familles botaniques y compris de plantes de grandes cultures et des arbres. C'est notamment l'agent de la pourriture brune de la pomme de terre, de la maladie de Moko du bananier et le flétrissement bactérien de la tomate.

Nous collaborateurs de l'équipe REACH ont réalisé la cartographie GWAS, en utilisant une population de 176 accessions naturelles d'*Arabidopsis*, pour identifier les QTL associés de réponse à la souche *R. solanacearum* GMI1000 à deux températures déférentes, 27°C et 30°C. Cette étude a permis d'identifier le gène *CELLULOSE SYNTHASE A3* (*CesA3*) sous-jacent à un QTL majeur impliqué dans la résistance quantitative précoce aux maladies (QDR) à 30°C (Figure XXXX).

*CesA3* appartient à une famille multigénique des gènes cellulose synthase (*CesA*) identifiés chez *A. thaliana* et code pour une sous-unité de cellulose synthase qui, avec autres sous-unités *CesA1*, -2, -5 et -6, compose la rosette de la cellulose synthase et participe à la production de cellulose lors de la formation de la paroi cellulaire primaire. L'importance du complexe primaire de cellulose synthase (CSC) de la paroi cellulaire dans l'établissement de mécanismes de défense contre *R. solanacearum* à des températures élevées a été démontrée par l'étude d'autres sous-unités primaires du complexe

*CesA*, *CesA1* (*rsw1-10*) et *CesA6* (*prc1-1*) mutants (Cano-Delgado et al., 2003; Menna et al., 2021). On a validé *CesA3* comme gène de susceptibilité impliqué dans le QDR à *R. solanacearum* à 30°C sur l'étude des mutants knock-down *cesa3 eli1-1* (Cano-Delgado et al., 2003) et *cesa3 je5* par nos collaborateurs. Mais jusqu'à présent, le rôle de la paroi cellulaire primaire en réponse à *Ralstonia* n'a pas été documenté, que ce soit à température ambiante ou à température élevée.

Dans ce projet nous souhaitons à rechercher l'implication possible de la paroi cellulaire primaire dans la réponse précoce des plantes à *R. solanacearum*, visant à identifier les bases génétiques des mécanismes robustes de résistance des plantes à cette bactérie pathogène dans un contexte de réchauffement climatique. Nos objectifs communs sont :

- De caractériser quelles modifications fines de la paroi des cellules pourraient être impliqués dans la résistance aux stress biotiques et abiotique combinés.

- De mieux comprendre le rôle des constituants de la paroi dans la résistance des plantes à l'infection par *Ralstonia*.

Afin de répondre à ces questions, les mécanismes de défense et la composition de la paroi des cellules racinaires sont en cours d'analyse, en comparant des conditions normales de croissance et des conditions de température élevée :

- (i) à partir de différents mutants pour le gène *CesA3*,

- (ii) à partir des accessions naturelles sensibles et résistantes au pathogène.

En guise d'étude préliminaire, Fabien Mounet et moi-même avons caractérisé la structure et la composition de la paroi des cellules au niveau des racines de plantes sauvages et des mutants *je5* et *ixr1-1* du gène *CesA3*. Chacun de ces mutants a une permutation d'un acide aminé qui touche des domaines différents de la protéine et laisse supposer des effets différents sur la fonction de *CesA3*. Les analyses histochimiques par des marquages spécifiques des celluloses (Calcofluor et rouge Congo) montrent que les mutants présentent des modifications de la structure de la paroi des cellules, surtout pour le mutant *je5* (Figure 11). Le Calcofluor marque la cellulose, la callose et d'autres  $\beta$ -glucanes non substitués ou faiblement substitués ; tandis que le rouge Congo se colore directement en  $\beta$ - (1  $\rightarrow$  4) -glucanes et en particulier en cellulose. Les lignées *je5* montrent des marquages beaucoup plus intenses de Calcofluor et rouge Congo. Les analyses chimiques ont montré que les mutants *je5* ont une réduction significative de la teneur en cellulose. Donc l'augmentation du marquage Calcofluor est due aux modifications de la structure de la paroi qui laissent une plus forte pénétration de colorant Calcofluor. Grâce à des analyses de FT-IR mené par mon collègue Fabien Mounet, nous avons pu observer une différence de composition entre les mutants et le sauvage, mais également entre les mutants eux-mêmes. Nos collègues ont pu montrer que les mutants présentent une réponse opposée à l'infection par *Ralstonia* : *je5* est plus résistant alors que *ixr1-1* est légèrement plus sensible. Tout porte à croire que la modification du dépôt de cellulose dans la paroi peut conditionner l'immunité des plantes d'*Arabidopsis* et nos recherches se poursuivent pour démontrer que ce paramètre est impliqué dans le phénotype des différentes accessions sauvages.



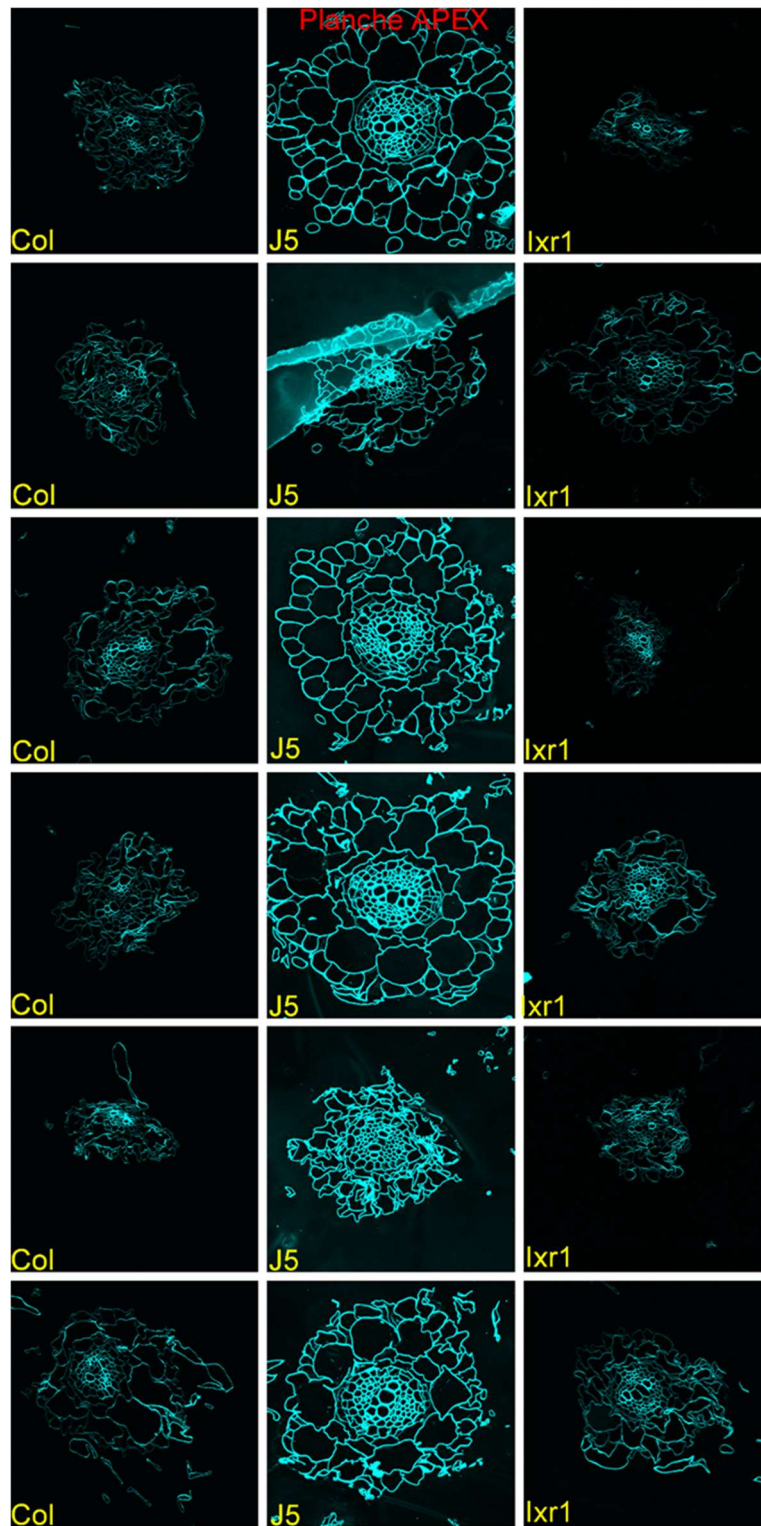


Figure 11. Comparaison de coloration au Calcofluor de coupes transversales des apex de racines d'*A. thaliana* entre wild-types et mutants *cesa3* (*je5* et *lrx1-1*).

## C2. Compréhension des modifications de la paroi associées au développement du gel loculaire : Formation et régulation des tissus gel loculaire, la partie manquante du développement du fruit de la tomate

La tomate est l'espèce modèle pour l'étude des fruits charnus. Différentes variétés de tomates ont été créées depuis sa domestication, permettant d'offrir des fruits plus ou moins fermes et plus ou moins juteux selon les utilisations finales. Une des caractéristiques des fruits charnus est la différenciation, après la fécondation, de cellules du placenta en tissu loculaire qui entourent les graines et qui contribuent à la qualité organoleptique du fruit de par sa composition et sa texture particulière. Bien que le processus de maturation soit associé depuis longtemps au tissu du péricarpe, des données récentes soutiennent l'hypothèse selon laquelle le tissu loculaire est une pièce maîtresse de ce processus. Par exemple, les changements transcriptionnels associés à la maturation semblent commencer d'abord dans le tissu du gel loculaire, avant de s'étendre au péricarpe. De plus, les caractères de ramollissement sont déterminés par le développement du tissu loculaire au début du développement du fruit et non pendant la maturation comme supposé précédemment. Parce que ce tissu a été sous-étudié, son origine et son rôle restent mal connus.

Les résultats récents de l'équipe GBF ont souligné que le facteur de transcription clé régulant la formation de tissus en gel modifiait intensément la composition de la paroi cellulaire ainsi que l'expression des gènes liés aux parois dans les tissus locaux.

Ce projet de recherche répondra à 2 objectifs :

- déterminer la nature de ce tissu gel loculaire puis établir un schéma de régulation chronologique et interactif de la différenciation de ces tissus par approche cytologique qui va révéler la modification de la paroi associée au développement du gel loculaire.
- en utilisant une approche génétique reverse, explorer la régulation de cette formation des tissus gel loculaire par des facteurs de transcription MADS box.

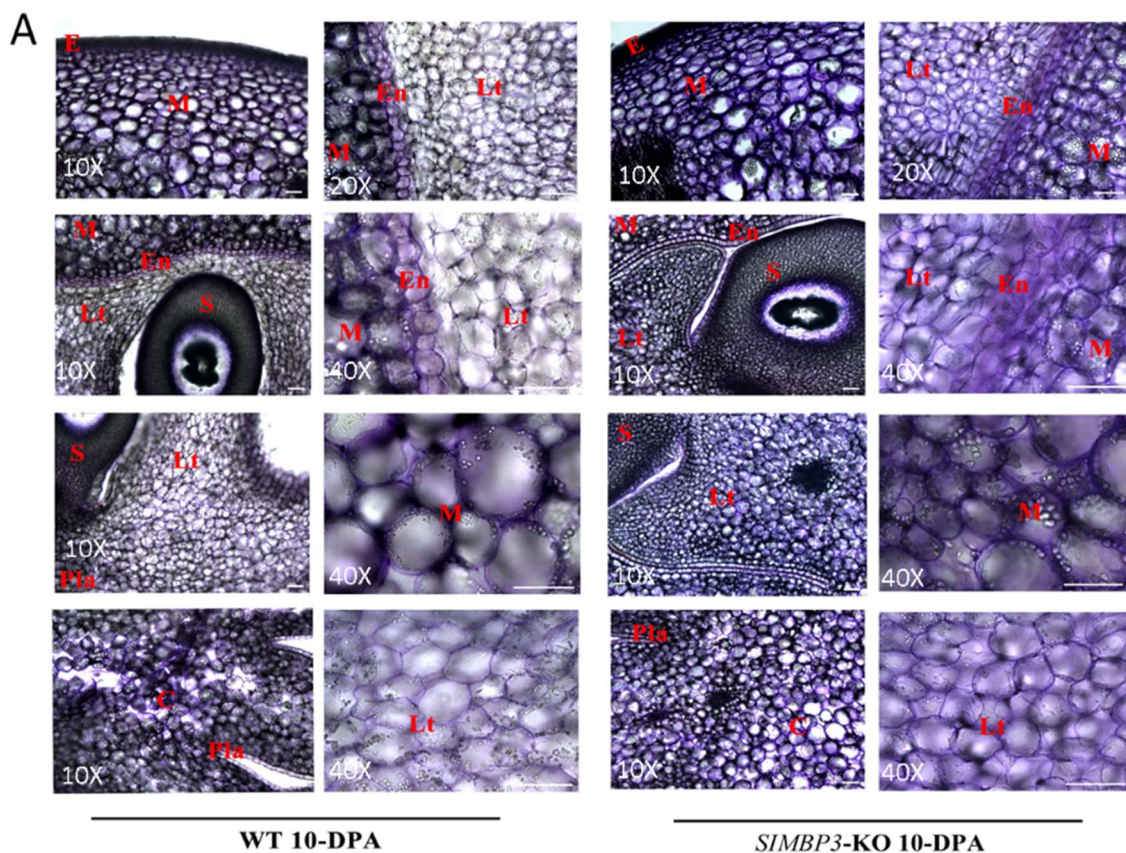
Dans ce projet, nous allons réaliser une caractérisation fine des modifications de la paroi associées à la mise place du gel loculaire, comme cela a pu être fait par exemple pour le développement du péricarpe. De plus, l'équipe GBF a généré des mutants avec des fruits fermes où les parois du gel demeurent rigides, ainsi que des plantes aux fruits mous aux parois peu cohésives (les sur-expresseurs de SI-AGL11 et SI-MBP3) (Huang et al., 2021). L'inactivation de SIMBP3 altère la formation de gel loculaire grâce à une reprogrammation transcriptomique massive lors des phases initiales du développement du fruit, telles que les gènes du cycle cellulaire et de l'expansion cellulaire (Huang et al., 2021). Cette expérience offre donc un matériel génétique de choix pour étudier les modifications de la paroi associée à la différenciation du gel loculaire. En effet, nous envisageons de réaliser une caractérisation histochimique et biochimique des parois au cours du développement du fruit sur les trois types de génotypes (WT, RNAi ou surexpresser). Au niveau (immuno) histochimique, il s'agirait de réaliser un ensemble de coupes de fruits à différents stades et de les tester avec des différentes méthodes pour révéler des différentes compositions des parois (Congo red et Calcofluor pour cellulose, Toluidine Blue pour les pectines etc.) afin de comprendre l'évolution des cellules et du tissu lors de la mise en place de gel loculaire ou lors de la différenciation des fruits transgéniques. Ces analyses seraient complétées par des dosages biochimiques des compositions associées aux parois.

Tout d'abord, pour mieux comprendre les processus de différenciation du tissu loculaire, aboutissant soit à un type All-Flesh (tout-chair et non liquéfié) de SIMBP3-KO, soit à un type de gel

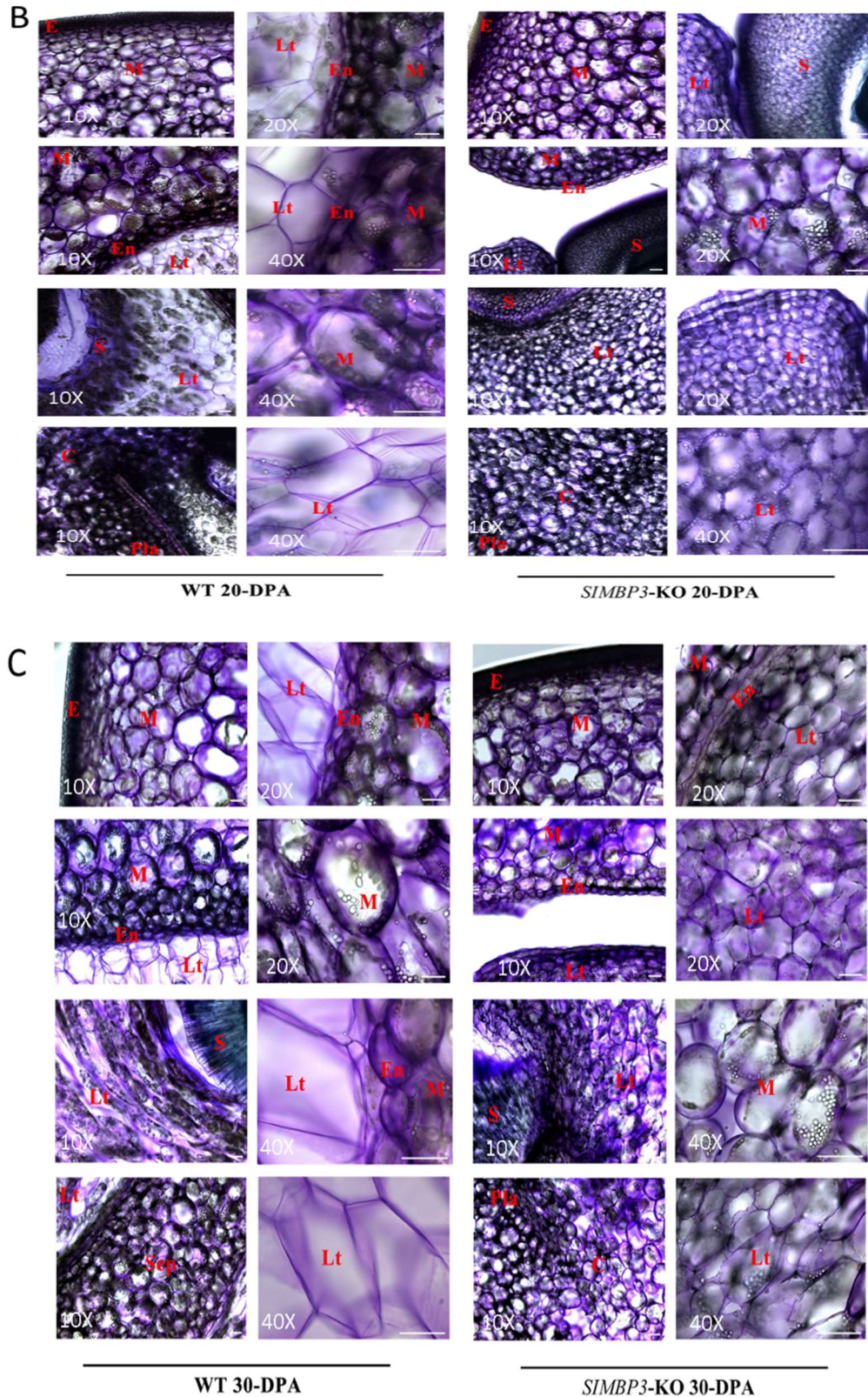
liquéfié de plante sauvage (WT), nous avons utilisé la coloration au bleu de toluidine comme moyen efficace de distinction des tissus liquéfiés des tissus non liquéfiés. Le bleu de toluidine est classé comme colorant polychromatique car il réagit différemment avec différents composants chimiques des cellules et donne un échantillon multicolore (Figure 12). Les couleurs générées peuvent renseigner sur la nature de la cellule et de ses parois. Une solution aqueuse de ce colorant est bleue, mais différentes couleurs sont générées lorsque le colorant se lie à différents groupes anioniques dans la cellule.

Les résultats ont montré que SIMBP3-KO ne diffère des fruits WT que par le tissu loculaire. La coloration de la paroi cellulaire du tissu loculaire WT était plus claire que celle du tissu cellulaire SIMBP3-KO. À 10 DPA (days post anthesis, jours après anthèse), les cellules du tissu loculaire de WT n'étaient que légèrement plus grandes que les cellules SIMBP3-KO et la coloration de la paroi cellulaire de WT était un peu plus claire. À 20 DPA, les cellules du tissu loculaire de WT sont devenues volumineuses avec des parois cellulaires minces et des changements de morphologie, tandis que les cellules de SIMBP3-KO n'ont vu que leur taille augmenter sans changement de morphologie ni d'épaisseur de la paroi. À 30 DPA, les cellules du tissu loculaire de WT sont devenues très grandes avec des parois cellulaires fines en forme de cristaux irréguliers, tandis que les cellules de SIMBP3-KO n'ont changé que leur taille (plus grand mais forme ronde semblable à l'original).

Ces analyses cytologiques montrent que la morphologie des cellules du tissu loculaire est bien différente entre les fruits SIMBP3-KO et WT. Les analyses précoces du développement avant 10 DPA pourraient nous renseigner davantage pour comprendre l'évolution de ce processus de mise en place et la différenciation du tissu loculaire.







**Figure 12. Comparaison des coupes des fruits de tomate entre SIMBP3-KO et le contrôle au cours du développement par la coloration au bleu de toluidine. A, transversales sections des fruits de 10 DPA (jours post anthèse) ; B, transversales sections des fruits de 20 DPA ; C, transversales sections des fruits de 30 DPA. E, exocarpe ; M, mésocarpe ; Fr, endocarpe ; Lt, tissu loculaire ; S, graine ; C, columelle ; Pla, placenta. Barres d'échelle : 50  $\mu$ m.**

## RÉFÉRENCES

- Akhtar, M., Jaiswal, A., Taj, G., Jaiswal, J. P., Qureshi, M. I., & Singh, N. K. (2012). DREB1/CBF transcription factors: their structure, function and role in abiotic stress tolerance in plants. *Journal of Genetics*, 91(3), 385-395. <https://doi.org/10.1007/s12041-012-0201-3>
- Aoun, N., Tauleigne, L., Lonjon, F., Deslandes, L., Vailleau, F., Roux, F., & Berthome, R. (2017). Quantitative Disease Resistance under Elevated Temperature: Genetic Basis of New Resistance Mechanisms to *Ralstonia solanacearum*. *Front Plant Sci*, 8, 1387. <https://doi.org/10.3389/fpls.2017.01387>
- Baima, S., Forte, V., Possenti, M., Peñalosa, A., Leoni, G., Salvi, S., Felici, B., Ruberti, I., & Morelli, G. (2014). Negative Feedback Regulation of Auxin Signaling by ATHB8/ACL5–BUD2 Transcription Module. *Molecular Plant*, 7(6), 1006-1025. <https://doi.org/10.1093/mp/ssu051>
- Barros, J., Serk, H., Granlund, I., & Pesquet, E. (2015). The cell biology of lignification in higher plants. *Ann Bot*, 115(7), 1053-1074. <https://doi.org/10.1093/aob/mcv046>
- Berleth, T., & Mattsson, J. (2000). Vascular development: tracing signals along veins. *Curr Opin Plant Biol*, 3(5), 406-411. [https://doi.org/10.1016/s1369-5266\(00\)00104-7](https://doi.org/10.1016/s1369-5266(00)00104-7)
- Berleth, T., Mattsson, J., & Hardtke, C. S. (2000). Vascular continuity and auxin signals. *Trends Plant Sci*, 5(9), 387-393. [https://doi.org/10.1016/s1360-1385\(00\)01725-8](https://doi.org/10.1016/s1360-1385(00)01725-8)
- Bhagia, S., Durkovic, J., Lagana, R., Kardosova, M., Kacik, F., Cernescu, A., Schafer, P., Yoo, C. G., & Ragauskas, A. J. (2022). Nanoscale FTIR and Mechanical Mapping of Plant Cell Walls for Understanding Biomass Deconstruction. *Acs Sustainable Chemistry & Engineering*, 10(9), 3016-3026. <https://doi.org/10.1021/acssuschemeng.1c08163>
- Boerjan, W., Ralph, J., & Baucher, M. (2003). Lignin biosynthesis. *Annu Rev Plant Biol*, 54, 519-546. <https://doi.org/10.1146/annurev.arplant.54.031902.134938>
- Bollhöner, B., Prestele, J., & Tuominen, H. (2012). Xylem cell death: emerging understanding of regulation and function. *Journal of Experimental Botany*, 63(3), 1081-1094. <https://doi.org/10.1093/jxb/err438>
- Boudet, A. M., Kajita, S., Grima-Pettenati, J., & Goffner, D. (2003). Lignins and lignocellulosics: a better control of synthesis for new and improved uses. *Trends Plant Sci*, 8(12), 576-581. <https://doi.org/S1360138503002486> [pii]
- Brackmann, K., Qi, J., Gebert, M., Jouannet, V., Schlamp, T., Grunwald, K., Wallner, E. S., Novikova, D. D., Levitsky, V. G., Agusti, J., Sanchez, P., Lohmann, J. U., & Greb, T. (2018). Spatial specificity of auxin responses coordinates wood formation. *Nat Commun*, 9(1), 875. <https://doi.org/10.1038/s41467-018-03256-2>
- Cano-Delgado, A., Penfield, S., Smith, C., Catley, M., & Bevan, M. (2003). Reduced cellulose synthesis invokes lignification and defense responses in *Arabidopsis thaliana*. *Plant J*, 34(3), 351-362. <https://doi.org/10.1046/j.1365-313x.2003.01729.x>
- Cao, P. B., Azar, S., SanClemente, H., Mounet, F., Dunand, C., Marque, G., Marque, C., & Teulières, C. (2015). Genome-Wide Analysis of the AP2/ERF Family in *Eucalyptus grandis*: An Intriguing Over-Representation of Stress-Responsive DREB1/CBF Genes. *PLoS One*, 10(4), e0121041. <https://doi.org/10.1371/journal.pone.0121041>
- Cao, P. B., Ployet, R., Nguyen, C., Dupas, A., Ladouce, N., Martinez, Y., Grima-Pettenati, J., Marque, C., Mounet, F., & Teulieres, C. (2020). Wood Architecture and Composition Are Deeply Remodeled in Frost Sensitive *Eucalyptus* Overexpressing CBF/DREB1 Transcription Factors. *Int J Mol Sci*, 21(8). <https://doi.org/10.3390/ijms21083019>
- Carocha, V., Soler, M., Hefer, C., Cassan-Wang, H., Fevereiro, P., Myburg, A. A., Paiva, J. A., & Grima-Pettenati, J. (2015). Genome-wide analysis of the lignin toolbox of *Eucalyptus grandis*. *New Phytol*, 206(4), 1297-1313. <https://doi.org/10.1111/nph.13313>

- Cassan-Wang, H., Goue, N., Saidi, M. N., Legay, S., Sivadon, P., Goffner, D., & Grima-Pettenati, J. (2013). Identification of novel transcription factors regulating secondary cell wall formation in Arabidopsis. *Front Plant Sci*, 4, 189. <https://doi.org/10.3389/fpls.2013.00189>
- Cassan-Wang, H., Soler, M., Yu, H., Camargo, E. L., Carocha, V., Ladouce, N., Savelli, B., Paiva, J. A., Leple, J. C., & Grima-Pettenati, J. (2012). Reference genes for high-throughput quantitative reverse transcription-PCR analysis of gene expression in organs and tissues of Eucalyptus grown in various environmental conditions. *Plant Cell Physiol*, 53(12), 2101-2116. <https://doi.org/10.1093/pcp/pcs152>
- Cassan-Wang, H., Soler, M., Yu, H., Plasencia, A., San Clemente, H., Ladouce, N., Hefer, C., Myburg, A. A., Paiva, J., & Grima-Pettenati, J. (2013). *New transcription factors regulating secondary xylem differentiation in Eucalyptus. (Oral presentation)* Plant Vascular Biology, 26-30 July 2013, Helsinki, Finland.
- Chaffey, N., Cholewa, E., Regan, S., & Sundberg, B. (2002). Secondary xylem development in Arabidopsis: a model for wood formation. *Physiol Plant*, 114(4), 594-600. <https://doi.org/10.1034/j.1399-3054.2002.1140413.x>
- Chai, G., Qi, G., Wang, D., Zhuang, Y., Xu, H., Bai, Z., Bai, M. Y., Hu, R., Wang, Z. Y., Zhou, G., & Kong, Y. (2022). The CCCH zinc finger protein C3H15 negatively regulates cell elongation by inhibiting brassinosteroid signaling. *Plant Physiol*, 189(1), 285-300. <https://doi.org/10.1093/plphys/kiac046>
- Chanoca, A., de Vries, L., & Boerjan, W. (2019). Lignin Engineering in Forest Trees. *Front Plant Sci*, 10, 912. <https://doi.org/10.3389/fpls.2019.00912>
- Chen, Y., Tong, S. F., Jiang, Y. Z., Ai, F. D., Feng, Y. L., Zhang, J. L., Gong, J., Qin, J. J., Zhang, Y. Y., Zhu, Y. Y., Liu, J. Q., & Ma, T. (2021). Transcriptional landscape of highly lignified poplar stems at single-cell resolution. *Genome Biology*, 22(1). <https://doi.org/ARTN> 319
- 10.1186/s13059-021-02537-2
- Coleman, H. D., Brunner, A. M., & Tsai, C. J. (2021). Synergies and Entanglement in Secondary Cell Wall Development and Abiotic Stress Response in Trees. *Front Plant Sci*, 12, 639769. <https://doi.org/10.3389/fpls.2021.639769>
- Conde, D., Triozzi, P. M., Balmant, K. M., Doty, A. L., Miranda, M., Boullosa, A., Schmidt, H. W., Pereira, W. J., Dervinis, C., & Kirst, M. (2021). A robust method of nuclei isolation for single-cell RNA sequencing of solid tissues from the plant genus Populus. *PLoS One*, 16(5), e0251149. <https://doi.org/10.1371/journal.pone.0251149>
- Couzigou, J. M., Laressergues, D., Becard, G., & Combier, J. P. (2015). miRNA-encoded peptides (miPEPs): A new tool to analyze the roles of miRNAs in plant biology. *RNA Biol*, 12(11), 1178-1180. <https://doi.org/10.1080/15476286.2015.1094601>
- Cuello, C., Marchand, P., Laurans, F., Grand-Perret, C., Laine-Prade, V., Pilate, G., & Dejardin, A. (2020). ATR-FTIR Microspectroscopy Brings a Novel Insight Into the Study of Cell Wall Chemistry at the Cellular Level. *Front Plant Sci*, 11, 105. <https://doi.org/10.3389/fpls.2020.00105>
- Dai, Y., Hu, G., Dupas, A., Medina, L., Blandels, N., Clemente, H. S., Ladouce, N., Badawi, M., Hernandez-Raquet, G., Mounet, F., Grima-Pettenati, J., & Cassan-Wang, H. (2020). Implementing the CRISPR/Cas9 Technology in Eucalyptus Hairy Roots Using Wood-Related Genes. *Int J Mol Sci*, 21(10). <https://doi.org/10.3390/ijms21103408>
- De Zio, E., Montagnoli, A., Karady, M., Terzaghi, M., Sferra, G., Antoniadi, I., Scippa, G. S., Ljung, K., Chiatante, D., & Trupiano, D. (2020). Reaction Wood Anatomical Traits and Hormonal Profiles in Poplar Bent Stem and Root. *Front Plant Sci*, 11, 590985. <https://doi.org/10.3389/fpls.2020.590985>
- Denance, N., Ranocha, P., Oria, N., Barlet, X., Riviere, M. P., Yadeta, K. A., Hoffmann, L., Perreau, F., Clement, G., Maia-Grondard, A., van den Berg, G. C., Savelli, B., Fournier, S., Aubert, Y., Pelletier, S., Thomma, B. P., Molina, A., Jouanin, L., Marco, Y., & Goffner, D. (2013).



- Arabidopsis *wat1* (walls are thin1)-mediated resistance to the bacterial vascular pathogen, *Ralstonia solanacearum*, is accompanied by cross-regulation of salicylic acid and tryptophan metabolism. *Plant J*, 73(2), 225-239. <https://doi.org/10.1111/tpj.12027>
- Denyer, T., Ma, X., Klesen, S., Scacchi, E., Nieselt, K., & Timmermans, M. C. P. (2019). Spatiotemporal Developmental Trajectories in the Arabidopsis Root Revealed Using High-Throughput Single-Cell RNA Sequencing. *Dev Cell*, 48(6), 840-852 e845. <https://doi.org/10.1016/j.devcel.2019.02.022>
  - Desaint, H., Aoun, N., Deslandes, L., Vaillau, F., Roux, F., & Berthome, R. (2021). Fight hard or die trying: when plants face pathogens under heat stress. *New Phytol*, 229(2), 712-734. <https://doi.org/10.1111/nph.16965>
  - Fahlgren, N., Montgomery, T. A., Howell, M. D., Allen, E., Dvorak, S. K., Alexander, A. L., & Carrington, J. C. (2006). Regulation of AUXIN RESPONSE FACTOR3 by TAS3 ta-siRNA affects developmental timing and patterning in Arabidopsis. *Curr Biol*, 16(9), 939-944. [https://doi.org/S0960-9822\(06\)01397-2](https://doi.org/S0960-9822(06)01397-2) [pii]
  - 10.1016/j.cub.2006.03.065
  - Farrell, J. A., Wang, Y., Riesenfeld, S. J., Shekhar, K., Regev, A., & Schier, A. F. (2018). Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science*, 360(6392). <https://doi.org/10.1126/science.aar3131>
  - Fischer, U., Kucukoglu, M., Helariutta, Y., & Bhalerao, R. P. (2019). The Dynamics of Cambial Stem Cell Activity. *Annu Rev Plant Biol*, 70, 293-319. <https://doi.org/10.1146/annurev-arplant-050718-100402>
  - Fukuda, H., & 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 Physiol*, 65(1), 57-60. <https://doi.org/10.1104/pp.65.1.57>
  - Garrett, K. A., Dendy, S. P., Frank, E. E., Rouse, M. N., & Travers, S. E. (2006). Climate change effects on plant disease: genomes to ecosystems. *Annu Rev Phytopathol*, 44, 489-509. <https://doi.org/10.1146/annurev.phyto.44.070505.143420>
  - Gierlinger, N. (2018). New insights into plant cell walls by vibrational microspectroscopy. *Appl Spectrosc Rev*, 53(7), 517-551. <https://doi.org/10.1080/05704928.2017.1363052>
  - Gomes, G. L. B., & Scortecci, K. C. (2021). Auxin and its role in plant development: structure, signalling, regulation and response mechanisms. *Plant Biol (Stuttg)*, 23(6), 894-904. <https://doi.org/10.1111/plb.13303>
  - Gutierrez, L., Bussell, J. D., Pacurar, D. I., Schwambach, J., Pacurar, M., & Bellini, C. (2009). Phenotypic plasticity of adventitious rooting in Arabidopsis is controlled by complex regulation of AUXIN RESPONSE FACTOR transcripts and microRNA abundance. *Plant Cell*, 21(10), 3119-3132. <https://doi.org/tpc.108.064758> [pii]
  - 10.1105/tpc.108.064758
  - Gutierrez, L., Mongelard, G., Flokova, K., Pacurar, D. I., Novak, O., Staswick, P., Kowalczyk, M., Pacurar, M., Demailly, H., Geiss, G., & Bellini, C. (2012). Auxin controls Arabidopsis adventitious root initiation by regulating jasmonic Acid homeostasis. *Plant Cell*, 24(6), 2515-2527. <https://doi.org/tpc.112.099119> [pii]
  - 10.1105/tpc.112.099119
  - HADJ BACHIR, I. (2022). *Etude de l'interaction entre la voie de signalisation de la réponse au froid impliquant les CBF et la régulation de la formation du bois chez l'Eucalyptus* Université Toulouse III - Paul Sabatier].
  - Hadj Bachir, I., Ployet, R., Teulières, C., Cassan-Wang, H., Mounet, F., & Grima-Pettenati, J. (2022). Chapter Nine - Regulation of secondary cell wall lignification by abiotic and biotic constraints. In R. Sibout (Ed.), *Advances in Botanical Research* (Vol. 104, pp. 363-392). Academic Press. <https://doi.org/https://doi.org/10.1016/bs.abr.2022.03.008>

- Hamann, T. (2012). Plant cell wall integrity maintenance as an essential component of biotic stress response mechanisms. *Front Plant Sci*, 3, 77. <https://doi.org/10.3389/fpls.2012.00077>
- Hendelman, A., Buxdorf, K., Stav, R., Kravchik, M., & Arazi, T. (2012). Inhibition of lamina outgrowth following *Solanum lycopersicum* AUXIN RESPONSE FACTOR 10 (SIARF10) derepression. *Plant Mol Biol*, 78(6), 561-576. <https://doi.org/10.1007/s11103-012-9883-4>
- Hernandez-Blanco, C., Feng, D. X., Hu, J., Sanchez-Vallet, A., Deslandes, L., Llorente, F., Berrocal-Lobo, M., Keller, H., Barlet, X., Sanchez-Rodriguez, C., Anderson, L. K., Somerville, S., Marco, Y., & Molina, A. (2007). Impairment of cellulose synthases required for Arabidopsis secondary cell wall formation enhances disease resistance. *Plant Cell*, 19(3), 890-903. <https://doi.org/10.1105/tpc.106.048058>
- Hong, Y. (2014). *Caractérisation des familles de facteurs de transcription ARF et Aux/IAA chez l'eucalyptus, rôles dans la formation du bois* Université Toulouse III]. <http://www.theses.fr/2014TOU30119>.
- Huang, B., Hu, G., Wang, K., Frasse, P., Maza, E., Djari, A., Deng, W., Pirrello, J., Burlat, V., Pons, C., Granell, A., Li, Z., van der Rest, B., & Bouzayen, M. (2021). Interaction of two MADS-box genes leads to growth phenotype divergence of all-flesh type of tomatoes. *Nat Commun*, 12(1), 6892. <https://doi.org/10.1038/s41467-021-27117-7>
- Huckelhoven, R. (2007). Cell wall-associated mechanisms of disease resistance and susceptibility. *Annu Rev Phytopathol*, 45, 101-127. <https://doi.org/10.1146/annurev.phyto.45.062806.094325>
- Itoh, J., Hibara, K., Sato, Y., & Nagato, Y. (2008). Developmental role and auxin responsiveness of Class III homeodomain leucine zipper gene family members in rice. *Plant Physiol*, 147(4), 1960-1975. <https://doi.org/10.1104/pp.108.118679>
- Jean-Baptiste, K., McFaline-Figueroa, J. L., Alexandre, C. M., Dorrity, M. W., Saunders, L., Bubb, K. L., Trapnell, C., Fields, S., Queitsch, C., & Cuperus, J. T. (2019). Dynamics of Gene Expression in Single Root Cells of *Arabidopsis thaliana*. *Plant Cell*, 31(5), 993-1011. <https://doi.org/10.1105/tpc.18.00785>
- Johnsson, C., Jin, X., Xue, W., Dubreuil, C., Lezhneva, L., & 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. *Physiol Plant*. <https://doi.org/10.1111/ppl.12766>
- Jones, B., Frasse, P., Olmos, E., Zegzouti, H., Li, Z. G., Latche, A., Pech, J. C., & Bouzayen, M. (2002). Down-regulation of DR12, an auxin-response-factor homolog, in the tomato results in a pleiotropic phenotype including dark green and blotchy ripening fruit. *Plant J*, 32(4), 603-613. <https://doi.org/10.1046/j.1365-313x.2002.01450.x>
- Karannagoda, N., Spokevicius, A., Hussey, S., Cassan-Wang, H., Grima-Pettenati, J., & Bossinger, G. (2022). *Eucalyptus grandis* AUX/INDOLE-3-ACETIC ACID 13 (EgrIAA13) is a novel transcriptional regulator of xylogenesis. *Plant Mol Biol*, 109(1-2), 51-65. <https://doi.org/10.1007/s11103-022-01255-y>
- Kim, J.-Y., Symeonidi, E., Pang, T. Y., Denyer, T., Weidauer, D., Bezrutczyk, M., Miras, M., Zöllner, N., Hartwig, T., Wudick, M. M., Lercher, M., Chen, L.-Q., Timmermans, M. C. P., & Frommer, W. B. (2021). Distinct identities of leaf phloem cells revealed by single cell transcriptomics. *The Plant Cell*, 33(3), 511-530. <https://doi.org/10.1093/plcell/koaa060>
- Kim, W. C., Kim, J. Y., Ko, J. H., Kang, H., Kim, J., & Han, K. H. (2014). AtC3H14, a plant-specific tandem CCCH zinc-finger protein, binds to its target mRNAs in a sequence-specific manner and affects cell elongation in *Arabidopsis thaliana*. *Plant J*, 80(5), 772-784. <https://doi.org/10.1111/tpj.12667>
- Lata, C., & Prasad, M. (2011). Role of DREBs in regulation of abiotic stress responses in plants. *Journal of Experimental Botany*, 62(14), 4731-4748. <https://doi.org/10.1093/jxb/err210>

- Lauressergues, D., Couzigou, J. M., Clemente, H. S., Martinez, Y., Dunand, C., Becard, G., & Combier, J. P. (2015). Primary transcripts of microRNAs encode regulatory peptides. *Nature*, 520(7545), 90-93. <https://doi.org/10.1038/nature14346>
- Lauressergues, D., Ormancey, M., Guillotin, B., San Clemente, H., Camborde, L., Duboe, C., Tourneur, S., Charpentier, P., Barozet, A., Jauneau, A., Le Ru, A., Thuleau, P., Gervais, V., Plaza, S., & Combier, J. P. (2022). Characterization of plant microRNA-encoded peptides (miPEPs) reveals molecular mechanisms from the translation to activity and specificity. *Cell Rep*, 38(6), 110339. <https://doi.org/10.1016/j.celrep.2022.110339>
- Lee, C., Zhong, R., & Ye, Z. H. (2012). Arabidopsis family GT43 members are xylan xylosyltransferases required for the elongation of the xylan backbone. *Plant Cell Physiol*, 53(1), 135-143. <https://doi.org/10.1093/pcp/pcr158>
- Li, A., Zhou, M., Wei, D., Chen, H., You, C., & Lin, J. (2017). Transcriptome Profiling Reveals the Negative Regulation of Multiple Plant Hormone Signaling Pathways Elicited by Overexpression of C-Repeat Binding Factors. *Front Plant Sci*, 8, 1647. <https://doi.org/10.3389/fpls.2017.01647>
- Li, H., Dai, X., Huang, X., Xu, M., Wang, Q., Yan, X., Sederoff, R. R., & Li, Q. (2021). Single-cell RNA sequencing reveals a high-resolution cell atlas of xylem in Populus. *Journal of Integrative Plant Biology*, 63(11), 1906-1921. <https://doi.org/10.1111/jipb.13159>
- Li, K., Liu, Z., Xing, L., Wei, Y., Mao, J., Meng, Y., Bao, L., Han, M., Zhao, C., & Zhang, D. (2019). miRNAs associated with auxin signaling, stress response, and cellular activities mediate adventitious root formation in apple rootstocks. *Plant Physiol Biochem*, 139, 66-81. <https://doi.org/10.1016/j.plaphy.2019.03.006>
- Li, Q., Yu, H., Cao, P. B., Fawal, N., Mathe, C., Azar, S., Cassan-Wang, H., Myburg, A. A., Grima-Pettenati, J., Marque, C., Teulieres, C., & Dunand, C. (2015). Explosive tandem and segmental duplications of multigenic families in Eucalyptus grandis. *Genome Biol Evol*, 7(4), 1068-1081. <https://doi.org/10.1093/gbe/evv048>
- Lin, Z., Long, J. M., Yin, Q., Wang, B., Li, H. L., Luo, J. Z., Wang, H. C., & Wu, A. M. (2019). Identification of novel lncRNAs in Eucalyptus grandis. *Industrial Crops and Products*, 129, 309-317. <https://doi.org/10.1016/j.indcrop.2018.12.016>
- Liu, P. P., Montgomery, T. A., Fahlgren, N., Kasschau, K. D., Nonogaki, H., & Carrington, J. C. (2007). Repression of AUXIN RESPONSE FACTOR10 by microRNA160 is critical for seed germination and post-germination stages. *Plant J*, 52(1), 133-146. <https://doi.org/TPJ3218> [pii]
- 10.1111/j.1365-313X.2007.03218.x
- Liu, Z., Duguay, J., Ma, F., Wang, T. W., Tshin, R., Hopkins, M. T., McNamara, L., & Thompson, J. E. (2008). Modulation of eIF5A1 expression alters xylem abundance in Arabidopsis thaliana. *J Exp Bot*, 59(4), 939-950. <https://doi.org/10.1093/jxb/ern017>
- Lopez-Anido, C. B., Vatén, A., Smoot, N. K., Sharma, N., Guo, V., Gong, Y., Anleu Gil, M. X., Weimer, A. K., & Bergmann, D. C. (2021). Single-cell resolution of lineage trajectories in the Arabidopsis stomatal lineage and developing leaf. *Developmental cell*, 56(7), 1043-1055.e1044. <https://doi.org/https://doi.org/10.1016/j.devcel.2021.03.014>
- Lu, S., Sun, Y. H., & Chiang, V. L. (2008). Stress-responsive microRNAs in Populus. *Plant J*, 55(1), 131-151. <https://doi.org/TPJ3497> [pii]
- 10.1111/j.1365-313X.2008.03497.x
- Mallory, A. C., Bartel, D. P., & Bartel, B. (2005). MicroRNA-directed regulation of Arabidopsis AUXIN RESPONSE FACTOR17 is essential for proper development and modulates expression of early auxin response genes. *Plant Cell*, 17(5), 1360-1375. <https://doi.org/tpc.105.031716> [pii]
- 10.1105/tpc.105.031716
- Marin, E., Jouannet, V., Herz, A., Lokerse, A. S., Weijers, D., Vaucheret, H., Nussaume, L., Crespi, M. D., & Maizel, A. (2010). miR390, Arabidopsis TAS3 tasiRNAs, and their AUXIN

RESPONSE FACTOR targets define an autoregulatory network quantitatively regulating lateral root growth. *Plant Cell*, 22(4), 1104-1117. <https://doi.org/tpc.109.072553> [pii]

- 10.1105/tpc.109.072553
- Menna, A., Dora, S., Sancho-Andres, G., Kashyap, A., Meena, M. K., Sklodowski, K., Gasperini, D., Coll, N. S., & Sanchez-Rodriguez, C. (2021). A primary cell wall cellulose-dependent defense mechanism against vascular pathogens revealed by time-resolved dual transcriptomics. *BMC Biol*, 19(1), 161. <https://doi.org/10.1186/s12915-021-01100-6>
- Mizoi, J., Shinozaki, K., & Yamaguchi-Shinozaki, K. (2012). AP2/ERF family transcription factors in plant abiotic stress responses. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, 1819(2), 86-96. <https://doi.org/https://doi.org/10.1016/j.bbagr.2011.08.004>
- Molina, A., Miedes, E., Bacete, L., Rodriguez, T., Melida, H., Denance, N., Sanchez-Vallet, A., Riviere, M. P., Lopez, G., Freydier, A., Barlet, X., Pattathil, S., Hahn, M., & Goffner, D. (2021). Arabidopsis cell wall composition determines disease resistance specificity and fitness. *Proc Natl Acad Sci U S A*, 118(5). <https://doi.org/10.1073/pnas.2010243118>
- Morey, P. R., & Dahl, B. E. (1975). Histological and Morphological Effects of Auxin Transport Inhibitors on Honey Mesquite. *Botanical Gazette*, 136(3), 274-280. <https://doi.org/Doi10.1086/336814>
- Mortimer, J. C., Miles, G. P., Brown, D. M., Zhang, Z., Segura, M. P., Weimar, T., Yu, X., Seffen, K. A., Stephens, E., Turner, S. R., & Dupree, P. (2010). Absence of branches from xylan in Arabidopsis gux mutants reveals potential for simplification of lignocellulosic biomass. *Proc Natl Acad Sci U S A*, 107(40), 17409-17414. <https://doi.org/10.1073/pnas.1005456107>
- Muller, C. J., Valdes, A. E., Wang, G., Ramachandran, P., Beste, L., Uddenberg, D., & Carlsbecker, A. (2016). PHABULOSA Mediates an Auxin Signaling Loop to Regulate Vascular Patterning in Arabidopsis. *Plant Physiol*, 170(2), 956-970. <https://doi.org/10.1104/pp.15.01204>
- Myburg, A. A., Grattapaglia, D., Tuskan, G. A., Hellsten, U., Hayes, R. D., Grimwood, J., Jenkins, J., Lindquist, E., Tice, H., Bauer, D., Goodstein, D. M., Dubchak, I., Poliakov, A., Mizrachi, E., Kullam, A. R., Hussey, S. G., Pinard, D., van der Merwe, K., Singh, P., . . . Schmutz, J. (2014). The genome of Eucalyptus grandis. *Nature*, 510(7505), 356-362. <https://doi.org/10.1038/nature13308>
- Nakano, T., Suzuki, K., Fujimura, T., & Shinshi, H. (2006). Genome-wide analysis of the ERF gene family in Arabidopsis and rice. *Plant Physiology*, 140(2), 411-432. <https://doi.org/10.1104/pp.105.073783>
- Navarro, M., Ayax, C., Martinez, Y., Laur, J., El Kayal, W., Marque, C., & Teulieres, C. (2011). Two EguCBF1 genes overexpressed in Eucalyptus display a different impact on stress tolerance and plant development. *Plant Biotechnology Journal*, 9(1), 50-63. <https://doi.org/https://doi.org/10.1111/j.1467-7652.2010.00530.x>
- Neumann, M., Xu, X., Smaczniak, C., Schumacher, J., Yan, W., Blüthgen, N., Greb, T., Jönsson, H., Traas, J., Kaufmann, K., & Muino, J. M. (2022). A 3D gene expression atlas of the floral meristem based on spatial reconstruction of single nucleus RNA sequencing data. *Nat Commun*, 13(1). <https://doi.org/10.1038/s41467-022-30177-y>
- Nguyen, H. C., Cao, P. B., San Clemente, H., Ployet, R., Mounet, F., Ladouce, N., Harvengt, L., Marque, C., & Teulieres, C. (2017). Special trends in CBF and DREB2 groups in Eucalyptus gunnii vs Eucalyptus grandis suggest that CBF are master players in the trade-off between growth and stress resistance. *Physiologia Plantarum*, 159(4), 445-467. <https://doi.org/https://doi.org/10.1111/ppl.12529>
- Nieminen, K. M., Kauppinen, L., & Helariutta, Y. (2004). A weed for wood? Arabidopsis as a genetic model for xylem development. *Plant Physiol*, 135(2), 653-659. <https://doi.org/10.1104/pp.104.040212>

- Nobori, T., Oliva, M., Lister, R., & Ecker, J. R. (2023). Multiplexed single-cell 3D spatial gene expression analysis in plant tissue using PHYTOmap. *Nat Plants*.  
<https://doi.org/10.1038/s41477-023-01439-4>
- Ohashi-Ito, K., & Fukuda, H. (2010). Transcriptional regulation of vascular cell fates. *Curr Opin Plant Biol*, 13(6), 670-676. <https://doi.org/10.1016/j.pbi.2010.08.011>
- Ouellet, F., Overvoorde, P. J., & Theologis, A. (2001). IAA17/AXR3: biochemical insight into an auxin mutant phenotype. *Plant Cell*, 13(4), 829-841. <https://doi.org/10.1105/tpc.13.4.829>
- Overvoorde, P. J., Okushima, Y., Alonso, J. M., Chan, A., Chang, C., Ecker, J. R., Hughes, B., Liu, A., Onodera, C., Quach, H., Smith, A., Yu, G., & Theologis, A. (2005). Functional genomic analysis of the AUXIN/INDOLE-3-ACETIC ACID gene family members in Arabidopsis thaliana. *Plant Cell*, 17(12), 3282-3300. <https://doi.org/tpc.105.036723> [pii]
- 10.1105/tpc.105.036723
- Pesquet, E., Ranocha, P., Legay, S., Digonnet, C., Barbier, O., Pichon, M., & Goffner, D. (2005). Novel markers of xylogenesis in zinnia are differentially regulated by auxin and cytokinin. *Plant Physiol*, 139(4), 1821-1839. <https://doi.org/10.1104/pp.105.064337>
- Pesquet, E., Wagner, A., & Grabber, J. H. (2019). Cell culture systems: invaluable tools to investigate lignin formation and cell wall properties. *Current Opinion in Biotechnology*, 56, 215-222. <https://doi.org/10.1016/j.copbio.2019.02.001>
- Pogorelko, G., Lionetti, V., Bellincampi, D., & Zabolina, O. (2013). Cell wall integrity: targeted post-synthetic modifications to reveal its role in plant growth and defense against pathogens. *Plant Signal Behav*, 8(9). <https://doi.org/10.4161/psb.25435>
- Polle, A., Chen, S. L., Eckert, C., & Harfouche, A. (2018). Engineering Drought Resistance in Forest Trees. *Front Plant Sci*, 9, 1875. <https://doi.org/10.3389/fpls.2018.01875>
- Popko, J., Hansch, R., Mendel, R. R., Polle, A., & Teichmann, T. (2010). The role of abscisic acid and auxin in the response of poplar to abiotic stress. *Plant Biol (Stuttg)*, 12(2), 242-258. <https://doi.org/10.1111/j.1438-8677.2009.00305.x>
- Reed, J. W. (2001). Roles and activities of Aux/IAA proteins in Arabidopsis. *Trends Plant Sci*, 6(9), 420-425. <http://www.ncbi.nlm.nih.gov/pubmed/11544131>
- Robischon, M., Du, J., Miura, E., & Groover, A. (2011). The Populus class III HD ZIP, popREVOLUTA, influences cambium initiation and patterning of woody stems. *Plant Physiol*, 155(3), 1214-1225. <https://doi.org/10.1104/pp.110.167007>
- Rodriguez-Villalon, A., & Brady, S. M. (2019). Single cell RNA sequencing and its promise in reconstructing plant vascular cell lineages. *Curr Opin Plant Biol*, 48, 47-56. <https://doi.org/10.1016/j.pbi.2019.04.002>
- Ryu, K. H., Huang, L., Kang, H. M., & Schiefelbein, J. (2019). Single-Cell RNA Sequencing Resolves Molecular Relationships Among Individual Plant Cells *Plant Physiology*, 179(4), 1444-1456. <https://doi.org/10.1104/pp.18.01482>
- Sakuma, Y., Maruyama, K., Osakabe, Y., Qin, F., Seki, M., Shinozaki, K., & Yamaguchi-Shinozaki, K. (2006). Functional analysis of an Arabidopsis transcription factor, DREB2A, involved in drought-responsive gene expression. *Plant Cell*, 18(5), 1292-1309. <https://doi.org/10.1105/tpc.105.035881>
- Shani, E., Salehin, M., Zhang, Y., Sanchez, S. E., Doherty, C., Wang, R., Mangado, C. C., Song, L., Tal, I., Pisanty, O., Ecker, J. R., Kay, S. A., Pruneda-Paz, J., & Estelle, M. (2017). Plant Stress Tolerance Requires Auxin-Sensitive Aux/IAA Transcriptional Repressors. *Curr Biol*, 27(3), 437-444. <https://doi.org/10.1016/j.cub.2016.12.016>
- Shulse, C. N., Cole, B. J., Ciobanu, D., Lin, J., Yoshinaga, Y., Gouran, M., Turco, G. M., Zhu, Y., O'Malley, R. C., Brady, S. M., & Dickel, D. E. (2019). High-Throughput Single-Cell Transcriptome Profiling of Plant Cell Types. *Cell Rep*, 27(7), 2241-2247 e2244. <https://doi.org/10.1016/j.celrep.2019.04.054>



- Smetana, O., Makila, R., Lyu, M., Amiryousefi, A., Sanchez Rodriguez, F., Wu, M. F., Sole-Gil, A., Leal Gavarron, M., Siligato, R., Miyashima, S., Roszak, P., Blomster, T., Reed, J. W., Broholm, S., & Mahonen, A. P. (2019). High levels of auxin signalling define the stem-cell organizer of the vascular cambium. *Nature*, *565*(7740), 485-489. <https://doi.org/10.1038/s41586-018-0837-0>
- Soler, M., Camargo, E. L., Carocha, V., Cassan-Wang, H., San Clemente, H., Savelli, B., Hefer, C. A., Paiva, J. A., Myburg, A. A., & Grima-Pettenati, J. (2015). The Eucalyptus grandis R2R3-MYB transcription factor family: evidence for woody growth-related evolution and function. *New Phytol*, *206*(4), 1364-1377. <https://doi.org/10.1111/nph.13039>
- Soler, M., Plasencia, A., Larbat, R., Pouzet, C., Jauneau, A., Rivas, S., Pesquet, E., Lapierre, C., Truchet, I., & 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 Phytol*, *213*(1), 287-299. <https://doi.org/10.1111/nph.14129>
- Strabala, T. J., & Macmillan, C. P. (2013). The Arabidopsis wood model-the case for the inflorescence stem. *Plant Sci*, *210*, 193-205. <https://doi.org/10.1016/j.plantsci.2013.05.007>
- Sundell, D., Street, N. R., Kumar, M., Mellerowicz, E. J., Kucukoglu, M., Johnsson, C., Kumar, V., Mannapperuma, C., Delhomme, N., Nilsson, O., Tuominen, H., Pesquet, E., Fischer, U., Niittyla, T., Sundberg, B., & Hvidsten, T. R. (2017). AspWood: High-Spatial-Resolution Transcriptome Profiles Reveal Uncharacterized Modularity of Wood Formation in Populus tremula. *Plant Cell*, *29*(7), 1585-1604. <https://doi.org/10.1105/tpc.17.00153>
- Tang, F., Wei, H., Zhao, S., Wang, L., Zheng, H., & Lu, M. (2016). Identification of microRNAs Involved in Regeneration of the Secondary Vascular System in Populus tomentosa Carr. *Front Plant Sci*, *7*, 724. <https://doi.org/10.3389/fpls.2016.00724>
- Tiwari, S. B., Hagen, G., & Guilfoyle, T. (2003). The roles of auxin response factor domains in auxin-responsive transcription. *Plant Cell*, *15*(2), 533-543. [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=12566590](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12566590)
- Tung, C.-C., Kuo, S.-C., Yang, C.-L., Yu, J.-H., Huang, C.-E., Liou, P.-C., Sun, Y.-H., Shuai, P., Su, J.-C., Ku, C., & Lin, Y.-C. J. (2023). Single-cell transcriptomics unveils xylem cell development and evolution. *Genome Biology*, *24*(1), 3. <https://doi.org/10.1186/s13059-022-02845-1>
- Turco, G. M., Rodriguez-Medina, J., Siebert, S., Han, D., Valderrama-Gomez, M. A., Vahldick, H., Shulse, C. N., Cole, B. J., Juliano, C. E., Dickel, D. E., Savageau, M. A., & Brady, S. M. (2019). Molecular Mechanisms Driving Switch Behavior in Xylem Cell Differentiation. *Cell Rep*, *28*(2), 342-351 e344. <https://doi.org/10.1016/j.celrep.2019.06.041>
- Ulmasov, T., Hagen, G., & Guilfoyle, T. J. (1999). Activation and repression of transcription by auxin-response factors. *Proc Natl Acad Sci U S A*, *96*(10), 5844-5849. [http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=10318972](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10318972)
- Vogel, J., & Somerville, S. (2000). Isolation and characterization of powdery mildew-resistant Arabidopsis mutants. *Proc Natl Acad Sci U S A*, *97*(4), 1897-1902. <https://doi.org/10.1073/pnas.030531997>
- Vogel, J. P., Raab, T. K., Schiff, C., & Somerville, S. C. (2002). PMR6, a pectate lyase-like gene required for powdery mildew susceptibility in Arabidopsis. *Plant Cell*, *14*(9), 2095-2106. <https://doi.org/10.1105/tpc.003509>
- Vogel, J. P., Raab, T. K., Somerville, C. R., & Somerville, S. C. (2004). Mutations in PMR5 result in powdery mildew resistance and altered cell wall composition. *Plant J*, *40*(6), 968-978. <https://doi.org/10.1111/j.1365-313X.2004.02264.x>
- Vogler, H., & Kuhlemeier, C. (2003). Simple hormones but complex signalling. *Curr Opin Plant Biol*, *6*(1), 51-56. [https://doi.org/10.1016/s1369-5266\(02\)00013-4](https://doi.org/10.1016/s1369-5266(02)00013-4)
- Wang, H., Jones, B., Li, Z., Frasse, P., Delalande, C., Regad, F., Chaabouni, S., Latche, A., Pech, J. C., & Bouzayen, M. (2005). The tomato Aux/IAA transcription factor IAA9 is involved in fruit



- development and leaf morphogenesis. *Plant Cell*, 17(10), 2676-2692. <https://doi.org/10.1105/tpc.105.033415>
- Wang, H., Schauer, N., Usadel, B., Frasse, P., Zouine, M., Hernould, M., Latche, A., Pech, J. C., Fernie, A. R., & Bouzayen, M. (2009). Regulatory features underlying pollination-dependent and -independent tomato fruit set revealed by transcript and primary metabolite profiling. *Plant Cell*, 21(5), 1428-1452. <https://doi.org/10.1105/tpc.108.060830>
  - Wang, H., Soler, M., Yu, H., Camargo, E., San Clemente, H., Savelli, B., Ladouce, N., Paiva, J. A. P., & Grima-Pettenati, J. (2011). Master regulators of wood formation in Eucalyptus. *BMC Proceedings* 5(Suppl 7), P110.
  - Wang, J. W., Wang, L. J., Mao, Y. B., Cai, W. J., Xue, H. W., & Chen, X. Y. (2005). Control of root cap formation by MicroRNA-targeted auxin response factors in Arabidopsis. *Plant Cell*, 17(8), 2204-2216. <https://doi.org/tpc.105.033076> [pii]
  - 10.1105/tpc.105.033076
  - Wendrich, J. R., Yang, B., Vandamme, N., Verstaen, K., Smet, W., Van de Velde, C., Minne, M., Wybouw, B., Mor, E., Arents, H. E., Nolf, J., Van Duyse, J., Van Isterdael, G., Maere, S., Saeys, Y., & De Rybel, B. (2020). Vascular transcription factors guide plant epidermal responses to limiting phosphate conditions. *Science*, 370(6518), eaay4970. <https://doi.org/doi:10.1126/science.aay4970>
  - Worley, C. K., Zenser, N., Ramos, J., Rouse, D., Leyser, O., Theologis, A., & Callis, J. (2000). Degradation of Aux/IAA proteins is essential for normal auxin signalling. *Plant J*, 21(6), 553-562. <http://www.ncbi.nlm.nih.gov/pubmed/10758506>
  - Wu, M. F., Tian, Q., & Reed, J. W. (2006). Arabidopsis microRNA167 controls patterns of ARF6 and ARF8 expression, and regulates both female and male reproduction. *Development*, 133(21), 4211-4218. <https://doi.org/dev.02602> [pii]
  - 10.1242/dev.02602
  - Xu, C., Shen, Y., He, F., Fu, X., Yu, H., Lu, W., Li, Y., Li, C., Fan, D., Wang, H. C., & Luo, K. (2019). Auxin-mediated Aux/IAA-ARF-HB signaling cascade regulates secondary xylem development in Populus. *New Phytol*, 222(2), 752-767. <https://doi.org/10.1111/nph.15658>
  - Xu, X., Crow, M., Rice, B. R., Li, F., Harris, B., Liu, L., Demesa-Arevalo, E., Lu, Z., Wang, L., & Fox, N. (2021). Single-cell RNA sequencing of developing maize ears facilitates functional analysis and trait candidate gene discovery. *Developmental cell*, 56(4), 557-568. e556. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7904613/pdf/nihms-1659061.pdf>
  - Yamaguchi-Shinozaki, K., & Shinozaki, K. (1994). A novel cis-acting element in an Arabidopsis gene is involved in responsiveness to drought, low-temperature, or high-salt stress. *Plant Cell*, 6(2), 251-264. <https://doi.org/10.1105/tpc.6.2.251>
  - Ye, Z. H., & Zhong, R. (2015). Molecular control of wood formation in trees. *J Exp Bot*, 66(14), 4119-4131. <https://doi.org/10.1093/jxb/erv081>
  - Yu, H., Liu, M., Zhu, Z., Wu, A., Mounet, F., Pesquet, E., Grima-Pettenati, J., & Cassan-Wang, H. (2022). Overexpression of EgriAA20 from Eucalyptus grandis, a Non-Canonical Aux/IAA Gene, Specifically Decouples Lignification of the Different Cell-Types in Arabidopsis Secondary Xylem. *Int J Mol Sci*, 23(9). <https://doi.org/10.3390/ijms23095068>
  - Yu, H., Soler, M., Mila, I., San Clemente, H., Savelli, B., Dunand, C., Paiva, J. A., Myburg, A. A., Bouzayen, M., Grima-Pettenati, J., & Cassan-Wang, H. (2014). Genome-wide characterization and expression profiling of the AUXIN RESPONSE FACTOR (ARF) gene family in Eucalyptus grandis. *PLoS One*, 9(9), e108906. <https://doi.org/10.1371/journal.pone.0108906>
  - Yu, H., Soler, M., San Clemente, H., Mila, I., Paiva, J. A., Myburg, A. A., Bouzayen, M., Grima-Pettenati, J., & Cassan-Wang, H. (2015). Comprehensive genome-wide analysis of the Aux/IAA gene family in Eucalyptus: evidence for the role of EgriAA4 in wood formation. *Plant Cell Physiol*, 56(4), 700-714. <https://doi.org/10.1093/pcp/pcu215>

- Zhang, T.-Q., Chen, Y., Liu, Y., Lin, W.-H., & Wang, J.-W. (2021). Single-cell transcriptome atlas and chromatin accessibility landscape reveal differentiation trajectories in the rice root. *Nat Commun*, 12(1), 2053. <https://doi.org/10.1038/s41467-021-22352-4>
- Zhang, T. Q., Chen, Y., & Wang, J. W. (2021). A single-cell analysis of the Arabidopsis vegetative shoot apex. *Dev Cell*, 56(7), 1056-1074 e1058. <https://doi.org/10.1016/j.devcel.2021.02.021>
- Zhang, T. Q., Xu, Z. G., Shang, G. D., & Wang, J. W. (2019). A Single-Cell RNA Sequencing Profiles the Developmental Landscape of Arabidopsis Root. *Mol Plant*, 12(5), 648-660. <https://doi.org/10.1016/j.molp.2019.04.004>
- Zinkgraf, M., Liu, L. J., Groover, A., & Filkov, V. (2017). Identifying gene coexpression networks underlying the dynamic regulation of wood-forming tissues in Populus under diverse environmental conditions. *New Phytologist*, 214(4), 1464-1478. <https://doi.org/10.1111/nph.14492>

## Annexe 10 : Projets financés en tant que co-porteurs du projet



## **Titre du projet**

# **La tomate, un nouveau système pour appréhender les mécanismes moléculaires auxine-dépendants impliqués dans la différenciation du xylème**

## **Porteurs du projet**

HUA CASSAN-WANG (MCF, UMR 5546 LRSV), JULIEN PIRELLO (MCF, UMR 990 GBF)

## **Description du projet**

### Contexte scientifique

L'auxine est connue pour être un facteur clé de la différenciation du xylème secondaire (bois) mais les mécanismes moléculaires qui régissent son action n'ont pas encore été élucidés. Plusieurs études suggèrent cependant que les facteurs de transcription Aux/IAA et ARF, éléments cruciaux de la voie de signalisation de l'auxine, sont impliqués dans ce processus.

L'équipe « Génomique fonctionnelle de l'Eucalyptus, GFE » a réalisé une analyse approfondie de ces deux familles chez l'Eucalyptus dont le génome a été récemment séquencé (Myburg et al, 2014). Des analyses phylogénétiques comparatives couplées à des études d'expression à grande échelle ont permis d'identifier plusieurs Aux/IAAs et ARFs fortement et préférentiellement exprimés dans le xylème d'Eucalyptus (Yu et al, 2014 ; Yu et al, 2015). La caractérisation fonctionnelle de plusieurs de ces gènes candidats a été réalisée par surexpression chez la plante modèle *Arabidopsis*, et ce en raison de la faible efficacité de la transformation génétique de l'*Eucalyptus*. La surexpression de plusieurs gènes candidats provoque des modifications de la structure du xylème secondaire dans les hampes florales et les hypocotyles. Par exemple, la surexpression du gène *EgrIAA9* augmente de façon très nette le nombre de vaisseaux et donc le ratio vaisseaux/fibres xylémiennes. La caractérisation fonctionnelle nécessite d'étudier également la perte de fonction de ce gène mais le mutant d'*Arabidopsis* correspondant ne présente pas de phénotype au niveau vasculaire. En revanche, le mutant *IAA9* de tomate étudié dans l'équipe « Génomique et Biologie des Fruits, GBF » présente une vasculature modifiée au niveau des feuilles et des pédoncules très larges (Wang et al 2005) offrant ainsi une alternative très intéressante à *Arabidopsis* pour étudier le rôle de ce gène dans la différenciation du xylème. Le système vasculaire joue également un rôle important dans le développement du fruit, thématique principale de l'équipe GFE.

### Objectifs et méthodologie

Le projet proposé vise à mieux comprendre les mécanismes moléculaires qui régissent l'action de l'auxine lors de la différenciation du système vasculaire et plus particulièrement du xylème via l'identification et la caractérisation fonctionnelle de nouveaux régulateurs auxine-dépendants. Il permettra également de développer un nouveau système pour l'étude de la croissance secondaire et le développement des tissus vasculaires. En effet, la tomate possède une véritable croissance secondaire avec la présence de xylème secondaire dont la présence est limitée à l'hypocotyle chez *Arabidopsis*. De plus, par rapport aux espèces utilisées comme références pour étudier le xylème secondaire (peuplier et eucalyptus), la tomate offre une large panoplie de ressources pour les approches de génétique inverse et de génomique fonctionnelle.

Il s'agira d'exploiter la collection de mutants de tomates transgéniques affectées dans les *Aux/IAA* et *ARF*, générée par l'équipe GBF. Les orthologues des candidats prometteurs identifiés chez l'*Eucalyptus* comme potentiellement impliqués dans le contrôle de la formation du xylème, seront ciblés et les mutants « perte-de-fonction » correspondants seront caractérisés au niveau phénotypique et histologique afin d'identifier les lignées présentant des altérations dans l'organisation vasculaire ou des perturbations liées aux parois cellulaires du xylème. Ce matériel constituera une source précieuse permettant dans un deuxième temps de caractériser le mode d'action et le rôle spécifique de ces gènes dans la formation du xylème.

Parallèlement à ce criblage, nous envisageons d'étudier plus en profondeur le gène *IAA9*, pour lequel d'une part, la surexpression chez *Arabidopsis* induit une modification du rapport vaisseaux/fibres et d'autre part, le mutant *IAA9* de tomate a déjà été repéré pour son phénotype vasculaire altéré. Ce mutant sera caractérisé au niveau histologique en examinant la structure et l'organisation du xylème, l'épaisseur

des parois et au niveau biochimique (composition des parois). Les modifications engendrées par la perte de fonction seront étudiées au niveau transcriptomique en évaluant au niveau du xylème, les niveaux de transcrits des gènes régulateurs et structuraux connus pour être impliqués dans la différenciation du xylème et la formation de la paroi secondaire, et ce à différents stades de développement de la plante. L'activité transcriptionnelle de la protéine IAA9 sera évaluée grâce au système « single cell ». La recherche de protéines partenaires dans le xylème sera réalisée par le criblage d'une banque double hybride construite à partir de xylème d'Eucalyptus par l'équipe GFE. Les orthologues des protéines partenaires d'IAA9 seront isolées chez la tomate et les interactions confirmées par FRET FLIM.

#### Résultats attendus

Ce projet inter-équipes permettra une meilleure compréhension du mécanisme d'action de l'auxine dans la régulation du xylème (i) via la mise en évidence du rôle et du mode d'action d'IAA9 et fera l'objet d'une publication commune (ii) via l'identification de nouveaux régulateurs dont l'étude pourra être poursuivie à travers une demande de projet (ANR, IDEX...)

Par ailleurs, le projet permettra la mise en place d'une nouvelle espèce, la tomate, plus adaptée qu'Arabidopsis à l'étude du développement des tissus vasculaires, notamment les xylèmes secondaires et pour laquelle il existe de nombreuses ressources génétiques et génomiques disponibles auprès du laboratoire GBF.

#### Complémentarité des partenaires, originalité, faisabilité

Ce projet qui vise à explorer le rôle des médiateurs de réponse à l'auxine dans la formation des tissus vasculaires, s'inscrit dans un champ de recherches très peu exploré et offre donc l'opportunité de générer des résultats très originaux. La faisabilité est excellente grâce aux résultats préliminaires déjà obtenus et aux compétences, ressources et expertises complémentaires des deux équipes. GBF possède une expertise reconnue dans la signalisation auxinique ainsi qu'une collection unique de mutants transgéniques de tomate perturbés dans l'expression des plusieurs gènes *Aux/IAA* et *ARF*. GFE est largement reconnue pour ses travaux sur la régulation transcriptionnelle de la formation du xylème et de la paroi secondaire. Les deux équipes bénéficieront de cette collaboration pour leurs thématiques propres car le système vasculaire permet d'alimenter le fruit, et des modifications du système vasculaire liées à la signalisation auxinique auront très certainement un impact sur le développement de ce dernier. Hormis une collaboration ponctuelle limitée à l'utilisation du système « single cell » maîtrisé par GBF, il s'agit d'une nouvelle collaboration pour laquelle aucun financement n'est acquis et pour laquelle un appui de la FRAIB permettrait de jeter les bases d'un projet plus ambitieux et à terme d'identifier des allèles prometteurs pour la sélection variétale.

#### Justification du financement demandé

Phénotypage : (Analyses histologiques, analyses biochimiques) 3000€

Caractérisations moléculaires : (qPCR, single cell, yeast two hybride, FRET FLIM) 3500€

Stagiaire (6 mois) : 3500€

#### Trois publications récentes associées aux projets de deux équipes partenaires

Wang, H. et al (2005). The tomato Aux/IAA transcription factor IAA9 is involved in fruit development and leaf morphogenesis. *Plant Cell* 17: 2676-2692.

Yu, H. et al , (2014) Genome-Wide Characterization and Expression Profiling of the AUXIN RESPONSE FACTOR (ARF) Gene Family in Eucalyptus grandis. *PLoS One* 9: e108906.

Yu, H. et al (2015) Comprehensive Genome-Wide Analysis of the Aux/IAA Gene Family in Eucalyptus: Evidence for the Role of EgrIAA4 in Wood Formation. *Plant Cell Physiol*.





## APPEL A PROJET 2021 INTER-UNITES - FR AIB

<b>Critères d'éligibilité</b>	Financement de 13 k€ maximum à partager entre co-porteurs. Projet réunissant au moins 2 partenaires de 2 Unités membres de la FR AIB Durée : 12 à 24 mois.
<b>Critères incitatifs (non obligatoires)</b>	Les collaborations entre Unités Ecologie/Biologie, bien que non obligatoires, sont encouragées. <ul style="list-style-type: none"> <li>• Unités d'écologie : EDB, SETE</li> <li>• Unités de biologie : CNRGV, LIPME, LRSV-GBF</li> </ul> Si le projet implique l'utilisation de plateformes, l'utilisation de celles du périmètre de la FR AIB est encouragée (Imagerie FR, Protéomique FR, TPMP, plateformes METATRON terrestre et aquatique, serre et volière de la « SETE », GeT-PlaGe, Agromix).
<b>Soumission</b>	Document téléchargeable ( <i>cf page suivante</i> ) Projets à envoyer à : <a href="mailto:dirfraib-Occitanie-Toulouse@inrae.fr">dirfraib-Occitanie-Toulouse@inrae.fr</a> <ul style="list-style-type: none"> <li>• Date limite de soumission le <b>29 mars 2021 midi</b>. Aucun projet ne sera accepté après cette limite.</li> <li>• Évaluation réalisée par le CS FR AIB selon les critères rapportés ci-dessous.</li> <li>• Retour d'évaluation mi-avril</li> </ul>
<b>Critères d'évaluation</b>	<b>Qualité du projet</b> <ul style="list-style-type: none"> <li>• Pertinence scientifique</li> <li>• Faisabilité en 12/24 mois</li> <li>• Caractère innovant ou original</li> <li>• Valeur ajoutée du financement demandé</li> </ul> <b>Composition du partenariat</b> <ul style="list-style-type: none"> <li>• Complémentarité/synergie/implication des 2 équipes</li> <li>• Collaboration entre Unités Ecologie/Biologie<sup>1</sup></li> <li>• Initiation d'une collaboration<sup>1</sup></li> </ul> <b>Utilisation d'une plateforme du périmètre de la FR AIB<sup>1</sup></b>
<b>Suivi des projets</b>	Une justification de l'utilisation des crédits pour l'utilisation des plateformes ou gratification d'étudiant sera demandée dans l'année qui suit l'obtention du financement.  Les porteurs de projets seront invités à présenter les résultats lors du forum de la FR AIB de l'année +1 ou +2  Les publications qui résultent d'un financement de la FR AIB devront porter la mention : "This study was supported by the Research Federation FR3450 CNRS-UT3 »

<sup>1</sup> critères non obligatoires



## APPEL A PROJET INTER-UNITES FR AIB 2021 / INTER-UNITS PROJECTS 2021

**Date limite de retour : lundi 29 mars 2021 midi / deadline : Monday, 2021 March 21**

- Titre du projet et son acronyme / Project title and acronym

**AWARE: Analysis of primary cell WALL in REsistance under combined stresses**

- Noms des deux porteurs du projet et de leurs laboratoires / Two principal investigators' names and laboratories

Marta MARCHETTI/Richard BERTHOME, REACH team, LIPME

Hua WANG-CASSAN, ReDyWood team, LRSV

- Le cas échéant: collaboration Ecologie-Biologie / If appropriate: collaboration Ecology-Biology

- Le cas échéant : Utilisation de plateforme / If appropriate: Use of a platform

Imagerie FR  Protéomique FR  TPMP  Plateformes "SETE"  GeT-PlaGe   
AgromiX

- Description du projet (contexte, objectifs, approches envisagées) / Project description (scientific context, objectives, approaches). 2 pages max

Terrestrial plants are subjected to many biotic and abiotic stresses during their lifetime and correspondingly evolved a wide range of mechanisms to develop and survive. Particularly, the plant immunity relies on both constitutive preformed barriers as well as induced mechanisms when they come into contact with invading pathogens leading either to resistance or disease development depending on the plant genotype and pathogens (or pathosystems) type of interactions. Moreover, a majority of plant resistances are negatively impacted by abiotic stresses.

Emerging evidences indicate that plant cells exploit sophisticated mechanisms of sensing the alteration of cell wall integrity during biotic stress (Hamann, 2012; Pogorelko et al., 2013). Due to its strategic localization at the plant–environment interface, cell wall probably contribute to the outcome of plant–pathogen interactions (Huckelhoven 2007). Actually, a set of *Arabidopsis* cell wall mutants have shown to have an impact on disease resistance and fitness phenotypes providing a link between plant cell wall composition and plant development/immunity phenotypes (Molina et al., 2021). Indeed, several cell wall *Arabidopsis* mutants that develop resistance to *Agrobacterium tumefaciens* (*rat*), powdery mildew pathogens (*pmr*) and the fungus, *Plectosphaerella cucumerina* have been identified (Vogel and Somerville 2000; Vogel et al. 2002, 2004; Zhu et al. 2003). As well, inactivation of *Arabidopsis* WAT1 (Walls Are Thin1), a gene required for secondary cell-wall deposition, or secondary wall-specific cellulose synthase (*CesA*) genes (*ixr*) conferred broad-spectrum resistance to pathogens (Denancé et al., 2013; Hernandez-Blanco et al., 2007).

Temperature elevation, one of the major abiotic stresses that plants have to cope with (Desaint et al., 2020), is predicted to favor the emergence of new pathogens and to increase the occurrence of epidemics (Garrett et al., 2006; Evans et al., 2008; Bebberet et al., 2013). Actually, temperature elevation have been shown to result in a drastic impact on plant defense responses and to inhibit the major source of resistance to *R. solanacearum* in *A. thaliana*, mediated by the immunoreceptor pair RSP4/RRS1-R (Aoun et al., 2017; Desaint et al., 2020). Interestingly, the possible implication of the primary cell wall in the early plant response to *R. solanacearum* have been pointed out in a project that aimed at identifying the genetic bases of plant **robust resistance mechanisms to this pathogenic bacterium in a global warming context**. A GWA mapping approach was realized to fine map QTLs associated to the variations of 176 natural accessions of *A.*



*thaliana* response to the *R. solanacearum* GMI1000 strain at two different temperatures, 27°C and 30°C. Specifically, inoculation were performed on unwounded roots to mimic natural infection condition. This study allowed the identification of *CesA3* underlying a major QTL involved in early quantitative disease resistance (QDR) at 30°C (Figure1).

*CesA3* belongs to a family of *CesA* genes identified in *A. thaliana* and encodes a cellulose synthase subunit which together with *CesA1*, -2, -5, and -6 compose the *CesA* complex and participate in cellulose production during the primary cell wall formation. So far, the role of the primary cell wall in response to *Ralstonia* was not documented, either at ambient or at elevated temperature.

*CesA3* has been validated as a susceptibility gene involved in QDR to *R. solanacearum* at 30°C through a reverse genetic approach focused on the study of the *cesa3 eli1-1* and *cesa3 je5* knock-down mutants. In addition, the importance of primary cell-wall cellulose synthase complex (CSC) in the establishment of defense mechanisms to *R. solanacearum* at elevated temperatures has been demonstrated through the study of other the primary *CesA* complex subunit *CesA1* (*rsw1-10*) and *CesA6* (*prc1-1*) mutants (unpublished, Marchetti M.).

### Research objectives

Among the 176 natural accession analyzed in the GWAs approach, 7 have shown resistance to *Ralstonia* at elevated temperature. In order to better understand and characterize **the links existing between the primary cell wall in resistance to the combined stress** we propose to exploit available *CesA3* mutants and the resistant natural accessions to decipher the mechanisms involved in *CesA3* depending defense responses at increased temperature. We propose to (i) identify, thanks to natural accessions, the haplotype of *CesA3* gene associated with the resistant phenotype at 30°C, (ii) characterize the cell wall structure/composition in the mutant and resistant natural accessions of *A. thaliana* subjected or not to biotic and abiotic stresses applied separately or combined and (iii) analyze the expression patterns of genes known to be involved in a broad network of stress signal perception regulating plant cell growth and immune response.

### Workpackage 1. Identification and study of *CesA3* haplotypes related to natural accessions resistant at 30°C

Our previous results support that a down regulation of the expression of *CesA 3* gene leads to resistance at elevated temperature (reverse genetic knock-down mutants). To determine if resistance observed in wild accessions is related to a transcriptional modulation of *CesA3* at 30°C or to a mutation in coding sequence leading to a structural modification, the genetic polymorphism associated with the resistant phenotype in *CesA3* coding region and promoter sequence will be analyzed. For this purpose, the *CesA3* promoter and genomic coding sequence will be amplified using high fidelity Tap polymerase in the 76 accessions and amplicons will be sequenced using a PacBio approach at the INRAE Gentyane sequencing platform (Clermont-Ferrand, France). These data will be implemented by the microassembly Atlas tool analysis to generate consensus of CDS from all public illumina sequences available (coll. J. Gouzy, bioinformatics team, LIPME) for 76 accessions belonging to the panel of the worldwide collection of natural accessions used in GWA mapping, including the 7 accessions that are resistant at 30°C. Combination of both strategy will allow to accurately defining specific haplotype of *CesA3* locus correlated with *R. solanacearum* resistance response. To enrich this approach, the expression profile of *CesA3* in the resistant natural accessions and the most susceptible accessions inoculated or not at 27°C and 30°C will be analyzed by quantitative qRT-PCR studies.

### Workpackage 2. Elucidating the putative contribution of *CESA3* dependent primary cell wall modifications in the elevated temperature resistance

The plant cell wall can work as a “passive/physical” defensive structure that many pathogens encounter before facing intracellular plant defenses (Lipka & Panstruga, 2005; Underwood & Somerville, 2008). Alteration in *CESA3*, or other

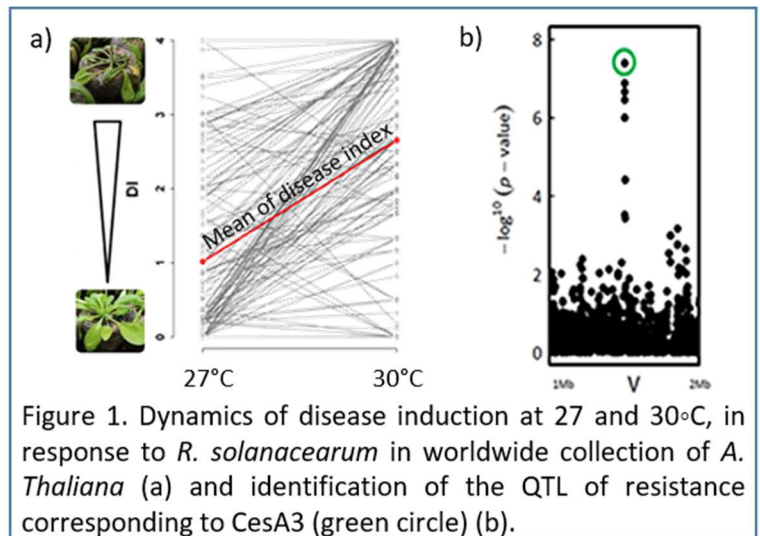


Figure 1. Dynamics of disease induction at 27 and 30°C, in response to *R. solanacearum* in worldwide collection of *A. Thaliana* (a) and identification of the QTL of resistance corresponding to *CesA3* (green circle) (b).





subunits of the CSC, induces a change in cell-wall structure, composition, or organization and plant development. These changes could modify the interaction between *R. solanacearum* and the cell-wall which would prevent the bacteria to break-down this first barrier and enter the cell. It can, therefore, be hypothesized that, for resistant natural accessions whose phenotypes depends on a specific *CesA3* expression, thermo-regulation or *CESA3* structure, the cell wall structure/composition is altered to prevent entry by *R. solanacearum*. These aspects will be analyzed on mutant lines and the 7 resistant and 7 susceptible natural accessions selected under two growth temperature 27°C and 30°C combined or not with bacterial inoculation. We plan to perform firstly the high throughput global cell wall chemical composition profiling by using a combination of FTIR spectroscopy and multivariate analysis (routinely performed by ReDyWood team, Dai et al 2020) as well as targeted measurement of cell wall composition such as saccharification. The most discriminant wave-numbers can further help us to target the particular cell wall components for the fine histochemical characterizations. We already have established a suitable *in vitro* pathosystem for *A. thaliana* – *R. solanacearum* allowing inoculation and observation of classical bacterial wilt symptoms on susceptible accessions. This system enables direct and repeated confocal microscopic observations of *Arabidopsis* roots in a non-destructive way. With the facility of local platform imagery TRI (<http://tri.ups-tlse.fr>), we will follow the dynamics of the cell wall and its components in planta by using the selection of a wide range of cell wall probes and CBM targeting different types of highly heterogeneous polymers such as cellulose, hemicelluloses and pectins. The promising perspective is to identify some particular aspects of cell wall microstructure *in muro* related to the thermo-stress and/or pathogen resistance.

**Workpackage 3.** Is a *CESA3*-dependent inducible signaling responsible for the robust resistance observed at high temperatures?

More than ten years of data mining has demonstrated that genes involved in cell wall formation present similar patterns of expression (co-expression) in response to environmental constraints (Ruprecht et al., 2012; Taylor-Teeple et al., 2015; Houston et al 2016). They likely contribute to extracellular compartment remodeling but the contribution of cell wall structure and composition to overall plant defense remain elusive. Recent advances in the field of cell wall signaling in plants, have point out the possible roles for cell wall-mediated signaling beyond the maintenance of cell wall integrity. The cell wall may thus relays information about the environment to the cell cytoplasm via signal transduction pathway(s) and prime a plant response to stresses. Even faint modifications of cell wall integrity by chemical treatments, mutations or abiotic constraints can promote plant resistance to pathogens infection.

A RNAseq analysis is going to be realized to identify genes differentially regulated in Col-0 susceptible accession compared to the *je5 cesA3* mutant infected or not by *Ralstonia*, either at 27°C or 30°C (ANR funding, REACH team). Since phenotypic studies have shown that the level of resistance in the mutant line is higher from 5 to 7 dai, the sampling will be realized at day6.

To enrich this approach, existing pipeline of analyses, such WGCNA or MixOmics (routinely performed by ReDyWood team) will be used to build gene co-expression networks from RNAseq data. This approach will provide a comprehensive landscape of co-expressed genes, and allow focusing on genes transcriptionally coordinated with *CesA3*. Modules enriched in CW-related genes and regulatory hubs can be correlated to phenotyping and chemotyping data to identify the DEGs potentially involved in a signaling pathway induced by cell wall remodeling in response to temperature, *Ralstonia* infection or both. To confirm the relevance of identified DEGs, expression profile of the most induced in infected plants under elevated temperature, will be profiled by RT-qPCR analysis in selected resistant and susceptible natural accessions.

In conclusion, we expect that this research project will help in understanding the contribution of *Arabidopsis* cell walls to disease-resistance responses in a combined stress condition, *R. solanacearum* infection and elevated temperature.



## References

- Aoun, N., Tauleigne, L., Lonjon, F., Deslandes, L., Vaillau, F., Roux, F., Berthomé, R. (2017). Quantitative disease resistance under elevated temperature: genetic basis of new resistance mechanisms. *Front Plant Sci* 8, 1387.
- Dai Y, Hu G, Dupas A, Medina L, Blandels N, Clemente HS, Ladouce N, Badawi M, Hernandez-Raquet G, Mounet F, Grima-Pettenati J, **Cassan-Wang H** (2020) Implementing the CRISPR/Cas9 Technology in Eucalyptus Hairy Roots using wood-related genes. *Int J Mol Sci* 21:3408.
- Denancé, N., Ranocha, P., Oria, N., Barlet, X., Rivière, MP., Yadeta, KA., Hoffmann, L., Perreau, F., Clément, G., Maia-Grondard, A., van den Berg, GC., Savelli, B., Fournier, S., Aubert, Y., Pelletier, S., Thomma, BP., Molina, A., Jouanin, L., Marco, Y., Goffner, D. (2013). *Arabidopsis* wat1 (walls are thin1)-mediated resistance to the bacterial vascular pathogen, *Ralstonia solanacearum*, is accompanied by cross-regulation of salicylic acid and tryptophan metabolism. *Plant J* 73 :225.
- Desaint, H., Aoun, N., Deslandes, L., Vaillau, F., Roux, F., Berthomé, R. (2021). Fight hard or die trying: when plants face pathogens under heat stress. *New Phytol.* 229:712.
- Hannan, T. (2012). Plant cell wall integrity maintainance as an essential component of biotic stress response mechanisms. *Front Plant Sci* 3: 77.
- Hernández-Blanco, C., Feng, D. X., Hu, J., Sánchez-Vallet, A., Deslandes, L., Llorente, F., Berrocal-Lobo, M., Keller, H., Barlet, X., Sánchez-Rodríguez, C., Anderson, L. K., Somerville, S., Marco, Y. & Molina, A. (2007). Impairment of cellulose synthases required for *Arabidopsis* secondary cell wall formation enhances disease resistance. *Plant Cell* 19 : 890.
- Houston, K., Tucker, M., Chowdhury, J., Shirley, N., and Little A. et al (2016) The plant cell wall: a complex and dynamic structure as revealed by the responses of genes under stress conditions. *Front Plant Sci* 7: 984.
- Hückelhoven R. . (2007) Cell wall associate mechanisms of disease resistance and susceptibility. *Annu Rev Phytopathol* 45 :101.
- Lipka, V., Panstruga, R. (2005) Dynamic cellular responses in plant-microbe interactions. *Curr Opin Plant Biol* 8:625.
- Molina, A., et al. (2021). *Arabidopsis* cell wall composition determines disease resistance specificity and fitness. *Proc Natl Acad Sci* 118.
- Pogorelko, G., Bellincampi, D., Zabolina, O. (2013). Cell wall integrity :targeted post-synthetic modifications to reveal its role in plant growth and defense against pathogen. *Plant Signal Behav* 8.
- Ruprecht, C., Persson, S. (2012) Co-expression of cell wall related genes: new tools and insights. *Plant Science* 3:83.
- Taylor-Teeples, M., Lin, L., De Lucas, M., Turco, G., Toal, TW., Gaudinier, A., Young, NF., Trabucco, GM. (2015). An *Arabidopsis* gene regulatory network for secondary cell wall synthesis. *Nature* 517: 571.
- Underwood, W., Somerville, S. (2008). Focal accumulation of defences at sites of fungal pathogen attack. *J Exp Bot* 59:3501.
- Vogel, J., Somerville, S. Isolation and characterization of powdery mildew-resistant *Arabidopsis* mutants. 2000 *Proc Natl Acad Sci U S A* 97:1897.

- Complémentarité, synergie et rôles respectifs des partenaires / Synergy of the partnership and respective implications

**Partner 1** : Partner1 has expertise on the *Ralstonia-Arabidopsis* pathosystem and heat stress, one of the central project of the REACH team. PI1 will provide the natural accession panel and mutants of the primary cell wall. PI1 will mainly implicated in WP1 and 2.

**Partner 2** : Partner2 has skills in histochemical of the cell wall and FTIR approaches. In concert with the ReDyWood team, PI2 will provide system biology approaches integrating 'omics (transcriptomics, metabolomics, epigenetics, and genetics) and phenotypic data. PI2 will be mainly implicated in WP2 and 3.

Both PIs had already established methods and tools and they have the relevant knowledge, necessary for the realization of the current proposal. This proposal requires a combination of their respective expertise including the characterization of phenotypical analysis of plant cell wall in close collaboration with the Imaging core facility of the FR AIB (TRI-Genotoul).

- Deux publications pour chaque porteur de projet / publications of PI's (2 for each one)

REACH Team :

- Aoun, N., Tauleigne, L., Lonjon, F., Deslandes, L., Vaillau, F., Roux, F. & **Berthomé, R.** (2017). Quantitative Disease Resistance under Elevated Temperature: Genetic Basis of New Resistance mechanisms. *Front Plant Sci* 8:1387.
- Desaint H, Aoun N, Deslandes L, Vaillau F, Roux F, **Berthomé R.** (2021). Fight hard or die trying: when plants face pathogens under heat stress. *New Phytol.* 229:712.

ReDyWood team:

- Dai Y, Hu G, Dupas A, Medina L, Blandels N, Clemente HS, Ladouce N, Badawi M, Hernandez-Raquet G, Mounet F, Grima-Pettenati J, **Cassan-Wang H** (2020) Implementing the CRISPR/Cas9 Technology in Eucalyptus Hairy Roots using wood-related genes. *Int J Mol Sci* 21:3408.
- Ployet, R., Veneziano, Lbate, MT., Regiani Cataldi, T., Christonana, M., Morel, M., San Clemente, H., Denis, M.? Fayreau, B., Tomazello Filho, M., Laclau, JP., Labate, CA., Chaix, G., Grima-Pettinati, J., **Mounet, F.**(2019). A system biology view of wood formation in *Eucalyptus grandis* trees submitted to different potassium and water regimes. *New Phytol* 223:766.



- Justification du budget et de la valeur ajoutée du financement demandé, en cas de financement déjà acquis sur le projet / Budget and complementarities with any previous funding of this topic

This project will complement and finalize the RNAseq study that is founded in the context of a current ANR. We integrate into this project the support for a Master 2 student that could help us in developing our project at the interphase of the two PIs. We will support her/his application for a PhD grant at SEVAB doctoral school under a co-supervision scheme. In addition to the mutual benefits for the two research niches of the PIs, the implementation of the current research program will be helpful in preparing applications for local and national research grants (e.g. SPE, TULIP New frontiers, ANR,..)

Budget (13000€):

**WP1** : Amplification system and Pac Bio sequencing and analysis 3500 €.

**WP2** : Cell wall chemical analysis (sample preparation, FTIR analysis, Saccharification, Acetyl bromide dosage) 2100 €  
Fluorescent and wild field microscopy, antibodies and consumables 1200 €  
24 confocal sessions 1800 €

**WP3** : RNA extraction, qPCR 600 €  
Molecular biology consumables 200 €

**Master 2 Internship (6 months 3600)**



# Annexe 11 : Citations des différents travaux de recherche

03/10/2023

<input type="checkbox"/>	TITRE	CITÉE PAR	ANNÉE
<input type="checkbox"/>	<a href="#">The genome of <i>Eucalyptus grandis</i></a> AA Myburg, D Grattapaglia, GA Tuskan, U Hellsten, RD Hayes, ... Nature 510 (7505), 356-362	830	2014
<input type="checkbox"/>	<a href="#">The Tomato <i>Aux/IAA</i> Transcription Factor <i>IAA<sub>9</sub></i> Is Involved in Fruit Development and Leaf Morphogenesis</a> H Wang, B Jones, Z Li, P Frasse, C Delalande, F Regad, S Chaabouni, ... The Plant Cell 17 (10), 2676-2692	590	2005
<input type="checkbox"/>	<a href="#">Regulatory features underlying pollination-dependent and-independent tomato fruit set revealed by transcript and primary metabolite profiling</a> H Wang, N Schauer, B Usadel, P Frasse, M Zouine, M Hernould, A Latche, ... The Plant Cell 21 (5), 1428-1452	304	2009
<input type="checkbox"/>	<a href="#">Characterization of the tomato ARF gene family uncovers a multi-levels post-transcriptional regulation including alternative splicing</a> M Zouine, Y Fu, AL Chateigner-Boutin, I Mila, P Frasse, H Wang, ... PLoS one 9 (1), e84203	185	2014
<input type="checkbox"/>	<a href="#">SI-IAA3, a tomato <i>Aux/IAA</i> at the crossroads of auxin and ethylene signalling involved in differential growth</a> S Chaabouni, B Jones, C Delalande, H Wang, Z Li, I Mila, P Frasse, ... Journal of experimental botany 60 (4), 1349-1362	161	2009
<input type="checkbox"/>	<a href="#">Genome-wide analysis of the lignin toolbox of <i>Eucalyptus grandis</i></a> V Carocha, M Soler, C Hefer, H Cassan-Wang, P Fevèreiro, AA Myburg, ... New Phytologist 206 (4), 1297-1313	129	2015
<input type="checkbox"/>	<a href="#">Identification of novel transcription factors regulating secondary cell wall formation in <i>Arabidopsis</i></a> H Cassan-Wang, N Goué, MN Saidi, S Legay, P Sivadon, D Goffner, ... Frontiers in plant science 4, 189	115	2013
<input type="checkbox"/>	<a href="#">The <i>Eucalyptus grandis</i> R2R3-MYB transcription factor family: evidence for woody growth-related evolution and function</a> M Soler, ELO Camargo, V Carocha, H Cassan-Wang, H San Clemente, ... New Phytologist 206 (4), 1364-1377	112	2015
<input type="checkbox"/>	<a href="#">Auxin-mediated <i>Aux/IAA-ARF-HB</i> signaling cascade regulates secondary xylem development in <i>Populus</i></a> C Xu, Y Shen, F He, X Fu, H Yu, W Lu, Y Li, C Li, D Fan, HC Wang, K Luo New Phytologist 222 (2), 752-767	74	2019
<input type="checkbox"/>	<a href="#">Reference Genes for High-Throughput Quantitative Reverse Transcription-PCR Analysis of Gene Expression in Organs and Tissues of <i>Eucalyptus</i> Grown in ...</a> H Cassan-Wang, M Soler, H Yu, ELO Camargo, V Carocha, N Ladouce, ... Plant and Cell Physiology 53 (12), 2101-2116	53	2012
<input type="checkbox"/>	<a href="#">Genome-Wide Characterization and Expression Profiling of the <i>AUXIN RESPONSE FACTOR (ARF)</i> Gene Family in <i>Eucalyptus grandis</i></a> H Yu, M Soler, I Mila, H San Clemente, B Savelli, C Dunand, JAP Paiva, ... PLoS One 9 (9), e108906	51	2014
<input type="checkbox"/>	<a href="#">Toward the discovery of maize cell wall genes involved in silage quality and capacity to biofuel production.</a> Y Barrière, V Méchin, F Lafarguette, D Manicacci, F Guillon, H Wang, ... Maydica 54 (2/3), 161-198	50	2009
<input type="checkbox"/>	<a href="#">Comprehensive Genome-Wide Analysis of the <i>Aux/IAA</i> Gene Family in <i>Eucalyptus</i>: Evidence for the Role of <i>EgrIAA4</i> in Wood Formation</a> H Yu, M Soler, H San Clemente, I Mila, JAP Paiva, AA Myburg, ... Plant and Cell Physiology 56 (4), 700-714	39	2015

<input type="checkbox"/>	TITRE	CITÉE PAR	ANNÉE
<input type="checkbox"/>	<b>Explosive Tandem and Segmental Duplications of Multigenic Families in <i>Eucalyptus grandis</i></b> Q Li, H Yu, PB Cao, N Fawal, C Mathé, S Azar, H Cassan-Wang, ... Genome Biology and Evolution 7 (4), 1068-1081	37	2015
<input type="checkbox"/>	<b>Transcriptional regulation of the lignin biosynthetic pathway revisited: new players and insights</b> J Grima-Pettenati, M Soler, ELO Camargo, H Wang Advances in botanical research 61, 173-218	36	2012
<input type="checkbox"/>	<b>Implementing the CRISPR/Cas9 Technology in <i>Eucalyptus</i> Hairy Roots Using Wood-Related Genes</b> Y Dai, G Hu, A Dupas, L Medina, N Blandels, H San Clemente, ... International Journal of Molecular Sciences 21 (10), 3408	27	2020
<input type="checkbox"/>	<b>Breeding grasses for capacity to biofuel production or silage feeding value: an updated list of genes involved in maize secondary cell wall biosynthesis and assembly</b> A Courtial, M Soler, AL Chateigner-Boutin, M Reymond, V Mechin, ... Maydica 58 (1), 67-102	19	2013
<input type="checkbox"/>	<b>Digging in wood: New insights in the regulation of wood formation in tree species</b> ELO Camargo, R Ployet, H Cassan-Wang, F Mounet, J Grima-Pettenati Advances in botanical research 89, 201-233	15	2019
<input type="checkbox"/>	<b>Identification of novel lncRNAs in <i>Eucalyptus grandis</i></b> Z Lin, J Long, Q Yin, B Wang, H Li, J Luo, HC Wang, AM Wu Industrial Crops and Products 129, 309-317	12	2019
<input type="checkbox"/>	<b>Overexpression of EgrIAA20 from <i>Eucalyptus grandis</i>, a Non-Canonical Aux/IAA Gene, Specifically Decouples Lignification of the Different Cell-Types in Arabidopsis Secondary Xylem</b> H Yu, M Liu, Z Zhu, A Wu, F Mounet, E Pesquet, J Grima-Pettenati, ... International Journal of Molecular Sciences 23 (9), 5068	2	2022
<input type="checkbox"/>	<b><i>Eucalyptus grandis</i> AUX/INDOLE-3-ACETIC ACID 13 (EgrIAA13) is a novel transcriptional regulator of xylogenesis</b> N Karannagoda, A Spokevicius, S Hussey, H Cassan-Wang, ... Plant Molecular Biology 109 (1-2), 51-65	2	2022
<input type="checkbox"/>	<b>Role of auxin transcription factors in the regulation of fruit development and ripening</b> JC Pech, A Latche, M Bouzayen, H Wang, B Jones Acta Physiologiae Plantarum 26 (3 suppl.)	2	2004
<input type="checkbox"/>	<b>Regulation of secondary cell wall lignification by abiotic and biotic constraints</b> IH Bachir, R Ployet, C Teulières, H Cassan-Wang, F Mounet, ... Advances in Botanical Research 104, 363-392		2022
<input type="checkbox"/>	<b>What MYB genes tell us about the specificities of woody plants?</b> M Soler, A Plasencia, J Lepikson-Neto, A Dupas, N Ladouce, H Yu, ... IUFRO Tree Biotechnology 2015 Conference		2015
<input type="checkbox"/>	<b>Auxin-Dependent Transcriptional Regulation During Fruit Development</b> H Cassan-Wang INP ENSAT Toulouse; INP Toulouse		2004
<input type="checkbox"/>	<b>Characterisation of Aux/IAA like genes expressed in tomato fruit</b> H Wang, Z Li, P Frasse, C Audran, B Jones, A Jauneau, M Bouzayen NATO SCIENCE SERIES SUB SERIES I LIFE AND BEHAVIOURAL SCIENCES 349, 309-310		2003