



Université de Lille 1, Sciences et Technologies UGSF - Unité de Glycobiologie Structurale et Fonctionnelle UMR Lille1/CNRS n°8576

Ecole Doctorale Sciences de la Matière, du Rayonnement et de l'Environnement

Spécialité : Ingénierie des Fonctions Biologiques

THESE

Julien Le Roy

Rôles des UDP-Glycosyltransférases dans la régulation de la lignification par glycosylation des précurseurs

Thèse dirigée par : Dr. Godfrey Neutelings

Soutenue le 30 Juin 2017 devant la commission d'examen :

Mme M. BAUCHER, Maître de Recherche, Université Libre de Bruxelles	Rapporteur
M J-C. MOLLET, Professeur, Université de Rouen	Rapporteur
Mme B. CHABBERT, Chargée de Recherche, INRA Reims	Examinateur
M F. MESNARD, Professeur, Université de Picardie	Examinateur
M S. HAWKINS, Professeur, Université de Lille 1	Examinateur
M G. NEUTELINGS, Maître de conférence, Université de Lille 1	Dir. de thèse

Remerciements

Tout d'abord, je tiens à remercier les membres du jury qui ont pris le temps d'évaluer mon travail de thèse : Mme Marie Baucher et M. Jean-Claude Mollet en tant que rapporteurs ainsi que Mme Brigitte Chabbert, M. François Ménard et M. Simon Hawkins comme examinateurs.

Je tiens aussi à remercier tout particulièrement Godfrey Neutelings pour m'avoir encadré durant la totalité de cette thèse, pour avoir fait preuve d'une très grande disponibilité à mon égard et m'avoir toujours beaucoup encouragé au cours de ces trois dernières années.

La thèse à débuter au sein de l'ex laboratoire UMR n°1281 INRA/Lille 1 SADV dont Jean-Louis Hilbert était le directeur. Je tiens à le remercier de m'avoir accueillie tout comme Christophe D'Hulst, directeur de l'UMR n°8576 CNRS/Lille 1 UGSF, pour la seconde moitié de la thèse.

Mes plus vifs remerciements vont à l'ensemble de l'équipe « fibres végétales » et surtout Simon Hawkins pour m'avoir accueilli au sein de son équipe. Plus particulièrement, Anne Créach et Brigitte Huss qui m'ont beaucoup soutenu et aidé tout au long de ma thèse. Merci à Anne-Sophie Blervacq pour les expériences d'hybridation *in situ* et de microscopie. Merci à Christophe Djemiel pour m'avoir permis de gagner du temps lors de travaux bioinformatique. Sans oublier Estelle Goulas, Sébastien Grec, Arnaud Day, Malika Chabi et Maxime Chantreau pour leur aide.

Merci à Marianne Delporte, David Gagneul et Jean-Louis Hilbert de leurs expertises concernant les aspects de production de protéines recombinantes et d'essais enzymatiques.

Merci à Brigitte Chabbert et David Cronier pour nous avoir initié au dosage de lignine par dosage au bromure d'acétyle.

Merci à Florian Veillet pour les analyses d'infection par Botrytis cinerea.

Merci à Coralie Bompard pour son expertise et son aide lors des modélisations tridimensionnelles.

Merci à la plateforme PAGéS et plus particulièrement à Dounia Mouajjah, Nao Yamakawa et Yann Guerardel qui ont permis les analyses en spectrométrie de masse et l'utilisation d'une chaine HPLC.

Merci à Delphine Fleury, Marie-Christine Quillet et Théo Hendricks pour leur encadrement et leur aide au cours de mes premiers pas dans le monde de la Recherche.

Je tiens aussi à remercier les étudiants de M1 ou de L3 qui m'ont aidés au cours de ce projet lors de stage : Marion Pillet, Yohann Cadart, Clément Gladysz et Aliénor Delsart.

Merci aux personnes, rencontrées au cours de la thèse, qui m'ont fait profiter de leurs travaux, de leur expérience et de leur sympathie.

Merci à ma famille qui m'a soutenue et à qui je dois énormément.

Et bien sur, merci à Marie qui m'a permis d'arriver où je suis.

Communications scientifiques dans le cadre de la thèse

Posters

<u>Régulation of monolignol glycosylation and lignification in flax and Arabidopsis</u> Julien Le Roy, Brigitte Huss, Anne Creach, Coralie Bompard, Marianne Delporte, Dounia Mouajjah, David Gagneul, Simon Hawkins et Godfrey Neutelings XIV Cell Wall Meeting à Chania (Grèce) Juin 2016</u>

<u>Monolignol glycosylation and lignification in Flax (Linum usitatissinum L.)</u> Julien Le Roy, Anne Creach, Brigitte Huss, Maria Cécilia Arias, Simon Hawkins, Godfrey Neutelings Lignin 2014 à Umeå (Suède) Aout 2014

<u>Identification et caractérisation des gènes de la voie de biosynthèse de la lignine chez le lin</u> Julien Le Roy, Anne Sophie Blervacq, Christophe Hano, Simon Hawkins et Godfrey Neutelings. 10e Congrès Réseau Français des Parois à Amiens, Juin 2014

<u>Impact de la glycosylation sur la diponibilité des monolignols chez le lin (Linum usitatissimum L.)</u> Julien Le Roy, Anne Creach, Brigitte Huss, Maria Cécilia Arias, Simon Hawkins, Godfrey Neutelings 10e Congrès Réseau Français des Parois à Amiens, Juin 2014</u>

Présentations orales

Impact of UDP-GlycosylTransferases (UGTs) on the availability of monolignols and their precursors during lignification Fiber quality workshop à Lille, Janvier 2016

<u>Relation between monolignols glycosylation and lignification in flax and Arabidopsis</u> Transatlantic flax seminar à Lille, Juin 2015

Articles scientifiques

<u>Glycosylation Is a Major Regulator of Phenylpropanoid Availability and Biological Activity</u> <u>in Plants</u> Julien Le Roy, Brigitte Huss, Anne Creach, Simon Hawkins et Godfrey Neutelings Front. Plant Sci. 2016; 7:735.

Spatial regulation of monolignol biosynthesis and laccase genes control developmental and stress-related lignin in flax Julien Le Roy, Anne Sophie Blervacq, Anne Créach, Brigitte Huss, Simon Hawkins et Godfrey Neutelings BMC Plant Biology, 2017; 17 :124.

<u>Résumé</u>

Mot clefs : Phénylpropanoïdes, lignines, parois secondaire, UDP-Glycosyltranférases, *Lin*, Arabidopsis

De nombreux processus liés à la régulation de la lignification dans les parois secondaires chez les plantes sont maintenant relativement bien connus. Toutefois, le rôle de la glycosylation des monolignols et de leurs précurseurs reste encore à établir dans cette voie de biosynthèse complexe. Les fibres primaires de lin (Linum usitatissimum) sont caractérisées par la présence d'une paroi secondaire épaisse contenant très peu de lignines contrairement aux cellules voisines du xylème dans la tige. Cette différence de composition s'accompagne également d'une accumulation de molécules apparentées aux lignines majoritairement présentes sous forme glycosylées dans les fibres. Le contexte général de ce travail était donc de rechercher une éventuelle influence entre la glycosylation des précurseurs et les quantités de lignines elles mêmes. Dans un premier temps nous avons caractérisé l'ensemble de la voie de biosynthèse des monolignols chez le lin en reconstituant les familles multigéniques à l'aide de la séquence de son génome puis notamment établi leurs profils d'expression. Dans un deuxième temps, la même démarche a été effectuée dans le but de d'identifier toutes les UDP GlycosylTransférases (UGTs) formant l'UGTome du lin. Parmi celles-ci, cinq orthologues d'UGTs d'Arabidopsis capables de glycosyler ces métabolites secondaires ont été caractérisés. Enfin, pour mieux comprendre la relation entre la glycosylation des monolignols et la lignification, un triple mutant *ugt72e1-2-3* d'*Arabidopsis* a été obtenu et a montré non seulement que la glycosylation influençait la lignification mais a aussi permis de mieux cerner le rôle probable de ces UGTs chez les plantes.

<u>Abstract</u>

Key words : Phenylpropanoids, lignins, secondary cell wall, UDP-Glycosyltransferases, flax, Arabidopsis

Many processes related to the regulation of plant secondary cell wall lignification are now relatively well known. However, the role of glycosylation of monolignols and their precursors within this complex biosynthetic pathway still remains to be established. Primary flax (Linum usitatissimum) fibers are characterized by the presence of a thick secondary cell wall containing very low amounts of lignins in contrast to the neighboring cells of the xylem in the stem. This difference in composition is also accompanied by an accumulation of molecules related to lignins, predominantly present in glycosylated forms within the fibers. The general context of this work was therefore to check if any relation between the glycosylation of the precursors and the amounts of lignins themselves exists. We first characterized the entire biosynthetic pathway of monolignols in flax by reconstituting the multigenic families using the sequence of its genome and then, among other approaches, established their expression profiles. In a second step, the same procedure was performed to identify all of the UDP GlycosylTransferases (UGTs) forming the UGTome of flax. Among these, five enzymes orthologous to Arabidopsis UGTs, which were able to glycosylate these secondary metabolites, were characterized. Finally, to gain more knowledge on the possible link between monolignol glycosylation and lignification, a triple Arabidopsis mutant ugt72e1-2-3 was obtained and showed that not only such link exists, but that UGTs play specific roles in plants.

Liste des abréviations	
Introduction	
I- Métabolisme secondaire chez les plantes vasculaires	
II- Phénylpropanoïdes	
1/- Généralités	
2/- Voie du shikimate	
3/- La voie générale des phénylpropanoïdes	
3-1/- Phenylalanine Ammonia-Lyase (PAL)	23
3-2/- Cinnamic acid 4-hydroxylase (C4H)	24
3-3/- 4-coumarate:CoA ligase (4CL)	25
4/- Les voies spécifiques	25
4-1/- Les formes simples	
4-2/- Les formes complexes	
III- Les lignines	
1/- Les lignines dans la paroi	
2/- Les cellules lignifiées	
2.1/- Les tissus conducteurs – Xylème	
2.2/- Les tissus de soutien – Sclérenchymes	
2.3/- Autres types cellulaires spécifiques	
3/- Diversité des lignines	
4/- La biosynthèse des lignines	
4-1/- Biosynthèse des monolignols	
4-2/- Transports des précurseurs des lignines	
4-3/- La polymérisation des lignines	
5/- La régulation de la lignification	
5-1/ La régulation transcriptionnelle	56
5-2/- La régulation post-transcriptionnelle	59
5-3/- Régulation traductionnelle & post-traductionnelle	60
IV- L'utilisation du lin comme plante modèle	61
1/- Botanique et utilisation	61
2- Formation des fibres primaires de lin	63
2-1/- Initiation	64
2-2/- Elongation	64
2-3/- Epaississement	65
V- Les activités UDP-Glycosyltransférases	67
1/- La glycosylation chez les plantes	67

2/- UDP-glycosyltransferase (ou GT1)	69
2-1/- Structure des UGTs	
2-2/- Régulation, localisation et interaction des UGTs	75
2-3/- Rôle des UGTs	79
VI- A la recherche d'un lien entre la glycosylation et la lignification	n chez le lin . 84
VII- Objectifs du travail de thèse	
Résultats	
I- Identification des gènes de biosynthèse des phénylpropanoïdes	chez le lin89
Avant propos	
Article : Spatial regulation of monolignol biosynthesis and laccase ge	enes control
developmental and stress-related lignin in flax	
II- Caractérisation de l'UGTome chez le lin	
1/- Identification et annotation des gènes UGTs	
2/- Analyse phylogénétique	
3/- Prédiction de la localisation subcellulaire	
4/- Expression des UGT chez le lin	
5/- Activités enzymatiques vis à vis de phénylpropanoïdes	
III- /- Identification des UGTs spécifiques des monolignols	146
1/- Sélection des gènes	
1-1/- Analyse phylogénétique	
1-2/- Expression	
1-3/- Production des protéines recombinantes	
1-4/- Activité enzymatique	
1-5/- Mutagénèse dirigée	
1-6/- Amélioration du protocole	
IV- La glycosylation des monolignols chez Arabidopsis	158
1/- Caractéristiques des gènes UGT72E1, UGT72E2 et UGT72E3	158
2- Obtention des triples mutants <i>ugt72e1/ ugt72e2/ ugt72e3</i>	
3/- Phénotypage des mutants ugt72e1, ugt72e2, ugt72e3	
3-1/- Modification de la floraison	
3-2/- Impact sur la sénescence	
3-3/ Modifications de la biomasse	
4/- Réponses aux stress des mutants	168
4-1/- Impact de l'infection par Botrytis cinerea	
4-2/- Carence en nutriments	
4-3/ Stress au froid	
5/ Relation entre les UGTs et la lignification	172

Discussion	175
I- Les UGT chez <i>Linum usitatissimum</i>	
1/-Caractéristiques de l'UGTome de lin	176
2/- Fonctions de UGT72N1, UGT72N2, UGT72Q1, UGT72R1 et UGT72R2	178
2-1/- Les informations obtenues à partir des profils d'expression	
2-2/- Les activités enzymatiques liées aux UGTs de lin	179
II- Relation entre la glycosylation des monolignols et la lignification	
1/- Glycosylation et quantités de lignines	180
2/- Glycosylation et composition des lignines	185
3/- Glycosylation et interactions protéines-protéines	185
4/- Glycosylation et types cellulaires	187
III- Rôle de UGT72E1 E2 et E3 dans la physiologie des plantes	
1/- Développement, sénescence et stress	189
2/- Métabolisme du fer	191
Matériel et méthodes	194
I- Matériel végétal	195
1/- Cultures de lin	195
2/- Cultures d'Arabidopsis thaliana	
3/- Obtention du triple mutant <i>ugt72e1-2-3</i>	
II- Biologie moléculaire	197
1/- Extraction d'ADN pour le génotypage	197
2/- Extraction d'ARN, synthèse d'ADNc et quantification de l'expression	197
3/- Clonages moléculaires	199
4/- Mutagénèse dirigée	200
III- Expression hétérologue et essais enzymatiques	200
1/- Expression et purification des protéines recombinantes	200
2/- Electrophorèse SDS-PAGE et Western Blot	201
3/- Essais enzymatiques sur les protéine recombinantes	202
4/- Essais enzymatiques sur des extraits protéiques de plantes	203
IV- Phénotypage	203
1/- Extraction des résidu pariétaux et dosage de lignines	203
2/- Microscopie et préparation des échantillons	204
3/- Infection par <i>Botrytis cinerea</i>	205
V- Analyses <i>in silico</i>	205
1/- Analyse phylogénétique	205
2/- Tests statistiques	205

3/- Analyse des racines	205
Bibliographie	.206
Annexes	.233
La glycosylation est un régulateur majeur de la disponibilité et de l'activité	
biologique des phénylpropanoïdes chez les plantes	234
Avant propos	234
Article: Glycosylation is a major regulator of phenylpropanoid availability and	
biological activity in plants	236

Liste des abréviations

4Cl : 4-coumarate : CoA ligase

ADN: Acide désoxyribonucléique

ADNc: ADN complémentaire

AHC : Acide hydroxycinnamique

ARN : Acide ribonucléique

ATP : Adénosine triphosphate

BAHD:BenzylalcoholO-acetyltransferase,AnthocyaninO-hydroxycinnamoyltransferase,N-Hydroxycinnamoyl/benzoyltransferaseetDeacetylvindoline 4-O-acetyltransferase

BLAST : Basic local alignment search tool

C3'H : *p*-Coumaroyl ester 3'-hydroxylase

C4H : Cinnamate 4-hydroxylase

CAD : Cinnamyl alcohol deshydrogenase.

CAZy : Carbohydrate-Active enZYmes

CCoAOMT : Caffeoyl CoA-O-methyltransferase.

CCR : Cinnamoyl CoA reductase.

CHI : Chalcone isomérase

CHS : Chalcone synthase

CNRS : Centre National de la Recherche Scientifique

CoA : Coenzyme A

COMT : Caféate 3-O-méthyltransférase

CSE : Caféoyl shikimate estérase

DO : Densité optique

EMS : Ethyl methanesulfonate

EST : Expressed sequence tag

F5H : Férulate 5-hydroxylase

GH : Glycosylhydrolase

GT: Gly cosyltransferase

HCT : Hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyltransférase

HPLC : Chromatographie liquide à haute performance

HQT : Hydroxycinnamoyl-CoA quinate hydoxycinnamoyltransférase

IPTG : Isopropyl β -D-1-thiogalactopyranoside

KO : Knock out

LAC : Laccase

LB : Lysogeny broth

MeJA : Méthyl-jasmonate

miRNA : microARN

MS : Milieu Murashige & Skoog

PAL : Phénylalanine ammonia-lyase

pb : Paire de bases

PCR : Réaction en chaîne par polymérase

PRX : Péroxydases

PSPG : Plant secondary product glycosyltransferase

qPCR : PCR quantitative en temps réel

SHT : Hydroxycinnamoyl-CoA : sérotonine N-hydroxycinnamoyltransferases

SDS-PAGE : Electrophorèse sur gel de polyacrylamide en présence de dodécylsulfate de sodium

SND : Secondary wall-associated NAC domain protein

STS : Stilbène synthase

TAL : Tyrosine ammonia-lyase

THT : Hydroxycinnamoyl-CoA: tyroamine N-hydroxycinnamoyltransferases

UDP: Uridinediphosphate

UGT : UDP-Glycosyltransférase

Unité 5H : 5-hydroxyguaiacyl

Unité C : catéchyles

Unité G : guaïacyle

Unité H : hydroxyphényle

Unité S : syringyle

UV : Ultraviolet

VND : Vascular-related NAC domain

WT : Sauvage

Introduction

I- Métabolisme secondaire chez les plantes vasculaires

Les plantes vasculaires accumulent une très large palette de molécules très secondaires" diverses connues sous le terme "métabolites ou plus récemment métabolites spécialisés (Pichersky et al., 2006; Schilmiller et al., 2008). Il en existe trois classes principales : les terpenoïdes, les alcaloïdes et les molécules phénoliques. Ils sont opposés au métabolites primaires qui incluent des molécules ubiquitaires jouant des rôles fondamentaux pour le vivant lors du développement cellulaire, de la respiration, du stockage de réserves et de la reproduction (Bourgaud et al., 2001). Ces métabolites spécialisés sont définis comme étant des composés produits par des familles phylogéniquement restreintes et constituent parfois un moyen de reconnaissance taxonomique. Ils ont longtemps été considérés comme non essentiels pour la plante, voire même des déchets issus du métabolisme primaire (Fraenkel, 1959; Hartmann, 2008). A l'aide des connaissances actuelles dans les domaines de la biochimie et de la biologie, il est maintenant clair que ces composés sont essentiels pour l'adaptation et les interactions des plantes à leurs environnements.

Au cours de l'évolution, la lignée verte a connu des bouleversements majeurs auxquels elle a fait face en adaptant sa morphologie, ses mécanismes physiologiques ainsi que son métabolisme. Il y a environ 450 million d'années, pendant la période de l'Ordovicien, les organismes vivants sont passés du milieu aquatique vers le milieu terrestre. Dans ce contexte, les voies métaboliques ont subi de profonds bouleversements essentiellement dus à la colonisation d'un environnement pauvre en eau. Ces organismes ont adopté des stratégies tendant à limiter les pertes en eau grâce à la mise en place d'une couche hydrophobe à la surface de l'ensemble des parties aériennes appelée cuticule. C'est une structure complexe composée principalement d'acides gras et de métabolites spécialisés appartenant aux familles des terpènes ainsi que des molécules phénoliques qui permettent de limiter les pertes en eau de la plante (Yeats and Rose, 2013). La connaissance des mécanismes de synthèse et de mise en place de la cuticule prend toute son importance dans le contexte actuel du changement climatique. En effet, la cuticule joue un rôle essentiel dans la résistance à la sècheresse (Kosma and Jenks, 2007) comme le montre par exemple le mutant d'Hordeum spontaneum (orge sauvage) eibi1 affecté dans le transport de cutine, l'un des deux constituants principaux de la cuticule (Chen et al., 2009a; Yang et al., 2013). Ceci à pour conséquence une diminution de l'épaisseur de la cuticule des feuilles conduisant à une perte d'eau plus importante lors d'un stress hydrique. Une meilleure compréhension de la mise en place de la cutine permettrait ainsi d'obtenir des plantes plus résistantes à la sècheresse et donc plus adaptées aux enjeux environnementaux actuels.

Au delà de son rôle dans l'homéostasie hydrique, la composition chimique de la cuticule peut également impacter la population microbienne présente dans la phyllosphère (Reisberg et al., 2013). Les interactions entre les métabolites spécialisés présents de manière générale sur parties aériennes de la plante et les microorganismes sont encore très largement méconnues mais suscitent un certain intérêt car ils pourraient permettre l'identification de relations positives/négatives influant sur la sélection variétale. Par contre, les rôles des métabolites spécialisés dans la partie racinaire sont mieux connus. Depuis la terrestrialisation des plantes, des symbioses entre les racines et des champignons sous la forme d'un organe mixte appelé mycorhize sont apparues. Il y a 416 à 350 million d'années, pendant le Dévonien, les racines ont évolué pour acquérir des nouvelles fonctionnalités permettant, entre autre, le développement de plantes toujours plus grandes (Kenrick and Strullu-Derrien, 2014). Ces changements majeurs ont également impliqué des acquisitions indirectes au niveau du développement et de la complexification de ces symbioses. Ce type d'association implique environ 90% des plantes (Wang and Qiu, 2006) et leur permet un apport supplémentaire en eau et en nutriments tels que l'azote et le phosphate (Parniske, 2008). Les symbioses se mettent en place suite à la sécrétion par la plante de sucres, hormones, enzymes ou des métabolites spécialisés comme des flavonoïdes (par exemple la rutine ou la quercétine). Lors de l'interaction entre L. bicolor (champignon) et P. trichocarpa (plante hôte), ces composés seraient sécrétés par les racines, transportés jusqu'au champignon pour activer l'expression du certains gènes. Une signalisation spécifique entre les deux individus va alors se mettre en place afin de diminuer la réponse immunitaire de la plante et donc faciliter la symbiose (Garcia et al., 2015). Ce type d'association peut avoir des effets très positifs sur la biomasse des plantes et donc sur le rendement, comme lors de l'apport supplémentaire en azote de la plante d'Arachis hypogaea (Arachide) inoculé avec Bradyrhibium (champignon symbiotique) (Wagner, 2011). L'étude de ces symbioses par le prisme des métabolites spécialisés pourra ainsi être un levier vers une agriculture nécessitant moins d'apport d'intrants et donc plus durable.

La colonisation du milieu terrestre a également impacté les stratégies liées aux modes de reproduction. Ainsi, la pollinisation est un type de coopération entre les plantes et les animaux mis en place il y a 298 à 252 millions d'années (période du Permien) (Bronstein et al., 2006). Cette association s'est considérablement développée il y a environ 100 million d'années (période du Jurassique) avec l'apparition des angiospermes (plantes à fleur). Les plantes à pollinisation entomophile utilisent un vaste panel de molécules pour attirer les insectes. Cette communication chimique peut être visuelle grâce à des molécules phénoliques de type flavonoïdes et/ou olfactive avec des molécules volatiles comme des terpènes. Un des exemples illustrant parfaitement la complexité de ces interactions est la pollinisation de *Ophrys sphegodes* (une orchidée) par une abeille mâle Andrena nigroaenea. Lorsque le pollen de la fleur est mature, la fleur émet une myriade de molécules volatiles (Dormont et al., 2014) dont certaines sont analogues à celles dégagées par l'abeille femelle sexuellement mature. De plus, la morphologie et la coloration de la fleur miment la forme de l'abeille femelle pour attirer le mâle. Celui-ci va alors « copuler » avec la fleur et permettre alors la pollinisation (Schiestl et al., 1999). Le plus surprenant vient ensuite car lorsque l'ovule est fécondé, la fleur émet un ester de terpène spécifique, présent chez l'abeille femelle qui n'est pas réceptive à l'accouplement (Schiestl and Ayasse, 2001). Cette stratégie permet à la fleur d'éviter des dommages potentiels, permettant ainsi aux graines de se développer correctement. Cette communication à distance à l'aide de molécules volatiles est un véritable langage à part entière dont la majorité des éléments reste encore à décrypter. Le métabolome d'un tissus/organe peut varier considérablement selon son stade de développement ou son âge comme le montre cet exemple de pollinisation. Avec les récentes avancées en chimie analytique et sa miniaturisation, il sera bientôt envisageable de réaliser des analyses globales non-ciblées de métabolites chez les plantes. Cela permettra une meilleure compréhension des « dialogues chimiques » entre la plante et son environnement mais aussi au sein de la plante elle-même.

De façon globale, les composés issus des métabolismes spécialisés sont retrouvés dans les interactions plante-environnement, des coopérations plante-champignon et -

insecte. Ils jouent également des rôles essentiels dans les mécanismes de défense de la plante contre les pathogènes et herbivores. Et comme nous l'avons vu, certains aspects peuvent être des leviers permettant l'amélioration variétale permettant de répondre aux enjeux planétaire actuels. Parmi les trois classes principales composant les métabolites spécialisés, nous développerons la classe des molécules phénoliques et plus spécifiquement la famille des phénylpropanoïdes.

II- Phénylpropanoïdes

1/- Généralités

Les phénylpropanoïdes font partie de la famille des molécules phénoliques dont le nombre est difficile à estimer à cause du peu de plantes analysées biochimiquement. Le dénominateur commun des composés phénoliques est la molécule de benzène accompagnée d'un groupement hydroxyle : le phénol (C_6) (Figure 1 en rouge). En plus de cette base, les phénylpropanoïdes contiennent un groupement propyl (C_3) (Figure 1





en bleu). Les membres de cette famille peuvent avoir des formes relativement simples lorsqu'elles sont libres (C₆-C₃) mais aussi former des polymères complexes tels que les tanins ou les lignines. Dans ce manuscrit, la distinction sera faite entre la famille des

C₆-C₃. Certains de ces composés sont ubiquitaires chez plantes vasculaires (comme les lignines et certains flavonoïdes) tandis que d'autres sont spécifiques de certains taxons (comme certains sinapoyl esters chez les *Brassicaceae* par exemple). Du fait de leur large diversité structurale, les phénylpropanoïdes sont impliqués dans un grand nombre d'interactions biotiques et abiotiques (Vogt, 2010).

On considère que le principal flux des phénylpropanoïdes provient de la phénylalanine. Cet acide aminé est lui même issu de la voie du shikimate. Les premières réactions que subit la phénylalanine constituent la voie générale des phénylpropanoïdes aboutissant à la formation du *p*-coumaroyl-CoA. Ce dernier est le composé principal d'amorçage de nombreuses voies spécifiques des phénylpropanoïdes comme les

flavonoïdes ou les précurseurs des lignines appelés monolignols (Figure 2). Ces différentes voies métaboliques ont toutes été caractérisées dans la hampe florale d'*Arabidopsis thaliana* (Vanholme et al., 2012) (Figure 3). Certaines de ces voies existent bien sûr dans d'autres organismes/organes et de ce fait, cette figure de synthèse fera office de référence pour la suite.



Figure 2 : Voie métabolique simplifiée conduisant aux différentes sous-familles de phénylpropanoïdes. La voie des acides shikimiques et chorismiques (en bleu) permet surtout l'alimentation en phénylalanine de la voie générale des phénylpropanoïdes (en rouge). Il y aura alors synthèse de molécules « activées » par l'ajout d'un groupement Coenzyme A (SCoA) comme le p-coumaroyl-CoA qui va être pris en charge par les nombreuses voies spécifiques des phénylpropanoïdes (en vert). Les flèches en pointillés symbolisent plusieurs réactions.

Nous allons nous intéresser à ces trois voies en décrivant la structure des familles de molécules ainsi que certains des gènes de biosynthèse. Cette partie n'a pas pour but d'être exhaustive, l'objectif étant de remettre la voie des phénylpropanoïdes dans un contexte global ainsi que d'appréhender la complexité de ces voies par des exemples choisis.



Figure 3 : Voies métaboliques générales des composés phénoliques dans la hampe florale d'Arabidopsis thaliana. La voie principale aboutissant aux lignines est délimitée par un trait épais. Les astérisques rouges n'ont pas d'importance ici. Les flèches continues symbolisent une réaction connue et les flêches en pointillées sont des réactions supposées. Les métabolites en noir sont présents dans la plante sauvage tandis que ceux en gris sont présents seulement chez certains mutants. AD, arogenate dehydrogenase; ADT, arogenate dehydratase; AS, anthranilate synthase; AS, anthocyanidin synthase; AT, amino transferase; ATs, acyltransferases; CHI, chalcone isomerase; CHS, chalcone synthase; CM, chorismate mutase; CS, chorismate synthase; DFR, dihydroflavonol 4-reductase; DHS, 3-deoxy-Darabino-heptulosonate 7-phosphate synthase; DHQD/ SD, 3-dehydroquinate dehydratase/shikimate dehydrogenase; DQS, 3-dehydroquinate synthase; EPSPS, 5-enolpyruvylshikimate-3-phosphate synthase; F3H, naringenin 3-dioxygenase; F3'H, flavonoid 3-hydroxylase; F6'H, feruloyl-CoA 6hydroxylase; FLS, flavonol synthase; HCALDH, hydroxy- cinnamaldehyde dehydrogenase; ICS, isochorismate synthase; IGPS, indole-3-glycerol phosphate synthase; PAI, phosphoribosylanthranilate isomerase; PAT, phosphoribosylanthranilate transferase; PD, prephenate dehydrogenase; SGT, SK, sinapate 1-glucosyltransferase; shikimate kinase; SMT, sinapoylglucose:malate sinapoyltransferase; SST, sinapoylglucose:sinapoylglucose sinapoylglucosetransferase; TSA, Trp synthase a-subunit; TSB, Trp synthase b-subunit; UGT, UDP-glucosyltransferase. D'après Vanholme et al. (2012).

2/- Voie du shikimate

Les phénylpropanoïdes dérivent de deux acides aminés différents, la phénylalanine et la tyrosine. Ils sont tous deux synthétisés à partir de la voie du shikimate issue elle-même de deux voies du métabolisme primaire : la glycolyse et la voie des pentoses phosphates représentées respectivement par l'acide phosphoénolpyruvique et l'érythose-4-phosphate (Figure 4). La conjugaison de ces deux molécules suivie de trois autres réactions enzymatiques entrainent la formation d'acide shikimique. Un des intermédiaires de cette voie peut former l'acide gallique qui constitue le squelette de base des tanins hydrolysable. Ces composés, importants pour repousser les herbivores, seront détaillés plus loin dans le manuscrit.

Figure 4 : Voies métaboliques des acides shikimiques et chorismiques conduisant aux phénylpropanoïdes. Mise en évidence d'autres voies métaboliques dérivant de celle-ci comme les tannins hydrolysables, des alcaloïdes ou l'acide salicylique. La transformation de prephenate en phénylalanine/tyrosine fait intervenir de nombreuses enzymes possédant des actions similaires nécessitant un fort niveau de régulation. ADH : arogenate dehydrogenase; ADT : arogenate dehydratase; CM : chorismate mutase; HPP-AT : 4-hydroxyphenylpyruvate aminotransferase; PDH : prephenate dehydrogenase; PDT : prephenate dehydratase; PPA-AT : prephenate aminotransferase; SPY-AT : phenylpyruvate aminotransferase ; PAL : phenylalanine ammonia lyase. Les flèches en pointillées symbolisent plusieurs réactions et les flèches continues une seule réaction.



Suite à trois nouvelles réactions enzymatiques, l'acide shikimique produit l'acide chorismique. Cette molécule est très importante car elle constitue le point de départ de voies métaboliques essentielles pour la plante comme la production de trois acides aminés aromatiques (phénylalanine, tyrosine et tryptophane) mais aussi d'acide salicylique. Ce dernier est un composé phénolique intervenant dans la transduction de signaux lors de nombreux stress abiotiques ainsi que lors d'attaques de pathogènes (Vicente and Plasencia, 2011). Sa synthèse peut également se faire plus en aval dans la voie générale des phénylpropanoïdes à partir de l'acide cinnamique (Figure 3 et Figure 4). Sa synthèse par ces deux voies suggère une régulation complexe de son homéostasie et sans doute un dialogue possible entre les différentes voies de production (Chen et al., 2009b). La synthèse de l'acide shikimique se fait entièrement dans le chloroplaste, et une fois synthétisés, les trois acides aminés aromatiques sont transportés dans le cytoplasme pour y être soit polymérisés en peptides ou subir de nouvelles transformations aboutissant à de nombreux métabolites secondaires.

Ainsi, la production de phénylpropanoïdes provient très majoritairement de la phénylalanine mais aussi, dans une certaine mesure, à cause de sa structure très proche, de la tyrosine, qui conduit vers des phénylpropanoïdes spécifiques tels que les acides hydroxycinnamiques et phénolamides (Tzin and Galili, 2010).

3/- La voie générale des phénylpropanoïdes

La phénylalanine est le point de départ à de la synthèse de nombreux métabolites secondaires tels que les stilbènes, les flavonoïdes ou les monolignols polymérisés en lignines que nous allons détailler par la suite (Figure 2 et Figure 3). Nous allons décrire certains gènes importants dans la biosynthèse du *p*-coumaroyl-CoA qui est l'intermédiaire commun des différents phénylpropanoïdes et de leurs dérivés. Ainsi, la phénylalanine subit d'abord une désamination par la phénylalanine ammonia-lyase (PAL). L'acide cinnamique ainsi formé est hydroxylé en position 4 du groupement aromatique par la cinnamic acid 4-hydroxylase (C4H) et enfin ce dernier est estérifié par la 4-coumarate:CoA ligase (4CL) (Figure 5).



Figure 5: La voie générale des phénylpropanoïdes (encadré en rouge) est alimentée par les voies du shikimate/chorismate (encadré en bleu). Les flèches en pointillées symbolisent plusieurs réactions et les flèches continues une seule réaction. PAL : phénylalanine ammonia-lyase ; TAL : tyrosine ammonia-lyase ; C4H : cinnamate 4-hydroxylase ; 4CL : 4-coumarate:CoA ligase.

3-1/- Phenylalanine Ammonia-Lyase (PAL).

La plupart des plantes possèdent plusieurs isoformes de PAL entrainant une possible redondance fonctionnelle mais qui, *in fine*, n'aboutit pas forcément aux mêmes produits. Par exemple chez *Populus tremuloides Michx*, PtPAL1 serait impliqué dans la synthèse des tanins condensés tandis que PtPAL2 est associé à la lignification (Kao et al., 2002). Chez Arabidopsis, il existe 4 gènes *PAL* (Raes et al., 2003) et de façon étonnante lorsqu'elles sont toutes inactivées, la plante peut tout de même se développer malgré une forte perturbation du phénotype (Huang et al., 2010). Ce quadruple mutant de taille extrêmement réduite, possède un caractère de stérilité mâle et malgré une diminution importante de sa teneur en lignines, ses hampes ont encore un port dressé. De plus, il réside une faible activité PAL dans les extraits protéiques de ce mutant suggérant une redondance fonctionnelle avec d'autres gènes restant à identifier.

La tyrosine et la phénylalanine ne différent que d'un groupement hydroxyle et peuvent donc être parfois pris en charge par les mêmes enzymes. La désamination de la tyrosine conduit directement à l'acide *p*-coumarique en court-circuitant l'activité C4H. Cette activité Tyrosine Ammonia-Lyase (TAL) (Figure 5) est importante chez les monocotylédones mais a également été détectée chez certaines dicotylédones (Barros et al., 2016; Hsieh et al., 2010; Rosler et al., 1997). Certaines PALs peuvent également utiliser la tyrosine comme substrat et sont, dans ce cas, appelées PTAL. Le rôle de ces enzymes bi-fonctionnelles n'est pas encore bien connu mais il est possible qu'elles participent à la régulation de l'incorporation de monomères dits non-conventionnels dans les lignines, notamment de *p*-coumaroyl-ester chez les graminées (Barros et al., 2016).

3-2/- Cinnamic acid 4-hydroxylase (C4H)

Tandis que la réaction PAL est réversible, la deuxième étape de la voie catalysée par C4H est considérée comme irréversible et donc constitue un important point de régulation du flux des phénylpropanoïdes généraux (Anterola and Lewis, 2002). C4H appartient à la famille des cytochromes P450 au même titre que les p-coumaroyl shikimate 3' hydroxylases (C3'H) et les ferulate-5-hydroxylases (F5H), intervenant toutes les deux plus en aval dans biosynthèse des phénylpropanoïdes. Elles ont aussi la particularité d'être ancrées dans la membrane du réticulum (Bayburt and Sligar, 2002; Karamat et al., 2012). Arabidopsis ne possède qu'un seul gène C4H. Son inactivation totale n'est pas viable pour la plante. Néanmoins, des mutants EMS ref (reduced epidermal fluorecence) avant des phénotypes plus ou moins forts selon la position et la nature de la mutation sur le gène ont été identifiés (Ruegger and Chapple, 2001; Schilmiller et al., 2009). Certains d'entre eux s'apparentent au quadruple mutant d'Arabidopsis pal1/2/3/4 par leur phénotype nain, la diminution des quantités de lignines et la stérilité mâle (Yoon et al., 2015). Chez d'autres espèces, C4H est retrouvé sous la forme d'une famille multigénique. Dans ce cas, au moins un des gènes peut être associé à la synthèse de lignines tandis que les autres sont vraisemblablement inductibles par certains stress (Carocha et al., 2015; Lu et al., 2006).

Il a été montré que C4H s'associerait avec certaines isoformes de PAL pour une action concomitante permettant d'alimenter les flux métaboliques par un mécanisme de canalisation métabolique (metabolic channeling). Cette association a été suggérée par le fait qu'un ajout exogène d'acide cinnamique radiomarqué ne diminue pas au cours du temps dans une suspension de cellules de tabac et n'est donc pas métabolisé (Rasmussen and Dixon, 1999). De plus, des expériences de co-localisation de deux isoformes de PAL (NtPAL1 et NtPAL2) avec NtC4H ont montré chez *Nicotiana tabacum* des associations localisées sur des endomembranes entre NtPAL1 et NtC4H, tandis que NtPAL2 serait vraisemblablement cytosolique (Achnine et al., 2004). L'hypothèse est donc que dans le cas de l'association PAL1-C4H, il y aura directement synthèse de l'acide *p*-coumarique qui sera utilisé dans la voie des phénylpropanoïdes. Par contre dans le cas de l'action de PAL2, c'est l'acide cinnamique que la plante produira pour alimenter les voies menant à l'acide salicylique ou des coumarines par exemple (Achnine et al., 2004).

3-3/- 4-coumarate:CoA ligase (4CL)

La famille multigénique 4-coumarate:CoA ligase a été historiquement nommée ainsi à cause de son activité préférentielle *in vitro* pour l'acide *p*-coumarique. La réaction catalysé par 4CL est très importante car elle intervient pour la formation de :

- *p*-coumaroyl-CoA qui est l'intermédiaire commun pour de nombreux phénylpropanoïdes et dérivés (Figure 2 et Figure 3)
- d'autres esters hydroxycinnamoyl-CoA, qui sous cette forme pourront entrer principalement dans la voie spécifique des monolignols (Figure 3)

De plus, les différentes isoformes peuvent être distinguées selon leurs activités enzymatiques, leur profil d'accumulation et leur phylogénie (Ehlting et al., 1999). De manière générale, une classe de 4CL est associée aux lignines et d'autres sont plutôt associées aux flavonoïdes et à d'autres molécules phénoliques solubles comme cela a été démontré chez certaines espèces telles qu'*Arabidopsis thalina* (Ehlting et al., 1999; Li et al., 2015), *Oryza japonica* (Sun et al., 2013a) ou *Morus notabilis* (Wang et al., 2016) par exemple.

La fonction de ces enzymes en termes de spécificité de substrats peut, par contre, être variable entre les espèces. Par exemple chez *Arabidopsis thaliana* ou *Magnolia kobus DC.*, l'activité contre l'acide sinapique n'est pas détectable par incorporation de précurseurs radiomarqués (Costa et al., 2005; Yamauchi et al., 2003) tandis que chez le robinier (*Robinia pseudoacacia L*), le laurier (*Nerium indicum Mill.*) et le soja (*Glycine max L.*), il est majoritaire (Lindermayr et al., 2002; Yamauchi et al., 2003). Après absorption d'acide sinapique radioactif par les plantes, seules celles ayant une activité 4CL spécifique de ce substrat présentent un marquage dans les lignines (Yamauchi et al., 2003). La différence dans la spécificité de substrats peut s'expliquer par des différences d'acides aminés clés dans la structure de l'enzyme (Schneider et al., 2003).

4/- Les voies spécifiques

Dans cette partie, nous allons nous intéresser aux différentes familles de molécules dérivant de la voie générale des phénylpropanoïdes. Elle sont de deux sortes :

les formes simples qui contiennent de 1 à 3 cycles benzoïques et les formes complexes telles que les tanins et les lignines.

4-1/- Les formes simples

a/- Les acides hydroxycinnamiques (AHC)

Les acides hydroxycinnamiques sont retrouvés chez toutes les plantes vasculaires plutôt sous forme de dérivés que sous leurs formes libres *in vivo* (Liu, 2010). Ils sont activés grâce à l'ajout d'un groupement SCoA pour devenir des précurseurs de nombreuses molécules, principalement les flavonoïdes et les lignines (Vogt, 2010) Bien qu'en proportion plus modeste, la conjugaison d'un AHC avec d'autres molécules organiques est d'importance majeure pour le développement de la plante ainsi que pour sa protection lors de stress biotiques et abiotiques (Liu, 2010). Dans cette partie nous allons voir en détail les deux conjugaisons majeures aboutissant aux esters et amides hydroxycinnamoyliques.

α /- Les esters hydroxycinnamoyliques

Cette famille de molécules provient de l'association d'une molécule activée d'AHC-CoA et de molécules acceptrices telles que des carbohydrates ou des acides organiques (Kroon and Williamson, 1999). Deux esters d'AHC synthétisés en grande quantités dans certains taxons de plantes vasculaires sont les acides chlorogénique et rosmarinique issues respectivement de la conjugaison avec un acide quinique et le 3,4dihydroxyphenyllactic (Liu, 2010; Petersen and Simmonds, 2003) (Figure 6). Ils sont produits par plusieurs organes de la plante et possèdent entre autre des rôles de protection contre les UV et les pathogènes (Clé et al., 2008; Leiss et al., 2009; Sánchez-Campillo et al., 2009). D'autres esters d'AHC sont eux accumulés spécifiquement dans les grains des céréales (riz, blé, orge) comme le campestanyl ferulate (Figure 6), un ester de stérol (Hakala et al., 2002), dont les fonctions *in planta* restent peu connus mais qui joue un rôle antioxydant et anticancéreux chez l'homme (Patel, 2012).



Figure 6 : Trois exemples d'esters d'AHC. La partie issue des phénylpropanoïdes (en noir) peut être associée à différents types de molécules (en rouge)

Les BAHD sont des enzymes catalysant un grand nombre de molécules activées par un groupement thioester CoA conduisant à la modification des acides ou alcools à chaines courtes ou longues chez les plantes (D'Auria, 2006). L'acronyme BAHD provient du nom de 4 enzymes de la même famille : benzyl alcohol *O*-acetyltransferase (BEAT), anthocyanin O-hydroxycinnamoyltransferase (AHCT), Nhydroxycinnamoyl/benzoyltransferase (HCBT) et deacetylvindoline 4-0acetyltransferase (DAT). Un des exemples les plus décrits est la HCT (Hydroxycinnamoyl-Coenzyme A Shikimate/Quinate Hydroxycinnamoyltransferase) grâce à son rôle dans la biosynthèse des lignines. Elle réalise la réaction réversible de conjugaison du hydroxycinnamoyl-CoA avec l'acide shikimique et/ou quinic pour donner respectivement l'hydroxycinnamoyl-shikimate et l'hydroxycinnamoyl-guinate in vitro (Figure 3), avec une préférence pour le shikimate (Hoffmann et al., 2003). Dans la voie des phénylpropanoïdes elle agit avant et après l'action de C3'H qui permet une hydroxylation en position 3 du groupement phényl (Figure 3). La réduction de son expression chez Medicago sativa, Nicotiana benthamiana ou Arabidopsis thaliana entraine une baisse de la quantité de lignines et une modification de leur composition ainsi qu'un déclin des teneurs en molécules phénoliques (Hoffmann et al., 2004; Shadle et al., 2007). Parce que ces plantes ne peuvent plus synthétiser les unités G et S correctement, leurs lignines sont enrichies en unités H, en accord avec la voie de biosynthèse des monolignols (Figure 3).

Une autre enzyme de la famille des BAHD, la HQT (Hydroxycinnamoyl-CoA: Quinate hydroxycinnamoylTransferase), conduit à la production d'acide chlorogénique (cafféoylquinate) (Escamilla-Treviño et al., 2014). Elle greffe spécifiquement le caffeoyl-CoA sur l'acide quinique pour synthétiser l'acide chlorogénique. De manière intéressante, sa modulation par RNAi et par sur-expression ne provoque pas de changement dans les lignines chez *Solanum lycopersicon, Solanum tuberosum* ou *Panicum virgatum* (Escamilla-Treviño et al., 2014; Niggeweg et al., 2004; Payyavula et al., 2015). Cette enzyme utilise le même substrat (à savoir le *p*-coumaroyl-CoA) que l'enzyme HCT pour la lignification et est exprimé dans l'ensemble de la plante incluant les tissus vasculaires (Payyavula et al., 2015). L'hypothèse actuelle est que la HCT utilise préférentiellement le caffeoyl-shikimate conduisant à la production de monolignols tandis que les hydroxycinnamoyl-quinates sont utilisés dans d'autres voies ou simplement compartimentés hors de portée des enzymes menant aux lignines.

β /- Les amides hydroxycinnamoyliques ou phénolamide

Tout comme les esters hydroxycinnamoyliques, les phénolamides sont issus d'un hydroxycinnamoyl-CoA. Celui-ci sera cette fois conjugué à un amide qui peut être aliphatique ou aromatique. Les phénolamides ainsi formés jouent des rôles assez variés dans la résistance aux pathogènes, la tolérance au froid ou dans certaines phases de développement telles que la senescence ou la reproduction (Macoy et al., 2015). Certains phénolamides comme le feruoyl-tyramine et le *p*-coumaroyl-tyramine sont accumulés lors d'infections ou de blessures (Ishihara et al., 2000; Keller et al., 1996). Ils seraient synthétisés dans le cytoplasme puis transportés dans la paroi afin d'y être polymérisés pour former une barrière physique contre les pathogènes (Hagel and Facchini, 2005; Ishihara et al., 2008). Leur production peut être catalysée par des hydroxycinnamoyl-N-hydroxycinnamoyltransferases (THTs) CoA: tyramine ou sérotonine Nhydroxycinnamoyltransferases (SHT) (Figure 7). Elles sont retrouvées chez les Solanaceaes, Papaveraceaes et Gramineaes (Jang et al., 2004). Elles utilisent principalement deux amines aromatique : la tyramine (issue de la voie de la tyrosine) et/ou la sérotonine (issues de la voie du tryptophane) (Figure 7). Ces enzymes sont fortement accumulées lors d'infections ou de stress dus aux UV (Hagel and Facchini, 2005; Schmidt et al., 1998, 1998). L'élicitation de cultures cellulaires de Solanum tuberosum et de Nicotiana tabacum, provoque la sur-expression de THT, mais induit aussi l'expression des gènes tels que les *PAL* et *4CL* qui permettent la synthèse de substrats de THT (Schmidt et al., 1998; Villegas et al., 1990). Les plantes sur-exprimant des gènes *THT* n'ont pour l'instant pas montré d'impact sur les lignines *in vivo* (Campos et al., 2014; Jang et al., 2004) ce qui suggère l'existence d'une voie indépendante.



Figure 7 : Synthèse de phénolamides par l'action de l'hydroxycinnamoyl-CoA: tyramine Nhydroxycinnamoyltransferases et sérotonine N-hydroxycinnamoyltransferases (THT et SHT). Il existe de nombreuses autres combinaisons décrites de façon plus exhaustive dans Macoy et al. (2015)).

γ /- Les acides hydroxycinnamiques associés avec des polymères pariétaux

Historiquement retrouvés chez les monocotylédones, les HCAs peuvent se lier à certains polymères de la paroi (Ishii, 1997). En effet, l'acide *p*-coumarique et férulique peuvent être assemblés par des liaisons covalentes aux arabinogalactanes, aux lignines et aux protéines structurales sous l'action de certaines péroxydases (Burr and Fry, 2009; Piber and Koehler, 2005). L'acide férulique se retrouve également associé aux parois de nombreux dicotylédones et gymnospermes (Carnachan and Harris, 2000; Ralph et al., 2008). Il serait principalement à l'origine des liaisons entre les polymères de pectines et les protéines structurales (Levigne et al., 2004). Par contre l'acide *p*-coumarique, qui peut composer jusqu'à 20% de lignines de *Zea mays* (Hatfield and Chaptman, 2009) se lie par une liaison ester principalement aux unités S des lignines, suggérant que celui-ci

pourrait jouer un rôle à part entière de précurseur de leur biosynthèse (Ralph et al., 2008). Ainsi, différents rôles pour ces HCA liés aux parois ont été proposés (Buanafina, 2009) notamment :

- la diminution de l'extensibilité des parois lors de la maturation des cellules
- la régulation de la croissance cellulaire
- l'adhésion cellule-cellule
- la prévention et la défense lors d'infections par des pathogènes
- un site de nucléation pour la synthèse de lignines chez les graminées

L'incorporation de ces unités d'AHC au sein des lignines serait réalisée par une voie spécifique. En effet, ces acides sous leurs forme d'ester de Coenzyme A sont transférés sur un monolignol par une enzyme de type BAHD (Voxeur et al., 2015). La sur-expression d'un gène *BAHD* chez *Oryza sativa* provoque une augmentation de 300% de l'incorporation d'acide *p*-couramique dans la paroi (Bartley et al., 2013). Mais cette spécificité n'est pas retreinte à quelques taxons, comme le prouve la détection d'acide férulique au sein des lignines dans un grand nombre d'angiospermes (Karlen et al., 2016). Les auteurs suggèrent même d'ajouter l'acide férulique comme un monomère de lignines au même titre que les trois monolignols « classiques ».

b/- Les monolignols

Les monolignols sont définis comme étant les (principaux) monomères des lignines, et de ce fait, sont indispensables au développement normal des plantes. Leur voie de biosynthèse est relativement bien décrite avec des enzymes issues d'une dizaine de familles généralement multigéniques (Figure 3). Des auteurs ont montré que les taux de lignines mesurés chez les plantes pourraient être obtenus avec seulement 10 à 50% des enzymes produites (Wang et al., 2014b). Ces résultats pourraient expliquer la difficulté d'obtenir des plantes transgéniques avec des quantités de lignines réduites (Wang et al., 2014b). Néanmoins, il existe tout de même de nombreux mutants pour des gènes de biosynthèse des monolignols qui ont une lignification altérée (Baxter and Stewart, 2013; Vanholme et al., 2010). L'homéostasie des phénylpropanoïdes de ces mutants va être perturbée, mais pour certains il y a également des dérégulations d'autres voies de biosynthèses. On peut ainsi citer le mutant *Atccr1* ou *Atcomt* qui vont accumuler respectivement du tryptophane et de l'acide salicylique (Vanholme et al., 2012). De manière générale, les mutants déficients en lignines vont connaitre :

- une accumulation de certains phénylpropanoïdes toxiques de manière dosedépendent pour la plante (Le Roy et al., 2016)
- une altération de l'intégrité de leurs parois cellulaires, et de ce fait, des anomalies de croissance et de développement
- des changements d'accumulation de certains métabolites essentiels (tryptophane et acide salicylique)

La voie de biosynthèse des monolignols sera rapportée de façon plus exhaustive dans l'article Le Roy et al., (2017) présenté dans le chapitre I de la partie "résultats".

c/- Les coumarines

La structure commune qui compose toutes les coumarines est un benzène couplé à un noyau lactone (Bourgaud et al., 2006). La forme la plus simple se nomme la coumarine et en fonction de ses ornementations, il existe quatre familles dont la plus décrite est celle des furocoumarines. Les coumarines sont spécifiques de certaines familles de plantes (*Rutaceae, Umbelliferae, Clusiaceae, Guttiferae, Caprifoliaceae, Oleaceae, Nyctaginaceae* et *Apiaceae*) et peuvent être synthétisées dans tous les organes (Venugopala et al., 2013). De nombreuses études montrent que ces molécules ont des propriétés antibiotiques et antifongiques et qu'elles peuvent également inhiber la germination de graines dans un contexte de compétition environnementale (Razavi, 2011).

La synthèse des coumarines se réalise à différents points de la voie des phénylpropanoïdes (Figure 8). Par exemple, la molécule de coumarine est dérivée de l'acide cinnamique qui subit une hydroxylation par une cinnamic acid 2-hydroxylase (C2H) suivie d'une lactonisation spontanée (Bourgaud et al., 2006) alors que l'umbelliferone et la scopoletine sont issues respectivement du *p*-coumaroyl-CoA et du feruoyl-CoA (Kai et al., 2008).



Figure 8 : Voie de biosynthèse simplifiée des coumarines à partir de la voie générale des phénylpropanoïdes (en rouge). PAL : phenylalanine ammonia lyase ; C2H : cinnamate/coumarate 2-hydroxylase ; F6H : Ferulate 6-Hydroxylase.

d/- Les lignanes et neo-lignanes

Les (néo)lignanes sont issus d'une dimérisation de deux phénylpropanoïdes de type C6-C3 chez un grand nombre de plantes (Suzuki and Umezawa, 2007; Umezawa, 2003). Cette famille compte plus de 3000 membres et molécules associées tels que les sesqui- ou flavano-lignanes dans le cas de conjugaison avec d'autres (néo)lignanes et flavonoïdes respectivement (Schmidt et al., 2010). Cette diversité vient des différents monomères utilisés et également d'une grande variabilité stéréochimique.

De manière générale, peu de choses sont connues sur leur fonctions *in planta*. Chez les ligneux, elles préviendraient des attaques fongiques et auraient une activité phytoalexine contre d'autres plantes non-ligneuses (Pohjamo et al., 2003). Par contre, on connaît bien leurs innombrables bienfaits sur l'homme notamment en tant que molécule anticancéreuse et antioxydante (Kitts et al., 1999). De ce fait, la connaissance de leurs voies de biosynthèse est très importante car une fois les gènes caractérisés, il est possible de modifier par bio-ingénierie l'accumulation de (néo)lignanes spécifiques chez certaines espèces (Satake et al., 2013). L'accumulation de (neo)lignanes au sein de taxons ou d'organes spécifiques suggère que les voies de biosynthèse sont finement régulées. Certaines d'entre-elles permettent une accumulation à de très fortes concentrations comme le secoisolariciresinol sous forme diglucoside (SDG) dans les graines de lin (Meagher et al., 1999) ou le pinorésinol et la sésamine dans les graines de sésame (Dar and Arumugam, 2013).



Figure 9 : Trois exemples de lignanes issues de la fusion d'au moins deux molécules d'alcool coniférylique. Le pinorésinol est un lignane, l'alcool dehydrodiconiféryle fait partie des néo-lignanes et le ligusinenoside est un oxyneolignane identifié chez Ligusticum sinensis (Ma et al., 2011).

On distingue les lignanes des néolignanes par la nature de leur assemblage (Figure 9). En effet, les lignanes s'associent par les deux groupement propyl en β - β (couplage queue-queue) (comme le pinorésinol ; Figure 9) tandis que les néolignanes sont formés par l'association des groupements phényl de l'un et propyl de l'autre (couplage tête-queue) ou par les deux groupement phényl (couplage tête-tête) (comme l'alcool dehydrodiconiférylique ; Figure 9) (Moss, 2009). Il existe également des

oxyneolignanes qui sont issues d'une liaison carbone-oxygène-carbone. Cette dernière famille est peu étudiée comparée aux deux autres mais peut être synthétisée par certaines plantes telles que *Ficus hispida L*. (Yap et al., 2015) ou *Juniperus brevifolia* (Seca and Silva, 2010) (comme le ligusinenoside ; Figure 9).

Il a été montré que la conjugaison des deux phénylpropanoïdes fait intervenir des protéines dirigentes qui contrôleraient régio- et stéréo-chimiquement ce couplage radicalaire (Davin and Lewis, 2000; Gang et al., 1999). Ces protéines dirigentes ont été identifiées pour la première fois chez Forsythia, ce qui a permis d'expliquer l'accumulation exclusive du stérioisomère (+)-pinorésinol (Davin et al., 1997). Il a également été montré que l'incubation des protéines recombinantes FiDIR1 et AtDIR6 en présence d'une laccase et d'alcool coniférylique conduit à la production préférentielle de la forme (+)- ou de la forme (-)-pinorésinol (Figure 10) (Pickel et al., 2010). La sécrétion de ces protéines vers l'apoplaste a pu être déterminée *in silico* par la présence d'un peptide signal (Pickel and Schaller, 2013) mais aussi de manière expérimentale (Davin and Lewis, 2000; Gang et al., 1999). Peu de choses sont connues sur la dimérisation des (néo)lignanes mais la réaction pourrait être catalysée par des peroxydases de classe III localisées dans les parois (Frías et al., 1991; Wang et al., 2013a). Ainsi, par une approche d'immunolocalisation, le secoisolariciresinol retrouvé chez la graine de lin a été localisé majoritairement dans les parois et, dans une moindre mesure, dans le cytosol (Attoumbré et al., 2010).



Figure 10 : Couplage préférentiel d'alcools coniféryliques sous l'action de protéines dirigentes. FvDIR1 favorise la synthèse de l'énantiomére (+)-pinorésinol tandis que AtDIR6 favorise l'énantiomère (-)-pinorésinol. D'après Pickel et al. (2010).

A la suite de ce couplage, le dimère va être modifié par différentes enzymes. Le pinoresinol-lariciresinol reductase (PLR) va permettre la réduction du (-)-pinorésinol

en (-)-lariciresinol et finalement en (+)-secoisolaricirésinol (Figure 11) (Hano et al., 2006). Il intervient juste après le couplage radicalaire et son inactivation provoque des changements importants dans le contenu en lignanes dans la graine de lin (Renouard et al., 2014). Chez le lin, les graines contiennent un mélange de lignanes de stéréochimie opposée, à savoir les (-) et (+)-secoisolariciresinol, tandis que les parties aériennes accumulent seulement des lignanes dérivant de l'énantiomère (-)-secoisolariciresinol (Hemmati et al., 2010). Cette différence serait due à l'action de deux PLRs : LuPLR1 et LuPLR2 qui agissent respectivement sur le (-)-pinorésinol et le (+)-pinorésinol. Les molécules pouvant ensuite être modifiées par des glycosyltransferases ou des oxydases/réductases. Si on considère que la dimérisation a lieu dans l'apoplaste, les (néo)lignanes ainsi formés doivent de nouveau rentrer dans le cytoplasme pour subir les autres réactions à l'instar de la glycosylation qui est effectuée par une UDP-Glycosyltransférase (UGT) cytosolique (Wang et al., 2013b).



Figure 11 : Biosynthèse de lignanes. DP, dirigent protein; PLR, pinoresinol/lariciresinol reductase; PrR, pinoresinol reductase ; SIRD, secoisolariciresinol dehydrogenase. D'après Nakatsubo et al. (2008).

e/- Les flavonoïdes

Parmi les familles de molécules issues de la voie des phénylpropanoïdes, les flavonoïdes représentent probablement la plus importante (Ferrer et al., 2008). A cause de leurs couleurs facilement identifiables, certains d'entre eux sont à l'origine de découvertes majeures en génétique comme les lois de ségrégation de Mendel, le phénomène d'épistasie ou l'identification des éléments transposables (Winkel-Shirley, 2001). Comme cela a été évoqué dans l'introduction, ces pigments sont déterminants dans le recrutement des pollinisateurs et la dissémination des graines chez de nombreuses plantes. Les flavonoïdes peuvent aussi jouer un rôle important dans les interactions plante-microorganisme, la fertilité mâle, la défense contre des pathogènes et la protection contre les UV (Ferrer et al., 2008). Leur structure générale en C15 inclut deux cycles benzoïques (A et B) reliés entre eux par une chaine C3 qui du fait de sa

structure, peut former un troisième cycle (C) (Figure 12). Cette famille regroupe une dizaine de sous-familles telles que les flavones (comme la flavone ou la tricine), les



flavonoles (comme la quercetin ou la rutine) ou les anthocyanidols (telles que la petunidine ou la cyanidine). Ces sous- familles se distinguent par le nombre d'hydroxylations, de méthoxylations et de glycosylations (Kumar et al., 2013).

Figure 12 : Structure générale des flavonoïdes avec la nomenclature usuelle des trois cycles A, B et C.

La chalcone synthase (CHS) est la première enzyme de cette voie de biosynthèse. Elle va faire réagir trois molécules de malonyl-CoA avec le *p*-coumaroyl-CoA pour former la chalcone (Figure 13). C'est une enzyme ubiquitaire parmi les plantes vasculaires et son action est souvent suivie d'une réaction d'isomérisation par la CHalcone Isomerase (CHI) pour former le troisième cycle C. Ensuite une série de transformations faisant intervenir des réductases, hydroxylases, glycosyltransférases et des acyltransférases permet d'aboutir à une myriade de flavonoïdes différents (Winkel-Shirley, 2001).

Des mutants d'Arabidopsis souvent obtenus par l'EMS, caractérisés par leurs déficiences/altérations de la pigmentation de la graine, sont appelés *tt* pour *transparent testa* (Shirley et al., 1995) et *tds* pour *tanin deficient seed* (Abrahams et al., 2002). Il en existe une vingtaine à ce jour et ils ont permis d'assez bien comprendre la voie de biosynthèse de certaines sous-familles de flavonoïdes (flavonol glycosides, anthocyanins et proanthocyanidins ou tanins condensés) grâce aux nombreux gènes à une seule copie (Appelhagen et al., 2014) (Figure 14).


Figure 13 : Actions de la chalcone synthase (CHS) et de la stilbène synthase (STS) après la conjugaison de pcoumaroyl-CoA et de trois molécules de malonyl-CoA. La voie générale de phénylpropanoïdes est encadrée en rouge. Trois exemples communs de modification de stilbènes sont indiqués. D'après Vannozzi et al. (2012)



Figure 14 : Panel de mutants d'Arabidopsis perturbés dans la synthèse de certains flavonoïdes contenus dans la graine. D'après Appelhagen et al. (2014).

f/- Les stilbènes

Ce type de molécule se compose de deux cycles benzéniques reliés entre eux par un pont éthylène (en C6-C2-C6). Il existe deux isoformes selon que la double liaison centrale est *cis* ou *trans* (Figure 15). Cette famille est accumulée de façon limitée et hétérogène au sein du règne végétal., On les retrouve essentiellement chez les *Vitaceae* qui comprend la vigne (*Vitis vinifera L.*) mais aussi chez les *Dipterocarpaceae*, *Gnetaceae* et *Fabaceae* (Rivière et al., 2012). Les stilbènes joueraient des rôles dans la défense en prévenant des attaques de champignons, bactéries, nématodes et herbivores (Chong et al., 2009; Jeandet et al., 2002). Il a également été démontré que le resvératrol et ses dérivés ont des activités anti-tumorales, anti-oxydantes ou neuroprotectrices (Jeandet et al., 2010) et sont donc utilisés dans le domaine de la santé.



Figure 15 : Structures chimiques des trans (E)- et cis (Z)-resvératrol.

La première étape de biosynthèse des stilbènes est réalisée par la stlibène synthase (STS) qui permettra la synthèse de resvératrol à partir de *p*-coumaroyl-CoA, mais selon la nature des précurseurs d'hydroxycinnamoyl-CoA utilisés, on retrouve d'autres intermédiaires (Chong et al., 2009). Cette enzyme possède une forte homologie de séquence avec la CHS chez *Arachis hypogea* avec environ 70% d'homologie protéique (Schröder et al., 1988) et probablement pour cette raison, les deux enzymes utilisent les même substrats (Figure 13). Par la suite, il peut y avoir des di- tri-mérisations, glycosylations, hydroxylations ou méthoxylations faisant intervenir de nombreuses enzymes qui sont à l'origine de la grande diversité des stilbènes (Chong et al., 2009) (Figure 13).

g/- Phenylpropènes

Parmi les molécules organique volatiles émises par la plante, celles qui dérivent de la phénylalanine appartiennent à la seconde famille la plus abondante derrière les terpénoïdes (Muhlemann et al., 2014). Les phenylpropènes, qui en font partie, ont la même structure en C6-C3 que les phénylpropanoïdes mais se distinguent par leur groupement propyl aliphatique (constitué seulement de carbones et d'hydrogènes) contenant une double liaison. Ce groupe de molécules intervient principalement dans la défense contre des animaux et des microorganismes ainsi que dans l'attraction de certains pollinisateurs (Koeduka et al., 2006). Le lieu de synthèse de ces molécules est de ce fait fortement dépendant de l'espèce considérée. Chez le basilic (*Ocimum basilicum*), ils sont accumulés dans les trichomes glandulaire des feuilles (Iijima et al., 2004) tandis que chez le pétunia (*Petunia hybrida*) ces composés se trouvent au niveau des fleurs (Verdonk et al., 2003).



Figure 16 : Exemple de biosynthèse de deux phénylpropènes : l'isoeugénol et l'eugénol. CFAT : coniferyl alcohol acetyltransferase; EGS : eugenol synthase; IGS : isoeugenol synthase

La voie de biosynthèse des phenylpropènes croise celle des monolignols chez le pétunia car la diminution de l'expression du gène *PhCCR1* impacte à la fois les lignines et les phenylpropènes (Muhlemann et al., 2014) La synthèse de ces molécules peut avoir lieu à partir de l'alcool coniférylique (Figure 16) mais il existe des voies utilisant d'autres monolignols tels que l'alcool *p*-coumarylique (Vassão et al., 2007). Les enzymes de biosynthèse sont spécifiquement exprimés dans les tissus floraux chez le pétunia, très pauvre en lignines (Dexter et al., 2007). Il n'y a donc typiquement pas de compétition pour les monolignols au sein de cet organe.

4-2/- Les formes complexes

a/- Les tanins hydrolysable et condensés

Les tanins sont des polyphénols qui ont la particularité de précipiter les protéines. On les trouve principalement dans les tissus épidermiques et les écorces des ligneux (Seigler, 1998). On distingue deux types de tanins par leur structure chimique : les tanins condensés (ou proanthocyanidines) dérivant des flavonoïdes et les tanins hydrolysés (ou gallotanins et ellagitanins) dérivant de la voie du shikimate (Figure 2 et Figure 17). Alors que les tanins condensés ont été identifiés chez de nombreux ptéridophytes, gymnospermes et angiospermes, les tanins hydrolysables sont présents chez un nombre limité d'eudicotylédones (Seigler, 1998). Ces derniers ne dérivant pas des phénylpropanoïdes ils ne seront pas détaillés ici. Les tanins condensés sont retrouvés chez Arabidopsis thaliana au niveau des téguments des graines et possèdent un rôle de protection de l'embryon contre des stress biotiques et abiotiques (Debeaujon et al., 2000). Comme évoqué pour les flavonoïdes, du fait du caractère épistatique des gènes de biosynthèse couplé avec un phénotype visuel (coloration des tégument des graines), la voie de biosynthèse des tannins condensés est relativement bien connue (Xu et al., 2014). Les précurseurs des tannins condensés migrent dans la vacuole grâce à des transporteurs comme MATE2 chez Medicago truncatula (Zhao et al., 2011). Mais à ce jour, peu de choses sont connues concernant leur polymérisation (Dixon et al., 2013).





b/- Cutine et subérine

Les AHC peuvent s'associer à des molécules lipidiques pour former deux types des polyesters: la cutine et la subérine. Comme évoqué précédemment, la synthèse de ces molécules a été un point clé dans la terrestrialisation des plantes. Ils jouent des rôles importants dans les échanges gazeux et liquides et forment une barrière physique pour prévenir l'attaque de pathogènes (Pollard et al., 2008).

La subérine se différencie de la cutine par sa localisation. Elle est présente sur la surface des racines, des écorces et des tubercules ou dans des tissus plus internes tels que l'endoderme, les zones d'abscission et de déhiscence des feuilles, les fruits ou les graines (Vishwanath et al., 2014). Elle est constituée de chaines lipidiques hydroxylées

en C16 à C24 et contrairement à la cutine, contient de hautes teneurs en glycérol et en molécules phénoliques, spécialement de l'acide férulique mais également des acides sinapique et *p*-coumarique (Vishwanath et al., 2014) (Figure 18). Les teneurs en phénylpropanoïdes de la subérine peuvent atteindre jusqu'à 10% de monomères insérés (Pollard et al., 2008).



Figure 18 : Modèle de représentation de subérine chez Solanum tuberosum *avec un lien potentiel avec les parois primaires. D'après Bernards (2002).*

La cutine est le principal constituant de la cuticule (de 40 à 80% de sa masse). Il est retrouvé sur les surfaces des parties aériennes des plantes comme les tiges, feuilles, fruits ou graines (Heredia, 2003). Sa structure varie selon le stade de développement et de l'organe considéré mais il se compose principalement d'acides gras hydroxylés C16 et C18 qui sont associés entre eux par des liaisons esters (Pollard et al., 2008). Les principaux AHC retrouvés sur la cuticule des feuilles sont des acides féruliques (Pollard et al., 2008) tandis que pour les fruits ce sont les acides *p*-coumarique et *m*-coumarique (Isaacson et al., 2009).

L'association des AHC à des acides gras chez Arabidopsis est due à une enzyme appartenant à la famille des BAHD. Elle a été caractérisée *in vitro* pour sa capacité à associer des hydroxycinnamoyl-CoA à des chaines lipidiques C16 (Gou et al., 2009). Lorsque cette enzyme est éteinte, il y a une diminution de la quantité d'acide férulique dans la subérine provoquant une plus grande sensibilité au stress salin (Gou et al., 2009).

c/- Les lignines

Les lignines (du latin lignum signifiant bois) sont des hétéro-polymères composés principalement (et non uniquement) des trois alcools phénoliques p-coumarylique, coniferylique et sinapylique. Ces unités sont baptisées respectivement unités H, G et S lorsqu'elles sont intégrées dans les lignines et leur proportion (essentiellement le rapport S/G) sert de marqueur. Elles sont considérées comme ayant joué un rôle essentiel dans l'adaptation des plantes à la vie sur la surface terrestre (Weng and Chapple, 2010). Elles sont présentes dans la plupart des parois secondaires et contribuent ainsi à la formation d'une structure hydrophobe et rigide. La synthèse et la déposition de ce polymère sont des mécanismes hautement régulés aussi bien spatialement que temporellement (Donaldson, 2001; Voxeur et al., 2015). Leur présence est essentielle chez les plantes vasculaires car elles contribuent à l'intégrité des cellules de conduction de la sève brute ainsi que des cellules participant au soutien. De plus, elles jouent un rôle dans la protection et la dissémination des graines (Figure 19). Le dépôt de lignines peut également être modifié lors de stress biotiques et abiotiques (Moura et al., 2010). Elles sont ainsi impliquées dans la défense contre certains pathogènes en formant une barrière très difficilement dégradable pour les microorganismes. En cas de blessure également, les lignines vont se déposer afin de former une barrière physique limitant la pénétration des pathogènes au sein de la plante (Bhuiyan et al., 2009).



Figure 19 : Différent types cellulaires aux parois lignifiées retrouvés chez les plantes supérieures au cours du développement. Une liste des stress impactant la lignification est indiquée en rouge. D'après Barros et al. (2015)

Puisque les connaissances concernant ce polymère sont importantes pour la compréhension des résultats de notre travail, une partie spécifique y sera consacrée dans la partie qui suit.

III- Les lignines

1/- Les lignines dans la paroi

Les lignines viennent s'imprégner dans de nombreuses parois secondaires entre les microfibrilles de cellulose et les hémicelluloses pour les souder et ainsi former un réseau hydrophobe et plus rigide (Zhong and Ye, 2009). Les parois de certains mutants déficients en lignines sont particulièrement épaisses comparées aux plantes sauvages (Jones et al., 2001). De ce fait, il est suggéré que les lignines permettraient l'attachement et le rapprochement des microfibrilles de celluloses entre elles (Kumar et al., 2016). De plus, les lignines pourraient être liées de manière covalente à certains groupements férulates contenus dans des hémicelluloses de type arabinoxylanes (Molinari et al., 2013; Ralph et al., 2004a) ainsi qu'à des protéines pariétales (Cong et al., 2013). De manière générale, l'analyse des liaisons entre les lignines et les polysaccharides est complexe, ce réseau pouvant varier au cours du développement, du type cellulaire et de l'espèce considérée. Une meilleure compréhension de cet agencement permettrait la mise en œuvre de stratégies de séparation de chaque constituant dans un but de valorisation industrielle optimale (Mottiar et al., 2016) des parois. En effet, les polysaccharides sont des sources importantes pour la production de papier et de carburants de seconde génération mais les lignines limitent leur extractabilité et donc réduit les rendements (Zeng et al., 2014).

2/- Les cellules lignifiées

2.1/- Les tissus conducteurs – Xylème

Le rôle du xylème chez les végétaux est à la fois de conduire la sève brute (eau et éléments minéraux prélevés dans le sol) et de participer au maintien du port dressé de la tige. Ces deux fonctions peuvent être assurées par des cellules uniques de type fibrestrachéides ou par des cellules distincts que sont les vaisseaux pour la conduction et les fibres pour le soutien. C'est par un mécanisme de différenciation cellulaire conduisant à une vidange totale de la cellule et à l'épaississement des parois que certaines cellules individuelles vont s'agencer entre elles pour former une trachée (Tyree and Zimmermann, 2002). Chez les angiospermes, lors de la différenciation des éléments de vaisseaux, les cellules vont s'associer pour former des vaisseaux mesurant jusqu'à 30 cm de long chez l'érable rouge par exemple (Tyree and Zimmermann, 2002). Les parois secondaires de ces cellules sont composées principalement de cellulose, d'hémicelluloses et de lignines. Afin de résister aux pressions dues au transport et à la gravité, les cellules mettent en place différents types d'ornementations optimisés pour ne pas collapser (Figure 20) (Oda and Fukuda, 2012). Au sein de ces structures, les lignines apparaissent comme essentielles car lorsque leur voie de biosynthèse est altérée (par le biais de la mutagenèse ou des inhibiteurs), cela conduit à des effondrements des vaisseaux dus aux pressions négatives appliquées lors du transport de la sève brute (Barros et al., 2015).



Figure 20 : Vaisseaux de xylème portant différentes ornementation de parois secondaires. De gauche à droite figurent des ornementations en spirales, réticulées, rayées et deux types ponctuées. D'après Oda and Fukuda (2012).

2.2/- Les tissus de soutien – Sclérenchymes

Les cellules du sclérenchyme sont retrouvées dans différentes parties de la plante, souvent à proximité du xylème, du phloème, de l'épiderme ou du cortex de certains fruits. Le sclérenchyme est un tissu mort constitué de cellules délimitées par une paroi épaisse lignifiée. Ces cellules peuvent être des sclérites de formes variables plutôt peu allongées, regroupées dans des parenchymes dont elles sont issues. Leur fonction est d'assurer la rigidité des organes. On trouve par ailleurs des fibres scléreuses, cellules très allongées en fuseau, à lumière étroite qui sont directement issues d'un méristème plutôt que d'une différenciation cellulaire comme pour les sclérites (Fahn, 1990). Dans certains cas, la distinction entre ces cellules est difficile et on va alors parler de fibres-sclerites, dont certaines sont retrouvées par exemple dans l'hypocotyle et l'épicotyle d'Arabidopsis (Lev-Yadun, 1997). Le degré de lignification des cellules du sclérenchyme varie beaucoup, certaines fibres scléreuses étant peu lignifiées tandis que la plupart des sclérites contiennent beaucoup de lignines (Evert, 2006). De manière générale les lignines contenues dans ces cellules sont riches en unités S (Fahn, 1990).

2.3/- Autres types cellulaires spécifiques

Il existe encore de nombreuses autres cellules lignifiées au sein des plantes qui sont en nombre plus faible que celles dédiées à la conduction ou au soutien mais qui sont tout de même importantes pour le développement de la plante.

Il faut par exemple citer la présence de lignines au sein de la paroi des cellules de l'endoderme et plus précisément dans des structures baptisées cadres de Caspary. Ce tissu joue un rôle essentiel dans la filtration des solutions absorbées par les racines et migrant vers le système vasculaire de la racine (Geldner, 2013). Les parois ayant ainsi été rendues hydrophobes, les solutions doivent transiter par la voie symplasmique et subir ainsi une filtration par l'endoderme. Certains mutants d'Arabidopsis pour des gènes de biosynthèse des lignines sont incapables de synthétiser correctement de telles barrières (Naseer et al., 2012).

On retrouve aussi des cellules lignifiées dans les téguments de certaines graines. Elles forment une barrière imperméable aux liquides et gaz et renforce mécaniquement la graine dans le but de protéger l'embryon (Barros et al., 2015). On note d'ailleurs que des graines mutantes d'Arabidopsis comportant des teneurs plus faibles en lignines ont un taux de germination inférieur (Liang et al., 2006). L'éventail très large des cellules lignifiées montre bien l'importance de ce polymère dans le développement de la plante. Les lignines de chaque type cellulaire sont différentes les unes des autres suggérant une régulation indépendante.

3/- Diversité des lignines

Les lignines sont souvent définies comme un assemblage de monolignols conduisant à la formation d'unités G, S et H. En fait leur agencement et leurs quantités relatives au sein des lignines conduisent à une diversité remarquable aussi bien parmi les familles taxonomiques, des types cellulaires ou même les zones pariétales spécifiques.

Dans les tissus conducteurs des gymnospermes, la proportion des différents monomères H/G/S est de l'ordre de 0–5/95–100/0, alors qu'elle est de 0–8/25–50/46–75 chez les angiospermes dicotylédones et 5–33/33–80/20–54 chez les monocotylédones (Ek et al., 2009). Une des différences majeures entre les gymnospermes et les plantes à fleurs est donc la présence d'unités S. L'apparition de cette nouvelle unité monomérique peut être vu comme un marqueur évolutif au sein de la lignée verte (Weng and Chapple, 2010). Il semble même que ce gain de fonction ait eu lieu deux fois de façon indépendante chez les angiospermes et les sélaginelles, l'un des premiers embranchements des plantes vasculaires (Weng et al., 2010). Il est intéressant de constater que pour le cas des angiospermes, il y a eu acquisition de deux enzymes (C3H et F5H) alors que pour le lycophyte, seulement une enzyme de type F5H-like permet la synthèse de cette nouvelle unité S.

La présence des monomères H, G et S n'est pas exclusive. D'autres unités noncanoniques telles que les unités catechyl (C) ou 5H (5-hydroxyguaiacyl) (Figure 21) sont présentes dans les téguments des graines de certaines orchidées et cactées (Chen et al., 2012). Les téguments d'un grand nombre de famille de genre *Cactoideae* contiennent même exclusivement des unités C tandis que d'autres accumulent des mélanges d'unités G/S/5H comme dans certaines espèces de la famille des *Escobaria* (Chen et al., 2013). Ce sont, à ce jour, les seuls exemples d'incorporation de C et de 5H dans les lignines de plantes non modifiées mais au regard de la grande diversité des espèces appartenant à ces familles et de la différence de leur répartition géographique, il ne semble pas qu'elles permettent une meilleure adaptation à l'environnement (Chen et al., 2013). Par contre, certaines plantes, chez qui les gènes de biosynthèse des monolignols ont été mutés, peuvent présenter des modifications similaires. Par exemple la perte de fonction d'un gène *COMT* permet l'incorporation d'unités 5H (Marita et al., 2003; Morreel et al., 2004a) et la suppression d'un gène *CCoAOMT* chez *Pinus radiata* mène à une incorporation d'unités C (Wagner et al., 2011). Lorsque les réactions enzymatiques permettant la synthèse d'unités G ou S sont bloquées, la plante va alors s'adapter en utilisant des monomères différents ce qui montre la grande plasticité de la lignification (Zhao, 2016).



Figure 21 : Structures des monolignols « classiques » (en noir) et non-canoniques (en vert) entrant dans la composition des lignines.

4/- La biosynthèse des lignines

Les précurseurs des lignines vont subir trois étapes majeures menant à sa formation. Ces étapes séquentielles sont d'abord (1) la biosynthèse des futurs monomères, puis (2) leur transport vers l'apoplaste et enfin (3) leur polymérisation en lignines.

4-1/- Biosynthèse des monolignols

La biosynthèse spécifique des monolignols a été décrite dans l'article Le Roy et al. (2017) présenté dans le chapitre I de la partie "résultats".

4-2/- Transports des précurseurs des lignines

A ce jour, les connaissances sur le transport des précurseurs restent fragmentaires. Afin d'y être polymérisé en lignines, les futurs monomères – mais peut être aussi des dimères (Petrik et al., 2014; Wilkerson et al., 2014) - synthétisés dans le

cytoplasme, ont besoin de traverser la membrane plasmique pour atteindre l'apoplaste. Actuellement, il existe trois théories qui s'opposent : (1) la diffusion passive, (2) l'exportation par des vésicules et (3) un transport actif via des transporteurs.

a/- Le transport passif

Le transport passif et non-sélectif pourrait expliquer la plasticité remarquable des lignines. En effet, lorsque l'expression de certains gènes de la voie de biosynthèse est altérée, les monomères « classiques » peuvent laisser la place à des monomères non canoniques (Chen et al., 2012, 2013; Marita et al., 2003; Morreel et al., 2004a; Wagner et al., 2011). Ce mode de transport a été modélisé à partir d'expériences de chromatographie sur colonne imprégnée de liposomes ou de disques bilipidiques (Boija and Johansson, 2006; Boija et al., 2007). Ces études ont mis en évidence que des analogues structuraux de monolignols traversent de manière passive ces membranes. Mais, afin de confirmer cette théorie, il est nécessaire de réaliser des analyses plus poussées ainsi que des expériences *in vivo*. De plus, la plante synthétise un grand nombre de molécules qui s'apparentent structurellement à celles testées lors de cette étude. Il est difficilement envisageable qu'elles puissent rester dans le cytoplasme tandis que les monolignols passent la membrane plasmique (Liu et al., 2011a).

b/- Le transport par des vésicules

Les hémicelluloses et les pectines sont exportées vers l'apoplaste via des vésicules issues de l'appareil de Golgi (Kim and Brandizzi, 2014). Il est donc logique de penser que les précurseurs des lignines pourraient être exportés de la même manière. Cette hypothèse est d'autant plus pertinente que lors de l'incorporation de précurseurs radioactifs dans la cellule, le marquage a été retrouvé dans la paroi cellulaire mais aussi dans les vésicules golgiennes et dans le réticulum endoplasmique rugueux (Fujita and Harada, 1979; Pickett-Heaps, 1968; Takabe et al., 1985). Malgré tout, l'hypothèse d'un système de transport des monomères par exocytose a depuis été remise en question. En effet, Kanada et al (2008) ont utilisé une combinaison d'inhibiteurs du métabolisme des phénylpropanoïdes et de la synthèse de protéines pour montrer que les marquages au sein du Golgi seraient issus de protéines (au moins partiellement) et non des précurseurs comme suggéré précédemment. Les auteurs de cette étude privilégient l'hypothèse d'un transporteur spécifique.

c/- Transport actif

L'implication du transport actif est supportée par des études plus récentes. Ce type de transport a été mis en évidence biochimiquement grâce à des fractions microsomales de membranes plasmiques issues de feuilles d'Arabidopsis et de racines de peuplier (Miao and Liu, 2010). Il a été montré que les alcools coniféryliques et sinapyliques (et dans une moindre mesure leurs dérivés aldéhydiques) sont transportés à l'intérieur de ces vésicules. Ce transport est fortement favorisé par l'ajout d'ATP (Adénosine TriPhosphate) ce qui va dans le sens de l'implication d'un transporteur de type ATP-binding cassette (ABC). Certains membres de cette famille avaient déjà était identités par leurs profils d'expression liés à certains tissus lignifiés (Egertsdotter et al., 2004; Paux et al., 2004; Whetten et al., 2001). Depuis, une étude à identifié le transporteur AtABCG29 comme étant un bon candidat pour l'export de monolignols (Alejandro et al., 2012). Par une approche de complémentation de mutants de levure, l'alcool p-coumarylique (précurseur des unités H) a été désigné comme substrat spécifique de AtABCG29. De manière intéressante, un mutant d'Arabidopsis pour ce gène montre une diminution non seulement des unités H mais également des deux autres unités monomériques des lignines ce qui suggère une régulation complexe. Plus récemment, la même démarche a été réalisée cette fois ci sur des tissus xylémiens de peuplier, de pin et de cyprès du Japon (Tsuyama et al., 2013). A l'aide d'une combinaison d'inhibiteurs de transporteurs, les auteurs imputent le transport de monolignols non pas à des transporteurs de type ABC mais à des transporteurs antiport à H⁺ Ces deux études suggèrent que le transport des précurseurs des lignines semble être régulé de manière complexe avec sans doute des différences entre les espèces, les types cellulaires ou les stades de développement considérés. Dans chacune de ces études, un aspect essentiel de la spécificité du transport est révélé par le fait que le comportement des monolignols dépend de leur statut de glycosylation. Les monolignols glycosylés seraient ainsi préférentiellement transportés vers la vacuole tandis que les monolignols aglycones passeraient la membrane plasmique vers l'apoplaste (Liu et al., 2011; Tsuyama et al., 2013). Cette caractéristique est développée de façon plus exhaustive dans l'article Le Roy et al. (2016) en annexe de ce manuscrit

4-3/- La polymérisation des lignines

L'assemblage des lignines est réalisé par un couplage oxydatif radicalaire qui met en place de nombreuses liaisons covalentes différentes entre les monomères. Généralement, on distingue les liaisons C-C dites « condensées » (telles que 5–5', ß–5', ß– ß and ß–1') des liaisons C-O-C dites « non condensées » (comme ß-O-4' par exemple) (Figure 22) (Terashima and Fukushima, 1988).



Figure 22 : Représentation des principales liaisons dans les lignines de gymnospermes. D'après Thakur and Thakur (2015).

Les liaisons condensées sont plus stables que les liaisons éthers (non condensées) et leur nature varie selon les espèces (Tableau 1). De nombreuses études tentent de remplacer les liaisons fortes par des liaisons éther dans le but d'une valorisation industrielle facilitée comme lors de la saccharification ou la fermentation (Ragauskas et al., 2014). En effet, il sera d'autant plus facile d'accéder aux polysaccharides pariétaux que les liaisons des lignines seront faibles. La nature des liaisons est principalement conditionnée par la nature des monomères incorporés. Des parois composées majoritairement d'unités G formeront de grande quantités de liaisons fortes telles que β -5 ou 5-5 à cause de la disponibilité élevée du C₅ (Boerjan et al., 2003). L'influence de la composition monomérique sur la nature des liaisons dans le polymère montre bien que l'association est de nature chimique plutôt qu'enzymatique (Dima et al., 2015; Ralph et al., 2004b).

Tableau 1 : Pourcentage des différents types de liaison dans les lignines d'organes de différentes plantes. β -O-4, beta-aryl ether; 5–5, biphenyl and dibenzodioxocin; β -5, phenylcoumaran; β -1, 1,2-diaryl propane; β - β , resinol; D'après Barros et al. (2015b).

Espèces	Echantillon	β- 0- 4	5-5	β-5	β-1	β-β	Référence
épicéa	bois broyé mature	48	10	10	7	2	Adler (1977)
Pin	bois broyé mature	62	-	20	-	18	Mansfield et al. (2012)
Bouleau	bois broyé mature	60	5	6	7	3	Adler (1977)
Peuplier	Bois	69	-	3	-	28	Mansfield et al. (2012)
Arabidopsis	Tige	79	-	7	-	14	Mansfield et al. (2012)
		77	1	15	1	6	Bonawitz et al. (2014)
Luzerne	Tige	81	1	8	-	6	Marita et al. (2003)
Maïs	Tige	100	-	-	-	-	Mansfield et al. (2012)
		60	-	27	3	10	Min <i>et al.</i> (2014)
Blé	Paille	75	3	11	3	4	Del Río et al. (2012)

A côté de cela, on trouve dans le cytosol des composés présentes sous forme de dimères, trimères ou plus, appelé communément oligolignols (Dima et al., 2015; Huis et al., 2012). Le mécanisme de synthèse de ces molécules n'est pas encore connu et donc il est difficile de dire si ces couplages sont de même nature que pour la polymérisation de la lignine ou s'ils sont réalisés par des enzymes cytosoliques de même type que les pcoumaroyl-CoA:monolignol transferases qui lient des alcools p-coumaryliques à des alcools sinapyliques (Petrik et al., 2014; Wilkerson et al., 2014). Dans cette partie, il sera détaillé la polymérisation qui a lieu dans la paroi par le biais des laccases et des péroxydases. Celle-ci se déroule en deux étapes. La première est une radicalisation oxydative entrainant la déstabilisation d'un substrat deux familles par d'oxydoreductases : les péroxydases et les laccases (Figure 23). Dans un deuxième temps, va avoir lieu un couplage radicalaire avec une autre molécule (Vanholme et al., 2010). Cette deuxième étape est purement chimique et obéit donc aux lois thermodynamiques et stériques classiques (Ralph et al., 2004b).



Figure 23 : Oxydation radicalaire des monolignols (ici l'alcool coniférylique) par les peroxydases et les laccases. D'après Wang et al. (2013a).

a/- Les laccases

Les laccases (ou blue copper-containing oxidoreductases) utilisent l'O2 pour oxyder les précurseurs des lignines. Avec 17 membres chez Arabidopsis thaliana et 39 chez Populus trichocarpa, les laccases constituent une famille multigénique (Weng and Chapple, 2010). La plupart des laccases possèdent un peptide signal compatible avec une action possible dans la paroi. La prédiction in silico montre par exemple, chez Arabidopsis, que seules AthLac14 et AthLAC16 en sont dépourvues et que c'est également le cas pour PtrLAC9 et PtrLAC27 chez le peuplier. Les laccases ont été identifiées comme étant potentiellement impliquées dans la polymérisation du fait de leur spécificité pour les monolignols in vitro (Mayer and Staples, 2002; Ranocha et al., 2002; Sterjiades et al., 1992) ainsi que par leurs profils d'expression notamment dans les tissus lignifiés, et leur co-expression avec les gènes de biosynthèse des monolignols (Brown et al., 2005; Gavnholt and Larsen, 2002). C'est seulement récemment que la première preuve de leur implication *in vivo* a été mise en évidence (Berthet et al., 2011; Schuetz et al., 2014; Turlapati et al., 2010; Zhao et al., 2013). A ce jour, les quatre laccases AthLAC4, 11, 15 et 17 ont été identifiées comme étant impliquées plus ou moins fortement dans la lignification au sein de la hampe florale d'Arabidopsis. Globalement, les mutants pour un seul gène ne possèdent pas un phénotype très fort mais les doubles mutants comme *lac4lac17* ont des vaisseaux collapsés et une réduction d'environ 40% des lignines totales (Berthet et al., 2011). Le triple mutant lac4lac11lac17 a un phénotype encore plus marqué avec de plus fortes altérations de croissance et également des tissus vasculaires collapsés (Zhao et al., 2013). Ces études mettent en évidence une redondance fonctionnelle des laccases entre-elles puisque les changements drastique dans les lignines ne sont visibles qu'à partir du double mutant.

b/- Les péroxydases

Les péroxydases appartiennent à une très grande famille multigénique. On les sépare en trois grandes classes en fonction de leurs structures tridimensionnelles, de leurs substrats et de leur répartition taxonomique (Cosio and Dunand, 2009). Cette partie traitera seulement des péroxydases de classe III qui sont retrouvées spécifiquement dans le règne végétal avec, par exemple, 73 membres chez *Arabidopsis thaliana* et 138 chez *Oryza sativa* (Welinder et al., 2002) avec une large palette d'actions physiologiques (Passardi et al., 2005). De manière générale, elles sont exprimées dans

l'ensemble de la plante et tout particulièrement dans les racines mais on ne connaît pas bien encore les fonctions précises de toutes les péroxydases. Elles sont pour la plupart sécrétées dans les parois bien qu'il existe des isoformes présentes dans la vacuole (Francoz et al., 2015). Les rôles supposés des péroxydases pariétales sont plutôt antagonistes car elles agiraient sur le relâchement des parois en clivant certaines liaisons polysaccharidiques comme cela a été décrit pour AtPRX36 (Kunieda et al., 2013) ou à l'inverse sur le renforcement de celles-ci soit par le biais de pontages moléculaires entre différents constituants pariétaux par l'intermédiaire de l'acide férulique (Burr and Fry, 2009), soit par leur action sur la polymérisation de précurseurs de lignines.

Cette activité de polymérisation a été identifiée initialement *in vitro* par leur spécificité d'action sur les monolignols (Ostergaard et al., 2000; Sterjiades et al., 1993; Tsutsumi et al., 1998). Par la suite, des caractérisations fonctionnelles de péroxydases par génétique inverse ont étaient entreprises. Par exemple, AtPRX72 a été identifiée grâce à sa forte homologie de séquence avec une péroxydase historique de *Zinnia elegans* (ZePrx) connue pour son activité de polymérisation de monolignols *in vitro* (Gabaldón et al., 2005). Un mutant pour ce gène a un retard de croissance vraisemblablement du au manque de lignines (Herrero et al., 2013). De manière intéressante, la lignification du cadre de Caspary dans l'endoderme de ce mutant ne parait pas être affectée (Lee et al., 2013). Par contre AtPRX64 semble être plus spécifique des lignines contenues dans ces renforts. Sa sous-expression provoque une diminution de la quantité de lignines dans les parois de l'endoderme (Lee et al., 2013). Cette spécificité serait peut être due à des micro-domaines dans la paroi permettant de retenir spécifiquement certaines péroxydases (Francoz et al., 2015).

La question de l'implication des péroxydases ou des laccases dans la polymérisation des lignines reste encore en suspens. La lignification de types cellulaires précis sont dépendants de péroxydases particulières comme dans le cas des cadres de Caspary (Lee et al., 2013). Mais en ce qui concerne les cellules du xylème les choses sont moins tranchées car si triple mutant *lac4/17/11* montre bien un phénotype important au niveau des tissus vasculaires (Zhao et al., 2013), il en va de même pour le mutant *prx72* qui montre un phénotype similaire (Herrero et al., 2013). Il existe plusieurs modèles qui tentent d'expliquer l'imbrication entre ces deux familles d'enzymes (Barros et al., 2015). Il est par exemple possible de penser qu'il existe des actions séquentielles

selon la spécificité de substrat / accepteur (O2 ou H_2O_2). A l'aide de substrats différents des monolignols comme le 4-methoxy- α -naphtol, la présence de peroxydases à été mise en évidence dans les tissus lignifiés de *Coleus blumei* (de Pinto and Ros Barceló, 1997). D'autre molécules comme l'aldhéhyde coniférylique et le ABTS (2,2-azinobis-(3ethylbenzothiazoline-6-sulfonate) montrent quant à elles les deux types d'activités (laccase et péroxydase) chez *Populus euramericana* (Ranocha et al., 1999). Une autre hypothèse serait une intervention spatio-temporelle différente des phénoloxidases car les laccases et les péroxydases ne sont pas exprimées uniquement dans les cellules lignifiées, ce qui suggère une possible coopération entre les cellules pour la lignification (Pesquet et al., 2013; Smith et al., 2013). Il est possible que certains phénoloxidases soient fournies par des cellules voisines et d'autres par les cellules en cours de lignification.

5/- La régulation de la lignification

5-1/La régulation transcriptionnelle

La déposition et les interactions entre les principaux constituants des parois secondaires (cellulose, hémicellulose et lignines) sont régulées de façon pyramidale. En effet les gènes de synthèse de ces bio-polymères sont gouvernés par des régulateurs communs qui contrôlent d'autres régulateurs plus fins permettant la synthèse spécifique de tel ou tel constituant. Dans cette configuration, une sous-famille de facteurs de transcription à domaine NAC (NAM, ATAF1/2 and CUC2) appelé VNS pour VND, NST/SND, and SMB, représente le premier niveau de régulation (Nakano et al., 2015a). Vient ensuite le second niveau qui va contrôler plus finement les différents composants de la paroi. Il est principalement composé de facteurs MYB qui vont soit directement activer des gènes de biosynthèse des composés pariétaux soit d'autres facteurs de transcription. Ces derniers sont considérés comme le troisième et dernier niveau de régulation transcriptionnelle de certains de ces gènes (Figure 24).



Figure 24 : Schéma simplifié de la régulation transcriptionnelle pyramidale des constituants de la paroi secondaire (lignines, cellulose et hémicelluloses). D'après Nakano et al. (2015a).

a/- Premier niveau de régulation transcriptionnelle : les facteurs de transcription VNS

Le domaine NAC des facteurs de transcription VNS est localisé dans la partie Nterminale très conservée de la protéine. Il va permettre la localisation nucléaire ainsi que la formation des homo- et hétéro-dimères (Yamaguchi and Demura, 2010). De nombreuses études montrent que la sur-expression des facteurs de transcription VND1 à 7 (VASCULAR-RELATED NAC-DOMAIN) chez Arabidopsis entraine la formation de parois secondaires ectopiques au sein de la plante (Endo et al., 2015; Kubo et al., 2005; Zhou et al., 2014). Chez la même espèce, la sur- et sous-expression des NST1 à 3 (NAC SECONDARY WALL THICKENING PROMOTING FACTOR) conduit à des altérations dans la déposition des parois secondaires (Mitsuda et al., 2007; Zhong et al., 2006, 2007a). Les VND et les NST peuvent être partiellement redondants comme le montrent des expériences de réversions de phénotypes « hampe non dressées » du mutant *snd1-nst1* par l'expression de VND7, NST1 et NST2 sous le contrôle d'un promoteur *SND1* (Zhong et al., 2010). D'autre peuvent être plus spécifiques, comme c'est par exemple le cas de VND6 qui permet d'activer certains gènes de mort cellulaire programmé (Ohashi-Ito et al., 2010).

Récemment, une étude a permis l'identification d'un « nouveau » facteur de transcription impliqué dans la synthèse de la parois secondaire : E2Fc (Taylor-Teeples et al., 2015). Il s'associerait avec au moins 23 promoteurs de gènes appartenant aux différents niveaux de régulation du plus général comme des VNS (VND6, 7) jusqu'aux gènes spécifiques des lignines (LAC4, C4H, CCoAOMT, CAD4,...). Ce gène connu

initialement pour réguler négativement l'endoréplication (del Pozo et al., 2002) met en lumière un niveau supérieur de complexité.

b/- Second niveau de régulation transcriptionnelle : les facteurs de transcription MYB46/83

En plus d'activer des gènes liés à la biosynthèse des parois, les VNS vont permettre l'expression d'autres facteurs de transcription essentiels pour la synthèse des parois secondaires comme *AtMYB46* et *AtMYB83* (McCarthy et al., 2009; Zhong et al., 2007b). Ces gènes sont exprimés préférentiellement dans le xylème et leur surexpression provoque une déposition ectopique de parois secondaires (Nakano et al., 2015b). Il semblerait ainsi exister une redondance fonctionnelle (tout au moins partielle) entre ces deux MYB (McCarthy et al., 2009). Des élément *cis* reconnus par AtMYB46 et AtMYB83 sont retrouvés dans de nombreux gènes impliqués dans la biosynthèse de la paroi et en particulier des lignines (Kim et al., 2012, 2013b; Zhong and Ye, 2012). Par ailleurs, ces deux MYBs vont moduler environ 40 autres gènes de facteurs de transcription importants pour la paroi secondaire.

c/- Troisième niveau de régulation transcriptionnelle

Parmi les nombreux facteurs de transcription impliqués dans cette dernière couche de régulation on peut citer *KNAT7* (Li et al., 2012a), *AtC3H14* (Ko et al., 2009), ainsi que des NACs et de nombreux autres facteurs MYB (Nakano et al., 2015b). Certains comme AtMYB58, AtMYB63 ou AtMYB85 vont réguler spécifiquement les gènes de biosynthèse des monolignols (Zhong et al., 2008; Zhou et al., 2009). La plupart de ces gènes possèdent dans leurs promoteurs une région riche en A et C appelée élément AC (aussi connu sous les noms PAL-box ou H-box) reconnue par ces facteurs MYBs (Raes et al., 2003). De plus, chez *Eucalyptus grandis*, un nouveau mécanisme de régulation à été identifié pour EgMYB1 (Soler et al., 2016). Ce MYB, connu pour réguler négativement la lignification (Legay et al., 2010), peut interagir physiquement avec l'histone linker EgH1.3. Cette interaction protéine-protéine se mettrait en place pour augmenter la répression des gènes *EgCCR* et *EgCAD2* dans des cellules dont les parois sont peu ou pas lignifiées (Soler et al., 2016). Ce type d'interaction semble anecdotique pour la régulation de la paroi secondaire mais serait indispensable pour que certains types cellulaires qui doivent rester non-lignifiés comme le cambium vasculaire.

L'homéostasie des phénylpropanoïdes semble également régulé par les facteurs des transcription généraux. En effet, les protéines MED5a et MED5b (aussi appelé REF4 et RFR1 respectivement) impactent la synthèse de lignines. Ces deux protéines sont des sous-unités du complexe mediator et co-régulent la transcription des gènes de biosynthèse des phénylpropanoïdes. Ce complexe, composé d'environ 20 à 30 sousunités, est présent de façon ubiquitaire chez les eucaryotes (Flanagan et al., 1991). La mutation de *MED5a/b* entraine une altération de l'homéostasie des phénylpropanoïdes (Bonawitz et al., 2012) et de façon intéressante, leur perte de fonction au sein du mutant *ref8* (codant une C3'H) permet une réversion partielle de la mutation (Bonawitz et al., 2014). De plus, MED5a/b sont impliqués dans la répression des anthocyanes et d'hydroxycinnamates d'esters contenus dans le mutant *f5h*. D'après ces résultats, les protéines MED5a/b semblent relativement importants pour la régulation des phénylpropanoïdes mais également pour la voie de l'indole glucosinolate qui provient également de la voie du shikimate (Kim et al., 2015)

5-2/- La régulation post-transcriptionnelle

Les régulations transcriptionnelles ne sont pas suffisantes à elles seules pour contrôler efficacement l'action de certains gènes de phénylpropanoïdes. Il existe en effet des mécanismes de régulations post-transcriptionnelles intervenant dans la lignification. Chez *P. trichocarpa*, les transcrits du gène de facteur de transcription *PtrSND1* peuvent subir un épissage alternatif formant ainsi un ARNm dit « long » et un autre « court » avec des action antagonistes (Li et al., 2012b; Zhao et al., 2014). Dans les fibres de tissus xylémiens, PtrSND1, sous sa forme longue, permet l'épaississement des parois comme le prouve sa sur-expression *in vivo* (Li et al., 2012b). Le transcrit court qui ne contient pas de domaine de liaison à l'ADN va s'associer au facteur PtrSND1 long. Cet hétérodimère ainsi formé ne pourra plus reconnaitre les éléments *cis* dans le promoteur des gènes impliqués dans la xylogenese et donc inhibera la croissance des parois secondaire (Zhao et al., 2014). Cette régulation peut être mise en place pour limiter et/ou stopper l'épaississement des parois secondaires des cellules.

La régulation post-transcriptionnelle peut également passer par l'utilisation de micro-ARN (miRNA). Ce sont des ARNs d'environ 21 nucléotides qui vont permettre *in fine* de conduire à la dégradation des ARNm (Voinnet, 2009) grâce à une reconnaissance séquence-spécifique. Ils régulent un grand nombre de processus physiologiques tels que

le développement (Rubio-Somoza and Weigel, 2011) ou la réponse aux stress (Kruszka et al., 2012; Sunkar et al., 2006). La surexpression de *miR397b* chez *Arabidopsis thaliana* réduit significativement les quantités de lignines (Wang et al., 2014a). Au delà des nombreuses prédictions *in silico*, le clivage de certains transcrits de laccases par des orthologues de *miR397* a également été montré chez le *Populus trichocarpa* (Lu et al., 2013), *Oryza sativa* (Zhang et al., 2013b) et chez un citrus (*Poncirus trifoliata*) (Jin et al., 2016). Dans un contexte de régulation de la biosynthèse de la paroi, il existe d'autres miRNAs qui impactent des caractères plus complexes. Par exemple la sous-expression de *miR156* provoque entre autre une diminution de la quantité de lignines mais au vu des effets pléiotropiques induits (augmentation de la longueur des racines, augmentation de la densité des trichome, diminution du diamètre de la tige,...) il est difficile d'y voir un ciblage directe d'un gène lié au métabolisme pariétal (Aung et al., 2015). *MiR156* régulerait en fait un certain nombre de gènes codant des facteurs de transcription.

5-3/- Régulation traductionnelle & post-traductionnelle

Il n'existe encore que peu d'informations sur l'importance des régulations posttraductionnelles dans la biosynthèse des lignines. C'est au niveau des enzymes PAL que les premières preuves de l'existence de telles régulations ont été suggérées, car leurs activités peuvent être modulés de manière extrêmement rapides (Jones, 1984; Lamb et al., 1979). C'est seulement plus tard que les mécanismes moléculaires qui en sont responsables, ont été mises en évidence (Zhang et al., 2013a). En effet, trois protéines Kelch repeat F-box (KFB01, 20 et 50) interagiraient avec les quatre isoformes de PAL chez Arabidopsis, afin qu'elles soient dégradées via le proteasome 26S (Zhang et al., 2013a). L'analyse de mutants déficients en KFB montre qu'une augmentation de la teneur en lignines existe dans ces plantes. De façon plus globale, il semblerait que ce type de régulation existe pour d'autre acteurs de la lignification comme C4H, HCT, CAD, et OMT1 (Kim et al., 2013a).

Un autre type de régulation post-traductionnelle connu pour cette voie de biosynthèse est la phosphorylation des enzymes par des kinases spécifiques. Cette modification réversible concerne environ 9000 protéines chez Arabidopsis et a pour effet de réguler leurs activités, localisation, stabilité ou interactions (Cheng et al., 2014). Des études ont décrit la phosphorylation d'une PAL chez le *Phaseolus vulgaris* (Allwood

et al., 1999; Paul Bolwell, 1992). Ceci a été confirmé par des analyses de phosphoproteomique globaux menés sur du xylème de *Populus trichocarpa* (Wang et al., 2015). Le rôle des phosphorylations sur l'enzyme PAL n'est pas encore bien connu car ils n'impactent visiblement pas leur activité *in vitro* (Allwood et al., 1999; Wang et al., 2015). Par contre l'activité de l'enzyme PtrAldOMT, qui permet la production de sinapaldehyde, est fortement dépendante de la phosphorylation sur des sérines en position 123 ou 125 entrainant une inhibition de son activité (Wang et al., 2015). La kinase responsable n'a pas encore été identifiée mais son étude fonctionnelle approfondie permettrait de mettre en lumière un nouveau niveau de régulation de la voie de la lignification. Ce sont, à ce jour, les deux seuls exemples de phosphorylation identifiés sur les enzymes de cette voie de biosynthèse mais avec le développement de nouveaux outils analytiques il est possible que cette facette de régulation soit de plus en plus étudiée dans le futur.

IV- L'utilisation du lin comme plante modèle

Notre équipe utilise le lin (*Linum usitatissimum L*.) comme l'un des modèles d'étude pour la lignification. En effet, les fibres présentes dans la périphérie de la tige et parfois qualifiées de primaires, longues ou (péri)phloémiennes dans la littérature, sont caractérisées par un profil pariétal extrêmement atypique en terme de lignines. Ce chapitre présente rapidement quelques caractéristiques de cette plante et nous reviendrons sur les aspects de lignification chez le lin à la fin de cette partie.

1/- Botanique et utilisation

Le lin est d'une eudicotyledone annuelle de l'ordre des Malpighiales. L genre *Linum* qui regroupe environ 180 espèces dont certaines sont pérennes et se différencient notamment par leur morphologie ainsi que la grande diversité de la couleur de leurs fleurs (Muravenko et al., 2010). Le lin possède une racine pivotante et une tige feuillée principale qui peut produire des ramifications partant de la base de la tige (Figure 25). La taille de la tige peut varier considérablement selon le génotype allant de 20 à 150 cm (Muir and Westcott, 2003). Elle porte des feuilles sessiles avec une phyllotaxie inversée, qui tomberont au fur et à mesure que les capsules se rempliront. La fécondation est essentiellement autogame et a lieu avant l'ouverture de la fleur pour donner un fruit de type capsule. Après la fécondation, on notera la présence de deux graines dans chacune des cinq loges carpellaires (Figure 25)



Figure 25 : Morphologie du lin et des organes à différents stades de développements. D'après wiki commons.

Le lin peut ainsi être cultivé pour la production de ses graines riches en huile (35 à 45% de leur poids sec). De manière générale les variétés optimisées pour la production des graines ont plutôt une taille réduite et un plus grand nombre de ramifications florales.

En Europe, pour des raisons historiques mais aussi climatiques, le lin est plutôt cultivé pour ses fibres. Les variétés à fibres sont généralement plus grandes et riches en faisceaux fibreux avec peu de ramifications secondaires. Après une centaine de jours de culture, le lin est arraché et déposé à même le sol pour y subir l'étape de rouissage. Cette étape consiste en une dégradation partielle des matrices reliant les fibres entre-elles. Elle est suivie du teillage qui consiste à broyer mécaniquement les tiges pour y récupérer des fibres de longueurs diverses. Les fibres longues sont traditionnellement utilisées dans le textile (Netravali and Pastore, 2014) et les fibres courtes ou « étoupes », autrefois utilisées dans la fabrication du papier ou la ficelle, sont aujourd'hui valorisées dans l'industrie automobile pour le rembourrage des sièges et le remplacement de la fibre de verre dans les pare chocs par exemple (Baley, 2002). Les « anas », sous-produits du broyage de la tige centrale, sont utilisés dans la fabrication de panneaux agglomérés, ou sont valorisés en litière pour chevaux ou animaux domestiques ou pour le paillage végétal des cultures.

2- Formation des fibres primaires de lin

En termes botaniques, on définit une fibre comme étant une cellule individuelle sclérenchymateuse ayant principalement quatre caractéristiques (Dam and Gorshkova, 2003; Esau, 1977; Fahn, 1990):

- un rapport longueur/diamètre élevé (supérieur à 1000)
- des parois cellulaires épaisses
- des extrémités en biseaux
- joue un rôle dans le soutien mécanique de la plante

Chez le lin, les fibres, du fait de leurs parois épaisses, interviennent dans le maintien de la plante mais également dans la formation d'une barrière contre les insectes suceurs en protégeant la sève élaborée du phloème (McDougall et al., 1993). Les fibres élémentaires longues du lin sont présentes sous forme de faisceaux fibreux situés entre le cortex et le phloème (Figure 26). Ces faisceaux sont au nombre de 20 à 50 par tige et chacun contient environ 10 à 30 fibres élémentaires (Gorshkova et al., 2003). Les principales étapes de formation de ces cellules spécifiques sont l'initiation, l'élongation et l'épaississement des parois



Figure 26 : Coupe transversale de lin coloré au phluoroglucinol-HCl montrant les fibres primaires aux parois épaisses. Les lignines apparaissent en rouge.

2-1/- Initiation

Contrairement aux fibres courtes présentes dans les tissus xylémiens qui dérivent d'un méristème secondaire, les fibres longues, qualifiées donc de primaires, dérivent d'une croissance primaire via le procambium (Esau, 1977; Fahn, 1990). Les mécanismes moléculaires qui gouvernent l'identité des fibres reste encore inconnue (Gorshkova et al., 2012). Récemment une étude de transcriptomique haut débit a comparé la zone apicale (contenant le méristème apical) et le reste de la tige (Zhang and Devholos, 2016). Cette étude met en lumière 12 000 gènes différentiellement exprimés dont 6 000 retrouvés plus abondamment dans la zone apicale. Cette approche haut débit permet non seulement d'apprécier l'ensemble de transcrits avant et après la différenciation des fibres primaires mais également dans la plupart des autres types cellulaires (phloème, xylème,...). C'est donc la première fois que ce type d'étude est réalisé sur une plante à fibre et parmi le grand nombre de gènes identifiés, certains pourraient être spécifiques de l'initiation des fibres (Zhang and Deyholos, 2016). Des travaux complémentaires seront toutefois nécessaires pour affiner cette étude car la zone d'initiation a une taille extrêmement limitée, rendant des études moléculaires difficiles.

2-2/- Elongation

L'élongation des fibres est relativement rapide car elles atteignent leur longueur finale au bout de deux à quatre jours avec une vitesse d'environ un à deux cm/jours (Gorshkova et al., 2003). Cette croissance est qualifiée de diffuse et anisotrope car c'est l'ensemble de la cellule qui s'allonge et non uniquement l'apex comme pour les tubes polliniques et les poils absorbants. Après une phase de croissance coordonnée, les fibres vont connaître une phase de croissance intrusive définie par un taux de croissance longitudinale plus élevé que celui des cellules environnantes. Ces cellules doivent donc s'insérer entre les cellules voisines au niveau des lamelles moyennes (Chernova and Gorshkova, 2007; Snegireva et al., 2015). Lorsque cette élongation est finie, la cellule peut contenir jusqu'à dix noyaux (Ageeva et al., 2005). D'une manière générale, on assiste à une croissance symplastique (Erickson, 1986) dont l'arrêt concommitant avec la formation de la paroi secondaire (Gorshkova et al., 2003, 2005). La séparation spatiotemporelle de ces deux étapes a permis de définir une zone particulière appelée « snappoint » qui se situe juste en dessous de l'apex de la tige (Figure 27). Au dessus du snappoint il y a croissance en longueur des fibres et en dessous il y a croissance secondaire de la paroi (Gorshkova et al., 2003).



Figure 27 : Développement des fibres primaires de Linum ussitatissimum. (A) : Représentation du snap-point (ligne rouge) au cours du développement. (B) : Représentation des trois étapes principales de formation des fibres de lin. D'après Gorshkova et al. (2003).

2-3/- Epaississement

Durant cette étape, les parois vont s'épaissir pour atteindre 5 à 15 µm d'épaisseur (Morvan et al., 2003). Cette nouvelle différenciation est réalisée de façon centripète (de l'extérieur de la tige vers l'intérieur) (Figure 28). L'épaississement va être progressif dans toute la plante pour être maximal au moment du brunissement des capsules en champ c'est à dire peu avant l'arrachage des plantes. Cette étape n'est pas immédiatement synonyme de mort cellulaire car la transcription est tout aussi active dans les fibres comparée aux tissus de l'apex de la tige (Gorshkov et al., 2016). La mort cellulaire des fibres primaires ressemble à celle des fibres de xylème (Ageeva et al., 2012). Lorsque la plante fructifie, les fibres contenues dans les tiges encore vertes accumulent des globules lipidiques ce qui suggère une digestion des membranes de certains organites (Ageeva et al., 2012). A des stades plus avancés il y a une forte vacuolisation et l'apparition de cyto-segresomes (des fragments vacuolaires autolytiques). Des organites tels que des chloroplastes ou les mitochondries se dégradent progressivement et enfin il y a rupture des tonoplastes qui est la dernière étape de la mort cellulaire (Ageeva et al., 2012). Ce type de mort cellulaire programmée par autophagie ressemble à celui qui a lieu lors de la senescence des pétales (van Doorn and Woltering, 2008) ou dans les fibres du xylème de peuplier (*Populus tremula x P. tremuloides*) (Courtois-Moreau et al., 2009).



Figure 28 : Différents stades de développement des parois de fibres primaires de Linum ussitatissimum. *D'après Gorshkova et al. (2003).*

V- Les activités UDP-Glycosyltransférases

1/- La glycosylation chez les plantes

Les réactions de (dé)-glycosylations permettent l'ajout ou le retrait d'un groupement glycane sur une large variété de molécules telles que des protéines, lipides, métabolites spécialisées ou d'autres carbohydrates. Chez les organismes vivants, ces réactions sont réalisées de façon enzymatique à l'aide principalement de deux familles d'enzymes classées selon leurs activités dans la base de donnée CAZy (Carbohydrate-Active enZymes) : les Glycosyltransférases (GT) et les Glycosylhydrolases (GH)(Lombard et al., 2014) (Figure 29).



Figure 29 : Mécanisme général de (dé)glycosylation réalisé par des glycosyltransférases et glycosylhydrolases. R = lipide, protéine, acides nucléiques, métabolites, hormones...

De façon générale, les plantes contiennent bien plus d'enzymes actives contre les carbohydrates (ou CAZymes) que les autres règnes (Coutinho et al., 2003). Ceci est particulièrement vrai pour les GH et GT avec une proportion plus importante quand elles sont rapportées aux nombres de gènes totaux comparé à d'autres eucaryotes tels que Saccharomyces cerevisiae, Drosophila melanogaster ou Homo sapiens mais également de nombreuses bactéries (Cao et al., 2015; Coutinho et al., 2003). Une des raisons de cette forte amplification pourrait être le développement et la complexification des parois cellulaires végétales. Il semblerait en effet que 48% des gènes en relation avec ces parois correspondent à des CAZymes chez Arabidopsis thaliana (Cao et al., 2015). De plus, la conversion des produits de la photosynthèse en di- oligo- et poly-saccharides nécessite également des enzymes CAZyme spécifiques (Keegstra and Raikhel, 2001). D'autre part, les plantes ont subi de nombreux événements de duplications chromosomiques (Bowers et al., 2003; Panchy et al., 2016). Ces duplications pourraient également expliquer que les gènes CAZymes sont retrouvés en plus grand nombre chez les plantes que dans d'autres organismes généralement sous forme de clusters sur les génomes (Coutinho et al., 2003).

Selon la nomenclature CAZy les glycosyltransferases (EC 2.4.x.y) peuvent être classés en 103 familles (en mai 2017) selon leurs activités enzymatiques. Ces enzymes vont permettre le transfert d'un groupement glycane vers une autre molécule acceptrice. Le substrat donneur est généralement activé par l'ajout d'un groupement nucléoside diphosphate de type UDP ou GDP mais également dans une moindre mesure par un nucléoside monophosphate (CMP), des lipides (dolochol phosphate) ou peut parfois être non substitué (Lairson et al., 2008). La liaison glycosidique ainsi formée est, dans la plupart des cas, associée à un atome d'oxygène du substrat accepteur mais elle peut également se faire sur des atomes d'azote, de soufre ou de carbone (Lairson et al., 2008). Selon, les familles d'enzymes, la réaction peut soit maintenir la stéréochimie de la liaison (NDP- α -sucre -> α -glycoside) soit l'inverser (NDP- α -sucre -> β -glycoside). Ces mécanismes sont spécifiques pour chaque famille de la base de données CAZy et sont indépendants de la conformation tri-dimensionnelle des enzymes. Bien que la structure de toutes les GT n'ait pas encore été résolue (Liang et al., 2015), trois conformations tridimensionnelles ont été identifiées : GT-A, GT-B et GT-C (Figure 30). Les enzymes à structure GT-A comprennent deux domaines de type Rossmann $\beta/\alpha/\beta$ adjacents et sont généralement dépendantes d'un ion métallique divalent. Les enzymes GT-B sont elles aussi constituées de deux domaines de type Rossmann $\beta/\alpha/\beta$ mais cette fois-ci, elles se font face. C'est dans le creux situé entre ces deux domaines que se trouve le site actif qui est la plupart du temps indépendant des ions métalliques. La troisième structure, GT-C, est intégralement hydrophobe afin d'intervenir sur des molécules lipidiques liées à des sucres au sein des membranes (Gloster, 2014).



Figure 30 : Trois représentations des structures tridimensionnelles retrouvées chez les glycosyltransférases. La structure GT-A est SpsA de Bacillus subtilis, GT-B : phage T4 β -glucosyltransférase et GT-C provient de l'oligosaccharyltransférases de Campylobacter lari. Les hélices sont en rouge, les tonneaux beta sont en jaunes et les boucles sont vertes. D'après Gloster (2014).

Ces enzymes peuvent être retrouvées libres dans la cellule ou associées à des membranes. Les GT contenant un domaine membranaire, sont essentiellement retrouvées sur l'appareil de Golgi dans lequel a lieu l'assemblage des polysaccharides pariétaux (Figure 31) (Scheible and Pauly, 2004). De plus, certaines GT peuvent s'associer à d'autres protéines ou entre elles comme par exemple AtGalT29A et AtGalT31A appartenant à la famille des GT29. Elles s'associeraient physiquement pour produire des chaines latérales de β -1,6-galactane sur des protéines à arabinogalactanes (AGPs) (Dilokpimol et al., 2014; Geshi, 2014).



Figure 31 : Représentation simplifiée d'une glycosyltransferase membranaire golgienne. Les acides aminés Nterminaux sont localisés sur le coté cytoplasmique tandis que la partie C-terminale contenue dans le domaine globulaire se situe dans la lumière du golgi. D'après Keegstra and Raikhel (2001)

Dans le contexte présenté dans ce mémoire, c'est la famille des GT1 ou Uridine diphosphate Glycosyltransférase (UGT) qui sera développée par la suite.

2/- UDP-glycosyltransferase (ou GT1)

De façon générale, les UGTs catalysent l'ajout d'un glycane sur une molécule de faible poids moléculaire. Les membres de cette famille multigénique interviennent dans de nombreux mécanismes physiologiques et plus particulièrement dans les métabolismes spécialisés. Ils vont considérablement augmenter la diversité chimique permettant à la plante une meilleure adaptation à de nombreux stress biotiques et abiotiques par exemple (Vogt, 2010). La famille des UGTs est généralement la plus abondante chez les plantes vasculaires parmi les GT (Cao et al., 2015). Du point de vue stéréochimique, ces enzymes vont former une liaison glycosidique différente de la liaison d'origine qui lie le glycane à l'UDP; ce sont des enzymes « d'inversion ». Les UGTs

vont ainsi permettre l'ajout d'un glycane qui peut être principalement un glucose, galactose, xylose, rhamnose ou un arabinose (Arimura and Maffei, 2016).

2-1/- Structure des UGTs

a/- Structure primaire

Les domaines C-terminaux des UGTs des plantes semblent relativement bien conservés dans les différentes espèces (Campbell et al., 1997). Ces enzymes ont été classées à l'aide d'une nomenclature permettant de les regrouper selon leur homologie de séquences (Mackenzie et al., 1997; Ross et al., 2001). Un numéro est attribué à chaque famille de séquences ayant en commun au moins 40% d'homologie suivi d'une lettre associé aux séquences dans cette famille ayant au moins 60% d'homologie. Dans cette configuration la proximité des séquences UGTs sont facilement appréciables par une simple comparaison de leurs noms.



Figure 32 : Nomenclature utilisée pour nommer des UDP-glycosyltransférase. D'après Ross et al. (2001)

Des analyses phylogénétiques ont permis de mettre en lumière plusieurs groupes évolutionnaires distincts. Les études les plus récentes ont identifiés 24 (Yonekura-Sakakibara and Hanada, 2011) et 17 groupes d'UGTs (Caputi et al., 2012) possédant un ancêtre commun, en comparant différents organismes de la lignée verte. Bien qu'il existe des exceptions, il est généralement possible de considérer qu'un groupe génétique catalyse la réaction de glycosylation plutôt sur un même type de substrat (Caputi et al., 2012; Nishimura et al., 2010; Yonekura-Sakakibara and Hanada, 2011). Ces études phylogéniques réalisées sur quelques représentants de la lignée verte mettent en lumière une large expansion du nombre d'UGTs lors de la terrestrialisation (Caputi et al., 2012). Cette expansion est également concomitante avec la mise en place de nombreux métabolismes spécialisés comme la protection contre certains facteurs abiotiques tels que les UV, la sécheresse, l'hypoxie ainsi que des facteurs biotiques comme les bactéries, virus ou herbivores (Gachon et al., 2005). Une autre illustration de cette expansion d'UGTs est le nombre important de pseudogènes correspondants qui reflète les nombreuses duplications de cette famille multigénique (Paquette et al., 2003).

Dans la séquence primaire, il existe un motif très conservé présent dans la plupart des organismes de la lignée verte : le motif « Plant secondary product glycosyltransferase » (ou PSPG) (Figure 33 et 34).



Figure 33 : Répartition des glycosyltransférases dans les différentes familles CAZy au sein de la lignée verte. Les angiospermes eucotylédones sont en bleu, les monocotylédones sont en vert, les bryophytes en rouge et les algues vertes sont en mauve. D'après Cao et al. (2015)
Chez les plantes, seules les UGT80 et UGT81 qui ne possèdent pas de PSPG semblent dériver d'un ancêtre commun qui est plus éloigné que les autre UGTs possédant ce motif conservé. Cette séquence de 44 acides aminés a été identifiée pour la première fois comme signature dans la partie C-terminale de certaines UGTs en 1994 (Hughes and Hughes, 1994) (Figure 34). Depuis, il est maintenant connu que ce motif particulier est impliqué dans l'interaction avec l'UDP-sucre (Osmani et al., 2009).



Figure 34: Web logo représentant le motif conssensus « Plant secondary product glycosyltransferase » (PSPG) spécifique à la plupart des UGT chez les plantes. D'après Osmani et al. (2009)

Les données cristallographiques de quatre UGTs révèlent que parmi les 44 acides aminés, 10 d'entre eux, particulièrement conservés, sont impliqués directement dans cette interaction (Brazier-Hicks et al., 2007a; Li et al., 2007; Offen et al., 2006; Shao et al., 2005). Les autres participeraient à la stabilisation intramoléculaire de l'enzyme (Li et al., 2007). Par contre, ce motif n'intervient que peu ou pas dans l'interaction avec le substrat accepteur. De manière générale, il est très difficile de prédire un substrat à partir de la séquence primaire de l'enzyme lorsque peu de chose sont connue sur ces orthologues d'autres espèces (Modolo et al., 2007; Osmani et al., 2009). Chez *Dorotheanthus bellidiformis*, UGT73A5 et UGT71F2, ne possèdent que 19% d'homologie de séquence mais pourtant agissent toutes les deux sur des anthocyanes (Vogt, 2002). A l'inverse, chez *Allium cepa*, UGT73G1 et UGT73J1, qui appartiennent donc à la même famille, ont des activités (principalement des flavonoïdes) tandis que le deuxième est capable de transférer un glycane sur seulement deux flavonoïdes avec une régiosélectivité sur le carbone 7 (Kramer et al., 2003).

b/- Structures secondaires et tertiaires

Au contraire de la structure primaire, la compréhension des repliements de l'enzyme dans l'espace est essentielle pour l'identification des substrats accepteurs. De façon générale, l'ensemble des études de cristallographie menées sur des UGTs de plantes démontre que la plupart des acides aminés qui interagissent avec le substrat accepteur sont dans le domaine N-terminal (Brazier-Hicks et al., 2007a; Hiromoto et al., 2013; Li et al., 2007; Offen et al., 2006; Shao et al., 2005; Wetterhorn et al., 2016). A ce jour, il y a neuf structures cristallographiques d'UGTs végétales référencées dans la base de données « The Protein Data Bank » (www.rcsb.org; (Berman et al., 2000)). Ces UGT sont connues pour glycosyler des flavonoïdes, des mycotoxines comme le deoxynivalenol (DON) ou des xénobiotiques in vitro en utilisant principalement de l'UDP-glucose. Parmi le nombre important d'acides aminés impliqués directement dans la reconnaissance du substrat accepteur, la majorité des études s'accordent sur le fait qu'une histidine située au début du domaine N-terminal permet la première étape catalytique de la glycosylation, à savoir la déprotonation du substrat accepteur (Osmani et al., 2009). D'autres acides aminés sont indispensables pour le positionnement correct du substrat accepteur comme le montre des résultats de mutagénèse dirigée contre l'asparagine en position 142 de AtUGT74F1 (Cartwright et al., 2008). Sa mutation en tyrosine va alors conduire à un changement de regio-sélectivité de la glycosylation contre la quercetine.

c/- La glycosylation par des glycosyl-hydrolases

Les GH sont connues pour catalyser la réaction opposée aux UGTs en hydrolysant une liaison glycosidique (Henrissat and Davies, 2000) (Figure 29). Malgré tout, certaines enzymes de type GH peuvent également aboutir aux mêmes produits que des UGTs. Ces enzymes sont appelées transglycosidases car au lieu d'utiliser de l'UDP-sucre comme substrat donneur, elles vont utiliser des sucres activés sous la forme acyl-sucre (Henrissat and Davies, 2000). Un des exemples de glycosylation de ce type concerne la biosynthèse d'anthocyanes au niveau des pétales chez certaines plantes (Matsuba et al., 2010). Ces glycosylations avaient déjà été mises en évidence sur le carbone 5 (Ogata et al., 2004) et le carbone 7 (Davies, 2008; Yonekura-Sakakibara et al., 2007) d'anthocyanes mais l'enzyme responsable n'avait pas été identifiée du fait que les auteurs utilisaient de l'UDP-glucose comme substrat donneur. Par la suite, il a été démontré que c'est en fait le D-vanillyl-glucose qui était le donneur de glucose (Matsuba et al., 2010). Un autre exemple a été retrouvé au sein d'une population d'*Arabidopsis thaliana* en ségrégation comportant des différences de teneurs en flavonols glycosylés (Keurentjes et al., 2006; Stracke et al., 2009). Ainsi un QTL (quantitative trait locus) lié à un phénotype d'accumulation de flavonols glycosylés a été identifié. Bien qu'il n'y avait pas d'UGTs prédits dans ce QTL, les auteurs ont spéculé qu'un gène lié à la glycosylation était tout de même en cause. Par la suite, il s'est avéré que *AtBGLU6*, codant une GH avec une activité de transglycosylation était bien à l'origine de variation génétique (Ishihara et al., 2015).

Le fait que la plante utilise plusieurs types d'enzymes pour générer les mêmes produits pourrait permettre une plus grande plasticité. Par exemple, les UGTs sont essentiellement prédits pour être cytoplasmiques tandis que certaines GH possèdent des peptides signaux leur permettant d'être transportées dans la vacuole ou la paroi. Ainsi, Os9BGlu31, qui est localisé dans la vacuole, est capable de glycosyler un très large panel de molécules comme des phénylpropanoïdes vrais, des coumarines, des flavonoïdes ou des phytohormone *in vitro (Luang et al., 2013)*. Il manque encore des analyses fonctionnelles permettant d'identifier le rôle de cette enzyme mais il semblerait qu'elle soit impliquée dans des mécanismes tels que la senescence ou la sécheresse (Komvongsa et al., 2015; Luang et al., 2013).

2-2/- Régulation, localisation et interaction des UGTs

a/- Régulation des UGTs

De façon assez étonnante, et au vu du nombre important de publications concernant les UGTs, peu de choses sont connues concernant leur régulation. Si l'on considère les régulations au niveau transcriptionnel, c'est pour les UGTs impliquées dans la glycosylation des flavonoïdes que l'on a le plus d'informations. En effet, les étapes de glycosylations de certains flavonoïdes font partie intégrante de leur voie de biosynthèse et sont donc potentiellement régulées de la même manière que les gènes de biosynthèse. Par exemple, deux facteurs de transcription MYB de *Vitis vinifera*, VvMybPA1 et VvMybPA2, sont connus pour réguler positivement la biosynthèse des proanthocyanidine (tannins condensés) (Terrier et al., 2009). En plus de leur rôle de régulateur sur les gènes de la voie de biosynthèse, leur sur-expression à mis en évidence des différences d'expression de trois gènes *UGTs* (Khater et al., 2012). Les auteurs suggèrent qu'ils seraient impliqués dans les étapes de galloylation (glycosylation des acides galliques) des tannins condensés (Khater et al., 2012). De la même manière, la sur-expression du facteur de transcription de type MYB TT2 (Transparent testa 2) dans des chevelus racinaires de *Medicago truncatula* conduit à l'accumulation de tannins

condensés (Pang et al., 2008). En comparant les données transcriptomiques de ces cultures avec celles des téguments des graines (tissus accumulant naturellement des tannins condensés), les auteurs ont identifié le gène *UGT72L1*. Celui-ci est un bon candidat pour la formation d'epicatechin 3'-O-glucoside et semble être régulé par le facteur de transcription TT2 (Pang et al., 2008, 2013). Un autre exemple qui décrit une régulation transcriptionnelle d'UGTs concerne la voie de biosynthèse des flavonoles (flavonoïde impliqué entre autre dans la protection aux UV). Le facteur de transcription MYB7 semble être un répresseur de cette voie de biosynthèse comme le montre le mutant d'*Arabidopsis thaliana* déficient en MYB7 qui accumule ces métabolites (Fornalé et al., 2014). En plus de sur-exprimer les gènes « classiques » de biosynthèse des flavonoles, le mutant sur-exprime *AtUGT73B2* (Fornalé et al., 2014).

La régulation de l'expression de la plupart des UGTs est modulée par de nombreux stress biotiques et abiotiques comme le montre une analyse transcriptomique sur l'ensemble de la famille multigénique des UGTs d'*Arabidopsis thaliana* (Saint von Paul et al., 2011). Pour des stress complexes et multifactoriels, les UGTs pourraient jouer un rôle central qui reste encore à définir. C'est le cas par exemple du stress salin qui provoque des changements dans l'expression de certaines UGTs chez *Gossypium davidsonii* (Zhang et al., 2016), *Brassica napus* (Long et al., 2015) et chez *Arabidopsis thaliana* (Sun et al., 2013b).

b/- Localisation cellulaire et sub-cellulaire des UGTs

De façon intéressante, la localisation sub-cellulaire des UGTs au sein de la lignée verte est différente des mammifères. En effet, chez ces derniers, les UGTs possèdent un domaine transmembranaire qui conduit à leur fixation au niveau de réticulum endoplasmique (Radominska-Pandya et al., 1999). La très grande majorité des UGTs de la lignée verte ne possède pas de domaines transmembranaires et seraient ainsi cytosoliques d'après les prédictions *in silico* (par des programme tel que TopPred2, SignalP et Psort) (Ross et al., 2001). Ces analyses ont été corroborées par des constructions couplant des gènes *UGTs* avec *GFP*. Parmi ces UGTs analysées, il y a UGT78D2 qui peut transférer un glycane sur certains flavonoïdes (Lee et al., 2005), UGT71B6 et UGT71C5 potentiellement impliquées dans la glycosylation de l'acide abscissique (Dong and Hwang, 2014; Liu et al., 2015) et enfin trois UGTs potentiellement impliquées dans la glycosylation de dérivés d'acides gibbérelliques chez *Stevia*

rebaudiana (Humphrey et al., 2006). A ce jour, bien que les analyses *in silico* se confirment *in vivo*, peu d'UGTs ont véritablement été caractérisées pour cet aspect et donc une certaine prudence s'impose à ce sujet. A l'inverse, certaines UGTs qui peuvent être prédites comme étant chloroplastiques ou retenues dans le réticulum endoplasmique comme UGT85A1 (Woo et al., 2007) se sont révélées, en réalité, être cytosoliques (Šmehilová et al., 2016).

D'autres localisations sont également possibles. Le mutant *sk1* (pour silkless-1) chez Zea mays possède des pistils altérés (Hayward et al., 2016). Le gène muté SK1 responsable de ce phénotype a été identifié comme tant un orthologue de AtUGT82A1 qui semble interagir avec la voie de biosynthèse de l'acide jasmonique. Des expériences de localisation sub-cellulaire ont permis d'identifier la protéine correspondante dans les péroxysomes qui sont un des lieux de synthèse de ce régulateur (León, 2013). Cette UGT possède une séquence de trois acide aminés (-SVL) dans sa séquence C-terminale proche de la séquence canonique d'adressage vers les péroxysomes (-SKL) (Hu et al., 2012). Un autre exemple concerne Arabidopsis thaliana avec l'UGT72E1 qui semblerait être exportée vers le noyau comme le montre des expériences d'expressions transitoires dans des feuilles de tabac (Huang et al., 2014). Dans ce cas, aucun peptide signal connu n'a été identifié mais de façon intéressante, UGT71E1 pourrait interagir physiquement avec une protéine MAPKKK (mitogen-activated protein kinase kinase kinase) nucléaire appelée SIS8 (pour Sugar InSensitive-8). Le fait que les mutants sis8 et ugt72e1 montrent les mêmes phénotypes de résistance à de fortes teneurs en saccharose suggère également que ces deux protéines interagissent ensemble (Huang et al., 2014). Il reste encore à déterminer comment UGT72E1 peut entrer dans le noyau.

c/- Interactions des UGTs avec d'autres protéines

Comme évoqué dans l'exemple de l'interaction entre UGT72E1 et SIS8 (Huang et al., 2014), certaines UGTs peuvent interagir physiquement avec d'autres protéines. Il y a encore peu d'interactions décrites à ce jour bien qu'elles semblent conditionner de façon importante l'activité de glycosylation. Par exemple, l'UGT75B1 (précédemment appelée UGT1) semble s'associer au sein d'un grand complexe protéique à l'origine de la synthèse de callose au niveau des plaques cellulaires lors de la mitose (Hong et al., 2001a). Ce dépôt de callose est contrôlé par des protéines qui interagissent physiquement avec UGT75B1 (Hong et al., 2001a, 2001b) (Figure 35). Les auteurs

suggèrent que l'élongation du polymère de callose serait issue de l'action coordonnée d'une sucrose synthase (SuSy), également présente dans ce complexe (Amor et al., 1995), qui serait chargée de décomposer le saccharose et ainsi de fournir de l'UDP-glucose à l'UGT75B1 pour l'ajouter à la chaîne de callose (Hong et al., 2001b).



Figure 35 : Modèle putatif du complexe protéique responsable de la formation de callose au niveau de la plaque cellulaire lors de la mitose. D'après Hong et al. (2001b).

Un autre exemple d'association avec d'autres protéines existe lors de la biosynthèse de dhurrin. Cette molécule, dérivée de la tyrosine, est impliquée dans la défense contre les herbivores (Tattersall et al., 2001). Deux cytochromes P450 CYP71E1 et CYP79A1 forment un hétérodimère au niveau de la membrane du réticulum endoplasmique et vont recruter UGT85B1 pour former un métabolon (Laursen et al., 2016; Nielsen et al., 2008) (Figure 36). Ce complexe protéique peut réaliser jusqu'à quatre réactions différentes sur la même molécule sans que celle-ci soit relarguée dans le cytoplasme.



Figure 36 : Modèle putatif du metabolon constitué de deux cytochromes P450 membranaires CYP79A1 et CYP71E1 et de l'UDP-Glycosyltransferase UGT85B1. D'après Nielsen et al. (2008).

2-3/- Rôle des UGTs

Les UGTs interviennent dans des phénomènes de glycosylation au sein d'un très grand nombre de voies de biosynthèse de métabolites spécialisés, hormones ou xénobiotiques dans la plante et donc participent activement à de nombreux mécanismes physiologiques (Gachon et al., 2005; Le Roy et al., 2016). L'objectif de cette partie n'est pas d'être exhaustif, et donc seuls quelques exemples liés au développement et à des stress seront traités. Ces informations nous ont orienté notamment dans la recherche des rôles biologique des UGTs dans nos travaux.

a/- Relation entre la coniférine et la lignification

La coniférine est un β -D-glucoside de l'alcool coniférylique. Elle est produite par la même voie de biosynthèse que les précurseurs de lignines, en l'occurrence l'alcool coniférylique, et finit par être glycosylée par des UGTs (Le Roy et al., 2016). Elle est accumulée de manière très importante dans la sève cambiale de nombreux gymnospermes et ce jusqu'à 10 mmol/g de matière fraiche (Freudenberg and Harkin, 1963; Leinhos and Savidge, 1993; Savidge, 1991; Schmid et al., 1982). Sa production atteint son maximum aux premiers stades de la formation des parois secondaires des fibres trachéides (formation de la couche S1) chez *Pinus thunbergii* (Fukushima et al., 1997). Récemment, une étude a pris avantage des dernières avancées technologiques de spectrométrie de masse pour suivre la localisation tissulaire et cellulaire de la coniférine dans la tige de *Ginkgo biloba* (Aoki et al., 2016). Les auteurs ont montré que la quantité maximale de coniférine se situait au niveau des fibres trachéides proche du cambium (Figure 37). Plus la différenciation des cellules de fibres trachéides était avancée, moins il y avait de coniférine, ce qui suggèrerait qu'elle pourrait participer à la formation des lignines (Aoki et al., 2016). Cette observation est en accord avec d'autres dosages de coniférine faites à la jonction du cambium et des fibres trachéides chez d'autres ligneux (Morikawa et al., 2009; Tsuyama and Takabe, 2014). L'analyse montre en plus que la coniférine n'est pas stockée dans les rayons xylémiens et donc suggère que leur synthèse est endogène (Aoki et al., 2016). Les auteurs font également état d'une localisation vacuolaire pour la coniférine, ce qui va dans le sens d'autres travaux publiés récemment (Alejandro et al., 2012; Dima et al., 2015; Liu et al., 2011).



Figure 37 : Coupes transversales de tige de Ginkgo obtenues lors de deux expériences. (a) : Visualisation de la coniférine endogène par cryo-TOF-SIMS (points mauves) au niveau des fibres trachéides en cours de différenciation. (b) : Autoradiographie montrant des lignines radioactives issues de l'administration de ¹⁴C-coniférine. X : zone cambiale ; Y : Début de lignification de la lamelle moyenne ; Z : Début de l'épaississement de la paroi secondaire. D'après Aoki et al. (2016)

Chez *Arabidopsis thaliana*, trois UGTs ont été identifiées comme étant des acteurs de la glycosylation *in vitro* et *in vivo* des monolignols et, de manière plus générale, des phénylpropanoïdes vrais (Lanot et al., 2006, 2008; Lim et al., 2001). Le Tableau 2 montre les variations des deux principaux monolignols glycosylés pour les mutants silencing/d'insertion et de sur-expression de *UGT72E1*, *UGT72E2* et *UGT72E3*. Mais ces lignées ne semblent pas révéler de modifications au niveau des quantités de lignines dans les parois cellulaires. Une étude approfondie de ces trois gènes sera rapportée dans le paragraphe III. Tableau 2 : Bilan des dosages de coniférine et de syringine dans les mutants ugt72e1, ugt72e2 et ugt72e3. D'après Lanot et al. (2006, 2008).

				Coniférine	Syringine		
				Comparé à Col-0			
RNAi ou KO	ugt72e1	SALK_078702	Racine à la lumière	-	-		
	ugt72e2kd	RNAi	Racine à la lumière	Baisse	Baisse		
	ugt72e3kd	RNAi	Racine à la lumière	-	-		
	ugt72e1-2-3kd	RNAi	Racine à la lumière	Forte baisse	Forte baisse		
Sur-expresseur	255 1107725105		Racine à la lumière	Augmentation	Augmentation		
	353_00172E10E		Feuille de 4 semaines	Augmentation	-		
	255 1107725205		Racine à la lumière	Forte augmentation	Augmentation		
	355_001722202		Feuille de 4 semaines	Forte augmentation	Augmentation		
	255 1167725205		Racine à la lumière	Augmentation	Augmentation		
	333_00172E30E	1	Fouillo do 4 somainos	Forte augmentation	Forte augmentation		

b/- Les UGTs dans la senescence

La senescence est un mécanisme complexe et hautement régulé qui précède la mort cellulaire programmée (Li et al., 2014b). Au sein de la plante, ce mécanisme fait intervenir jusqu'à quatre voies liées à des phytohormones : l'acide jasmonique, l'éthylène, les cytokinines et l'acide salicylique (Buchanan-Wollaston et al., 2005). Dans le but de contrôler les conditions expérimentales, certains auteurs ont développé des systèmes pour induire la senescence par un traitement à l'obscurité. La senescence provoquée par ce traitement est très similaire à la senescence naturelle mais possède quelques différences notamment au niveau de l'implication de l'acide salicylique qui devient plus mineure dans ce cas (Buchanan-Wollaston et al., 2005). Certaines UGTs apparaissent comme importantes dans ce processus comme c'est le cas du gène SDG8i (pour Sporobolus drought gene 8i) identifié chez Sporobolus stapfianus du fait de son activation lors d'un stress hydrique sévère (Le et al., 2007). Lorsque ce gène est surexprimé chez *Arabidopsis thaliana*, la plante montre une augmentation de taille, une plus forte tolérance au froid et à la sécheresse mais également un retard de senescence lorsque celle-ci est induite (Islam et al., 2013). Les auteurs suggèrent que du fait de l'implication de cette UGT dans de nombreux mécanismes, celle-ci participerait à la division et l'expansion cellulaire (Islam et al., 2013). Chez Arabidopsis thaliana, il y a de nombreux gènes UGTs sur-exprimés lors de la senescence comme UGT73B3, UGT73B5 ou UGT73D1 (Langlois-Meurinne et al., 2005) mais à ce jour aucun lien direct entre leur activité et ce phénomène physiologique n'a pu être montré. En revanche, le mutant *ugt76b1* possède une plus forte résistance aux pathogènes biotrophiques et on a pu noter que la senescence était accélérée sur des feuilles détachées à l'obscurité (Saint von Paul et al., 2011). Des analyses globales de métabolites sur ce mutant, mais également sur le sur-expresseur du gène correspondant, a montré que l'acide isoleucique (2hydroxy-3-methyl-pentanoic acid) pouvait être le substrat de cette UGT (Saint von Paul et al., 2011).

c/- Le métabolisme du fer

Lorsqu'il y a un déficit en fer dans le milieu, il y a globalement trois phénomènes qui seront observés (Curie and Mari, 2016) : une altération du pool de fer lié aux parois des cellules racinaires, la sécrétion de flavines (molécules à hétérocycle azotées), coumarines, coumarino-lignanes et dérivés glycosylés pour capter et solubiliser le fer (Sisó-Terraza et al., 2016) et une inhibition de la subérisation des cellules de l'endoderme pour faciliter le transport de fer.

Ces trois actions qui font intervenir de nombreux gènes régulés en partie par l'éthylène et l'acide abscissique (Curie and Mari, 2016). Parmi ces gènes, on peut citer ceux codant des enzymes de biosynthèse des phenylpropanoides comme PAL, C4H, 4CL ou *F6'H-1* (Feruloyl-CoA 6'-hydroxylase) (Rodríguez-Celma et al., 2013). Cette dernière intervient lors des premières étapes de synthèse des coumarines (Figure 8) et semble essentielle pour leur sécrétion lors d'un déficit de fer (Kai et al., 2008; Schmid et al., 2014). La glycosylation de ces composés semble aussi être importante pour cette sécrétion. En effet, chez le mutant KO pour le gène de ß-glucosidase 42 d'Arabidopsis thaliana (AtBGLU42) qui cliverait les coumarines glycosylées, cette sécrétion de coumarine est très limitée (Zamioudis et al., 2014). Il semblerait donc que ces métabolites soient déjà présents au sein de la plante sous des formes glycosylées et c'est seulement lors d'un stress (manque de fer) qu'ils seraient déglycosylés puis exportés dans le milieu extérieur (Clemens and Weber, 2016). Les acteurs de ces glycosylations sont pour l'instant inconnus mais il est intéressant de noter que UGT72E1 est surexprimé lors d'un déficit en fer (Mai et al., 2016). Cette UGT est connue pour son action sur des phénylpropanoides vrais (Lanot et al., 2006, 2008; Lim et al., 2003) mais la protéine recombinante correspondante n'a pas été testée in vitro contre les différentes coumarines. On sait par ailleurs que les hydroxycoumarines sont reconnues par une très grande majorité d'UGTs chez Arabidopsis car sur 90 protéines testées, 48 étaient capables de les reconnaître (Lim et al., 2003).

d/- Les UGTs sont liés aux stress

Les phytohormones sont des acteurs essentiels intervenant lors d'un stress abiotique. Elles vont participer à la signalisation afin de permettre une réponse adaptée lors des stimuli internes et externes (Wani et al., 2016). De plus, elles peuvent également dialoguer entre elles lors d'un même stress (Figure 38). Bien qu'agissant en très faible concentration, les phytohormones sont indispensables pour la plante et leur glycosylation par des UGTs peut avoir des impacts importants lors de stress abiotiques (Le Roy et al., 2016; Ostrowski and Jakubowska, 2014). Ainsi, les UGTs liées à la glycosylation des phytohormones sont essentielles pour l'adaptation des plantes aux stress abiotiques et les mutants correspondants peuvent avoir un phénotype très fort (Liu et al., 2015; Wang et al., 2011).



Figure 38 : Les rôles possibles des phytohormones lors de stress abiotiques et leurs interactions. CK : Cytokinines, Aux : Auxines, SA : acide salicylique, ABA : Acide abscissique, ET : Ethylène, BR : Brassinostéroides, SL : Strigolactone, GA : Acide gibbérilique. D'après Wani et al. (2016).

Le Tableau 3 met en évidence une sélection de mutants pour différentes UGTs qui présentent des modulations de tolérance lors de stress biotique et abiotique. De toute évidence, certaines UGTs sont impliquées dans la glycosylation de phytohormones mais pour d'autres, les substrats restent encore inconnus. Il va de soit que toutes les UGTs d'Arabidopsis n'ont pas été testées pour l'ensemble des stress du tableau. Par exemple, l'UGT74C1 qui aurait pour substrat l'auxine et ses dérivés est essentielle pour le développement correct de la plante (Grubb et al., 2014). Du fait du phénotype très marqué de la plante mutante *ugt74c1*, aucun stress n'a été appliqué mais il n'est pas exclu qu'UGT74C1 soit impliqué dans ces stress abiotiques.

Tableau 3 : Sélection d'UGTs impliquées dans des stress abiotiques et biotiques. Les plantes mutantes (perte de fonction ou sur-expresseur) ont été identifiées par des modulations des résistances/tolérances aux stress. IBA : Indole-3-butyric acid ; ABA : acide abscissique ; TNT : 2,4,6-trinitrotoluene.

Nom	Substrat putatif	Stress osmotique	Stress salin	Sécheresse	Froid	Xénobiotique	Photopériode	Glucose / saccharose	Biotique	Références
UGT71C5	ABA			х			х	х		(Z. Liu et al. 2015; Zhen Liu et al. 2015; Yan et al. 2010)
UGT72E1	Phenylpropanoïdes vrais							х		(Huang et al. 2014)
UGT74E2	IBA, TNT		х	х		Х			Х	(Tognetti et al. 2010; Gandia-Herrero et al. 2008; Park et al. 2011)
UGT75D1	IBA	Х	Х							(Gui-Zhi Zhang et al. 2016)
UGT76C2	Cytokinines	Х		Х						(Wang et al. 2011; Li et al. 2015)
UGT79B2/B3	Anthocyanes		Х	Х	Х					(Li et al. 2017)
UGT85A5	?		Х							(Sun et al. 2013b)
UGT87A2	?	Х	Х	Х						(Li et al. 2016)

Par ailleurs, le rôle des UGTs dans la réponse aux stress peut se refléter de manière différente. Par exemple, la surexpression du gène *AtUGT85A5* chez le tabac permet à la plante d'être plus résistante à la salinité (Sun et al., 2013b). Ces plantes vont ainsi voir leur taux de germination et leur teneur en chlorophylles augmenter comparé aux plantes sauvages. De plus, elles vont accumuler plus de prolines et sucres (important pour l'homéostasie des cellules) et induire l'expression de gènes codant des transporteurs de sucres ou de protéines LEA.

VI- A la recherche d'un lien entre la glycosylation et la lignification chez le lin

Si l'on compare la composition des parois secondaires rencontrées dans les différents types cellulaires chez les plantes, les parois des fibres primaires de lin font partie des exceptions. Elles sont en effet qualifiées de parois hypolignifiées du fait qu'elles contiennent de très faibles quantités de lignines qui sont de l'ordre de 2 à 4 %, par opposition aux parois des cellules du xylème qui en contiennent environ 25 % (Day et al., 2005). Les tiges de lin intègrent ainsi deux types cellulaires aux parois secondaires contrastées mais physiquement proches les uns des autres et l'on considère donc cette espèce comme étant un modèle d'étude d'intérêt pour la lignification. Les analyses sont d'autant plus facilitées qu'il est relativement aisé de séparer mécaniquement les tissus externes contenant les fibres primaires des tissus internes contenant le xylème.

En plus d'avoir une quantité de lignines moindre, leur composition dans les fibres primaires est également différente de celle des cellules du xylème (Chantreau et al., 2014; Day et al., 2005; del Río et al., 2011). Des analyses par thioacidolyse montrent des taux molaires des unités H : G : S de l'ordre de 7 : 71 : 21 pour les lignines des fibres comparé à 1 : 82 : 17 pour les tissus xylémiens des tiges de lin (Chantreau et al., 2014). De manière générale, les lignines de lin contiennent donc de fortes proportions en unités G entrainant la forte présence de lignines condensées (40%–50%) incluant des liaisons β - β (resinol; 9%) et β -5 (phenylcoumaran; 14%) (del Rio et al., 2011). De plus, les lignines des fibres primaires contiennent des taux inhabituels d'unités H (environ 25%), ce qui est très particulier pour des eudicotylédones.

Des analyses de métabolomique ciblée sur les molécules phénoliques qui composent les lignines, qualifiés ainsi de "lignomics", ont été menées par des approches UHPLC-FT-ICR-MS (Morreel et al., 2004b, 2010). Cette approche a été utilisée pour analyser indépendamment les tissus internes et externes de la tige de lin (Huis et al., 2012). Sur les 81 composés aromatiques identifiés lors de cette analyse, la grande majorité était accumulée préférentiellement dans les tissus externes. Par ailleurs, et de façon intéressante, la quasi-totalité des molécules (monolignols, oligolignols et (neo)lignanes) présentes sous leurs formes glycosylées ont été retrouvées plus abondamment dans ces mêmes tissus. Il a alors été possible d'envisager un lien entre les faibles quantités de lignines et l'accumulation de leurs précurseurs sous forme glycosylés.

Une population de mutants de lin obtenue par mutagenèse EMS à été créée récemment, ce qui constitue une base intéressante pour des analyses de génétique formelle (Chantreau et al., 2013). Certains mutants, obtenus par cette approche, sont impactés au niveau des fibres primaires comme c'est le cas du mutant *lbf1* (pour *lignified bast fiber1*) qui possède des fibres hyperlignifiées (Chantreau et al., 2014). Bien que la mutation à l'origine de ce phénotype n'ait pas encore été caractérisé, il est maintenant clair que l'hypolignification observée dans les plantes sauvages n'est pas due à des contraintes physiques et qu'il serait bien de nature génétique. Des analyses de transcriptomique globale montrent qu'il existe chez ce mutant, une dérégulation de certains gènes de la voie de biosynthèse des monolignols mais aussi de peroxydases ou encore de NADPH-oxydases dont l'expression est augmentée chez le mutant lbf1. Ces analyses mettent également en évidence une diminution significative de l'expression d'un gène UGT orthologue d'AtUGT72E1, AtUGT72E2 et AtUGT72E3 responsables de la glycosylation d'un certain nombre de précurseurs de lignines. Encore une fois, il existe des indications allant dans le sens d'un lien et plus précisément d'un système de régulation négatif de la lignification par la glycosylation dans les fibres de lin.

VII- Objectifs du travail de thèse

L'action des phénylpropanoïdes est essentielle pour le développement et l'adaptation des plantes à l'environnement. Parmi les nombreux mécanismes impliquant ces molécules, la lignification est sans doute l'un des plus importants. La déposition de lignines au niveau des parois cellulaires permet de les renforcer pour le transport d'eau et de lutter contre la gravité. La biosynthèse de leurs précurseurs est hautement régulée sur plusieurs niveaux moléculaires et c'est sans doute une combinaison de plusieurs types de régulations qui est à l'origine du caractère hypolignifié des fibres primaires de *Linum usitatissimum*. Mes travaux de thèse se placent donc dans un contexte de compréhension des mécanismes moléculaires mises en œuvre lors du développement de la paroi des fibres primaires chez cette espèce.

Afin d'étudier les aspects liés à la lignification chez le lin, il nous a tout d'abord paru indispensable de distinguer clairement les gènes liés à la biosynthèse des monolignols des autres gènes conduisant à la production des phénylpropanoïdes cités précédemment. L'objectif de la première partie de ce travail de thèse a donc été d'identifier ces gènes afin de constituer un socle solide pour de futures études fonctionnelles. Ce travail a également inclus un aspect de compréhension de la régulation post-transcriptionnelle de la phase de polymérisation par l'étude de l'action d'un microARN, miR397, qui va cliver certains transcrits de gènes laccases comme cela a été montré chez différentes espèces (Jin et al., 2016; Lu et al., 2013; Wang et al., 2014a; Zhang et al., 2013b).

Dans la seconde partie, les acteurs de la glycosylation des phénylpropanoïdes, et plus particulièrement des monolignols, ont été étudiés. Précédemment, il a été observé une accumulation de monolignols et d'oligolignols glycosylés dans les tissus externes hypolignifiés de la tige de lin (Huis et al., 2012). L'objectif de cette partie était donc de rechercher de possibles liens entre ces deux caractéristiques dans ce tissu.

Les gènes *UGTs* (UGTome) ont tout d'abord été identifiés dans le génome de *Linum usitatissimum* (Wang et al., 2012b) par une approche *in silico* et leur profil d'expression déterminé par une approche qRT-PCR haut débit. Par la suite, nous avons caractérisé les profils d'activités de glycosylation d'un certain nombre de phénylpropanoïdes dans les organes de la plante. L'étape suivante consistait à rechercher les acteurs spécifiques, responsables de la glycosylation des monolignols et de leurs précurseurs. Ainsi, cinq enzymes ont plus particulièrement été étudiées d'un point de vue biochimique et structurale.

Enfin la dernière partie du travail a été consacrée à la recherche du rôle biologique des UGTs en lien avec la lignification. Dans la mesure où l'obtention de plantes transformées est très difficile chez le lin, nous avons logiquement choisi d'utiliser le modèle Arabidopsis. Chez cette espèce, trois UGTs ayant une activité en lien avec les précurseurs de lignines sont connues. Nous avons ainsi produit des triples mutants pour ces gènes que nous avons caractérisés. Un phénotypage approfondi de ces plantes a été mené et nous avons pu montrer un lien entre ces gènes et le métabolisme des lignines.

Résultats

I- Identification des gènes de biosynthèse des phénylpropanoïdes chez le lin

Avant propos

Comme évoqué dans la première partie de ce manuscrit, le lin est un modèle pertinent pour l'étude de la synthèse des lignines et de leur régulation. La biosynthèse des monomères qui vont être intégrés dans le polymère se déroule en plusieurs étapes faisant intervenir plus d'une dizaine d'enzymes différentes. Ces enzymes sont en général issues de familles multigéniques qui possèdent souvent la même fonction catalytique mais elles peuvent parfois conduire à la production de molécules très différentes. Dans le cadre de notre travail, nous nous sommes essentiellement intéressés à la production de précurseurs de lignines et il est donc primordial pour nous de savoir quels sont les gènes/enzymes qui sont spécifiques de leur production par rapport à d'autres acteurs impliqués dans la biosynthèse des autres phénylpropanoïdes. Cette étape est essentielle pour cibler ultérieurement des gènes permettant de modifier la production de lignines ou de leurs précurseurs glycosylés. Ainsi, dans cette première partie du travail, l'ensemble des gènes intervenant dans la biosynthèse générale des phénylpropanoïdes ont été identifiés en mettant à profit la disponibilité de la séquence du génome de lin. Ensuite, afin de sélectionner ceux véritablement impliqués dans la biosynthèse des précurseurs jusqu'aux monolignols, l'expression de ces gènes dans différents organes a été analysée. Enfin pour ceux qui semblaient être spécifiques des tissus lignifiés, des expériences d'hybridation in situ ont été réalisées pour confirmer leur rôle dans cette voie de biosynthèse.

A partir des données de transcriptomique obtenues, des analyses de coexpression de gènes ont montré que la plupart d'entre eux étaient co-régulés. L'éventualité d'un rôle des facteurs MYBs a été évoquée et la régulation de la polymérisation des monolignols par une famille de microARNs a également été montrée.

En plus du suivi de l'expression de ces gènes dans différents tissus, leur modulation dans des conditions de stress a été analysée par l'ajout d'acide salicylique et de methyl-jasmonate et en réalisant des stress abiotiques en conditions de sécheresse ou en blessant les plantes. L'ensemble de ces travaux a conduit à la rédaction d'un article soumis à la revue BMC Plant Biology. Il a été accepté sous conditions de révisons mineures et au moment de l'écriture de ce manuscrit, nous sommes en attente de son acceptation définitive.

Article : Spatial regulation of monolignol biosynthesis and laccase genes control developmental and stress-related lignin in flax

Julien Le Roy, Anne Sophie Blervacq, Anne Créach, Brigitte Huss, Simon Hawkins and Godfrey Neutelings*

Univ. Lille, CNRS, UMR 8576 - UGSF - Unité de Glycobiologie Structurale et Fonctionnelle, F-59000 Lille, France

Running title: Flax lignin gene regulation

*Corresponding author: Godfrey Neutelings godfrey.neutelings@univ-lille1.fr Tel: 00 33 320 434 029

Abstract

Background

Bast fibres are characterized by very thick secondary cell walls containing high amounts of cellulose and low lignin contents in contrast to the heavily lignified cell walls typically found in the xylem tissues. To improve the quality of the fiber-based products in the future, a thorough understanding of the main cell wall polymer biosynthetic pathways is required. In this study we have carried out a characterization of the genes involved in lignin biosynthesis in flax along with some of their regulation mechanisms.

Results

We have first identified the members of the phenylpropanoid gene families through a combination of *in silico* approaches. The more specific lignin genes were further characterized by high throughput transcriptomic approaches in different organs and physiological conditions and their cell/tissue expression was localized in the stems, roots and leaves. Laccases play an important role in the polymerization of monolignols. This multigenic family was determined and a miRNA was identified to play a role in the posttranscriptional regulation by cleaving the transcripts of some specific genes shown to be expressed in lignified tissues. *In situ* hybridization also showed that the miRNA precursor was expressed in the young xylem cells located near the vascular cambium. The results obtained in this work also allowed us to determine that most of the genes involved in lignin biosynthesis are included in a unique co-expression cluster and that MYB transcription factors are potentially good candidates for regulating these genes.

Conclusions

Target engineering of cell walls to improve plant product quality requires good knowledge of the genes responsible for the production of the main polymers. For bast fiber plants such as flax, it is important to target the correct genes from the beginning since the difficulty to produce transgenic material does not make possible to test a large number of genes. Our work determined which of these genes could be potentially modified and showed that it was possible to target different regulatory pathways to modify lignification.

Keywords

Flax, laccase, lignin, microRNA, monolignol, phenylpropanoids, stress

Background

Lignin is among the most abundant biological polymers on earth. It is a complex aromatic molecule synthesized during the onset of the secondary cell wall (SCW) formation in plants providing stiffness proprieties for mechanical strength, hydrophobicity for water transport and contributes to defense against pests and pathogens. Lignin is produced by a complex biosynthetic pathway (Fig. 1) which also leads to the production of a wide range of phenylpropanoids with diverse and sometimes unknown functions such as hydroxycinnamic acids, flavonoids, coumarins, chalcones, phenylpropenes and stilbenes [1]. In brief, phenylalanine derived from the shikimate pathway is used as an initial substrate [2]. This amino acid is first deaminated by phenylalanine ammonia-lyase (PAL; EC 4.3.1.5), hydroxylated by cinammate 4-hydroxylase (C4H; EC 1.14.13.11) and then esterified with CoA by 4-coumarate:CoA ligase (4CL; EC 6.2.1.12) forming *p*-coumaroyl CoA. This metabolite can lead to the formation of *p*-coumaryl alcohol by the action of cinnamoyl CoA reductase (CCR; EC 1.2.1.44) and cinnamyl alcohol dehydrogenase (CAD; EC 1.1.1.195). p-coumaroyl CoA can also be successively transformed by hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl transferase (HCT; EC 2.3.1.133) and *p*-coumarate 3-hydroxylase (C3'H; EC 1.14.14.1) to form caffeoyl CoA that can be methoxylated by caffeoyl CoA 3-O-methyltransferase (CCoAOMT; EC 2.1.1.104) into feruloyl CoA. This CoA ester can be reduced into coniferaldehyde by CCR and further transformed into coniferyl alcohol by CAD. Both molecules are hydroxylated by ferulate 5hydroxylase (F5H; EC 1.14.13) and methoxylated by caffeate/5-hydroxyferulate O-methyl-transferase (COMT; EC 2.1.1.68) forming sinapaldehyde and sinapyl alcohols respectively. In addition to *p*-coumaryl alcohol, both alcohols are named monolignols. These can be transported to the apoplast and transformed into radicals by laccases and peroxidases [3, 4]. The oxidative coupling of these activated monolignols leads to the production of *p*-hydroxyphenyl (H), syringyl (S) and guaiacyl (G) units respectively, present in different proportions within the lignin fractions depending on the taxonomic families, species, tissues or environment [5]. The lignin biosynthesis pathway has been comprehensively characterized in model species such as Arabidopsis and the identified gene models then used to search for orthologous sequences in more economically-important plants including food crops as well as woody or fiber species.

Bast fibers are present in bundles within the stem cortex between the epidermis and the xylem core in some non-woody plants. These bundles may contain only primary fibers derived from the procambium in species such as flax (*Linum usitatissimum*) or ramie (*Boehmeria nivea*) while others such as jute (*Corchorus species*), kenaf (*Hibiscus cannabinus*) or hemp (*Cannabis sativa*) contain additional

secondary fibers derived from the vascular cambium. Fiber species are considered as interesting biological models because their stems contain two main populations of cells showing highly contrasted SCW compositions [6]. The xylem cells from the inner-stem tissues have a typical SCW structure with up to 30% of the cell wall weight represented by ligning whereas the bast fibers possess cellulose-rich thick SCWs with lignin contents ranging from less than 1 % (ramie) to 19 % (kenaf) [7]. The cell wall structure and composition is very important in these species because they determine the properties of the extracted raw materials. Flax, for instance, has been used for several thousands of years to make ropes and linen tissues and more recently, for the production of environmentally friendly fiber-based composite materials [8, 9]. To optimize the use of these materials in the future, targeted engineering of the cell wall to improve product quality will require a thorough understanding of the biosynthetic pathways leading to the production of the main cell wall polymers. Over the past years, significant efforts including genome sequencing [10], development of molecular tools [11-13] and databases [14, 15] have allowed us not only to improve our understanding of flax fiber cell wall biology, but also to identify the different gene models potentially associated with the biosynthesis of different cell wall polymers. Nevertheless, many cell wall genes are part of multigenic families and a major challenge to engineering is the identification of the individual family members involved in lignification. For non-model plants such as flax, it is important to target the correct genes from the beginning since rapid and easy transformation to obtain high amounts of transgenic material is particularly difficult [16].

To clearly identify the genes that are indeed involved in flax lignin production, we first conducted a high throughput reverse-transcriptase quantitative PCR (HT-RT-qPCR) approach to establish the expression profile of *in silico*-identified genes using organs/tissues with contrasted lignin contents. To further characterize these genes and because it was not possible to use large-scale reporter gene transformation, we used *in situ* hybridization to confirm the expression of the identified genes in lignified tissues. This type of approach has already been used to detect some lignin gene transcripts in woody plants [17-19] but never on such a large scale. We also benefited from the development of this technique in flax to show that a microRNA, *miR397*, already shown to be involved in the cleavage of several laccase transcripts [20, 21] is co-expressed in the same cells as the lignin biosynthetic genes confirming the role of miRNAs on laccase regulation in flax. In the context of a possible impact of global climate changes on the quality of plant products, and because lignin is an important component in stress responses [5], the expression of the identified genes was also analysed under a range of different environmental conditions.

The high amount of data obtained by HT-RT-qPCR has enabled us to show that the expression of most of these genes is coordinated, with a probable implication of MYB transcription factors.

Results

Identification of genes related to phenylpropanoid biosynthesis in flax

The whole genome shotgun assembly of flax [10] was first surveyed to identify the genes potentially related to phenylpropanoid biosynthesis. BLAST analyses were carried out to search against the 88,420 scaffolds including 43,471 gene models. A total of 69 genes were identified and organized in gene families containing between 3 and 15 members (Table 1) and the presence of the corresponding transcripts for each gene was checked among the different public EST databases. Only *Lus4CL8*, *Lus4CL9* and *LusCCoAOMT5* have no corresponding ESTs (E value <e-50) and the expression of *LusHCT1*, *LusCCR3*, *LusCCR8*, *LusCAD8* and *LusCAD11* was also undetectable when using whole genome microarrays [22]. In addition to these 8 genes, no expression data were obtained for *LusCCR9*, *LusCCR11*, *LusCCR12* and *LusCAD9* using EST-based microarrays [11, 23]. The intron/exon structure of the genes is graphically represented in Additional file 1: Fig. S1.

In silico and expression approaches to identify monolignol biosynthetic genes

To identify the genes implicated in lignin precursor biosynthesis, a phylogenetic analysis was first performed by combining the protein sequences derived from the flax phenylpropanoid genes with those from *Arabidopsis thaliana, Vitis vinifera* and *Populus trichocarpa* (Additional file 2: Fig. S2). Different *bona fide* proteins characterized at the biochemical level or by forward/reverse genetic approaches previously used to identify lignin genes in *Eucalyptus grandis* [24] were also added to the data. In addition to this *in silico* sequence comparison, we also performed HT-RT-qPCR on whole stems, roots and leaves as well as on inner stem xylem-rich tissues and outer stem bast fiber-rich tissues (Fig. 2). Gene expression was also determined in leaves and whole stems under different stress conditions (Additional file 3: Fig. S3). Due to the high conservation of several gene sequences and intron/exon positions in some clades, but also the lack of annotation of the more specific 5'/3'-UTRs portions in the genome, we sometimes had to design primers targeting several close-related genes. Taken together, we performed amplifications on 35 single, 12 double, 2 triple and 4 quadruple gene groups (Additional file 4: Table S1). In our experimental conditions, we were unable to detect the expression of *LusPAL4*, *LusC4H4*, *Lus4CL3*, *Lus4CL6*, *LusCCR10*,

LusCCR12, *LusF5H5*, *LusF5H6*, *LusF5H7*, *LusF5H8* and *LusCAD6* genes whatever the organs or environmental conditions. We then used both the phylogenetic analyses (Additional file 2: Fig. S2) and the HT-RT-qPCR (Fig. 2) data to identify the potential lignin-specific genes in flax. The major conclusions are summarized in the following subchapters.

PAL: The deamination step catalysed by the PAL enzymes leads to the formation of cinnamic acid, which may further undergo a series of esterification, hydroxylation and methylation steps. There are no clear reports on a single "lignin-specific" *PAL* gene in the usual plant models such as *Arabidopsis* or *Populus* because of a probable functional redundancy between the members of this gene family. For this reason, it was not easy to highlight *PAL* genes that are essential for lignin biosynthesis in flax when using only a phylogenetic approach. The 4 flax *PAL* genes identified in this work were all located among the *bona fide* genes. Although the bootstrap values were sometimes low among the subdivisions in this part of the tree, *LusPAL1* and *LusPAL2* were closely related to the 3 xylem-specific poplar genes *PtrPAL2*, *PtrPAL4* and *PtrPAL5* [25] while *LusPAL3* and *LusPAL4* were closer to the *Arabidopsis* orthologs *AthPAL1* and *AthPAL2* involved in lignin metabolism [26]. *LusPAL1* and *LusPAL2* are highly expressed in stems and roots compared to leaves. Their respective expression levels in the inner stem (xylem tissues) are close to five- and nine-fold higher than that observed in the hypolignified external stem tissues. Consequently, *LusPAL1* and *LusPAL2* are the most likely candidates for lignin biosynthesis in flax.

C4H: C4H (CYP73) is the first of the three cytochrome P450 monooxygenases involved in the hydroxylation of phenylpropanoid precursors. In flax, five distinct genes located on different scaffolds were annotated by KEGG as trans-cinnamate 4-monooxygenases (K00487). In the neighbour joining tree, 2 distinct clades were present, one contained all of the *bona fide* functionally characterized orthologs, together with an additional *Vitis* sequence and the flax proteins LusC4H1-4. In the other clade, LusC4H5 was located with PtrC4H3 and the eucalyptus EgrC4H1, suggested to be the main C4H involved in lignin biosynthesis [24] even though *EgrC4H2* was also preferentially expressed in the xylem. The HT-RT-qPCR profiles of the flax genes showed that *LusC4H1, LusC4H2, LusC4H3* and *LusC4H5* were expressed in lignified tissues with highest levels for *LusC4H1-2* making them the best candidates for a major role in lignification.

4CL: The 4CL enzymes produce specific CoA thioesters of 4-hydroxycinnamic acids at an

97

important crossroad in the phenylpropanoid pathway. Nine *4CL* gene models were identified in the genome of *Linum usitatissimum*. When the major known sequences were assembled in the phylogenetic tree, a delimited *bona fide* clade containing two separated classes was identified as shown in previous studies [27, 28]. The class I, predicted to be involved in monolignol biosynthesis, contained four flax proteins Lus4CL1-4. The *Lus4CL3* gene remained undetectable in our conditions whereas the other three genes had higher expression in roots and in stems. The relative expression of *Lus4CL4* was 6 fold higher in the xylem as compared to the hypolignified external stem tissues and 15- and 25- fold higher in the stems and roots as compared to the leaves and so is likely involved in lignin biosynthesis.

HCT: In the lignin biosynthetic pathway, HCTs first catalyse the formation of *p*-coumaroyl shikimate from *p*-coumaroyl CoA and shikimic acid. To produce caffeoyl CoA, the product is first hydroxylated by C3'H and then converted again by HCT acting in the reverse direction. The closely related acyltransferase hydroxyl-cinnamoyl CoA:quinate hydroxycinnamoyl transferase (HQT) can lead to the formation of chlorogenic acid by the transfer of quinic acid on the same *p*-coumaroyl CoA substrate, forming p-coumaroylquinic acid. The distinction between HCT and HQT genes without further biochemical characterization is not always very easy because of their very similar sequences. We identified five different genes in the flax genome closely related to the previous identified HCT/HQT genes. The expression profiles show that *LusHCT1* and *LusHCT2* are active in the roots and highly expressed in stem-internal tissues. Their position among the *bona fide* HCT sequences in the phylogenetic tree also suggests their implication in lignin biosynthesis. The *LusHCT4* gene is also among the true HCTs but the expression of this gene is approximately 50 fold times higher in the hypolignified external tissues of the stem.

C3'H: The C3'H genes generally belong to small hydroxylase subfamilies involved in monolignol biosynthesis. The enzyme (CYP98A3) catalyses the 3'-hydroxylation of 4-coumaroyl shikimate and 4-coumaroyl quinate into the corresponding caffeoyl-conjugated form. In flax, we identified only 3 gene models corresponding to C3'H. A gene duplication event possibly occurred because *LusC3'H2* and *LusC3'H3* are both located on the same scaffold in a 14 kb region. When considering the structure (Additional file 1: Fig. S1) and homology between the 3 genes, it is possible to speculate that a first duplication event of *LusC3'H1* formed the *LusC3'H2* gene, which in turn was duplicated forming *LusC3'H3*. This gene was shorter at the 3'-end but still had between 92 and 98% amino acid sequence identity with

the 2 former corresponding proteins. In *Arabidopsis* only one *C3'H* gene was identified whereas an expansion of this family by lineage-specific tandem duplications in *Populus* [29] and *Eucalyptus* [24] was described. The *LusC3'H1* gene is more strongly expressed in the stem and root tissues and therefore probably involved in lignin biosynthesis.

CCoAOMT: Among both methylation enzyme families, CCoAOMTs can transfer a –CH3 group from a donor to a hydroxycinnamoyl CoA ester, namely caffeoyl CoA. Among the 5 genes identified in this work, *LusCCoAOMT4* was previously characterized in knockdown flax plants [30]. The four genes *LusCCoAOMT1-*4 have high sequence identity, similar gene structures and could not be individualized by the HT-RT-qPCR approach. The amplicon was detected at high levels in lignin-containing organs such as stems and roots compared to leaves and preferentially in the internal stem tissues. In the phylogenetic tree, they collocate with the sequences identified as *bona fide* enzymes in *Arabidopsis, Eucalyptus, Nicotiana* and *Populus*.

CCR: This enzyme is important because it catalyses the first specific step of the monolignol production by reducing hydroxycinnamoyl CoA esters. Taking into account the large size of this gene family as determined by genome sequencing, they are likely to be involved in numerous other metabolic pathways. In flax, we identified 12 genes with very different intron/exon patterns. The *LusCCR3*, *LusCCR5* and *LusCCR6* genes do not fit the 3 patterns described for *Arabidopsis* [31]. Although most of the genes encode proteins with similar sizes, *LusCCR3* appears smaller and thus may be truncated. In *Arabidopsis*, 12 *CCR* genes were present in duplicated chromosome regions including 4 genes distributed in tandem [31]. No similar situation exists in flax although *LusCCR2* and *LusCCR5* were located on the same scaffold but at a distance of 360 kb. The differences in the structure of these 2 genes are however not in favour of a recent duplication event. The *bona fide* lignin biosynthesis genes were grouped in a common clade also containing *LusCCR1*, *LusCCR2*, *LusCCR1* and *LusCCR12*. The expression profiles of these four genes in the lignified tissues showed that *LusCCR1* and *LusCCR11* are the best lignin-associated candidates.

COMT: As indicated by the enzyme name, the caffeate/5 hydroxyferulate *O*-methyl-transferase was first thought to act as a bifunctional enzyme able to transfer a methyl group from S-adenosyl methionine to the C₃ and C₅ positions of the aromatic rings of caffeic and 5-hydroxyferulic acids [32]. In fact, the COMT enzyme promotes the biochemical pathway leading to the formation of sinapyl alcohol by catalysing the methylation of 5-hydroxyconiferyl alcohol [33]. Only 3 gene models were identified within

the flax *COMT* gene family, which is low compared to many other species [34]. *LusCOMT1* and *LusCOMT2* have close exon/intron patterns but are present on different scaffolds. They both are expressed at higher rates in lignified tissues. The position of the corresponding proteins within the *bona fide* group in the phylogenetic tree seems to support the hypothesis of their implication in the lignin biosynthetic pathway.

F5H: The F5H cytochrome P450-dependent monooxygenase is required for the hydroxylation of coniferaldehyde and coniferyl alcohol leading to the formation of S lignin through the synthesis of sinapyl alcohol [35]. In flax, 8 gene models were identified and compared to orthologous models expressed in lignin-rich tissues. In the phylogenetic tree, these proteins are located on a branch separated from the *bona fide* enzymes by a node predicted by a low bootstrap value. The flax proteins were further subdivided into 2 groups of sequences with similar exon/intron patterns. The first group contained *F5H1*, *F5H2*, *F5H3* and *F5H4*, which could not clearly be separated by HT-RT-qPCR during the design of the primers except for *F5H2*. The common *F5H1_2_3_4* and specific *F5H2* amplicons were detected at high levels in lignin-containing organs such as roots and whole stems and also in xylem-rich outer tissues. The second group with *F5H5*, *F5H6*, *F5H7* and *F5H8* genes had undetectable expression.

CAD: The NADPH-dependent CAD enzymes play an important role in monolignol biosynthesis by catalyzing the reduction of hydroxycinnamyl aldehydes into their corresponding monolignols which is the last step in the biosynthetic pathway [36]. They are encoded by a multigenic family and their homologs have been detected widely in bacteria and eukaryota but not in animals [37]. Classification by several authors highlighted 3 to 7 groups/classes based on phylogenetic approaches [38-41]. Recently an accurate classification proposed a separation in 3 functional clades for both dicots and monocots [42] with a lignin specific class I and a possible lignin-associated class II. In the present study we identified 15 flax gene models predicted in the published genome. Both *LusCAD12* and *LusCAD13* were located on scaffold120 in a 5.5 kb fragment and have the same exon/intron pattern associated with a high (84%) amino acid sequence identity, reflecting a possible duplication event. *LusCAD1, LusCAD2, LusCAD3* and *LusCAD4* segregate with the known *bona fide* genes within the phylogenetic tree. Because of their very high sequence identity (98,6% at the protein level), it was not possible to separate *LusCAD1* and *LusCAD2* when designing the primers for qRT-PCR. Their expression profiles show that the amplicons were mostly present in the lignified tissues such as roots and whole stems and also in xylem-rich outer tissues. To a

lesser extent, *LusCAD8* showed the same expression profile in these organs but revealed an opposite behaviour when comparing outer and inner stem tissues. This gene is located outside the *bona fide* clade.

Cell-/tissue-specific expression of lignin and selected lignin biosynthetic genes

The likely implication of identified candidate genes in the flax lignification process was further confirmed using *in situ* hybridization that established a close spatial correlation between lignifying cells determined by the Wiesner reaction and gene expression in stems, roots and leaves (Fig. 3 and summarized in Additional file 5: Table S2). These genes were *LusPAL1*, *LusC4H2*, *Lus4CL4*, *LusHCT1*, *LusC3'H1_2_3*, *LusCCoAOMT1_2_3_4*, *LusCCR1_2*, *LusF5H1_2_3_4*, *LusCOMT1_2* and *LusCAD1_2*. For most genes, the strongest expression was observed in the stems and the roots in the first 2-5 secondary xylem cell layers from the cambial zone towards the inner part of the organs. In the leaves, the stain was in each case restricted to the primary xylem. For all genes, signal intensity was generally higher in stems and roots compared to leaves except for *CCR* for which equal intensity was observed in all 3 organs. In addition, the *4CL4* probe failed to reveal the presence of the corresponding transcripts in the stems and roots despite the clear organ specificity revealed by the HT-RT-qPCR results. The observed lack of signal might be due to poor probe efficiency since several fragments of the *4CL4* gene were tested but did not produce a positive signal in roots or stems. Overall, these results confirm that the expression of the selected genes is intimately associated with the actively differentiating secondary xylem zone thereby confirming their probable implication in the lignification process.

Identification of laccase genes and their regulation by microRNAs

Laccases play an important role in controlling lignin polymerization by catalysing the oxidation of the precursors [43]. Both *in silico* prediction and high throughput sequencing in different species [20, 21] have shown that laccase transcript levels are regulated by microRNAs. We identified 45 corresponding gene models in flax and used them to construct a phylogenetic tree (Fig. 4A). In comparison to previous published data reporting the classification of laccases in 6 clades [21], our tree contained an additional clade with 7 flax sequences. The expression profiles (Fig. 4B) of these laccase genes were extracted from public high throughput transcriptomic data [11, 15]. Of the 45 genes, 21 showed higher expression in 'high lignin' inner tissues compared to 'low lignin' outer tissues and are preferentially expressed on the top part of the stem where the development of the SCW takes place.

Based on previous studies reporting predictive and experimental data on the cleavage of LAC4 and LAC17 transcripts in Arabidopsis [20] and potentially 29 laccase transcripts in Populus [21] by miR397, we first searched for a flax orthologous miRNA precursor sequence in the EST and genome databases as previously described [44]. The c3244 EST from the genolin database [11] and scaffold1999 extracted from the phytozome genomic database [45] both contained the pre-miRNA sequence (Fig. 5). The deduced mature *lus-miR397* sequence was first used to search for targets among the flax laccases and among these, 11 genes were predicted with penalty mismatches included between 0 and 2.5 within the duplex. The amplification and sequencing of the 5'-end of the degraded transcripts by RLM-RACE allowed us to experimentally validate the cleavage of LusLAC1, LusLAC5, LusLAC9, LusLAC38, LusLAC39 and LusLAC44 among the predicted targets (Fig. 5). These 6 genes were present among the 21 genes highly expressed in the xylem-rich tissues. The five first laccase transcripts are close to the Arabidopsis ligninrelated AthLAC4 gene [43] and all belong to the clade 1 (Fig. 4A). LusLAC44 belongs to clade 4 including PtrLAC13 which was shown to be cleaved by Ptr-miR397 [46]. These results demonstrate that flax laccase genes are regulated by microRNAs suggesting that this mechanism is implicated in controlling lignification in this species. Further evidence supporting this idea was provided by *in situ* hybridisation data (Fig. 6) indicating a close correlation between tissue-/cell-specific expression of the *lus-miR397* precursor, monolignol specific genes and lignified tissues in stems and roots.

Stress-related flax phenylpropanoid gene expression

Since lignin is an important component in stress responses, we analysed the expression of the phenylpropanoid genes with a special focus on the previous identified lignin-associated genes to search for those regulated by environmental modifications. Flax plants and isolated organs were submitted to a range of different stress conditions and gene expression determined by HT-RT-qPCR. The expression profiles were shown in Additional file 36: Fig. S3 and summarized in Table 2. Salicylic acid (SA) plays an important role in plant defense signaling and accumulates in response to pathogen infection [47]. In flax stems, SA significantly down-regulated the lignin-related genes *LusC4H1* and *LusCCoAOMT1_2_3_4* but had an opposite effect on *LusPAL2* and the *LusF5H1_2_3_4* group. Higher expression levels were also observed for *Lus4CL8* and *LusCAD10*. In parallel, SA had a moderate effect on the expression of lignin genes in treated leaves. Only the up-regulation of the *LusCAD4_9_11* group and *LusCAD8* were evident. Methyljasmonate is a precursor of the active jasmonate form in plants [48] and mimics biotic as well as

abiotic stresses [49]. MeJA application led to a very strong down-regulation of most of the phenylpropanoid genes including all of the identified *bona fide* lignin genes within the leaves except for *LusCCR11*. On the contrary, this signal regulator has a low overall effect on gene expression in flax stems. It is also interesting to note that some genes show no detectable expression levels in either the presence or absence of SA and MeJA, possibly due to the *in vitro* conditions used for organ incubation.

The effect of abiotic stresses was further tested by inducing water deficit on whole flax plants grown in soil. Under these conditions, *LusC4H1* and *LusCCR11* were the only lignin-associated genes showing higher expression in leaves while most of the other *bona fide* genes were slightly upregulated in the stems. The effect of wounding was also tested on flax stems and leaves. In these organs, the effect on the lignin-associated genes was much stronger in the stems compared to the leaves. The modulation of phenylpropanoid gene expression was also evaluated after 48 hours of continuous light illumination. When compared to 16h/8h light/dark conditions, only the *LusPAL1* and *Lus4CL4* genes were downregulated and 7 genes or gene-groups were significantly upregulated. When taken together, our results showed that the strongest impact on the previously genes identified as involved in lignin biosynthesis, occurred in the stem during dehydration, wounding and in the presence of SA.

Gene clustering

Genes involved in the same biological process are often co-regulated. To further confirm the implication of the former identified genes in the lignification pathway, we searched for possible common expression profiles among the gene set. A k-means clustering approach was used to identify correlated groups exhibiting similar expression profiles in the different tissues and/or stress conditions. The number of flax genes in each cluster ranged between 2 and 14 (Fig. 7A). Among the flax phenylpropanoid associated genes, all the potential *bona fide* genes identified previously, except for *LusCAD1_2* and *LusCCR11*, were included in cluster 9 (Fig. 7B) showing that the expression of these genes were closely co-regulated in flax. Altogether, the clustering results provide further support for the involvement of the identified genes in lignification.

Gene co-regulation requires the activity of common transcription factors and since MYB proteins are known to function as master switches of secondary cell wall and lignin gene transcription [50, 51] we examined the promoters of our identified genes for MYB consensus motifs named MBS (MYB binding site) (C/T)AAC(A/T)A(A/C)C and MBSIIG (or SMRE: secondary wall MYB-responsive element)

(C/T)ACC(A/T)A(A/C)C. Our results (Additional file 6: Table S3) showed that the MBSIIG site was present in the 500bp proximal promoter fragments of all the *bona fide* genes present in the cluster 9 whereas both MBS sites are absent from the *LusCCR11* promoter in cluster 1. Interestingly, both MYB sites were absent from the proximal promoters of the *LusF5H1_2_3_4* genes which is in accordance with previous results showing that *F5H* is directly regulated by a NAC SND1 transcription factor [52].

Discussion

Identification of genes involved in the lignin biosynthetic pathway in flax

In the past 25 years, a tremendous amount of information about lignin biosynthesis has been obtained using model plants such as *Arabidopsis* and *Populus* species. Currently, the possibility to obtain comprehensive genomic sequences in other species is allowing the scientific community to gather additional data on other species. Secondary cell walls (SCWs), whether they are specialized in water conduction (xylem) or mechanical support (sclerenchyma), usually contain around 25% lignin. However, in some specialized tissues such as bast fibers, the cell wall contains very low amounts of lignin together with high cellulose levels [53, 54]. Although these cells only contain low amounts of lignins, the presence of these phenolic molecules can have an important impact on the quality of the final products such as linen tissues and biocomposites.

In this context, we carried out a comprehensive genome-wide analysis of the gene set involved in phenylpropanoids, and more specifically in lignin biosynthesis in flax. The size of individual phenylpropanoid gene families is difficult to compare between species due to differences in the stringency of the sequence annotations. This is especially true for widespread enzymatic activities such as hydroxylases or methyltransferases. On the other hand, gene expression profiling often gives clear information about the function of the more specific *bona fide* lignin genes [24]. Usually, individual qRT-PCR experiments are performed to validate lignin transcript level pattern analysis [55] but HT-RT-qPCR approaches are increasingly used as they are more appropriately sized for characterizing the members of multigene families in a single step [56]. In flax, we first identified the genes potentially involved in lignification by examining the expression profiles in roots and stems compared to the leaves and also by comparison between the internal xylem stem tissues (high lignin) and the outer fiber-bearing tissues (low lignin).

104

Further confirmation of the likely implication of selected phenylpropanoid genes in lignification was provided by *in situ* hybridization, allowing the precise localization of the lignin biosynthetic gene expression. Both *in situ* hybridization and also promoter-reporter gene studies conducted on a low number of genes previously showed that *bona fide* lignin genes were expressed in the first few cell layers of differentiating secondary xylem [17-19, 57, 58]. To our knowledge, our work reports for the first time an *in situ* hybridization approach on a large number of lignin genes in a higher plant species. As was shown by the results of the high throughput expression analyses, the very high expression levels in the xylem-rich tissues compared to the bast fiber cells is in accordance with our previous results suggesting that lignification of flax fibers is at least in part regulated at the transcriptional level [15].

Finally, when the expression patterns of all the phenylpropanoid genes were considered, most of these lignin-associated genes were found to be co-expressed within a same cluster confirming that these genes were indeed involved in a common metabolic pathway. The presence of common MYB *cis*-element motifs in the promoters of these genes associated to their co-expression was already described for the orthologous genes in *Arabidopsis* [59]. In this species, MYB58 can directly activate most lignin biosynthesis genes but not *F5H* showing that the syringyl lignin biosynthesis is activated by a different regulatory pathway probably controlled by NAC transcription factors [52].

Selection of candidates for lignin engineering in flax

The success of targeted engineering of cell wall lignin genes to improve the quality of different flax-derived products depends upon the correct identification of individual family gene members involved in the trait to be modified. Some genes potentially involved in lignin biosynthesis have already been partially characterized in flax. The first flax lignin cDNA sequence was deposited in the GenBank database in 2005 and the corresponding *CCoAOMT* gene functionally characterized [30]. It was chosen because it is responsible for the synthesis of feruloyl CoA used for the formation of both G- and S-lignin units and, as confirmed by down-regulation, because it plays a central role in maintaining SCW integrity by regulating the quantity and the S/G proportion in the non-condensed lignin fraction. This gene corresponds to the *LusCCoAOMT4* identified in our study and is included in the *bona fide* group of genes expressed in lignified tissues. In another study, a *CAD* gene fragment sharing 100% homology with *LusCAD4* was targeted by a downregulation strategy [60]. This gene associated with *LusCAD3* was not retained in our work as a major

lignin gene because of its much lower differential expression between lignified and poorly-lignified organs or tissues when compared to *LusCAD1_2*. Although the RNAi transformed plants contained lower amounts of lignins compared to the control, it is interesting to note that M Wrobel-Kwiatkowska, M Starzycki, J Zebrowski, J Oszmianski and J Szopa [60]) did not observe the typical brown-midrib phenotype of *CAD* mutants observed in other species. In contrast, TILLed flax chemical mutants of *LusCAD1* possess the characteristic orange-coloured xylem phenotype [16] confirming the likely involvement of this gene in the lignification process. Flax lignin has also received interest because of the very low amounts of S-units [53] so we were therefore interested to examine the *F5H* gene family. This enzyme is specifically required for the synthesis of S lignin since it catalyses the 5-hydroxylation of coniferaldehyde and/or coniferyl alcohol [35]. Its down-regulation in *Arabidopsis* leads to a reduced S/G value [61]. The expression ratio of *LusF5H2* in the lignified organs and tissues is comparable to those of other identified genes so the biosynthesis of low amounts of S units in flax does not seem to be controlled by *F5H* gene expression. However, further computational modelling and biochemical characterization of the members of the F5H family have to be done in order to provide more clues on the specificity of these enzymes towards various substrates.

lus-miR397 controls several laccase genes in flax

Laccases are mono- or multimeric copper-oxidases in eukaryotes and procaryotes with a broad range of aromatic or non-aromatic substrates [62] and have been implicated in the oxidative polymerization of monolignols. Among the 17 gene models described in *Arabidopsis, AthLAC4, AthLAC11* and *AthLAC17* were clearly shown to contribute to the constitutive lignification of the floral stems [43, 63]. In *Populus trichocarpa*, 49 gene models have been identified but it was not yet possible to clearly identify those involved in lignification [21]. In flax, we identified 45 laccase gene models of which 23 were highly expressed in lignified tissues. MicroRNA targeting (*miR397*) of some laccase transcripts, including *AthLAC4* and *AthLAC17*, contribute to the overall gene regulation [20, 21]. This microRNA has been identified in many gymnosperms, monocots and dicots with 1 to 3 copies per species [64]. In this study, we experimentally demonstrated that transcripts corresponding to six flax laccase genes were targeted by *lus-miR397*. On the basis of our phylogenetic analysis, five of these genes were found in the clade 1, which also contained *AthLAC4* whereas six non-targeted genes were present in clade 2 along with *AthLAC17* showing that either no functional ortholog of this gene is present in flax or that these genes have lost their

ability to be cleaved by *lus-miR397*. Strong evidence for a likely role of microRNA regulation of laccasemediated lignification in flax was provided by the cell-/tissue-specific localization of the corresponding transcript precursors. Our results are in agreement with those obtained with a *GUS-GFP* fusion driven by the *miR397b* promoter, showing that this gene is expressed in the stem vascular tissues of *Arabidopsis* [65]. Further analyses should allow a better understanding of whether *miR397* plays a role in maintaining constant transcript levels or is implicated in a finer control of different laccase paralogs.

Stress related phenylpropanoid genes

Flax plants selected for fiber production are mostly cultivated around the Mediterranean and temperate climate zones. In the context of climate change, cell wall metabolism, structure and fiber quality are likely to be affected in the future and it is therefore informative to know which genes are affected by biotic or abiotic stresses. We took advantage of the high throughput gene expression approach to try and distinguish between developmental lignin genes and stress-related phenylpropanoid gene expression.

Defence-related hormones such as SA and MeJA can control lignin gene activation. SA is an important signalling molecule for systemic acquired resistance [66] and triggers the expression of lignin genes in many different species [67]. In flax, several lignin genes showed opposite expression profiles in the presence of SA suggesting that the composition of lignin may change during the defence mechanism or that more specific metabolic pathways including these genes are activated. The effect of MeJA on this metabolic pathway seems more subtle. This component has been shown to participate in the signal transduction pathway in response to different forms of stresses [68]. A dose effect was described in Fragaria fruits [69] as well as an opposite CAD gene expression pattern in tea plants [70]. In our experimental conditions, most of the flax genes were underexpressed indicating that the synthesis of lignin is probably not part of the response to this elicitor. Abiotic stresses can also have an influence on lignification. Mechanical stresses can be responsible for the deposition of various polymers including lignin that may function as a physical barrier around the injury zone [71]. In this context, lignin genes are activated in different species including sweet potato [72], maize [73] or Arabidopsis [74]. Most of the flax lignin genes were activated in the stem but less in the leaves showing that in this species, lignin is also synthesized in reaction to wounding. Outsides these bona fide genes, several common phenylpropanoid genes including CADs were also activated in both organs showing that they may potentially be involved in the production of other abiotic stress responsive metabolites as was already proposed in poplar [38]. The

effect of drought stress on plants is more difficult to understand because it can have positive or negative roles on lignification depending on the species, the organs/tissues and the intensity and duration of the stress period [75]. In our experimental conditions (12 days after watering stop) most of the lignin genes were upregulated in the stems and the leaves. The increase in lignification has been described as one of the reactions included in a general adaptation strategy of plants faced with water loss and may result in an increase of mechanical strength and/or water impermeability [76].

Conclusion

We have identified, in this study, the individual members of the multigenic families most likely involved in lignin biosynthesis (and polymerisation) in flax. These data provide the knowledge base necessary to undertake targeted engineering and/or marker-based selection of lignin biosynthesis in flax, as well as important information on the behaviour of these genes in a number of different stress conditions. The latter information will be particularly useful for farmers and breeders in the current context of increasing variability in weather conditions associated with climate change.

Methods

Plant material

For HT-RT-qPCR expression analysis, flax plants (*Linum usitatissimum* L. cv Diane) were grown in 12x12 cm soil-containing pots in a greenhouse under 16h/20°C day and 8h/18°C night conditions for 35 days. For developmental expression studies, (i) leaves, ii) 2-cm root fragments taken under the stem base, (iii) top 4 cm of stem, (iv) fiber-rich outer-tissues and (v) xylem-rich inner-tissues from stem base were harvested. For stress expression studies, the following samples were taken from these 35-day-old plants: (i) leaves and stems harvested 24h after wounding with a scalpel, (ii) leaves and stems incubated for 24h in Petri dishes containing 200 μ M salicylic acid (SA) or methyljasmonate (MeJA) dissolved in ½ MS medium [77], (iii) leaves and stems from plants submitted to a drought stress imposed by withholding water for 12 days, (iv) stems from plants first placed in phytotrons during 7 days for acclimation and then submitted to continuous light for 48 hours at 20°C. All the samples were triplicates and always frozen immediately in liquid nitrogen before storage at -80°C prior to RNA extraction. For *in situ* hybridization, the median stem, leaves and 2-cm root fragments were collected and fixed as described below.
In silico identification and analysis of phenylpropanoid genes in L. usitatissimum

The flax phenylpropanoid protein and gene sequences were identified in the annotated v1.0 of the *Linum usitatissimum* genome hosted on the Phytozome website [45]. The entire database derived from the published genome sequences [10] was queried with the names of the gene families used as keywords and also by BLASTp interrogation with previously lignin-related identified sequences from *Arabidopsis thaliana* [78] and *Populus trichocarpa* [25]. When it was necessary, unknown portions of some sequences were replaced by the corresponding sequences found in public EST databases and truncated sequences were discarded. For each gene model, the presence of a corresponding EST was searched in the dbEST [79] and genolin [11] databases. The intron/exon structures were also retrieved and the identities of the protein sequences compared. The k-means medians clustering function from the TIGR MultiExperiment Viewer (TM4 Mev v4.8.1) was used to cluster the genes according to their expression profiles. The distance metric was Pearson correlation and the default parameters were used for the k-means calculation.

Phylogenetic analyses

For each gene family, the protein sequences were aligned with ClustalW and the phylogenetic analyses obtained with the MEGA 6.06 [80] software for performed by maximum likelihood method based on the JTT matrix-based model [81]. Consensus trees were generated with 1000 bootstrap replicates.

HT-RT-qPCR using a 96.96 Dynamic array

Total RNA was extracted from the different organs and tissues using the TriReagent method (Molecular Research Center). RNA integrity and concentration were evaluated with RNA Standard Sens Chips in the Experion automated capillary electrophoresis system (Bio-Rad). cDNA was then synthesized using the High Capacity RNA-to-cDNA Kit (Applied Biosystems) according to the manufacturer's instructions. The large scale quantitative RT-PCR was performed with a BioMark HD System using a Fluidigm 96.96 dynamic array (IntegraGen, Evry France) according to the Fluidigm Advanced Development Protocol with EvaGreen using primer pairs listed in Additional file 4: Table S1. A first preamplification reaction (1 cycle: 95°C 10 min; 14 cycles: 95°C 15 seconds, 60°C 4 min) was performed for each sample in 10 µl by pooling primer pairs (final concentration, 50 nM), 3.3 µl cDNA, and 5 µl 2X PreAmp Master Mix (Applied Biosystems). For each assay, 5 µl 10X Assay Mix containing 9 µM forward

primer, 9 µM reverse primer, and 1X Assay Loading Reagent was loaded into the chip. The following thermal cycles were executed: 1 cycle: 95°C 10 min; 40 cycles: 95°C 15 seconds, 60°C 1 min. The amplifications were always conducted in triplicate on individual samples including controls and on a mix of cDNAs to determine the amplification efficiencies. The Ct values were analysed with the qBase+ software (Biogazelle, Belgium). The data normalization was performed using the previous experimentally validated reference genes *LusETIF5A1*, *LusUBI1* and *LusEF1A* [82].

Histochemical analysis

Flax stems, roots and leaves were fixed with 4% paraformaldehyde in phosphate buffer (0.2 M, pH 7) for 16 h. They were then washed in phosphate buffer containing 0.1 M glycine and dehydrated through a series of ethanol solutions and progressively embedded in paraffin (ParaplastPlus; Oxford Labware). Tissue sections were obtained with a Leica RM2065 microtome and placed on silanised-coated slides. Before staining, tissue sections were deparaffinised with Histoclear (Labonord) and rehydrated with decreasing concentrations of ethanol. The presence of lignin was determined by staining with the Wiesner reagent (phloroglucinol–HCl) [83].

Spatial gene expression determination by in situ hybridization

To obtain DNA templates for RNA probe synthesis, PCR amplifications were performed with genespecific primers tailed with a T7 RNA polymerase binding site Additional file 4: Table S1. The amplicons (1 µg) were used as templates to synthesize sense and antisense probes for each gene, with the incorporation of UTP-digoxigenin as the label using the DIG labelling mix (Roche). Deparaffinised sections were treated according to P Roongsattham, F Morcillo, C Jantasuriyarat, M Pizot, S Moussu, D Jayaweera, M Collin, ZH Gonzalez-Carranza, P Amblard, JW Tregear, et al. [84], hydridized and probe detected with NBT (nitro-blue tetrazolium chloride) / BCIP (5-bromo-4-3'-indolylphosphate *p*-toluidine salts) solutions according to the manufacturer's instructions. Controls without probe or with sense probe were performed to check for endogenous alkaline phosphate activity.

miRNA target prediction and experimental validation

The sequence of the *lus-miR397* precursor was identified among the known flax ESTs using stringent rules described elsewhere [44]. The laccase transcript models were analysed with psRNATarget [85] to predict the targets that can potentially match with the microRNA without any gaps. When referring

to the miRNA sequence, there were not more than one mismatch between the nucleotides number 1 and 9, none between 10 and 11, and no more than 2 consecutive mismatches after position 11. The experimental validation of the laccase targets was carried out by using a modified protocol of the RNA ligation mediated (RLM) RACE included in the GeneRacer kit (Invitrogen). RNA was extracted from a mix of flax organs and the cleaved transcripts (without 5'-cap) were selectively ligated to a 5'-RNA adaptor. After reverse transcription, PCR amplification was performed with a 5' adaptor-specific primer and a reverse gene specific primer Additional file 4: Table S1 located downstream of the predicted cleavage site. The fragments were then gel purified and sequenced.

Abbreviations

4CL: 4-coumarate:CoA ligase **C3'H:** *p*-coumarate 3-hydroxylase **C4H:** cinammate 4-hydroxylase **CAD:** cinnamyl alcohol dehydrogenase **CCoAOMT:** caffeoyl CoA 3-*O*-methyltransferase **CCR:** cinnamoyl CoA reductase **COMT:** caffeate/5-hydroxyferulate *O*-methyl-transferase **F5H:** ferulate 5-hydroxylase **HCT:** hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl transferase **MeJA:** methyljasmonate **PAL:** Phenylalanine ammonia-lyase **SA:** salicylic acid

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and material

The flax DNA sequences mentioned in this article can all be retrieved from the Phytozome database: <u>https://phytozome.jgi.doe.gov/pz/portal.html</u>

Competing interests

The authors declare that they have no competing interests

Funding

This research was supported by the french Ministère de l'Enseignement Supérieur et de la Recherche

Authors' contributions

JLR and ASB performed the sample preparations and experimental procedures, GN participated in the experiments, AC performed data analysis, BH and SH critically revised the manuscript and GN designed the study, supervised the work and wrote the manuscript.

Acknowledgements

We are indebted to the Research Federation FRABio (Univ. Lille, CNRS, FR 3688, FRABio, Biochimie Structurale et Fonctionnelle des Assemblages Biomoléculaires) for providing the scientific and technical environment conducive to achieving this work. JLR gratefully acknowledges the french Ministère de l'Enseignement Supérieur et de la Recherche for financial support.

References

1. Le Roy J, Huss B, Creach A, Hawkins S, Neutelings G. Glycosylation is a major regulator of phenylpropanoid availability and biological activity in plants. Frontiers in Plant Science 2016; 7.

2. Boerjan W, Ralph J, Baucher M. Lignin biosynthesis. Annual review of plant biology 2003; 54:519-546.

3. Freudenberg K: Lignin biosynthesis. Berlin, Heidelberg, New York: Springer; 1968.

4. Sterjiades R, Dean JF, Eriksson KE. Laccase from Sycamore Maple (Acer pseudoplatanus) Polymerizes Monolignols. Plant Physiology 1992; 99(3):1162-1168.

 Moura JCMS, Bonine CAV, De Oliveira Fernandes Viana J, Dornelas MC, Mazzafera P. Abiotic and Biotic Stresses and Changes in the Lignin Content and Composition in Plants. Journal of Integrative Plant Biology 2010; 52(4):360-376.

6. Neutelings G. Lignin variability in plant cell walls: contribution of new models. Plant Science 2011; 181(4):379-386.

7. Mohanty AK, Misra M, Hinrichsen G. Biofibres, biodegradable polymers and biocomposites: An overview. Macromolecular Materials and Engineering 2000; 276-277(1):1-24.

8. Zohary D, Hopf M: Domestication of Plants in the Old World, Third edition edn. New York: Oxford University Press; 2004.

9. Summerscales J, Dissanayake NPJ, Virk AS, Hall W. A review of bast fibres and their composites. Part 1
 – Fibres as reinforcements. Composites Part A: Applied Science and Manufacturing 2010; 41(10):1329-1335.

112

Wang Z, Hobson N, Galindo L, Zhu S, Shi D, McDill J, Yang L, Hawkins S, Neutelings G, Datla R et al.
 The genome of flax (Linum usitatissimum) assembled de novo from short shotgun sequence reads. Plant Journal 2012;
 72(3):461-473.

Fenart S, Ndong Y-P, Duarte J, Riviere N, Wilmer J, van Wuytswinkel O, Lucau A, Cariou E, Neutelings
 G, Gutierrez L. Development and validation of a flax (Linum usitatissimum L.) gene expression oligo microarray. BMC
 Genomics 2010; 11(1):592.

12. Chantreau M, Chabbert B, Billiard S, Hawkins S, Neutelings G. Functional analyses of cellulose synthase genes in flax (Linum usitatissimum) by virus-induced gene silencing. Plant Biotechnology Journal 2015; 13(9):1312-1324.

13. Ragupathy R, Rathinavelu R, Cloutier S. Physical mapping and BAC-end sequence analysis provide initial insights into the flax (Linum usitatissimum L.) genome. BMC Genomics 2011; 12:217.

14. Day A, Fenart S, Neutelings G, Hawkins S, Rolando C, Tokarski C. Identification of cell wall proteins in the flax (Linum usitatissimum) stem. Proteomics 2013; 13(5):812-825.

Huis R, Morreel K, Fliniaux O, Lucau-Danila A, Fenart S, Grec S, Neutelings G, Chabbert B, Mesnard F,
 Boerjan WA. Natural hypolignification is associated with extensive oligolignol accumulation in flax stems. Plant Physiology 2012; 158(4):1893-1915.

16. Chantreau M, Grec S, Gutierrez L, Dalmais M, Pineau C, Demailly H, Paysant-Leroux C, Tavernier R, Trouve JP, Chatterjee M et al. PT-Flax (phenotyping and TILLinG of flax): development of a flax (Linum usitatissimum L.) mutant population and TILLinG platform for forward and reverse genetics. BMC Plant Biology 2013; 13:159.

17. Harding SA, Leshkevich J, Chiang VL, Tsai C-J. Differential Substrate Inhibition Couples Kinetically Distinct 4-Coumarate:Coenzyme A Ligases with Spatially Distinct Metabolic Roles in Quaking Aspen. Plant Physiology 2002; 128(2):428-438.

18. Hawkins S, Boudet A, Grima-Pettenati J. Characterisation of caffeic acid O-methyltransferase and cinnamyl alcohol dehydrogenase gene expression patterns by in situ hybridisation in Eucalyptus gunnii Hook. plantlets. Plant Science 2003; 164(2):165-173.

19. Lacombe E, Hawkins S, Van Doorsselaere J, Piquemal J, Goffner D, Poeydomenge O, Boudet AM, Grima-Pettenati J. Cinnamoyl CoA reductase, the first committed enzyme of the lignin branch biosynthetic pathway: cloning, expression and phylogenetic relationships. Plant Journal 1997; 11(3):429-441.

20. Abdel-Ghany SE, Pilon M. MicroRNA-mediated systemic down-regulation of copper protein expression in response to low copper availability in Arabidopsis. Journal of Biological Chemistry 2008; 283(23):15932-15945.

21. Lu S, Li Q, Wei H, Chang M-J, Tunlaya-Anukit S, Kim H, Liu J, Song J, Sun Y-H, Yuan L et al. PtrmiR397a is a negative regulator of laccase genes affecting lignin content in Populus trichocarpa. Proceedings of the Natlional Academy of Science USA 2013; 110(26):10848-10853. 22. Chantreau M, Portelette A, Dauwe R, Kiyoto S, Cronier D, Morreel K, Arribat S, Neutelings G, Chabi M, Boerjan W et al. Ectopic Lignification in the Flax lignified bast fiber1 Mutant Stem Is Associated with Tissue-Specific Modifications in Gene Expression and Cell Wall Composition. Plant Cell 2014; 26(11):4462-4482.

23. Roach MJ, Deyholos MK. Microarray analysis of flax (Linum usitatissimum L.) stems identifies transcripts enriched in fibre-bearing phloem tissues. Molecular Genetics and Genomics 2007; 278(2):149-165.

24. Carocha V, Soler M, Hefer C, Cassan-Wang H, Fevereiro P, Myburg AA, Paiva JA, Grima-Pettenati J. Genome-wide analysis of the lignin toolbox of Eucalyptus grandis. New Phytologist 2015.

25. Shi R, Sun YH, Li Q, Heber S, Sederoff R, Chiang VL. Towards a systems approach for lignin biosynthesis in Populus trichocarpa: transcript abundance and specificity of the monolignol biosynthetic genes. Plant and Cell Physiology 2010; 51(1):144-163.

26. Rohde A, Morreel K, Ralph J, Goeminne G, Hostyn V, De Rycke R, Kushnir S, Van Doorsselaere J, Joseleau JP, Vuylsteke M et al. Molecular phenotyping of the pal1 and pal2 mutants of Arabidopsis thaliana reveals farreaching consequences on phenylpropanoid, amino acid, and carbohydrate metabolism. Plant Cell 2004; 16(10):2749-2771.

27. De Azevedo Souza C, Barbazuk B, Ralph SG, Bohlmann J, Hamberger B, Douglas CJ. Genome-wide analysis of a land plant-specific acyl:coenzymeA synthetase (ACS) gene family in Arabidopsis, poplar, rice and Physcomitrella. New Phytologist 2008; 179(4):987-1003.

28. Ehlting J, Buttner D, Wang Q, Douglas CJ, Somssich IE, Kombrink E. Three 4-coumarate:coenzyme A ligases in Arabidopsis thaliana represent two evolutionarily divergent classes in angiosperms. Plant J 1999; 19(1):9-20.

29. Hamberger B, Ellis M, Friedmann M, de Azevedo Souza C, Barbazuk B, Douglas CJ. Genome-wide analyses of phenylpropanoid-related genes in Populus trichocarpa, Arabidopsis thaliana, and Oryza sativa: the Populus lignin toolbox and conservation and diversification of angiosperm gene families. Canadian Journal of Botany 2007; 85(12):1182-1201.

30. Day A, Neutelings G, Nolin F, Grec S, Habrant A, Cronier D, Maher B, Rolando C, David H, Chabbert B et al. Caffeoyl coenzyme A O-methyltransferase down-regulation is associated with modifications in lignin and cell-wall architecture in flax secondary xylem. Plant Physiology and Biochemistry 2009; 47(1):9-19.

31. Barakat A, Yassin NB, Park JS, Choi A, Herr J, Carlson JE. Comparative and phylogenomic analyses of cinnamoyl-CoA reductase and cinnamoyl-CoA-reductase-like gene family in land plants. Plant Science 2011; 181(3):249-257.

32. Bugos RC, Chiang VL, Campbell WH. cDNA cloning, sequence analysis and seasonal expression of lignin-bispecific caffeic acid/5-hydroxyferulic acid O-methyltransferase of aspen. Plant molecular biology 1991; 17(6):1203-1215.

33. Osakabe K, Tsao CC, Li L, Popko JL, Umezawa T, Carraway DT, Smeltzer RH, Joshi CP, Chiang VL. Coniferyl aldehyde 5-hydroxylation and methylation direct syringyl lignin biosynthesis in angiosperms. Proceedings of the Natlional Academy of Science USA 1999; 96(16):8955-8960.

114

34. Xu Z, Zhang D, Hu J, Zhou X, Ye X, Reichel KL, Stewart NR, Syrenne RD, Yang X, Gao P et al. Comparative genome analysis of lignin biosynthesis gene families across the plant kingdom. BMC Bioinformatics 2009; 10 Suppl 11:S3.

35. Humphreys JM, Hemm MR, Chapple C. New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase. Proceedings of the Natlional Academy of Science USA 1999; 96(18):10045-10050.

36. Mansell RE, Stöckigt J, Zenk H. Reduction of Ferulic Acid to Coniferyl Alcohol in a Cell free System from a higher Plant. Zeitschrift für Pflanzenphysiologie 1972; 68:286-288.

37. Guo DM, Ran JH, Wang XQ. Evolution of the Cinnamyl/Sinapyl Alcohol Dehydrogenase (CAD/SAD) gene family: the emergence of real lignin is associated with the origin of Bona Fide CAD. Journal of Molecular Evolution 2010; 71(3):202-218.

38. Bagniewska-Zadworna A, Barakat A, Lakomy P, Smolinski DJ, Zadworny M. Lignin and lignans in plant defence: insight from expression profiling of cinnamyl alcohol dehydrogenase genes during development and following fungal infection in Populus. Plant Science 2014; 229:111-121.

39. Chao N, Liu SX, Liu BM, Li N, Jiang XN, Gai Y. Molecular cloning and functional analysis of nine cinnamyl alcohol dehydrogenase family members in Populus tomentosa. Planta 2014; 240(5):1097-1112.

40. Jin Y, Zhang C, Liu W, Qi H, Chen H, Cao S. The cinnamyl alcohol dehydrogenase gene family in melon (Cucumis melo L.): bioinformatic analysis and expression patterns. PLoS One 2014; 9(7):e101730.

41. Saballos A, Ejeta G, Sanchez E, Kang C, Vermerris W. A genomewide analysis of the cinnamyl alcohol dehydrogenase family in sorghum [Sorghum bicolor (L.) Moench] identifies SbCAD2 as the brown midrib6 gene. Genetics 2009; 181(2):783-795.

42. van Parijs FR, Ruttink T, Boerjan W, Haesaert G, Byrne SL, Asp T, Roldan-Ruiz I, Muylle H. Clade classification of monolignol biosynthesis gene family members reveals target genes to decrease lignin in Lolium perenne. Plant Biology (Stuttg) 2015; 17(4):877-892.

43. Berthet S, Demont-Caulet N, Pollet B, Bidzinski P, Cezard L, Le Bris P, Borrega N, Herve J, Blondet E, Balzergue S et al. Disruption of LACCASE4 and 17 results in tissue-specific alterations to lignification of Arabidopsis thaliana stems. Plant Cell 2011; 23(3):1124-1137.

44. Neutelings G, Fenart S, Lucau-Danila A, Hawkins S. Identification and characterization of miRNAs and their potential targets in flax. Journal of Plant Physiology 2012; 169(17):1754-1766.

45. Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N et al. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res 2012; 40(Database issue):D1178-1186.

46. Lu S, Yang C, Chiang VL. Conservation and diversity of microRNA-associated copper-regulatory networks in Populus trichocarpa. J Integr Plant Biol 2011; 53(11):879-891.

47. Lu H. Dissection of salicylic acid-mediated defense signaling networks. Plant Signaling & Behavior 2009; 4(8):713-717.

Denness L, McKenna JF, Segonzac C, Wormit A, Madhou P, Bennett M, Mansfield J, Zipfel C, Hamann
 T. Cell Wall Damage-Induced Lignin Biosynthesis Is Regulated by a Reactive Oxygen Species- and Jasmonic Acid Dependent Process in Arabidopsis. Plant Physiology 2011; 156(3):1364-1374.

49. Kim EH, Kim YS, Park S-H, Koo YJ, Choi YD, Chung Y-Y, Lee I-J, Kim J-K. Methyl Jasmonate Reduces Grain Yield by Mediating Stress Signals to Alter Spikelet Development in Rice. Plant Physiology 2009; 149(4):1751-1760.

50. Rahantamalala A, Rech P, Martinez Y, Chaubet-Gigot N, Grima-Pettenati J, Pacquit V. Coordinated transcriptional regulation of two key genes in the lignin branch pathway--CAD and CCR--is mediated through MYB- binding sites. BMC Plant Biology 2010; 10:130.

51. Zhong R, McCarthy RL, Haghighat M, Ye Z-H. The Poplar MYB Master Switches Bind to the SMRE Site and Activate the Secondary Wall Biosynthetic Program during Wood Formation. PLoS One 2013; 8(7):e69219.

52. Zhao Q, Wang H, Yin Y, Xu Y, Chen F, Dixon RA. Syringyl lignin biosynthesis is directly regulated by a secondary cell wall master switch. Proceedings of the Natlional Academy of Science USA 2010; 107(32):14496-14501.

53. Day A, Ruel K, Neutelings G, Crônier D, David H, Hawkins S, Chabbert B. Lignification in the flax stem: evidence for an unusual lignin in bast fibers. Planta 2005; 222(2):234-245.

54. Lapierre C, Pollet B, Petit-Conil M, Toval G, Romero J, Pilate G, Leplé J-C, Boerjan W, Ferret V, De Nadai V et al. Structural Alterations of Lignins in Transgenic Poplars with Depressed Cinnamyl Alcohol Dehydrogenase or Caffeic Acid O-Methyltransferase Activity Have an Opposite Impact on the Efficiency of Industrial Kraft Pulping. Plant Physiology 1999; 119(1):153-164.

55. Shen H, Mazarei M, Hisano H, Escamilla-Trevino L, Fu C, Pu Y, Rudis MR, Tang Y, Xiao X, Jackson L et al. A genomics approach to deciphering lignin biosynthesis in switchgrass. Plant Cell 2013; 25(11):4342-4361.

56. Pinzon-Latorre D, Deyholos MK. Characterization and transcript profiling of the pectin methylesterase (PME) and pectin methylesterase inhibitor (PMEI) gene families in flax (Linum usitatissimum). BMC Genomics 2013; 14:742.

57. Feuillet C, Lauvergeat V, Deswarte C, Pilate G, Boudet A, Grima-Pettenati J. Tissue- and cell-specific expression of a cinnamyl alcohol dehydrogenase promoter in transgenic poplar plants. Plant Mol Biol 1995; 27(4):651-667.

58. Hawkins S, Samaj J, Lauvergeat V, Boudet A, Grima-Pettenati J. Cinnamyl Alcohol Dehydrogenase: Identification of New Sites of Promoter Activity in Transgenic Poplar. Plant Physiol 1997; 113(2):321-325.

59. Ma S, Shah S, Bohnert HJ, Snyder M, Dinesh-Kumar SP. Incorporating Motif Analysis into Gene Coexpression Networks Reveals Novel Modular Expression Pattern and New Signaling Pathways. PLoS Genetics 2013; 9(10):e1003840. 60. Wrobel-Kwiatkowska M, Starzycki M, Zebrowski J, Oszmianski J, Szopa J. Lignin deficiency in transgenic flax resulted in plants with improved mechanical properties. Journal of Biotechnology 2007; 128(4):919-934.

61. Ohman D, Demedts B, Kumar M, Gerber L, Gorzsas A, Goeminne G, Hedenstrom M, Ellis B, Boerjan W, Sundberg B. MYB103 is required for FERULATE-5-HYDROXYLASE expression and syringyl lignin biosynthesis in Arabidopsis stems. Plant Journal 2013; 73(1):63-76.

62. Berthet S, Thevenin J, Baratiny D, Demont-Caulet N, Debeaujon I, Przemysław B, Leplé J, Huis R, Hawkins S, Gomez LD et al. Role of Plant Laccases in Lignin Polymerization. Chapter 5. Advances in Botanical Research 2012; 61:145-172.

63. Zhao Q, Nakashima J, Chen F, Yin Y, Fu C, Yun J, Shao H, Wang X, Wang ZY, Dixon RA. Laccase is necessary and nonredundant with peroxidase for lignin polymerization during vascular development in Arabidopsis. Plant Cell 2013; 25(10):3976-3987.

64. Kozomara A, Griffiths-Jones S. miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Research 2013; 42(D1):D68-D73.

65. Wang CY, Zhang S, Yu Y, Luo YC, Liu Q, Ju C, Zhang YC, Qu LH, Lucas WJ, Wang X et al. MiR397b regulates both lignin content and seed number in Arabidopsis via modulating a laccase involved in lignin biosynthesis. Plant Biotechnology Journal 2014; 12(8):1132-1142.

66. Yalpani N, Raskin I. Salicylic acid: a systemic signal in induced plant disease resistance. Trends Microbiology 1993; 1(3):88-92.

67. Miedes E, Boerjan W, Vanholme R, Molina A. The role of the secondary cell walls in plant resistance to pathogens. Frontiers in Plant Science 2014; 5.

68. Sharma M, Laxmi A. Jasmonates: Emerging Players in Controlling Temperature Stress Tolerance. Frontiers in Plant Science 2015; 6:1129.

69. Concha CM, Figueroa NE, Poblete LA, Onate FA, Schwab W, Figueroa CR. Methyl jasmonate treatment induces changes in fruit ripening by modifying the expression of several ripening genes in Fragaria chiloensis fruit. Plant Physiology and Biochemistry 2013; 70:433-444.

70. Deng WW, Zhang M, Wu JQ, Jiang ZZ, Tang L, Li YY, Wei CL, Jiang CJ, Wan XC. Molecular cloning, functional analysis of three cinnamyl alcohol dehydrogenase (CAD) genes in the leaves of tea plant, Camellia sinensis. Journal of Plant Physiology 2013; 170(3):272-282.

71. Hawkins S, Boudet A. Wound-induced lignin and suberin deposition in a woody angiosperm (Eucalyptus gunnii Hook.): Histochemistry of early changes in young plants. Protoplasma 1996; 191(1):96-104.

72. Kim YH, Bae JM, Huh GH. Transcriptional regulation of the cinnamyl alcohol dehydrogenase gene from sweet potato in response to plant developmental stage and environmental stress. Plant cell reports 2010; 29(7):779-791.

73. de Obeso M, Caparros-Ruiz D, Vignols F, Puigdomenech P, Rigau J. Characterisation of maize peroxidases having differential patterns of mRNA accumulation in relation to lignifying tissues. Gene 2003; 309(1):23-33.

117

74. Delessert C, Wilson IW, Van Der Straeten D, Dennis ES, Dolferus R. Spatial and temporal analysis of the local response to wounding in Arabidopsis leaves. Plant molecular biology 2004; 55(2):165-181.

75. Cabané M, Afif D, Hawkins S: Lignins and Abiotic Stresses. In: Advances in Botanical Research. vol. 61: Academic Press Elsevier; 2012: 220-262.

76. Hu Y, Li WC, Xu YQ, Li GJ, Liao Y, Fu FL. Differential expression of candidate genes for lignin biosynthesis under drought stress in maize leaves. Journal of Applied Genetics 2009; 50(3):213-223.

77. Murashige T, Skoog F. A revised medium for rapid growth and bio assays with tobacco tissue cultures. Physiologia Plantarum 1962; 15.

78. Raes J, Rohde A, Christensen JH, Van de Peer Y, Boerjan W. Genome-wide characterization of the lignification toolbox in Arabidopsis. Plant Physiol 2003; 133(3):1051-1071.

79. Boguski M, Lowe T, Tolstoshev C. dbEST database for expressed sequence tags. Nat Genet 1993; 4(4):332-333.

80. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution 2013; 30(12):2725-2729.

81. Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. Bioinformatics 1992; 8(3):275-282.

82. Huis R, Hawkins S, Neutelings G. Selection of reference genes for quantitative gene expression normalization in flax (Linum usitatissimum L.). BMC Plant Biology 2010; 10:71.

83. Clifford MN. Specificity of acidic phloroglucinol reagents. Journal of Chromatography A 1974; 94:321 324.

84. Roongsattham P, Morcillo F, Jantasuriyarat C, Pizot M, Moussu S, Jayaweera D, Collin M, Gonzalez-Carranza ZH, Amblard P, Tregear JW et al. Temporal and spatial expression of polygalacturonase gene family members reveals divergent regulation during fleshy fruit ripening and abscission in the monocot species oil palm. BMC Plant Biology 2012; 12:150.

Bai X, Zhao PX. psRNATarget: a plant small RNA target analysis server. Nucleic Acids Research 2011;
 39(Web Server issue):W155-W159.

86. Babu PR, Rao KV, Reddy VD. Structural organization and classification of cytochrome P450 genes in flax (Linum usitatissimum L.). Gene 2013; 513(1):156-162.

87. Day A, Addi M, Kim W, David H, Bert F, Mesnage P, Rolando C, Chabbert B, Neutelings G, Hawkins S. ESTs from the Fibre Bearing Stem Tissues of Flax (Linum usitatissimum L.): Expression Analyses of Sequences Related to Cell Wall Development. Plant Biology 2005; 7(1):23-32.

88. Venglat P, Xiang D, Qiu S, Stone SL, Tibiche C, Cram D, Alting-Mees M, Nowak J, Cloutier S, Deyholos M et al. Gene expression analysis of flax seed development. BMC Plant Biology 2011; 11:74.

118

Table 1 Characteristics of the flax phenylpropanoid genes identified in this work.

Gene name	Gene model	Genome	Microarray EST	GenBank EST	Ref
LusPAL1	Lus10040416	scaffold86	c247	JG065006	5
LusPAL2	Lus10023531	scaffold1216	c247	EH791565	4
LusPAL3	Lus10026518	scaffold617	c904	JG216766	5
LusPAL4	Lus10013805	scaffold618	c904	JG216766	5
LusC4H1	Lus10034449	scaffold310	c1220	JG081734	5
LusC4H2	Lus10019110	scaffold30	c32549	JG085936	1;5
LusC4H3	Lus10021671	scaffold208	c6570	GW864597	2
LusC4H4	Lus10035011	scaffold43	c4373	JG075414	1;5
LusC4H5	Lus10027598	scaffold2	c11079		
Lus4CL1	Lus10026143	scaffold319	c3640	CA483243	
Lus4CL2	Lus10008677	scaffold1635	c3640	CA483243	
Lus4CL3	Lus10024123	scaffold353	c680	JG220816	5
Lus4CL4	Lus10005390	scaffold547	c680	JG218659	5
Lus4CL5	Lus10002791	scaffold125	c15833	JG239276	5
Lus4CL6	Lus10016135	scaffold344	c3980	EH791222	4
Lus4CL7	Lus10021431	scaffold612	c3980	GW866203	2
Lus4CL8	Lus10025842	scaffold605			
Lus4CL9	Lus10038259	scaffold28			
LusHCT1	Lus10002321	scaffold120		JG213400	5
LusHCT2	Lus10026097	scaffold319	c3481	JG235206	5
LusHCT3	Lus10010786	scaffold18	c6233	JG202486	5
LusHCT4	Lus10026123	scaffold319	c8096	CV478884	2
LusHCT5	Lus10022163	scaffold342	c6233	JG202486	5
LusC3'H1	Lus10033524	scaffold701	c581	JG230193	5
LusC3'H2	Lus10020847	scaffold711	c5820	JG226878	5
LusC3'H3	Lus10020850	scaffold711	c5820	CA483318	
LusCCoAOMT1	Lus10027888	scaffold1143	c395	JG228418	5
LusCCoAOMT2	Lus10002837	scaffold810	c395	JG228418	5
LusCCoAOMT3	Lus10019841	scaffold1491	c2503	JG094269	5
LusCCoAOMT4	Lus10014074	scaffold1247	c2503	JG231505	3;5
LusCCoAOMT5	Lus10034584	scaffold9			
LusCCR1	Lus10041651	scaffold272	c5445	JG227337	5
LusCCR2	Lus10024068	scaffold353	c2262	GW866594	2
LusCCR3	Lus10008774	scaffold729		JG203738	5
LusCCR4	Lus10022239	scaffold225	c11788	GW865040	2
LusCCR5	Lus10024138	scaffold353	c13273	JG149727	5

			1 · · · · · · · · · · · · · · · · · · ·		
LusCCR6	Lus10042399	scaffold123	c20639	JG103050	5
LusCCR7	Lus10026273	scaffold898	c20639	JG103050	5
LusCCR8	Lus10006885	scaffold329		JG097906	5
LusCCR9	Lus10003780	scaffold806		JG088940	5
LusCCR10	Lus10012930	scaffold434	c6761	JG215733	5
LusCCR11	Lus10030973	scaffold261		JG240727	5
LusCCR12	Lus10035369	scaffold151		JG240727	5
LusF5H1	Lus10028361	scaffold413	c772	JG214534	5
LusF5H2	Lus10041811	scaffold272	c772	JG214534	5
LusF5H3	Lus10014273	scaffold275	c2179	GW865908	2
LusF5H4	Lus10025975	scaffold319	c2179	GW867449	2
LusF5H5	Lus10034300	scaffold310	c56836		
LusF5H6	Lus10012582	scaffold6	c56836		
LusF5H7	Lus10022303	scaffold225	c56836		
LusF5H8	Lus10041511	scaffold272	c56836		
LusCOMT1	Lus10015576	scaffold233	c2253	JG229091	5
LusCOMT2	Lus10032929	scaffold51	c629	JG229091	5
LusCOMT3	Lus10009442	scaffold981	c4476	GW865137	2
LusCAD1	Lus10027864	scaffold1143	c2456	JG020400	5
LusCAD2	Lus10002812	scaffold810	c2456	JG020400	5
LusCAD3	Lus10019811	scaffold1491	c3049	GW865020	2
LusCAD4	Lus10014104	scaffold1247	c3049	JG215109	5;6
LusCAD5	Lus10014363	scaffold275	c4852	JG019640	5
LusCAD6	Lus10010149	scaffold587	c6705	GW867690	2
LusCAD7	Lus10026070	scaffold319	c4852	JG102096	5
LusCAD8	Lus10023268	scaffold98		JG062652	5
LusCAD9	Lus10035956	scaffold76		EH792205	4
LusCAD10	Lus10002089	scaffold575	c4394	JG255907	5
LusCAD11	Lus10025706	scaffold605		GW866334	2
LusCAD12	Lus10002302	scaffold120	c5057	JG282203	5
LusCAD13	Lus10002300	scaffold120	c6631	JG216469	5
LusCAD14	Lus10009955	scaffold200	c3693	JG140771	5
LusCAD15	Lus10039595	scaffold15	c3758	JG217534	5

1: PR Babu, KV Rao and VD Reddy [86]); 2: A Day, M Addi, W Kim, H David, F Bert, P Mesnage, C Rolando, B Chabbert, G Neutelings and S Hawkins [87]); 3: A Day, G Neutelings, F Nolin, S Grec, A Habrant, D Cronier, B Maher, C Rolando, H David, B Chabbert, et al. [30]); 4: MJ Roach and MK Deyholos [23]); 5: P Venglat, D Xiang, S Qiu, SL Stone, C Tibiche, D Cram, M Alting-Mees, J Nowak, S Cloutier, M Deyholos, et al. [88]); 6: M Wrobel-Kwiatkowska, M Starzycki, J Zebrowski, J Oszmianski and J Szopa [60])

 Table 2
 Summary of the stress response of flax phenylpropanoid genes. Asterisks indicate

 potential lignin-specific flax genes

Gene	Stress								
	S	A	Me	eJA	Dro	ught	Wo	und	Light
	Stem	Leaf	Stem	Leaf	Stem	Leaf	Stem	Leaf	Stem
LusPAL1*				-	-		+		-
LusPAL2*	+			-	-		+	-	
LusPAL3	-			-	-			+	
LusPAL4									
LusC4H1*	-		-	-		+	+	+	
LusC4H2*				-	+		+		
LusC4H3	-	+					+		
LusC4H4									
LusC4H5									-
Lus4CL1_2				-		+	+	+	
Lus4CL3									
Lus4CL4*				-	-		+		-
Lus4CL5	+	-	-	-	+		+	+	+
Lus4CL6									
Lus4CL7				-	-		-		
Lus4CL8	+			-					
Lus4CL8_9	+	-		-				+	
LusHCT1*				-	-		+		
LusHCT2*				-	-		+	-	
LusHCT3_5	+	-		-	+	+	+	+	+
LusHCT4			-	-			-	-	
LusHCT5									
LusC3'H1*				-	-		+		+
LusC3'H1_2_3				-	-		+		

LusCCoAOMT1_2_3_4*	-			-	-				
LusCCoAOMT5									
LusCCR1_2*				-	-				
LusCCR2		-		-					+
LusCCR3_4						+			
LusCCR4						+			
LusCCR5	+		-	-	+	+	+		+
LusCCR6			-	-					
LusCCR7			-	-					
LusCCR8				-					
LusCCR8_9				-				+	+
LusCCR10									
LusCCR11*					+	+	+		
LusCCR12									
LusF5H1_2_3_4*	+		+	-	+		+	+	
LusF5H2	+		-	-	+		+	+	
LusF5H5									
LusF5H6									
LusF5H7									
LusF5H8									
LusCOMT1_2*				-	+		+	+	
LusCOMT2*	+			-				+	
LusCOMT3	-			-	+				
LusCAD1_2*				-	-		+	+	
LusCAD3_4		+	-	-	+		+		
LusCAD4_9_11	-	+	-	+	+		+		
LusCAD5_7_12_13				-					+
LusCAD6									
LusCAD8	-	+	-		+	+		+	
LusCAD10	+			-	+	+	+		
LusCAD12		-		-	+		+	+	
LusCAD14					+		+	+	+
LusCAD15	+			-	+		+		+

Figure legends

Figure 1 The monolignol and lignin biosynthetic pathway. 4CL: 4-coumarate:CoA ligase; BGLU: beta glucosidase; C3'H: *p*-coumarate 3-hydroxylase; C4H: cinnamate 4-hydroxylase; CAD: cinnamyl alcohol dehydrogenase; CCoAOMT: caffeoyl CoA 3-*O*-methyltransferase; CCR: cinnamoyl CoA reductase; COMT: caffeate/5-hydroxyferulate *O*-methyl-transferase; CSE: caffeoyl shikimate esterase; F5H: ferulate 5-hydroxylase; HCT: hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl transferase; PAL: phenylalanine ammonia-lyase; UGT: UDP glycosyltransferase.

Figure 2 Expression of the flax phenylpropanoid genes determined by HT-RT-qPCR. A: gene expression in leaves, roots and whole stems. B: gene expression in the inner xylem-rich tissues (IT) and outer fiber-rich stem tissues (OT) of the stem. Values are means ± SD (n=9). Asterisks indicates values that were determined significantly different from their control (A: Leaf and B: OT) using a Mann & Whitney's test (* P<0.01). The heat maps under the histograms represent the expression values in the roots and stems compared to the leaves (A) or in the IT compared to the OT (B). The colour codes and relative expression ranges in log2 values are represented on the left. * = significant difference in expression.

Figure 3 Lignin and lignin gene transcript localization by *in situ* hybridization in flax stems, roots and leaves. A set of paraffin stem sections (10 μ m-thick) were first hybridized with the corresponding sense probes (left column) in order to confirm the specificity of the signal. Bar = 50 μ m. The top panels show lignin localization (in red) by phloroglucinol-HCl staining. phl I: primary phloem; xyl I: primary xylem; xyl II: secondary xylem; cz: cambial zone; bf: bast fiber. Bar = 50 μ m.

Figure 4 Laccase gene family in flax. A: A phylogenetic tree was produced in MEGA 6 using the maximum likelihood method based on the JTT matrix-based model. The analysis was performed with 45 flax, 17 *Arabidopsis*, and 49 poplar laccase sequences. The arrows indicate the targets predicted by psRNAtarget and the asterisks, the experimentally validated targets. B: Heat map showing laccase gene expression in different flax organs and tissues.

Figure 5 *Lus-miR397* structure and validation of predicted targets. The sequences of the *pre-lus-miR397* and the mature *lus-miR397* are indicated. The association with laccase targets are shown only for the six transcripts validated by RLM 5'-RACE. The cleavage sites are indicated above the corresponding position and the number of cloned RACE products sequenced is shown above each sequence (number of sequenced fragments at this position/total number of sequenced clones).

Figure 6 Spatial localization of *lus-mir397* transcripts in flax organs. Paraffin sections (10 μ m) of the stems, roots and leaves were hybridized with the sense controls (left) and the antisense specific probes (right). cz: cambial zone; end: endodermis; phl I: primary phloem; xyl I: primary xylem. Bar = 20 μ m.

Figure 7 Clustering of flax phenylpropanoid genes according to their expression profile in different organs and stress conditions. A: the k-means medians clustering function from the TIGR MultiExperiment Viewer (TM4 Mev v4.8.1) was used to cluster the genes according to their expression profiles. The distance metric was Pearson correlation and the default parameters were used for the k-means calculation. The expression profiles are represented as heat maps. The colour code and expression range in log2 values are represented on top of the figure. B: the biosynthetic pathway leading to monolignol production and further to H, G and S units is summarized. Three expression ratios are indicted as heat maps beside the genes. C: control; CL: continuous light; DEHYD: drought stressed stems; IT: internal stem tissues; L: leaf; MeJA: methyljasmonate; OT: outer stem tissues; R: root; S; stem; SA: salicylic acid.





Figure 2









Figure 5

Pre-lus-miR397	
U U 5'- AU GAG GCAGCG 3'- UA CUC CGUCGC G U	A U CUUG UGAAG UCGUUCA AAC ACUUC AG U G U ACUCU
Mature lus-miR397	
AUUGAGUGCA	GCGUUGAUGAA
LUSLAC1 7/7 A GU 3' A AGUIGCGACGUGAGUUA 5' 5'CUJAGG U UCAACGCUGCACUGAAUGAAGAA3' C GA	LUSLAC5 8/8 A 3' AGUAGUGCGACGUGAGUUA 5' 5'CUCAGG UCAUCAACGCUGCACUCAAUGACGAG3' A
LusLAC9 8/8 3' An UAGUUGGACGUAGUU 5' 5'CUAAGGUU AUCAACGCUGCACUCAA GAAGAG3' A C	LUSLAC38 7/10 A G A 3' A UAGUUGCGACGUGAGUU 5' 5'CUAAGG U AUCAACGUGCACUCAA GAAGAG3' C A C
LUSLAC39 4/6 G 3' AN UAGUUGCGACGUAAGUUA 5' 5'CUAAGAUU AUCAACGCUGCACUCAAUGAAGAG3' A	LUSLAC44 8/8 A A 3' AGUAGUUGCGACGUGAGUU 5' 5'UUGAGG UCAUCAACGCUGCACUCAA AACCAG3' A C

Figure 6



Figure 7



Additional files:

https://bmcplantbiol.biomedcentral.com/articles/10.1186/s12870-017-1072-9

Additional file 1: Fig. S1. Flax phenylpropanoid gene structures. The name of the gene is indicated followed by the length in bp between the start and stop codons in parentheses. The exons are shown as boxes and the introns as lines. The scale in bp is shown on the top of the figure. (JPG 432 Ko)

Additional file 2: Fig. S2. Molecular phylogenetic analysis of the phenylpropanoid genes. (PDF 2 Mo)

Additional file 3: Fig. S3. Expression of the flax phenylpropanoid genes under stress conditions as determined by HT-RT-qPCR. OH: start of the light period on the first day; 48H: end of the night on the second day; 48H_CL: 48 hours after continuous light; CONT: control; Dehyd: dehydration; L: leaf; MeJA: methyljasmonate; S: stem; SA: salicylic acid. (PDF 1.4 Mo)

Additional file 4: Table S1. Primers used for HT-RT-qPCR, RLM RACE and in situ hybridization. (XLSX 15 Ko)

Additional file 5: Table S2. Description of the gene expression localization in the tissues of flax roots, stems and leaves. (DOC 41 Ko)

Additional file 6: Table S3. Position of the MBSII and MBSIIG specific *cis* elements on both strands (+/-) of the flax phenylpropanoid gene promoters. (DOC 79 Ko)

II- Caractérisation de l'UGTome chez le lin

1/- Identification et annotation des gènes UGTs

Dans le but de caractériser l'ensemble des UGTs de lin, il est nécessaire de les identifier de manière la plus exhaustive possible. En prenant avantage de la publication de la séquence du génome du lin (Wang et al., 2012c), une étude a mis en évidence 137 gènes UGTs (Barvkar et al., 2012). Pour identifier ces séquences, les auteurs ont utilisé le motif conservé PSPG qui est présent sur la séquence de la plupart des UGTs. Le fait de n'utiliser qu'une seule séquence consensus pour obtenir les membres d'une famille multigénique peut paraitre assez réducteur et il existe une possibilité non négligeable de ne pas identifier des UGTs portant un motif qui diffère un peu de ce consensus. C'est le cas par exemple des UGT80 et 81 qui n'apparaissent pas dans cette étude. Ces deux familles d'UGTs diffèrent des autres par leur PSPG modifié, leur nombre d'introns importants (de 5 à 13 chez Arabidopsis thaliana) et semblent actifs contre les molécules lipidiques (Paquette et al., 2003). Afin d'avoir un socle plus solide sur lequel travailler, il a été entrepris de ré-identifier l'ensemble des UGTs chez le lin. Pour cela, plusieurs séquences consensus issues de la base de données CAZy ont été utilisées et recherchées sur les gènes prédits issus du séquençage du lin (Wang et al., 2012c). Au total, 282 UGTs ont été identifiés par cette stratégie (Figure 39).



Figure 39: Processus d'identification et de sélection de l'ensemble des UGTs à partir du génome du lin. Les séquences ont été obtenues dans le génome en recherchant des motifs CAZy spécifiques des GT1. Un tri a ensuite été effectué selon la taille des séquences pour arriver à un total de 188 UGTs.

Tableau 4: Identifiants des UGTs de lin dont la taille des séquences peptidiques prédites est supérieure à 300 acides aminés. Certains identifiants de la base de données phytozome, ont été fusionnées et sont ainsi séparés par un tiret. Les identifiants précédemment assignés par Barvkar et al., (2012) ont été indiqués et ceux modifiés dans notre travail surlignés en gris.

Phytozome_Id	ld Barvkar et al., (2012)	LusUGT	Phytozome_Id	ld Barvkar et al., (2012)	LusUGT
Lus10000309		LusUGT712D3	Lus10019364	UGT85Q4	LusUGT85Q4
Lus10000632	UGT85S1	LusUGT85S1	Lus10019808	UGT84G3	LusUGT84G3
Lus10000643	UGT74E1	LusUGT718A2	Lus10019809	UGT84G1	LusUGT84G1
Lus10000991		LusUGT718A1	Lus10019831	UGT73Z4	LusUGT73Z4
Lus10000993		LusUGT85R6	Lus10019832	UGT73Z2	LusUGT73Z2
Lus10001905	UGT72M2	LusUGT72M2	Lus10019833	UGT73B12	LusUGT73I2
Lus10001906	UGT72M1	LusUGT72M1	Lus10019835	UGT73Y1	LusUGT73Y1
Lus10003154	UGT79A3	LusUGT79A3	Lus10019989	UGT75N2	LusUGT75N2
Lus10003322	UGT73X5	LusUGT73X5	Lus10020124	UGT97B2	LusUGT97B2
Lus10003323	UGT73X3	LusUGT73X3	Lus10020381		LusUGT80A7
Lus10003386		LusUGT719A1	Lus10020382		LusUGT80A5
Lus10003453	UGT712B3	LusUGT712B3	Lus10020556	UGT74T2	LusUGT74T2
Lus10003454	UGT712B1	LusUGT712B1	Lus10020559	UGT74T1	LusUGT74L1
Lus10003455		LusUGT712B8	Lus10021163		LusUGT716A1
Lus10003805	UGT71R1	LusUGT71R1	Lus10021437	UGT86A9	LusUGT86A9
Lus10003809		LusUGT74Q1	Lus10021438	UGT86A7	LusUGT86A7
Lus10003900	UGT72M3	LusUGT72M3	Lus10021610	UGT75N3	LusUGT75N3
Lus10003944	UGT72Q1	LusUGT72Q1	Lus10021718		LusUGT89J1
Lus10004381	UGT79B18	LusUGT79D1	Lus10022220	UGT90A10	LusUGT90B1
Lus10004670	UGT709E1	LusUGT709E1	Lus10022221	UGT90D1	LusUGT90D1
Lus10004671	UGT76P1	LusUGT76P1	Lus10022222	UGT90A8	LusUGT90D2
Lus10004672		LusUGT76P3	Lus10022628		LusUGT73X6
Lus10004673_4		LusUGT76Q1	Lus10022815		LusUGT80A4
Lus10005950	UGT72B15	LusUGT72B15	Lus10023893	UGT73W4	LusUGT73W4
Lus10005951	UGT72T1	LusUGT72T1	Lus10023894	UGT73W1	LusUGT73W1
Lus10006280	UGT94H1	LusUGT94H1	Lus10023939	UGT73X1	LusUGT73R1
Lus10006351	LuUGT74S5-1	LusUGT74S5_1	Lus10024035	UGT72N1	LusUGT72N1
Lus10006352	UGT74S6	LusUGT74S6	Lus10024037	UGT72R2	LusUGT72R2
Lus10006353	UGT74S3	LusUGT74S3	Lus10024117	UGT74V1	LusUGT74V1
Lus10006721	UGT74S5	LusUGT74S5	Lus10024118	UGT74V2	LusUGT74V2
Lus10007421	UGT85Q1	LusUGT85Q1	Lus10024486	UGT74S3-1	LusUGT74S3-1
Lus10007597		LusUGT713A1	Lus10024544		LusUGT80A8
Lus10007972	UGT89E1	LusUGT89E1	Lus10024583	UGT85A30	LusUGT85A30
Lus10008453	UGT79B20	LusUGT79C1	Lus10024584	UGT85A31	LusUGT85A31
Lus10008742	UGT74S2	LusUGT74S2	Lus10024834		LusUGT74L2
Lus10008955	UGT94G2	LusUGT94G2	Lus10025513	UGT92G2	LusUGT92G2
Lus10008956	UGT94G3	LusUGT94G3	Lus10025711	UGT91J2	LusUGT91J2
Lus10009412		LusUGT74T3	Lus10025741	UGT85K7	LusUGT85K7
Lus10009556		LusUGT80A6	Lus10025854	UGT78A13	LusUGT78A13
Lus10009876	UGT709E2	LusUGT709F1	Lus10026394	UGT87H1	LusUGT87H1
Lus10010239	UGT73U1	LusUGT73U1	Lus10026793	UGT71P1	LusUGT71P1
Lus10010468		LusUGT74Q2	Lus10026795	UGT71A24	LusUGT71A24
Lus10010476		LusUGT71R2	Lus10026926		LusUGT97B4
Lus10010665	UGT85Q3	LusUGT85Q3	Lus10026927	UGT97B3	LusUGT97B3
Lus10010712	UGT74B3	LusUGT74J1	Lus10027542	UGT88A12	LusUGT88K1
Lus10010943	UGT85A32	LusUGT85A32	Lus10027584		LusUGT717A1
Lus10011882		LusUGT80A3	Lus10027739	UGT73T1	LusUGT73T1
Lus10011883		LusUGT80C1	Lus10027850	UGT91C2	LusUGT91C2
Lus10012011	UGT73X4	LusUGT73X4	Lus10027862		LusUGT84G4
Lus10012148		LusUGT713A2	Lus10028572		LusUGT80B3
Lus10012726	UGT91J3	LusUGT91J3	Lus10028863	UGT94G4	LusUGT94G4
Lus10013337	UGT79B17	LusUGT79B17	Lus10028864		LusUGT94G5
Lus10013367_8		LusUGT79C2	Lus10028865	UGT94G1	LusUGT94G1
Lus10013500	UGT89E2	LusUGT89E2	Lus10029452	UGT72B16	LusUGT72B16
Lus10013652_3		LusUGT85Q5	Lus10029453	UGT72T2	LusUGT72T2

			_		
Lus10013920	UGT85R5	LusUGT85R5	Lus10031068	UGT82A2	LusUGT82B1
Lus10013922		LusUGT85R7	Lus10031248	UGT92A4	LusUGT92B1
Lus10013923	UGT85R4	LusUGT85R4	Lus10031388	UGT85A22	LusUGT85A22
Lus10013924	UGT85R3	LusUGT85R3	Lus10031819	UGT92A3	LusUGT92B2
Lus10013925	UGT85R1	LusUGT85R1	Lus10032218	UGT85A27	LusUGT85A27
Lus10014079	UGT73B7	LusUGT73L1	Lus10032220	UGT85A26	LusUGT85A26
Lus10014080	UGT73Y2	LusUGT73Y2	Lus10034650	UGT89B4	LusUGT89J2
Lus10014082	UGT73B13	LusUGT73I1	Lus10034699	UGT712C1	LusUGT712C1
Lus10014083	UGT73Z3	LusUGT73Z3	Lus10035903	UGT85K6	LusUGT85K6
Lus10014084	UGT73Z5	LusUGT73Z5	Lus10035951	UGT91J1	LusUGT91J1
Lus10014086	UGT73Z1	LusUGT73Z1	Lus10036086	UGT71Q1	LusUGT71Q1
Lus10014105	UGT84G2	LusUGT84G2	Lus10036087	UGT71A25	LusUGT71A25
Lus10014148	UGT74S4	LusUGT74S4	Lus10036088	UGT71Q2	LusUGT71Q2
Lus10014401	UGT73S1	LusUGT73S1	Lus10036837	UGT89A2	LusUGT89H1
Lus10014402	UGT73V1	LusUGT73V1	Lus10036840		LusUGT89H2
Lus10014403	UGT73W3	LusUGT73W3	Lus10036967		LusUGT715A3
Lus10014404	UGT73W2	LusUGT73W2	Lus10036969		LusUGT715A1
Lus10014437	UGT73X2	LusUGT73X2	Lus10037117		LusUGT715A2
Lus10014831		LusUGT709F2	Lus10037268	UGT76P2	LusUGT76P2
Lus10015515	UGT75N1	LusUGT75N1	Lus10039037	UGT71M1	LusUGT71M1
Lus10015744	UGT712B2	LusUGT712B2	Lus10039041	UGT74U1	LusUGT74U1
Lus10015745	UGT712B6	LusUGT712B6	Lus10039237		LusUGT79D3
Lus10015746		LusUGT712F1	Lus10039277	UGT85R2	LusUGT85R2
Lus10015749	UGT712B5	LusUGT712B5	Lus10039301	UGT88A11	LusUGT88K2
Lus10015750		LusUGT712D2	Lus10039302	UGT88G1	LusUGT88G1
Lus10015751	UGT712B4	LusUGT712B4	Lus10039588	UGT72S1	LusUGT72S1
Lus10015752	UGT712B7	LusUGT712B7	Lus10040178	UGT79B19	LusUGT79D2
Lus10015753	UGT712D1	LusUGT712D1	Lus10040242	UGT709E3	LusUGT709E3
Lus10016126	UGT86A8	LusUGT86A8	Lus10040246		LusUGT76Q2
Lus10016127	UGT86A6	LusUGT86A6	Lus10040533		LusUGT716A2
Lus10016128	UGT86A10	LusUGT86A10	Lus10040725		LusUGT76N4
Lus10016268		LusUGT73X7	Lus10041055	UGT85A25	LusUGT85A25
Lus10016459	UGT76M1	LusUGT76M1	Lus10041713	UGT72R1	LusUGT72R1
Lus10016460	UGT76N3	LusUGT76N3	Lus10041715	UGT72N2	LusUGT72N2
Lus10016461	UGT76N1	LusUGT76N1	Lus10042242		LusUGT91E1
Lus10017201	UGT709D1	LusUGT709D1	Lus10042260	UGT87G1	LusUGT87G1
Lus10017542	UGT85Q2	LusUGT85Q2	Lus10042261	UGT87J1	LusUGT87J1
Lus10017825	UGT74S1	LusUGT74S1	Lus10042262	UGT87J2	LusUGT87J2
Lus10018878		LusUGT80B2	Lus10043445	UGT94K1	LusUGT94K1

Au delà de la simple comparaison *in silico*, une inspection manuelle a parfois permis de fusionner des séquences qui avaient été séparées lors de l'annotation du génome probablement à cause d'erreurs de séquençage/assemblage (Tableau 4). Par exemple, Lus10013652 et Lus10013653 identifiés tous deux comme des UGTs d'après la base de donnée CAZy sont en réalité un seul gène (Figure 40). Après cette analyse, seules les UGTs ayant une taille supérieure à 300 acides aminés ont été conservées pour la suite. Il est vraisemblable que les séquences de tailles inférieures sont issues de duplications tronquées et ne sont donc pas actifs. Il ne faut pas non plus exclure que pour certains d'entre eux, des problèmes de séquençage aient eu lieu et que ces séquences aient en réalité une taille supérieure. Au final, 188 UGTs ont été sélectionnées parmi les 282 identifiés initialement pour la suite de notre étude (Figure 39). Figure 40: Exemple d'association de deux fragments d'UGTs sur le scaffold 317. Ces deux fragments de gènes annotés séparément dans la base de données Phytozome sont en réalité un seul gène UGT rebaptisé Lus10013652_3.



Les séquences ont ensuite été annotées sur la base de la nomenclature recommandée pour les UGTs (Figure 32) en utilisant tout d'abord les appellations attribuées aux gènes d'*Arabidopsis*. Si aucune famille ou sous famille ne correspondait chez cette espèce, une recherche sur d'autres UGTs annotées a été effectuée, et si là encore aucun homologue n'avait pu être identifié, un nouveau numéro ou lettre a été attribué. Il faut également noter que nous avons du changer un certain nombre d'annotations qui avaient été précédemment définies (Barvkar et al., 2012) car elles ne respectaient pas tout à fait la nomenclature usuelle. (Tableau 4)

2/- Analyse phylogénétique

L'ensemble de ces UGTs identifiées a été analysé afin de déterminer les différents groupes phylogénétiques. Aux séquences de lin ont été associées celles d'*Arabidopsis thaliana* ainsi que des UGTs issues de groupes phylogénétiques non présents chez cette espèce (Figure 41).



Figure 41: Arbre phylogénétique portant l'ensemble des UGTs de Linum usitatissimum et Arabidopsis thaliana ainsi que des UGTs d'autres espèces représentant le groupe O (AAL92460.1 de Zea mays et ONH95178.1 de Prunus persica), le groupe P (Os07g14630.1 d'Oryza sativa et NP_001148991.1 de Zea mays) et le groupe Q (NP_001136887.1 et NP_001151546.1 de Zea mays). Arbre réalisé par maximum likelihood 100 bootstraps.

Les familles d'UGTs 80 et 81 actifs contre les lipides et les stérols, n'ont pas été inclues dans cette analyse car elles sont beaucoup trop divergentes. De manière logique, l'arbre phylogénétique compilant les UGTomes de ces deux espèces met en évidence très clairement les groupes déjà identifiés lors de précédentes études (Caputi et al., 2012; Li et al., 2001). On note également qu'un autre groupe phylogénétique, Q, spécifiquement identifié chez *Zea mays* n'est pas présent chez le lin (Li et al., 2014a). Par contre, le lin

possède des UGTs issues d'un groupe phylogénétique, P, absent chez *Arabidopsis thaliana* mais bien présent chez d'autres plantes comme *Vitis vinifera, Glycine max, Oryza sativa, Populus trichocarpa, Prunus persica* et *Malus x domestica* (Caputi et al., 2012) (Tableau 5). De façon intéressante, l'arbre phylogénétique montre aussi un nouveau groupe constitué de cinq UGTs se dégageant clairement des autres groupes.

Tableau 5: Répartition des groupes phylogéniques au sein de différent Angiospermes. Données pour A. thaliana, M. domestica, V. vinifera, G. max, P. trichocarpa et O. sativa issues de Caputi et al. (2012); P. persica de Wu et al. (2017); L. usitatissimum de Barvkar et al. (2012) et Z. mays de Li et al. (2014a)

Groupe phylogénique	Arabidopsis thaliana	Glycine max	Linum usitatissimum (2012)	Linum usitatissimum (cette étude)	Malus x domestica	Oryza sativa	Populus trichocarpa	Prunus persica	Vitis vinifera	Zea mays
A	14	25	16	19	33	14	12	10	23	8
В	3	3	5	8	4	9	2	2	3	3
С	3	1	6	6	7	8	6	4	4	5
D	13	43	21	25	13	26	14	19	8	18
E	22	36	22	24	55	38	49	29	46	34
F	3	1	1	1	6	-	-	4	5	2
G	6	15	19	23	40	20	42	34	15	12
н	19	3	6	9	14	7	5	9	7	9
1	1	18	9	15	11	9	5	5	14	9
I	2	3	4	4	12	3	6	7	4	3
К	2	2	5	5	6	1	2	7	2	1
L	17	19	19	26	16	23	23	18	31	23
м	1	4	3	3	13	5	6	14	5	3
N	1	1	1	1	1	2	1	1	1	4
0	-	5	-	-	5	6	3	1	2	5
Ρ	-	3	-	5	5	9	2	4	11	1
Q	-	-	-	-	-	-	-	-	-	7
R?	-	-	-	5	-	-	-	-	-	-
Total	107	182	137	179	241	180	178	168	181	147

3/- Prédiction de la localisation subcellulaire

Afin de mieux caractériser l'ensemble des UGTs de Linum usitatissimum, des analyses in silico de prédiction de domaines transmembranaires et de signaux de sécrétion permettant d'obtenir une idée de la localisation sub-cellulaire possible des protéines ont été réalisées (Tableau 6). Les UGTs chez les plantes sont généralement caractérisées comme étant cytoplasmiques (Ross et al., 2001) et nous avons ainsi voulu vérifier s'il en était de même chez le lin. Tableau 6: Analyse in silico des séquences protéiques d'UGTs de Linum ussitatissimum. Les paramètres par défaut ont été utilisés pour TMHMM v 2.0 (<u>http://www.cbs.dtu.dk/</u>) permettant d'identifier les domaines transmembranaires ; SignalP v 4.1 (<u>http://www.cbs.dtu.dk/</u>) pour identifier les peptides signaux; TargetP v 1.1 (<u>http://www.cbs.dtu.dk/</u>) pour la localisation sub-cellulaire (scores entre 1 et 5 ; plus le score est bas plus le résultat est robuste) ; SecretomeP v 2.0 pour identifier les protéines secrétées par les voies non-canoniques (si le score est > 0,7 la protéine est secrétée).

				Tara	etP	
Nom	Taille (pb)	тмнмм	SignalP 4.1	Localisation	Score	SecretomeP (NN-score)
LusUGT709D1	480	0	Non	_	4	0.643
LusUGT709E1	487	0	Non	S	4	0.455
LusUGT709E3	367	0	Non	S	4	0.587
LusUGT709F1	486	0	Non	_	5	0.386
LusUGT709F2	481	0	Non	-	5	0.437
LUSUGT712B1	4/6	0	Non	-	4	0.603
LusUGT712B2	458	0	Non	-	4	0.685
LusUGT712B4	475	0	Non	-	3	0.729
LusUGT712B5	479	0	Non		4	0.720
LusUGT712B6	476	0	Non	_	5	0.553
LusUGT712B7	328	0	Non	-	2	0.707
LusUGT712B8	466	0	Non	-	3	0.696
LusUGT/12C1	481	0	Non		4	0.596
LusUGT712D1	464	0	Non	M	5	0.525
LusUGT712D2	318	0	Non	M	5	0.628
LusUGT712F1	462	0	Non		5	0.626
LusUGT713A1	505	0	Non		3	0.619
LusUGT713A2	505	0	Non		3	0.452
LusUGT715A1	626	0	Non		3	0.459
LusUGT715A2	365	0	Non	_	3	0.554
LusUGT715A3	465	0	Non	_	5	0.590
LusUGT716A1	360	0	Non		4	0.586
LusUGT716A2	428	0	Non		5	0.543
LUSUGI/1/A1	310	0	Non	C	2	0.154
LUSI GT718A1	562	0	Non	L.	5	0.009
LusUGT719A1	434	0	Non	C	2	0.625
LusUGT71A24	466	0	Non	<u> </u>	3	0.729
LusUGT71A25	467	0	Non	-	3	0.704
LusUGT71M1	493	1	Non	S	5	0.952
LusUGT71P1	495	0	Non		3	0.618
LusUGT71Q1	490	0	Non	S	5	0.496
LusUGT71Q2	506	0	Non	С	5	0.511
LusUGT71R1	500	0	Non	-	4	0.648
LusUGT71R2	501	0	Non	-	5	0.588
LusUGT72B15	484	0	Non	-	4	0.4/1
LusUGT72M1	465	0	Non		2	0.467
LusUGT72M1	485	0	Non	-	3	0.478
LusUGT72M3	485	0	Non		2	0.605
LusUGT72N1	491	0	Non	_	4	0.614
LusUGT72N2	491	0	Non	_	3	0.508
LusUGT72Q1	484	0	Non	S	4	0.383
LusUGT72R1	501	0	Non	_	2	0.539
LusUGT72R2	506	0	Non		2	0.520
LusUGT72S1	474	0	Non	S	5	0.432
LUSUG17211	473	0	Non	-	5	0.536
LusUGT73I1	4/4	0	Non	Ē	5	0.570
LusUGT73I2	464	0	Non	M	5	0.621
LusUGT73L1	487	0	Non	S	3	0.647
LusUGT73R1	490	0	Non	_	5	0.558
LusUGT73S1	530	0	Non	_	3	0.444
LusUGT73T1	487	0	Non	_	5	0.343
LusUGT73U1	512	0	Non	-	4	0.474
LusUGT73V1	497	0	Non	-	3	0.615
LUSUGT73W1	510	0	Non	-	3	0.447
LusUGT73W2	488	0	Non	-	3	0.531
LusUGT73W4	488	0	Non	S	3	0,506
LusUGT73X2	513	0	Non	M	5	0.490
LusUGT73X3	510	0	Non		5	0.514
LusUGT73X4	482	0	Non	С	4	0.347
LusUGT73X5	500	0	Non		2	0.509
LusUGT73X6	403	0	Non		4	0.457
LusUGT73X7	504	0	Non		4	0.429
LusUGT73Y1	498	0	Non	-	3	0.367
LUSUGT/3Y2	342	0	Non	-	3	0.190
LusUG1/321	492	0	Non	5	5	0.358
LusUGT7373	496	0	Non	5	4	0.464
LusUGT73Z4	459	0	Non	,	5	0.275
LusUGT73Z5	473	0	Non	S	5	0.315
LusUGT74J1	462	0	Non	-	3	0.344
LusUGT74L1	475	0	Non		3	0.429
LusUGT74L2	509	0	Non	C	4	0.423
LusUGT74Q1	456	0	Non		3	0.572
LusUGT74Q2	455	0	Non		3	0.453
LusUGT74S1	468	0	Non		3	0.449
LusUGT74S2	465	0	Non	С	5	0.486
LUSUGI/453	4/1	0	Non	C	2	0.493
LusUGT7454	456	0	Non	~	3	0.542
LusUGT74S5	437	0	Non	M	4	0.384
LusUGT74S5 1	305	0	Non		2	0.326
LusUGT74S6	458	0	Non	C	5	0.355
LusUGT74T2	457	0	Non	М	4	0.565
LusUGT74T3	453	0	Non	M	5	0.643
LusUGT74U1	463	0	Non	_	3	0.534
ll usUGT74V1	500	0	Non		3	0.443

LusUGT74V2	476	0	Non		2	0.493
UELIGT7ENI1	505	C C	Non	c	Ā	0.762
LUSUGI75N1	505	0	NON	3	4	0.762
LusUG175N2	509	0	Non	_	5	0.687
LusUGT75N3	473	0	Non	S	4	0.673
usUGT76M1	471	0	Non	С	5	0.564
LucliGT76NI1	452	0	Non	NA NA	2	0.502
LUSUGI / DINI	452	0	NOT	IVI	3	0.595
usUGT76N3	452	0	Non	_	3	0.472
usUGT76N4	451	0	Non		2	0.465
usLIGT76P1	456	0	Non	M	5	0.362
USUGI70F1	430	0	NUIT	IVI	3	0.302
usUGT76P2	482	1	Non	M	2	0.609
usUGT76P3	500	0	Non		5	0.492
usUGT7601	332	0	Non	M	4	0 544
	332	0	Non	IVI		0.544
usUG176Q2	460	0	Non	-	3	0.590
usUGT78A13	464	0	Non	_	5	0.623
usLIGT7943	474	0	Non	_	3	0.510
UGT70047	4/4	0	Non	-	5	0.310
USUG1/9B1/	468	0	NON	_	5	0.349
usUGT79C1	462	0	Non	S	4	0.592
usUGT79C2	459	0	Non	S	4	0.588
usliGT70D1	472	1	Non		5	0.456
.us00179D1	472	1	NUIT	-	3	0.430
usUGT79D2	474	0	Non	-	4	0.584
usUGT79D3	320	0	Non	C	4	0.877
ucl IGT80A3	618	0	Non	C	4	0328
USUGTOOAS	507	0	Non	C C	-	0.520
usUG180A4	507	0	Non	_	4	0.521
usUGT80A5	594	0	Non	C	3	0.382
usUGT8046	564	0	Non		5	0 404
UCLICT20A7	534	0	Nee	-	2	0.503
usUG180A7	5/0	U	NON		2	0.503
usUGT80A8	650	0	Non		2	0.564
usUGT80B2	656	0	Non		2	0,507
UCLICT2002	600	r c	Nee	-	2	0.527
usUG16UB3	000	U	INON		2	0.527
usUGT80C1	547	0	Non		4	0.370
usUGT82B1	376	0	Non		2	0.591
USLIGT84C1	457	0	Nor	c	5	0.499
	43/	0	NOT	3	3	0.466
usUGT84G2	462	0	Non	S	2	0.540
usUGT84G3	468	0	Non	S	2	0.626
UELIGT24C4	165	0	Non		2	0.422
usUG16464	400	U	INON		3	0.432
usUGT85A22	491	0	Non		3	0.519
usUGT85A25	501	0	Non		4	0.284
usLIGT85A26	490	0	Non	_	4	0.499
USUGT05A20	400	0	Non	-		0.455
usUG185A27	489	0	Non	C	5	0.543
usUGT85A30	484	0	Non	_	4	0.502
ucliGT85A31	478	0	Non		4	0.421
	470	2	Non	-		0.721
USUG185A32	655	2	NON	-	2	0.769
usUGT85K6	479	0	Non	_	5	0.710
usUGT85K7	481	0	Non		4	0.604
usUGT8501	/02	0	Non	_	3	0.524
	432	0	Non	-		0.524
.usuG185Q2	501	0	NON	-	5	0.525
usUGT85Q3	487	0	Non	C	2	0.557
usUGT8504	492	0	Non		3	0.525
	400	0	Non	-	2	0.525
USUG185Q5	400	0	NON	L	3	0.680
usUGT85R1	478	0	Non	C	5	0.461
usUGT85R2	495	0	Non	С	3	0.464
ucliCT9ED2	400	0	Non	-	4	0.425
.050018585	490	0	INUIT	-	4	0.433
LusUG185R4	487	0	Non	-	4	0.460
usUGT85R5	462	0	Non	M	4	0.457
usUGT85R6	488	0	Non		4	0.374
	400	0	Non	-		0.574
usug185K/	492	0	NON	-	4	0.520
usUGT85S1	480	0	Non	_	4	0.464
usUGT86A10	489	0	Non		2	0.468
usLIGT86A6	380	0	Non	s	5	0.252
0300100A0	500	0	NON	5	5	0.252
usUG186A7	487	0	Non	-	4	0.580
usUGT86A8	487	0	Non		5	0.537
USUGT8649	494	0	Non		4	0.409
	522		N	-		0.703
usUG18/G1	532	U	NON	_	3	0.761
usUGT87H1	449	0	Non	M	5	0.610
usUGT87J1	473	0	Non		2	0.386
usl IGT8712	473	0	Non	M	5	0.473
	423		NUIT	141	5	0.473
.usUG188G1	481	0	NON	_	3	0.673
usUGT88K1	464	0	Non	M	5	0.520
usUGT88K2	480	0	Non		4	0,526
UELIGT20E1	300	C C	Non	-		0.277
.u300189E1	530	J	INUTI		4	0.277
usUGT89E2	493	1	Non		3	0.767
usUGT89H1	481	0	Non	С	5	0.713
USLIGT80H2	481	0	Non	C	5	0.713
	310		NUC		3	0.713
usUG189J1	518	U	NON		4	0.767
usUGT89J2	481	0	Non		5	0.731
USUGT90R1	465	0	Non	м	5	0.528
UCLICT00D1	402	C C	Nee		2	0.520
usUG190D1	492	U	NON		3	0.519
usUGT90D2	477	0	Non	S	5	0.560
usUGT91C2	471	2	Non	С	5	0.913
USUGT91F1	317	0	Non		Δ	0.214
UCLICTO 111			Nee	-		0.001
020013111	477	0				0.801
11070410	477	0	NOT	S		
USUG19112	477 471	0	Non	S	3	0.841
USUGT9112	477 471 464	0	Non	S	3	0.841
usUGT91J2	477 471 464	0	Non	S S S	3	0.841
usUGT91J2 usUGT91J3 usUGT92B1	477 471 464 485	0 1 0 0	Non Non Non	5 5 5 -	3 3 5	0.841 0.522 0.429
usUGT91J2 usUGT91J3 usUGT92B1 usUGT92B2	477 471 464 485 489	0 1 0 0 1	Non Non Non Non	\$ \$ 	3 3 5 5	0.841 0.522 0.429 0.318
usUGT91J2 usUGT91J3 usUGT92B1 usUGT92B2 usUGT92G2	477 471 464 485 489 522	0 1 0 0 1 1	Non Non Non Non	\$ \$ 	3 3 5 5	0.841 0.522 0.429 0.318 0.394
usUGT91J2 usUGT91J3 usUGT92B1 usUGT92B2 usUGT92G2	477 471 464 485 489 522	0 1 0 0 1 0 1	Non Non Non Non	S S - - - - - - - - - - - - - - - - - -	3 3 5 5 5 5	0.841 0.522 0.429 0.318 0.394
usUGT91J2 usUGT92B1 usUGT92B2 usUGT92G2 usUGT94G1	477 471 464 485 489 522 446	0 1 0 0 1 0 0 0	Non Non Non Non Non Non	S S - - S -	3 3 5 5 5 5 2	0.841 0.522 0.429 0.318 0.394 0.671
usUGT9112 usUGT9113 usUGT92B1 usUGT92B2 usUGT92G2 usUGT94G1 usUGT94G2	477 471 464 485 489 522 446 472	0 1 0 0 1 0 0 0 0 0	Non Non Non Non Non Non Non	S S - - - - - - - - - - - - - - - - - -	3 3 5 5 5 2 1	0.841 0.522 0.429 0.318 0.394 0.671 0.856
usUG19112 usUGT91J3 usUGT92B1 usUGT92B2 usUGT92G2 usUGT94G1 usUGT94G2 usUGT94G3	477 471 464 485 489 522 446 472 456	0 1 0 1 0 1 0 0 0 0 0	Non Non Non Non Non Non Non	S S S 	3 3 5 5 5 2 1 3	0.841 0.522 0.429 0.318 0.394 0.671 0.856 0.716
usUGT91J2 usUGT91J3 usUGT92B1 usUGT92B2 usUGT92G2 usUGT94G1 usUGT94G3 usUGT94G3	477 471 464 485 489 522 446 472 456	0 1 0 1 0 0 0 0 0 0 0	Non Non Non Non Non Non Non	S S S S S S S S	3 3 5 5 5 2 1 3	0.841 0.522 0.429 0.318 0.394 0.671 0.856 0.716
usUGT91J2 usUGT91J3 usUGT92B1 usUGT92B2 usUGT92G2 usUGT94G1 usUGT94G2 usUGT94G3 usUGT94G4	477 471 464 485 489 522 446 472 456 462	0 1 0 1 0 0 0 0 0 0 0 0	Non Non Non Non Non Non Non Non Non	S S S S - - - - - - - - - - -	3 3 5 5 2 1 3 5	0.841 0.522 0.429 0.318 0.394 0.671 0.856 0.716 0.603
usUGT91J2 usUGT91J3 usUGT92B1 usUGT92B2 usUGT92G2 usUGT94G1 usUGT94G3 usUGT94G4 usUGT94G5	477 471 464 485 489 522 446 472 456 462 348	0 1 0 0 1 0 0 0 0 0 0 0 0	Non Non Non Non Non Non Non Non Non Non	S S S - - - - - - - - -	3 3 5 5 2 1 3 5 5 5	0.841 0.522 0.429 0.318 0.394 0.671 0.856 0.716 0.603 0.692
usUGT91J2 usUGT91J3 usUGT92B1 usUGT92B2 usUGT92G2 usUGT94G1 usUGT94G2 usUGT94G3 usUGT94G4 usUGT94G5 usUGT94H1	477 471 464 485 489 522 446 472 456 462 348 474	0 1 0 1 0 0 0 0 0 0 0 0 0 0 0	Non Non Non Non Non Non Non Non Non Non	S S - - - - - - - - - - - - - - - - - -	3 3 5 5 2 1 3 5 5 5 5 5 5	0.841 0.522 0.429 0.318 0.394 0.671 0.856 0.716 0.603 0.692 0.489
usUGT9112 usUGT92B1 usUGT92B1 usUGT92B2 usUGT92G2 usUGT94G1 usUGT94G3 usUGT94G4 usUGT94G4 usUGT94G5 usUGT94H1	477 471 464 485 489 522 446 472 456 462 348 474 474	0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0	Non Non Non Non Non Non Non Non Non Non	S S - - - - - - - - - - - - - - - - - -	3 3 5 5 2 1 3 5 5 5 5 5 5 5 5	0.841 0.522 0.429 0.318 0.394 0.671 0.856 0.716 0.603 0.692 0.489
usUG19112 usUG192B1 usUG192B2 usUG192B2 usUG194G1 usUG194G1 usUG194G3 usUG194G4 usUG194G4 usUG194G4 usUG194H1 usUG194K1	477 471 464 485 522 446 472 456 462 348 474 474	0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0	Non Non Non Non Non Non Non Non Non Non	S S S - - - - - - - - - - - - - - - - -	3 3 5 5 2 1 3 3 5 5 5 5 2 2	0.841 0.522 0.429 0.318 0.394 0.671 0.856 0.716 0.603 0.692 0.489 0.429
USUG19112 USUG19213 USUG19281 USUG19282 USUG19262 USUG19463 USUG19463 USUG19464 USUG19464 USUG19464 USUG19461 USUG19461 USUG19481 USUG19782	477 471 464 485 489 522 446 472 456 462 348 474 482 479	0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	Non Non Non Non Non Non Non Non Non Non	S S - - - - - - - - - - - - - - - - - -	3 3 5 5 2 1 3 5 5 5 5 2 3	0.841 0.522 0.429 0.318 0.394 0.671 0.856 0.716 0.603 0.692 0.489 0.429 0.660
LISUG19112 LISUG192B1 LISUG192B2 LISUG192B2 LISUG192B2 LISUG192G2 LISUG194G1 LISUG194G3 LISUG194G3 LISUG194G5 LISUG194H1 LISUG194H1 LISUG194H1 LISUG194F83	477 471 464 485 489 522 446 472 456 462 348 474 482 479 479	0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	Non Non Non Non Non Non Non Non Non Non		3 3 5 5 2 1 3 5 5 5 5 2 3 3 5 5 5 5 5 5 5 5 5 5 5 5	0.841 0.522 0.429 0.318 0.394 0.671 0.856 0.716 0.603 0.692 0.489 0.429 0.660 0.653

De façon intéressante, cette prédiction montre que huit UGTs de lin possèdent une ou deux hélices transmembranaires d'après la base de données TMHMM (Krogh et al., 2001). En guise de comparaison, seule une UGT d'*Arabidopsis thaliana* (UGT92A1), dont la fonction reste encore inconnue, possède ce type de domaine d'après cette même base de données.

D'après la prédiction SignalP 4.1 (Petersen et al., 2011), il semblerait qu'aucune des UGTs de lin ne possède de peptide signal. Par contre, une analyse de la localisation sub-cellulaire des UGTs par TargetP, permettant de prédire l'adressage des protéines soit vers les chloroplastes, la mitochondrie ou le système de sécrétion, indique des résultats différents. En effet, TargetP met en évidence 20 UGTs chloroplastiques, 16 UGTs mitochondriales et 27 UGTs secrétées (Tableau 6). Les résultats de ce type d'analyse restent tout de même prédictifs et doivent être confirmés par des analyses expérimentales plus robustes. Une étude précédente avait montré l'implication du motif –SVL en position C-terminale pour l'adressage d'une UGT de Zea mays dans les peroxysomes (Hayward et al., 2016) mais aucune prédiction ne le montre pour les UGTs de lin.

La plupart des protéines qui sont secrétées dans le milieu extracellulaire sont dirigées vers le réticulum endoplasmique par un peptide signal présent sur leur séquence N-terminale. Il existe toutefois des protéines de type leaderless secretory proteins (LSP) (Krause et al., 2013) situées sur la surface des cellules et qui ne possèdent pas ce type de peptide. Chez les plantes, cette sécrétion non conventionnelle pourrait intervenir pour la moitié des protéines secrétées (Krause et al., 2013). Par ailleurs, il existe la de bases de données SecretomeP, qui héberge des séquences issues de mammifères et de bactéries (Bendtsen et al., 2004). Même si celle-ci n'a pas été conçue spécifiquement pour les plantes, de nombreuses études de protéomique végétale l'ont utilisé malgré ces limitations (Lonsdale et al., 2016). Au total, 21 UGTs ont un score supérieur à 0,7 ce qui est le score préconisé pour que la protéine soit prédite comme étant sécrétée (Bendtsen et al., 2004).

4/- Expression des UGT chez le lin

Afin d'avoir une première idée générale du rôle possible des 188 UGTs, leur expression a été mesurée dans différents organes, tissus et lors de stress. Comme lors de la caractérisation des gènes impliqués dans la biosynthèse des monolignols, nous avons utilisé la technologie de qRT-PCR à haut débit Fluidigm qui permet d'obtenir les profils d'expression d'un très grand nombre de gènes en même temps. Cette approche est parfaitement adaptée au nombre de gènes que nous souhaitions étudier et a permis d'obtenir ces profils à la fois dans différents organes de la plante mais aussi après l'application d'un stress hydrique d'une blessure ou de cultures en lumière continue. Sur les 188 gènes UGTs identifiés à partir du génome, l'expression de 125 d'entre-eux a été avérée. Il n'a pas été possible de distinguer efficacement certains d'entre eux au moment de la conception des amorces. De ce fait, certains gènes ont été groupés par 2, 3 ou 4 (Figure 42).



Figure 42: Expression des UGTs dans différents organes et lors de stress chez Linum usitatissimum par qRT-PCR. A gauche, le nombre de gène(s) correspond aux groupements d'UGT réalisé lors de la désignation des amorces. Les gènes sous-exprimés sont en bleu et sur-exprimés en rouge. Les cases noires correspondent à des analyses qui n'ont pas pu être réalisées. F : feuilles ; T : tiges.

Il est notable que lors des stress (lumière continue, blessures, racines à la lumière et stress hydrique), l'expression d'un grand nombre d'*UGTs* change. Mais globalement,

ce changement n'est pas aussi fort que les différences d'expression observées au sein des organes/tissus (racines, feuilles, tiges, tissus internes et externes de la tige). De façon intéressante, pour les stress liés à la lumière, la plupart des *UGTs* dont l'expression change sont sous-exprimés (Tableau 7). En effet, les échantillons obtenus pour les racines à la lumière ou les tiges en lumière continue, connaissent une réduction de l'expression de 46 et 47% de leurs gènes *UGTs* respectivement alors que seuls 9 et 6% des gènes UGTs sont sur-exprimés. Excepté pour les feuilles lors d'un stress hydrique, les autres conditions de stress montrent une expression relativement stable pour environ 40% des UGT (Tableau 7). Par contre, à peine moins de 20 % des UGT ont une expression stable au sein de différents organes/tissus ce qui suggère l'existence d'une régulation transcriptionnelle spatiale importante pour la plupart de ces UGTs.

Tableau 7: Variation de l'expression d'UGTs ou groupe d'UGTs dans différentes conditions. Les pourcentages sont rapportés au nombre total de couples d'amorces utilisés lors de la RT-qPCR digitale. Les gènes reportés ont montré des expressions significativement différentes des échantillons de référence par le test statistique de Kruskal-Wallis (N>3 répétitions ; $p_{value} < 0,05$).

Echantillon		Express	Expression des UGT ou groupe d'UGT (en %)						
		Sous-exprimé	Sur-exprimé	Stable	Non déterminé				
	Lumière continue (48h)	47	6	44	3				
	Blessure	28	37	33	2				
Stress	Racine à la lumière	46	9	39	6				
	Stress hydrique (T)	18	36	40	6				
	Stress hydrique (F)	40	36	17	7				
sns	Tissues externe/interne	55	28	13	5				
s/tis	Racine/feuille	53	32	9	6				
gane	Tige/feuille	32	49	13	6				
ō	Bourgeon/fleur	57	15	23	5				

5/- Activités enzymatiques vis à vis de phénylpropanoïdes

Dans un contexte de caractérisation des activités de glycosylation vis à vis de différents phénylpropanoïdes précurseurs de monolignols, nous avons dans un premier temps mesuré l'activité globale au sein de différents organes chez le lin. Pour cela, une

extraction douce permettant de conserver l'activité enzymatique des UGTs a été réalisée puis dix phénylpropanoïdes vrais ont été utilisés comme substrats (quatre acides, trois aldéhydes et trois alcools). Au bout de deux heures d'incubation, l'extrait a été injecté dans une colonne HPLC (en phase inverse) pour y être analysé. Si l'activité de glycosylation était avérée, les produits glycosylés plus hydrophiles que les substrats, présentaient un temps de rétention plus faible. Une partie importante de ce travail a été consacrée à l'optimisation des conditions expérimentales et à la détection des produits glycosylés.

Nos résultats montrent que pour la plupart des substrats, il y a apparition d'un pic au bout de 1h et 2h d'incubation comme par exemple pour l'acide férulique (Figure 43).





Afin de confirmer que les pics apparus après les incubations correspondaient bien aux substrats glucosylés, la molécule de coniférine a été utilisée comme référence pour les extraits enzymatiques incubés avec l'alcool coniférylique (Figure 44). Les chromatogrammes montrent que le standard de la coniférine sort avec le même temps de rétention que le pic qui apparaît à l'issu des incubations (Figure 44).



Figure 44: Chromatogramme (λ = 260 nm) obtenu après injection d'un essai enzymatique contenant des protéines totales de tige de lin (35 jours) incubées pendant 1 et 2 heures avec l'alcool coniférylique comparé au standard coniférine seul.

Une analyse en spectrométrie de masse a confirmé que les produits obtenus correspondaient bien aux substrats sous leur forme glycosylé.

Après avoir effectué toutes ces vérifications, les essais enzymatiques ont été menés sur des extraits protéiques de tiges, de feuilles et de racines (Figure 45). De manière générale, aucune glycosylation n'a été identifiée sur le groupement propyl des acides et ainsi, nous en avons déduit qu'aucun ester n'a été formé. L'acide cinnamique qui ne possède pas de groupement hydroxyle libre sur son noyau aromatique n'est pas glycosylé. La même constatation a été faite pour les acides *p*-coumarique et l'alcool sinapylique.



Figure 45: Comparaison d'activités UGTs dans des extraits de protéines totales de tige, de feuilles et de racines de lin (35 jours). N=3 répétitions ; test statistique Kruskal-Wallis ; * : $p_{value} < 0,05$.
Toutes les autres molécules ont été glycosylées à des degrés différents. La procédure expérimentale permet uniquement de comparer une même activité entre les organes. De manière globale, il est donc très clair que les activités glycosyltransférases dans la racine sont nettement en deçà des deux autres organes. A l'inverse, les feuilles ont une activité enzymatique bien supérieure tandis que la tige possède une activité intermédiaire. En dehors des trois substrats mentionnés précédemment, tous ont été transformés dans au moins un des organes (toujours au moins les feuilles).

Dans le contexte d'un contrôle de la disponibilité des phénylpropanoïdes pour la lignification, les activités UGTs ont été mesurées dans les tissus internes et externes de la tige (Figure 46). Ces résultats montrent qu'il existe moins de différences dans les activités qu'entre les tiges/feuilles/racines. Comme attendu, les trois substrats utilisés précédemment ne sont pas glycosylés. Par ailleurs, les activités enzymatiques vis à vis des aldéhydes coniféryliques et sinapyliques sont significativement plus fortes dans les tissus riches en fibres primaires alors que l'activité de glycosylation de l'alcool *p*-coumarylique est plus forte dans les tissus xylémiens. Cette différence de spécificité dans les tissus internes et externes suggère qu'il existe des UGTs qui vont spécifiquement glycosyler un substrat particulier et que ce n'est pas une glycosylation non-spécifique.



Figure 46: Comparaison d'activités UGTs dans des extraits de protéines totales de tissus externes et internes de la tige de lin (35 jours). N=3 répétitions ; test statistique Kruskal-Wallis ; * : p_{value} < 0,05.

III- /- Identification des UGTs spécifiques des monolignols

1/- Sélection des gènes

1-1/- Analyse phylogénétique

Dans le but d'identifier les UGTs responsables de l'accumulation de phénylpropanoïdes glycosylés dans les tissus externes de la tige du lin (Huis et al., 2012), certaines UGT ont été sélectionnées pour être caractérisées fonctionnellement. La sélection s'est faite à partir d'un arbre phylogénétique regroupant des orthologues des UGTs 72E1, 72E2 et 72E3 d'*Arabidopsis thaliana* préalablement identifiés pour leur capacité à glycosyler des monolignols et/ou certains de leurs précurseurs (Figure 47). Les cinq UGTs sélectionnées pour cette étude sont organisées phylogénétiquement en deux paires (LusUGT72N1 et N2 // LusUGT72R1 et R2) et LusUGT72Q1.



Figure 47: Détail du groupe phylogénétique E de la Figure 41. Les UGTs étudiées dans ce travail sont soulignées en vert et les UGTs d'Arabidopsis thaliana spécifiques des précurseurs de lignines en rouge.

1-2/- Expression

Afin d'avoir une première idée de l'action de ces cinq gènes, leurs expressions ont été analysées par qRT-PCR. Ces expressions sont uniquement à comparer pour un même gène entre différents organes. Nous avons ainsi montré que les profils sont très différents d'un organe à l'autre (Figure 48). En effet, tandis que *LusUGT72N1* et *LusUGT72N2* sont préférentiellement exprimés dans les racines, les autres UGTs sont plutôt spécifiques des feuilles (Figure 48).



Figure 48: Expression relative de LusUGT72N1, N2, Q1, R1 et R2 dans différents organes par qRT-PCR chez des plantes de 35 jours. N>3 répétitions ; test statistique Kruskal-Wallis réalisé contre les valeurs obtenues dans les feuilles; *, ** et *** correspondent à une p_{value} de < 0,05 ; <0,001 et <0,0001 respectivement.

L'analyse de l'expression de ces gènes a également été réalisée sur les tissus internes et externes des tiges (Figure 49). Il semble ici qu'il y ait une différence significative uniquement pour *UGT72Q1* en faveur des tissus externes contenant les fibres primaires hypolignifiées.



Figure 49: Expression relative de LusUGT72N1, N2 et Q1 entre différents tissus de la tige de 35 jours. N>3 répétitions ; test statistique Kruskal-Wallis ; *** correspondent à une $p_{value} < 0,0001$.

Nous avons également voulu savoir si les expressions de ces gènes sont impactées par des stress (Figure 50). Tout d'abord, il est notable que parmi les différentes conditions de stress testées ici, l'expression des gènes *UGTs* varie beaucoup, aussi bien en étant sur- ou sous-exprimés. De manière intéressante, excepté pour *LusUGT72Q1*, un stress hydrique (arrêt d'arrosage), des blessures et de la lumière continue vont provoquer une sur-expression des quatre autres UGTs (Figure 50). En revanche, les profils obtenus suite à l'incubation d'organes avec l'acide salicylique sont plus nuancés. Ainsi *LusUGT72R1* et *LusUGT72Q1* sont sur-exprimés dans les tiges tandis qu'ils verront leur expression diminuer dans les feuilles. Pour *LusUGT72N2*, c'est le contraire : il est sur-exprimé dans les feuilles mais sous-exprimé dans les tiges. L'autre régulateur testé ici est le methyl jasmonate qui va entrainer une diminution globale de l'expression de l'ensemble des cinq gènes UGTs aussi bien dans les feuilles que dans les tiges (Figure 50).



Figure 50: Expression des gènes LusUGT72N1, N2, Q1, R1 et R2 lors de différent stress. Chaque condition de stress est comparée à une condition non stressée adaptée. Le niveau d'expression est représenté par un dégradé dont l'échelle est représentée en haut à droite. Test statistique Kruskal-Wallis N>3 répétitions ; p_{value} < 0,05. ns : non significatif (blanc) ; nd : non déterminé (noire).

1-3/- Production des protéines recombinantes

Afin de caractériser les substrats potentiellement catalysés dans la plante par les cinq UGTs sélectionnés, les protéines recombinantes correspondantes ont été produites dans un système hétérologue. Dans le but de choisir l'hôte optimal pour cette production, des analyses *in silico* ont été conduites pour savoir si ces UGTs possédaient ou non des modifications post-traductionnelles. Les cinq enzymes candidates ne semblent pas être glycosylées (N-, O- et C-glycosylation) d'après la base de données GlycoEP (Chauhan et al., 2013). Par ailleurs, la littérature témoigne de nombreuses études relatant la production efficace d'UGTs dans des systèmes procaryotes et c'est ainsi que nous l'avons également sélectionné. Il a été initialement choisi d'utiliser le vecteur pDEST17 inductible à l'IPTG (isopropyl β -D-1-thiogalactopyranoside) dans la souche bactérienne BL21-CodonPlus-(DE3)-RIL.



Figure 51: Gel SDS-PAGE d'extraits protéiques solubles dans différentes conditions de température et de concentration en IPTG pour la production de LusUGT72N2.

Dans un premier temps, des tests d'expression ont été réalisés en faisant varier le temps d'incubation du substrat avec l'enzyme, la température ainsi que la concentration d'IPTG afin d'identifier les conditions optimales de production. Pour illustrer cette démarche, la Figure 51 montre des résultats obtenus avec différentes conditions d'induction après une nuit de culture pour l'enzyme LusUGT72N2. Après avoir testé

plusieurs conditions (temps de d'induction : 1h, 15h et 24h ; Concentration d'IPTG : 0,1, 0,5 et 1 mM ; température : 21, 28 et 37°C) ces tests ont permis d'identifier les conditions optimales pour les cinq UGT de lin (Tableau 8). De façon générale, la production des UGTs est meilleure pour la température utilisée la plus basse : 21°C. Les concentrations d'IPTG optimales varient entre 0,1 et 0,5 mM et les temps d'induction entre 15h et 24h.

	Température	mpérature [IPTG] (en mM)	
UGT72N1	21°C	0,1	16-24h
UGT72N2	21°C	0,5	16-24h
UGT72Q1	21°C	0,1	16h
UGT72R1	21°C	0,1	16h
UGT72R2	21°C	0,5	16h

Tableau 8: Optimisation des conditions de production des protéines recombinantes.

Une fois ces conditions déterminées, la production de chaque UGT a été suivie par la purification de l'enzyme (portant une étiquette 6-His en position N-terminale) par une colonne au Nickel de type HisTrap (GE Healtcare). L'extrait protéique brut soluble obtenu après lyse bactérienne y a été injecté et ainsi, l'enzyme produite mais aussi certaines protéines artéfactuelles, ont alors été retenues dans la colonne. Puis avec des solutions contenant une concentration d'imidazole croissante (molécule qui va prendre la place des protéines dans la colonne), les protéines ont alors été éluées. Le tampon de base avec l'extrait protéique contenait déjà 25 mM d'imidazole. Afin d'optimiser la purification, différentes concentrations d'imidazole ont été utilisées pour séparer de façon spécifique les UGTs par rapport aux autres protéines. De manière générale, les cinq enzymes de cette étude se comportaient de la même manière. La Figure 52 montre un exemple avec LusUGT72N1. Après avoir élué les protéines avec une solution de 100 mM d'imidazole, une élution à 500 mM a permis de décrocher de façon plus spécifique les protéines d'intérêt.



Figure 52 : Gel SDS-PAGE d'extraits protéiques solubles éluées à différentes concentrations d'imidazole lors de la purification de LusUGT72N1 par colonne HisTrap.

De façon à confirmer l'identité des protéines ainsi purifiées, des westerns blots ont été conduits avec des anticorps spécifiques des étiquettes 6X-His sur les extraits protéiques des fractions solubles après induction à l'IPTG (Figure 53).



Figure 53 : Western blot obtenu à l'aide d'anticorps anti-6XHIS tag sur des extraits protéiques solubles lors de la production de LusUGT72N1, N2, Q1, R1 et R2 ainsi que AtUGT72E2.

1-4/- Activité enzymatique

Une fois purifiées, les UGTs ont été mises en présence des différents substrats afin de déterminer la spécificité des activités enzymatiques. En plus d'avoir utilisé les mêmes phénylpropanoïdes vrais que pour les activités enzymatiques des protéines extraites des organes de la plante (excepté l'aldéhyde et l'alcool p-coumarylique), d'autres substrats comme l'acide salicylique, l'acide gibbérellique et la coumarine ont été utilisés mais pour ceux ci, aucune glycosylation n'a pu être détectée.

Afin de connaître la quantité exacte d'UGTs produites et donc de pouvoir comparer l'action des différentes enzymes, une quantification des bandes sur gel a été réalisée. Pour cela, les extraits protéiques contenant les protéines d'intérêt ont été déposés à différentes concentrations et analysés par le logiciel de traitement d'images Fiji (Schindelin et al., 2012). De cette façon, il a été possible de quantifier spécifiquement les UGTs contenues dans les extraits protéiques (Figure 54). Il a ensuite été possible de comparer les activités rapportées aux quantités d'UGTs. A titre de référence, AtUGT72E2 a également été produit car son activité de glycosylation *in vitro* contre les phénylpropanoïdes avait déjà été caractérisée (Lim et al., 2005).



Figure 54 : Quantification par densitomètre de LusUGT72N1 après séparation d'extraits solubles en SDS-PAGE. Cette analyse a été réalisée par le logiciel Fiji.

En fonction de leurs activités enzymatiques, les cinq UGTs de lin peuvent être classées en trois groupes comprenant (1) LusUGT72R1 et R2 qui ne montrent aucune activité sur les substrats testés, (2) LusUGT72N1 et N2 qui possèdent une forte activité en comparaison avec AtUGT72E2 et (3) LusUGT72Q1 qui a une plus faible activité (Figure 57). De façon intéressante, malgré des séquences primaires très proches, les membres du deuxième groupe semblent avoir chacun un substrat préférentiel. Ainsi,

LusUGT72N2 possède une activité plus importante que les autres UGTs pour l'alcool coniférylique tandis que pour LusUGT72N1 c'est l'acide férulique qui est préféré. Par ailleurs, l'utilisation d'un témoin de production sans plasmide montre bien l'absence d'activité de glycosylation endogène à *E. coli*.



Figure 55 : Essais enzymatiques des différentes protéines recombinantes LusUGT72N1, N2, Q1, R1, R2 et AtUGT72E2 contre des phénylpropanoïdes. Aucune activité n'a été montrée pour LusUGT72R1 et R2. nt : substrat non testé ; nd : activité non détecté. N=3 répétitions ; test statistique Kruskal-Wallis ; * : p_{value} < 0,05.

1-5/- Mutagénèse dirigée

Au vu des résultats d'activités enzymatiques obtenus avec les protéines recombinantes, nous avons tenté d'expliquer les différences de spécificité de substrats en analysant la conformation tridimensionnelle des enzymes. Pour cela, les cinq UGTs ont été modélisées à l'aide du serveur I-TASSER (Iterative Threading ASSEmbly Refinement) et comparés entre eux. I-TASSER va analyser les protéines tronçon par tronçon pour les comparer aux bases de données de protéines déjà identifiées. Comme attendu, les cinq UGTs prennent une conformation de type GT-B comportant deux domaines de type Rossmann $\beta/\alpha/\beta$ qui se font face (Figure 56).



Figure 56 : Modélisation tridimensionnelle de LusUGT72N1 par le logiciel I-TASSER (<u>http://zhanglab.ccmb.med.umich.edu/I-TASSER/</u>). Les trois acides aminés de la triade catalytique sont colorés.

Lorsque les modélisations 3D des UGTs ont été superposées, certaines différences sont apparues au niveau de l'interaction avec le substrat. La superposition des UGTs qui ne possèdent pas d'activité (LusUGT72R1 et R2) avec les UGTs qui ont une forte activité (LusUGT72N1 et LusUGT72N2) montre des divergences au niveau de certains acides aminés (Figure 57) permettant de formuler des hypothèses sur leur fonction. En effet, pour stabiliser la chaine propyl des différents substrats, LusUGT72N1 et N2 possèdent deux acides aminés de type histidines. Ces deux histidines (H84 et H183), du fait de leur caractère basique, pourraient permettre aux protons des fonctions –OH, -CHO ou –COOH en bout de chaine du substrat d'être stabilisés tandis que l'acide glutamique ou l'alanine de LusUGT72R1 et LusUGT72R2 qui ne possèdent pas les mêmes caractères chimiques pourraient au contraire repousser le substrat. Cette hypothèse permettrait donc d'expliquer les différences d'activités de ces UGTs.



Figure 57 : Vue partielle de la superposition des modèles 3D de LusUGT72N2 (en bleu) et LusUGT72R1 (en orange) au niveau de la zone de reconnaissance du substrat (ici l'alcool coniférylique). Les acides aminés qui interagissent passivement avec le substrat sont mis en évidence. Au second plan, l'histidine de la triade catalytique est visible par un astérisque.

LusUGT72Q1 qui a lui une activité de glycosylation faible, semble posséder une configuration tridimensionnelle intermédiaire. D'un coté, il y a présence d'une boucle qui correspond à l'H84 et de l'autre coté, il y a un acide glutamique à la position de l'H183. Cette boucle qui apparaît en retrait ne permet plus de présenter d'acide aminé clé au niveau de la zone de fixation du substrat accepteur. La boucle pourrait donc expliquer ces faibles activités de glycosylation car il est possible que cette absence d'acide aminé pourrait perturber la fixation du substrat (Figure 58).



Figure 58 : Vue partielle de la superposition des modèles 3D de LusUGT72N2 (en bleu) et LusUGT72Q1 (en orange) au niveau de la zone de reconnaissance du substrat (ici l'alcool coniférylique). Les acides aminés qui interagissent passivement avec le substrat sont mis en évidence. Au second plan, l'histidine de la triade catalytique est visible par un astérisque



Figure 59: Alignement des séquence protéiques de LusUGT72N1, N2, Q1, R1 et R2. Le motif PSPG est indiqué en mauve. Les trois acides aminés composant la triade catalytique sont encadré en rouge et les acides aminés supposé interagir avec les substrat accepteur sont encadré en jaune.

Dans le but de tester ces hypothèses, des expériences de mutagénèses dirigées ont été menées sur ces deux histidines (Figure 59). Sur LusUGT72N1, elles ont été mutées afin de les remplacer par les même acides aminées retrouvés chez LusUGT72R1 et LusUGT72R2 à savoir une alanine en position 84 et un acide glutamique en position 183. Cette mutation a été réalisée par amplification du plasmide entier à l'aide d'amorces contenant un mésappariement à la position souhaitée. Puis, les protéines mutantes ont été produites chez *E. coli* et des essais enzymatiques ont été conduits (Figure 60). Excepté pour H183D-LusUGT72N1, de manière générale, les deux mutants montrent une diminution d'activité comparée à la protéine LusUGT72N1 sauvage.



Figure 60 : Essais enzymatiques à partir de LusUGT72N1 et de deux protéines mutées : H84A-LusUGT72N1 et H183D-LusUGT72N1 contre des phénylpropanoïdes. nt : substrat non testé ; nd : activité non détecté. N=3 répétitions ; test statistique réalisé en comparant avec la protéine non mutée (Kruskal-Wallis ; * : p_{value} < 0,05).

1-6/- Amélioration du protocole

Lors de la production et purification d'enzymes recombinantes, les fractions protéiques contenant les UGTs sortant de la colonne de nickel (HisTrap) subissaient deux autres traitements. Elles ont tout d'abord été dessalées à l'aide d'une colonne d'exclusion stérique (de type sephadex - PD-10) afin d'éliminer l'imidazole. Puis l'extrait a été à la fois purifié et concentré par une colonne de concentration (Vivaspin, Sartorius) qui va spécifiquement éliminer les protéines de moins de 30 kDa. Après ces traitements, l'extrait protéique a été séparé sur gel SDS-PAGE pour quantifier les UGTs par densitomètrie. Afin d'éviter cette dernière étape qui peut sembler peu précise, des efforts supplémentaires doivent être menés dans la purification des UGTs. Pour cela, différentes approches sont actuellement conduites comme par exemple l'utilisation de colonnes d'exclusion stérique à grand volume pour séparer de façon plus efficace les protéines selon leur taille. A l'aide d'un détecteur en sortie de colonne, il est possible de visualiser la présence de protéines et de les collecter de manière spécifique. Les premiers essais ont été encourageants comme le prouve la purification de LusUGT72N1 (Figure 61). Le chromatogramme permet de visualiser les protéines séparées selon leur taille. Plus elles sont grandes plus elles vont sortir rapidement. Ici les trois pics majoritaires en sortie de colonne ont été déposés sur gel SDS-PAGE. Il est clairement visible que la fraction 3 en sortie de gel filtration contient l'UGT à une plus forte concentration que la fraction en sortie de HisTrap.



Figure 61 : Purification de LusUGT72N1 par colonne d'exclusion stérique (gel filtration). La fraction recueillie après la colonne Nikel (HisTrap) est passée sur une colonne d'exclusion stérique. Les fractions éluées en sortie de colonne sont analysées par un détecteur UV (λ =230nm) comme rapporté sur le chromatogramme. Trois fractions correspondant aux pics majoritaires du chromatogramme ont été déposées sur gel SDS-PAGE. La fraction n°3 en mauve contient la protéine LusUGT72N1.

IV- La glycosylation des monolignols chez Arabidopsis

1/- Caractéristiques des gènes UGT72E1, UGT72E2 et UGT72E3

Dans la première partie de ce chapitre, nous avons caractérisé chez le lin les orthologues des gènes *UGT72E1*, *UGT72E2* et *UGT72E3* d'Arabidopsis. Au début de ce travail, seuls ces trois gènes avaient été identifiés comme étant liés à la glycosylation de précurseurs de lignines et notamment les monolignols pour les deux derniers (Lim et al. 2001; Lanot et al. 2008; Lanot et al. 2006). Le Tableau 2 montre les variations des deux principaux monolignols glycosylés pour les lignées sur- ou sous-exprimant ces gènes. Comme nous l'avons spécifié précédemment, notre but était de rechercher un lien entre les activités UGTs et la lignification. Au delà de cet aspect, il nous a également semblé intéressant de tenter d'attribuer un rôle biologique aux membres de cette sous famille d'enzymes. Dans la mesure où il est difficile d'obtenir des plantes transformées chez le lin, nous avons donc focalisé la suite de notre travail sur le modèle Arabidopsis.

Les trois gènes liés au métabolisme des phénylpropanoïdes étudiés ici appartiennent à la même famille et sous famille 72E. Les séquences protéiques déduites sont donc logiquement proches (Tableau 9). UGT72E2 et E3 possèdent une similarité forte de l'ordre de 86%, UGT72E1 partage quant à lui une homologie plus faible de l'ordre de 65% avec les deux autres.

Tableau 9 : Pourcentage d'homologie entre les séquences protéiques de UGT72E1, UGT72E2 et UGT72E3.



Les expressions spatio-temporelles au sein de la plante ont également été recherchées sur la base de données eFP Browser (http://bar.utoronto.ca/efp/cgibin/efpWeb.cgi) (Figure 62). Le gène *ugt72e1* est principalement exprimé dans la graine mature ainsi que dans les feuilles en senescence, *ugt72e2* est lui fortement exprimé dans la racine tandis que *ugt72e3* l'est plutôt dans les fleurs, principalement au niveau des anthères. Si on se réfère à d'autres données transcriptomiques réalisées indépendamment sur les principaux tissus de la racine, les transcrits de *ugt72e2* sont essentiellement retrouvés dans le cortex de la racine (Birnbaum et al. 2003).



Figure 62 : Localisation de l'expression d'UGT72E1, UGT72E2 et UGT72E3 dans différents organes d'Arabidopsis thaliana. Les figures ont été obtenues avec eFP Browser (bar.utoronto.ca/eplant) (Waese et al., 2016) à l'aide de données issues de puces à ADN de type Affymetrix ATH1 normalisées par la méthode de GCOS.

2- Obtention des triples mutants ugt72e1/ ugt72e2/ ugt72e3

Afin de pouvoir déterminer le rôle biologique des trois gènes *UGT72E1*, *UGT72E2* et *UGT72E3* et éviter une redondance de fonctions possible, nous avons procédé à l'obtention du triple mutant correspondant. Les simples mutants ont été commandés à partir de la plateforme T-DNA express (http://signal.salk.edu/cgi-bin/tdnaexpress) qui regroupe plusieurs collections de mutants d'insertions d'*Arabidopsis thaliana* (Figure 63).



Figure 63 : Position des ADN-T dans les mutants ugt72e1, ugt72e2, et ugt72e3. Les gènes correspondants sont constitués de gauche à droite de la région 5' UTR (trait fin), sa séquence codante (trait épais), région 3' UTR (trait fin) à l'échelle. pb : paires de bases.

Pour chacun des mutants commandés, l'insertion était positionnée dans l'exon unique des gènes. Les mutants étant livrés sous forme de mélange de graines en ségrégation, nous avons d'abord procédé à la sélection des homozygotes par PCR avec un système à trois amorces (Figure 64). La première PCR a été réalisée avec un couple d'amorces situées de part et d'autre de l'ADN-T et la seconde PCR avec une amorce dans l'ADN-T et une amorce dans l'ADN génomique proche de la zone d'insertion.



Figure 64 : Sélection des individus homozygotes par PCR à trois amorces. Exemple avec le mutant ugt72e2. Les positions des trois amorces sont représentées sur le gène UGT72E2 contenant l'insert en haut. Les différents profils d'amplifications pour des individus sauvages (WT), homozygotes (HomoZ) et hétérozygotes (HétéroZ) sont illustrés en bas.

Les lignées homozygotes ainsi sélectionnées ont été croisées afin d'obtenir dans un premier temps des doubles mutants *ugt72e1-2, ugt72e1-3* et *ugt72e2-3*. La difficulté majeure rencontrée dans l'obtention de ces plantes était due au fait que les gènes *UGT72E2* (At5g66690) et *UGT72E3* (At5g26310) étaient positionnés sur le même chromosome. Ainsi, le criblage des descendants a nécessité l'utilisation d'une population plus importante que les croisements de ces mutants avec *ugt72e1*. Enfin les doubles mutants ont été à leur tour croisés et un dernier criblage a permis d'identifier les triples mutants *ugt72e1-2-3*.



Figure 65 : Expression d'UGT72E1, UGT72E2 et UGT72E3 dans les mutants d'insertion correspondants par RT-PCR. Les ARNs ont été extraites de plantes entières de 25 jours cultivées sur MS ½.

L'effet des mutations a été vérifié par RT-PCR semi-quantitative (Figure 65) sur des ARNs extraits de plantules entières. Dans les plantes sauvages (Col-0), l'expression était bien présente avec un niveau relativement faible pour le gène *UGT72E3*. Ceci est en accord avec des études d'expression effectuées précédemment (Lanot et al. 2006). De façon intéressante, l'expression du gène *UGT72E3* est légèrement plus importante dans les mutants *ugt72e2* et dans une moindre mesure chez *ugt72e1*. La même analyse faite au niveau du triple mutant *ugt72e1-2-3* montre bien que l'insertion des ADN-T a effectivement totalement supprimé l'expression de ces gènes.

3/- Phénotypage des mutants ugt72e1, ugt72e2, ugt72e3.

Afin de déterminer les rôles potentiels de ces UGTs dans le développement chez *Arabidopsis thaliana*, nous avons procédé à une analyse détaillée de leur phénotype en les comparant aux plantes sauvages. Pour certaines expériences, une lignée obtenue par Lanot et al. (2006) a également été utilisée. Elle a été obtenue par l'introduction de fragments de 250 bp correspondant à des portions des séquences codantes des gènes *UGT72E1, UGT72E2 et UGT72E3* afin de favoriser leur inactivation épigénétique par RNAi. Cette lignée *ugt72e1-2-3kd* possède une expression très limitée mais non nulle de ces trois gènes *UGTs*. Par ailleurs, elle a la particularité d'avoir un contenu en monolignols glycosylés très faible comparé au WT (Lanot et al. 2006) (Tableau 2).

De manière générale et quelque soit la mutation, les différentes lignées ne présentent à première vue aucune modification majeure de leur phénotype dans des conditions optimales de croissance (Figure 66). Il nous a donc paru nécessaire d'aborder le phénotypage par des analyses plus précises. Dans la mesure où il était difficile de mener ces démarches de manière exhaustive, nous avons utilisé la bibliographie qui est notamment citée dans le chapitre introductif de ce mémoire pour nous guider vers des paramètres pertinents.



Figure 66 : Phénotype des différents mutants utilisés lors de ce travail. Les plantes ont été cultivées durant 30 jours. TM : Triple mutant ; ugt72e123kd : triple RNAi obtenue par Lanot et al. (2006).

3-1/- Modification de la floraison

Des expériences ont été menées en serre afin de mesurer plus finement les différents caractères liés au développement des plantes. Parmi les caractères pris en compte, le nombre de jours entre la germination et l'émergence de la hampe ou de l'anthèse a été mesuré (Figure 67). Les résultats montrent que l'on peut séparer les différentes lignées en 2 catégories. En effet, les mutants *ugt72e2* et la lignée *ugt72e123kd* se comportent comme les plantes sauvages alors que les mutants *ugt72e1, ugt72e3* et *ugt72e1-2-3* présentent un retard d'environ 2 jours pour les 2 paramètres.



Figure 67 : Comparaison des délais (nombre de jours) compris entre la germination et l'émergence (1 cm) des hampes florales (en haut) ou l'anthèse (en bas) des différents mutants. N>53 individus par génotype. Différence significative avec Col-0 (Test t de Student ;*** p-value < 0,001).

Chez Arabidopsis, le nombre de feuilles dans la rosette est lié à la vitesse de floraison et donc ce caractère a également été pris en compte. Dans ce cas, les mêmes groupes apparaissent en accord effectivement avec le retard de floraison (Figure 68). Ces phénotypes seraient donc en lien avec les deux gènes *ugt72e1* et *ugt72e3* mais pas *ugt72e2*. Il est aussi intéressant de noter qu'il ne semble pas y avoir d'effet cumulatif entre les mutations au niveau du triple mutant.



Figure 68 : Comparaison du nombre de feuilles à l'anthèse entre les différents mutants. N>53 individus par génotype. Différence significative avec Col-0 (test t de Student ;*** p-value < 0,001 ; * p-value < 0,05).

Enfin la lignée *ugt72e123kd* n'a pas le même comportement que le triple mutant *ugt72e1-2-3*. Dans ce cas, on ne peut exclure à la fois la possibilité que cette différence provienne soit de l'expression résiduelle qui est de l'ordre de 10% (Lanot et al., 2006), soit d'un manque de spécificité des séquences utilisées pour l'inactivation génique à travers le RNAi.

Le phénotypage lié à la floraison nous a également conduit à rechercher des différences éventuelles dans les tailles des hampes ou le nombre de siliques (Tableau 10). Sur ces deux critères, aucune différence significative n'a pu être notée.

Tableau 10: Comparaison de la longueur des hampes florales et du nombre de siliques des différents mutants. Ces caractères ont été mesurés au stade 9.70 (plante en senescence) après 60 jours de culture (Boyes et al., 2001). N>20 individus par génotype ; Test t de Student ; pas de différences significatives

Génotype	Longueur de la hampe principale	Nombre de siliques sur hampe principale
Col-0	44,8 ± 5,9	39,6 ± 3,1
ugt72e1-2-3	$40,9 \pm 6,5$	39,8 ± 8,6
ugt72e1-2-3kd	44,5 ± 4,8	36,4 ± 6,2
ugt72e1	43,3 ± 6,2	40,2 ± 2,7
ugt72e2	41,9 ± 5,5	34,1 ± 8
ugt72e3	46,9 ± 5,3	43,5 ± 6,9

3-2/- Impact sur la sénescence

Un autre caractère mesuré sur des plantes cultivées en serre est le nombre de feuilles qui ont débuté leur senescence (changement de couleur allant du vert au jaune/rouge) au moment de l'anthèse. Dans ce cas, le gène *UGT72E1* semble avoir une influence sur ce phénomène car à la fois le mutant correspondant et le triple mutant *ugt72e1-2-3* disposent de plus de feuilles en senescence que le sauvage (Figure 69).



Figure 69 : Comparaison du nombre de feuilles en senescence à l'anthèse. N>53 individus par génotype. Différence significative avec Col-0 (test t de Student ;*** p-value < 0,001).

Afin de confirmer le rôle du gène *UGT72E1* dans la senescence, d'autres expériences ont été conduites *in vitro*. La senescence a été induite artificiellement en plaçant les plantes à l'obscurité pendant 7 jours et des dosages des chlorophylles a et b ont été réalisés (Figure 70). Les résultats de l'expérience confirment bien que *UGT72E1* est lié à ce phénomène puisque les plantes mutantes correspondantes sont significativement plus sensibles au traitement à l'obscurité (Figure 70), ce qui se traduit par une baisse de la quantité des pigments. Cette expérience a été conduite sur deux lignées indépendantes pour le gène *UGT72E1*.



Figure 70 : Sénescence artificielle provoquée sur des plantules de 25 jours incluant les 7 derniers jours de traitement à l'obscurité. A : Aspect des mutants ugt72e1-1 et des témoins Col-0. B : Dosage de chlorophylles des parties aériennes. N>6 individus par génotype. Différence significative avec Col-0 (test t de Student ;*** p-value < 0,001 ; ** p-value < 0,01 ; * p-value < 0,05).

3-3/ Modifications de la biomasse

D'après les données trancriptomiques, certains gènes *UGTs* sont exprimés dans la racine, notamment *UGT72E2* (Figure 62). Il semble donc pertinent de mesurer des caractères liés à cet organe. Sur des plantes qui ont été cultivées en boite de pétri sur milieu nutritif, les longueurs des racines ont été mesurées en fonction du temps (Figure 71). Aucune différence de croissance entre les différents mutants et Col-0 dans ces conditions n'a pu être mesurée.



Figure 71 : Evolution de la longueur des racines au cours du temps chez les différents mutants cultivés en milieu MS ½. Le nombre de jours inclut 2 jours de stratification. N>11 individus par génotype ; Pas de différences significatives par test t de Student.

De plus, les racines ont également été pesées au cours du temps et de nouveau, aucune différence significative n'a été notée (Tableau 11). Il en va de même pour les parties aériennes.

Tableau 11 : Comparaison des masses des parties racinaires et aériennes des différents mutants. Les plantes ont été cultivées in vitro en boite de pétri. N>3 individus par génotype ; Pas de différences significatives par test t de Student.

	18 jours		21 jours		29 jours	
	Partie aérienne	Partie racinaire	Partie aérienne	Partie racinaire	Partie aérienne	Partie racinaire
Col-0	241,6 ± 63,1	53,7 ± 26,3	364,3 ± 116,6	64,9 ± 26,7	559,7 ± 138,1	81,6 ± 39,1
ugt72e1-2-3	145,2 ± 39,9	29,6 ± 12,6	378,2 ± 143,1	85,8 ± 17,2	486,0 ± 109,5	86,0 ± 6,3
ugt72e1	222,8 ± 23,4	47,4 ± 4,2	270,6 ± 118,8	54,2 ± 18,2	510,4 ± 213,0	112,2 ± 33,3
ugt72e2	230,0 ± 18,4	45,2 ± 5,2	264,0 ± 61,1	55,2 ± 20,3	559,6 ± 78,1	76,4 ± 8,2
ugt72e3	217,3 ± 110,0	45,5 ± 29,9	355,0 ± 38,8	58,0 ± 17,7	331,0 ± 236,6	58,4 ± 42,6

4/- Réponses aux stress des mutants

La plupart des UGTs peuvent glycosyler des métabolites secondaires qui sont induits lors de stress biotiques ou abiotiques. Pour ces raisons, l'impact de différents stress a été évalué sur les plantes mutées en comparaison aux témoins.

4-1/- Impact de l'infection par Botrytis cinerea

Certains phénylpropanoïdes pourraient jouer un rôle lors de l'infection des plantes par des pathogènes (D'Auria and Gershenzon 2005). Récemment, une étude a

montré que la sur-expression du gène *UGT72E2* entraînait une tolérance à l'infection de *Verticillium longisporum* (Konig et al. 2014). Dans notre cas, nous avons mené des expériences de stress biotique par un autre pathogène d'*Arabidopsis thaliana* : le champignon nécrotique *Botrytis cinerea*. Pour cela, des spores de ce champignon en suspension ont été appliquées sur les feuilles des différents mutants puis après trois jours d'infection, la taille des lésions a été mesurée. La première expérience a mis en évidence une plus forte tolérance du mutant *ugt72e1* comparé à Col-0 mais cette observation ne s'est pas confirmée par la suite lors de répétitions (Figure 72). De plus, une quantification de *Botrytis cinerea* par PCR quantitative n'a pas non plus permis de mettre en évidence des différences significatives par rapport aux témoins.



Figure 72: Infection des mutants ugt72e1, ugt72e2 et ugt72e3 par Botrytis cinerea. (A): mesure du diamètre des lésions (en mm) après 3 jours d'infection; N>45; test t de Student;*** p-value < 0,001. (B) répétition de la première expérience pour ugt72e1; N>82. Pas de différences significatives par test t de Student

4-2/- Carence en nutriments

Afin d'évaluer le comportement des mutants lors de stress abiotiques liés à des carences, les plantes ont été cultivées *in vitro* avec des concentrations variables de trois nutriments essentiels : l'azote, le phosphore et le fer. Concernant les deux premiers nutriments, les témoins ainsi que les triples mutants *ugt72e1-2-3* se sont développés en carence totale ou en présence de concentrations 5 fois supérieures à celle du milieu MS standard (Murashige and Skoog, 1962). L'aspect général de l'ensemble des plantes est homogène et lorsque l'on considère en particulier les racines, aucune différence entre

les témoins et mutants n'a pu être établie (Figure 73). Les mêmes observations ont été faites sur les parties aériennes et également sur les simples mutants.



Figure 73 : Phénotypes des triples mutants ugt72e1-2-3 comparés à Col-0 lors d'une carence et d'un excès d'azote et de phosphore dans le milieu de culture. Les concentrations [N]=30mM et [P]=5mM correspondent à cinq fois la concentration dans le milieu MS½.

En revanche, lors d'une carence totale en fer, les racines du triple mutant vont avoir une morphologie particulière comparée au sauvage. Ces racines vont en effet avoir tendance à zigzaguer (phénomène qualifié de "waving") de façon très intense (Figure 74). Ce phénotype est lié au gène *UGT72E2* car le mutant correspondant ainsi que le mutant *ugt72e1-2-3* présentent tous deux ce phénotype. De façon intéressante, de fortes concentrations en fer dans le milieu ne provoquent pas de différences phénotypiques visibles.



Figure 74 : Phénotypes des simples mutants ugt72e2 et triples mutants ugt72e1-2-3 comparés à Col-0 âgés de 7 jours (stratification incluse) sur milieu gélosé ne contenant pas de fer (A gauche). Le ratio entre les longueurs réelles des racines (en vert) et la longueur de la ligne droite entre le collet et l'apex racinaire (en rouge) ont été comparés aux témoins. N>8 individus par génotype. Test t de Student ;*** p-value < 0,001.

Ce phénotype avait déjà été décrit chez Arabidopsis lorsqu'il était cultivé verticalement sur un milieu gélosé avec une inclinaison d'environ 20° de la boite de pétri (plante vers le haut) (Mullen et al. 1998). Ce phénomène semble être lié au gravitropisme et à la perception mécanique de la plante. La rigidité du milieu impacte directement la croissance de la plante et va moduler ce phénotype de zigzag (Rutherford and Masson 1996). Afin de savoir si le phénotype observé chez *ugt72e2* a un lien direct avec la disponibilité en fer ou non, des cultures avec différentes concentrations d'agar ont été effectuées. Aucune différence n'a été observée entre le mutant *ugt72e2* et Col-0, cultivés sur des milieux (contenant du fer) avec des taux de 0,5%, 1%, 1,5% et 3% d'agar. Ces résultats suggèrent donc que le phénotype des racines du mutant *ugt72e2* n'est pas en relation avec des contraintes mécaniques. Par contre, il n'est pas exclu que le manque de fer implique de nouvelles contraintes encore inconnues sur le gravitropisme de la plante.

4-3/ Stress au froid

Lorsque que les plantes d'Arabidopsis sont placées à 4°C, elles vont réagir par de nombreux mécanismes dont l'accumulation de phénylpropanoïdes de type flavonoïdes (Schulz et al. 2016). Les plantes vont ainsi accumuler des anthocyanes au niveau des feuilles jusqu'à atteindre une couleur mauve (Bradbury et al. 2012). Afin de tester l'influence des gènes UGTs sur ce phénomène, le triple mutant préalablement cultivé à température 22/19°C a été placé à 4°C sous la même photopériode de 16h/8h. En comparaison avec Col-0, le triple mutant semble montrer un certain retard d'accumulation de ces molécules (Figure 75). Une analyse plus approfondie nécessiterait cependant d'être menée afin de confirmer ces résultats.



Figure 75 : Phénotypes des triples mutants ugt72e1-2-3 comparés à Col-0 âgés de 16 jours (stratification incluse) suite à un stress au froid. Les plantes ont été cultivées initialement entre 22°C/19°C (jour/nuit) pendant 5 jours puis ont été placées à 4°C durant 9 jours.

5/ Relation entre les UGTs et la lignification

Au vu des spécificités de substrat des trois UGTs étudiés ici et de leur influence dans l'homéostasie des monolignols glycosylés in vivo (Lanot et al. 2006; Lanot et al. 2008), il était bien entendu nécessaire de mesurer des paramètres liés à la lignification. Dans un premier temps, différentes coupes transversales de tiges et de racines à différents stades ont été colorées par le test de Wiesner (phloroglucinol-HCl) (Adler et al., 1949). Ce test biochimique permet de mettre en évidence les lignines (en rouge) par la présence de groupement aldéhydiques de certains monomères (Pomar et al., 2002). De plus, des colorations au Toluidine Blue O (TBO) permet de colorer différentiellement les polysaccharides et les lignines (O'Brien et al., 1964). Bien que nous n'ayons pas procédé à des quantifications précises (nombre d'assises lignifiées, diamètre des cellules du xylème,...) il ne semblait pas y avoir de différences entre les mutants et les plantes sauvages Col-0. Par contre, en regardant plus attentivement au niveau de l'hypocotyle et de l'épicotyle de plantes de 52 à 54 jours (stade 8.00 selon Boyes et al. (2001)) il semblait évident que le mutant ugt72e1-2-3 ainsi que ugt72e1 accumulaient plus de fibres sclérifiées (appelé sclérites) au niveau du cortex (Figure 76). Une analyse fine de la cinétique d'apparition de ces sclérites sera nécessaire pour voir si elles apparaissent simplement de manière plus précoce dans ces mutants ou si, *in fine*, elles seront plus nombreuses dans les stades plus avancés du développement.



Figure 76 : Apparition de sclérites sur les mutants ugt72e1/ ugt72e1-2-3 comparés à Col-0. (A) : coupe transversale d'hypocotyles de plantes de 53 jours colorés au phloroglucinol-HCl. (B) : vue grossie sur une coupe transversale d'un mutant ugt72e1 au niveaux des épicotyles montrant des amas de sclérites.

Afin de quantifier précisément les lignines, un dosage au bromure d'acétyle a été effectué sur la base de la hampe florale du triple mutant. Ce dosage réalisé par spectrophotométrie est une façon indirecte et fiable de quantifier les lignines dans des extraits de parois cellulaires (Moreira-Vilar et al., 2014). Cette méthode va entrainer une substitution des groupements –OH libres des lignines par des groupement acétyles de manière à les solubiliser en condition acide (Hatfield et al., 1999). Il est alors possible de mesurer l'absorbance à 280 nm et donc de les quantifier. Cette méthode a été choisie pour sa bonne reproductibilité ainsi que pour sa robustesse face à des échantillons peu lignifiés (Moreira-Vilar et al., 2014).

Ainsi, des plantes ont été cultivées, tout d'abord en jours courts afin de favoriser le développement de la rosette, puis en jours longs. Cette méthode permet de maximiser le développement secondaire et ainsi d'obtenir une hampe florale épaisse unique sur chaque plante (Vanholme et al. 2012). Lorsque les hampes ont atteint 24 cm, la partie située entre 1 et 9 cm au dessus du collet a été récoltée et utilisée pour le dosage de lignine. Les quantités de lignines étaient significativement plus importantes chez le triple mutant comparé au génotype sauvage (Figure 77). Cette augmentation est relativement importante car elle est de l'ordre de 25%.



Figure 77 : Dosage de lignines par la méthode du bromure d'acétyle sur des hampes du triple mutant ugt72e1-2-3 comparés à Col-0. Les plantes ont d'abord été cultivées en jours courts puis en jours longs pour favoriser la croissance secondaire de la hampe. N>8 individus par génotype (une analyse correspond à 3 plantes mélangées). Test du t de Student ;* p-value < 0,05.

Discussion

Nous allons, dans ce chapitre, tenter de remettre nos résultats dans le contexte de la littérature, donner des pistes de recherches futures qui nous semblent pertinents et enfin identifier les questions qui restent ouvertes. Une première partie va être consacrée, de manière générale, aux UGTs chez *Linum usitatissimum* avec une attention particulière pour les cinq UGTs orthologues des UGTs 72E1, 72E2 et 72E3 d'*Arabidopsis thaliana* qui ont une activité de glycosylation sur les monolignols et/ou leurs précurseurs. Dans la deuxième partie, nous ferons une synthèse des résultats obtenus lors de ce travail ainsi que dans les publications récentes pour tenter d'expliquer le lien possible entre la glycosylation des monolignols et la lignification. Enfin la dernière partie sera consacrée aux rôles biologiques des trois UGTs d'Arabidopsis.

I- Les UGT chez Linum usitatissimum.

L'origine de ce projet vient des résultats de la thèse de Rudy Huis (soutenue dans notre équipe en 2012) qui, suite à une collaboration avec l'équipe de W. Boerjan (VIB, Gand), avait montré qu'il existait une accumulation de monolignols et d'oligolignols sous forme glycosylés dans les tissus externes de la tige de *Linum usitatissimum* (Huis et al., 2012). Ces tissus incluent des fibres délimitées par des parois secondaires ayant des quantités de lignines particulièrement basses comparées aux parois secondaires du xylème par exemple. Il paraissait donc intéressant de savoir s'il existait une relation entre cette accumulation de constituants des lignines sous forme glycosylés et le manque de lignification dans ces parois. De manière plus générale, ce travail précédent avait déjà montré que la transcription des gènes liés à la production de lignines était moins active dans ces tissus et donc notre but était de savoir si la glycosylation des précurseurs pouvait également avoir un impact sur les taux de lignines dans les parois de ces fibres.

Ainsi, ce travail s'intéresse aux acteurs de cette glycosylation : les UGTs.

1/-Caractéristiques de l'UGTome de lin

La difficulté majeure d'un travail consacré aux UGTs est la taille de leur famille multigénique. Ces enzymes sont d'ailleurs souvent utilisées pour explorer une partie de l'histoire de l'évolution des génomes (Bock, 2016). En ce qui nous concerne, le nombre de gènes était plutôt d'un inconvénient car il nous a fallu tout d'abord évaluer l'importance de la famille en isolant les différentes séquences, puis aller rechercher celles qui pouvaient spécifiquement être impliquées dans notre problématique. Au début de ce travail, seul un article décrivant différentes UGTs de lin avait été publié (Barvkar et al., 2012). Les auteurs avaient recherché les séquences dans une version du génome qui n'était pas encore totalement annotée. Quelques temps plus tard, l'ensemble des données du séquençage a été déposé sur la base Phytozome et il nous a paru pertinent de voir s'il était possible de compléter cette annotation. La recherche de séquences UGTs est d'autant meilleure qu'elle est faite avec plusieurs modules spécifiques de cette famille d'enzymes. Au delà de la simple utilisation de la séquence PSPG comme le rapporte en général la littérature, nous avons voulu cribler le génome du lin à l'aide de différents motifs répertoriés dans la base de données CAZy. Ce travail a été effectué en collaboration avec MC Arias de l'équipe « Génétique microbienne » de l'UGSF. Ainsi, 282 UGTs putatifs ont été retrouvés grâce à ces motifs conservés. Puis une sélection manuelle a conduit à l'élimination de fragments supposés tronqués mais aussi à mettre en avant certaines erreurs de séquençage/annotations. Au final, nous avons conservé 188 séquences ayant une taille supérieure à 300 acides aminés incluant donc 56 UGTs nouvelles. Cet UGTome complet est relativement comparable à celui d'autres angiospermes d'un point de vue du nombre de séquences et de son organisation phylogénétique (Caputi et al., 2012; Li et al., 2014a; Sharma et al., 2014). Des analyses complémentaires pourraient s'avérer intéressantes car nous avons montré l'existence d'un nouveau groupe phylogénétique (groupe R?) encore inconnu chez les angiospermes. Il serait possible dans un premier temps d'inclure dans une nouvelle analyse un plus grand nombre d'UGTomes pour évaluer une potentielle spécificité de ce groupe.

Sans vouloir nous atteler à une trop grande inflation d'analyses bioinformatiques sur ces séquences, il nous a quand même paru important de rechercher des indices pouvant nous informer sur la localisation éventuelle de ces UGT à partir de leurs séquences protéiques. Ainsi, l'interrogation de banques de données a bien confirmé les profils « classiques » des UGTs de lin par rapport aux autres plantes, à savoir qu'elles semblaient bien être cytosoliques et solubles. Néanmoins, parmi les 188 UGTs analysées *in silico*, certaines possèdent tout de même des propriétés potentiellement atypiques. En effet la prédiction fait état de vingt-huit UGTs sécrétées, huit ayant un ou deux domaines transmembranaires et deux, qui sont LusUGT71M1 et LusUGT91J2, ayant les deux caractéristiques. Il semble dans ce cas intéressant de déterminer le compartiment cellulaire hébergeant ces UGTs particulières par des approches *in vivo* avec une attention toute particulière pour le réticulum endoplasmique car chez les mammifères, les UGTs y sont ancrées.

Enfin, les nombreux résultats d'expression des UGTs de lin obtenus par RT-qPCR digital dans différents tissus mais également lors de stress constituent une base solide pour de possible futures approches de caractérisation fonctionnelle de ces gènes.

2/- Fonctions de UGT72N1, UGT72N2, UGT72Q1, UGT72R1 et UGT72R2

Parmi l'ensemble des UGTs identifiés chez le lin, cinq candidats ont été sélectionnés en raison de leur implication possible dans la glycosylation des monolignols et leurs précurseurs dans les tissus externes de tiges (Huis et al., 2012). Les gènes *LusUGT72N1*, *LusUGT72N2*, *LusUGT72Q1*, *LusUGT72R1* et *LusUGT72R2* ont donc été choisis car ce sont des orthologues d'*AtUGT72E1*, *AtUGT72E2* et *AtUGT72E3* (Lanot et al., 2006, 2008; Lim et al., 2005) connus pour leur action sur ces mêmes molécules.

2-1/- Les informations obtenues à partir des profils d'expression

Dans un premier temps, les profils d'expression des cinq gènes de lin ont été recherchés dans différents organes par qRT-PCR. *LusUGT72N1* et *LusUGT72N2* qui semblent être issus d'une duplication récente du fait de leur homologie forte, possèdent un profil relativement similaire avec une expression préférentielle dans les racines. En revanche, nous n'avons pas mesuré de différences significatives entre les tissus du xylème et les tissus contenant les fibres hypolignifiées de la tige. Le niveau global de l'expression de *LusUGT72R1* et de *LusUGT72R2* dans la tige en général étant extrêmement faible, il n'a pas été possible de déterminer la différence de leur niveau d'expression entre ces deux tissus. Seul *LusUGT72Q1* est plus exprimé dans les tissus externes comparés aux tissus internes. Les profils d'expression de ces UGTs que nous avons déterminé dans la tige sont en accord avec les résultats des puces à ADN de type Agilent obtenus à partir du génome complet (Chantreau et al., 2014). En effet, ces puces n'ont pas non plus réussi à mettre en évidence d'expression de *LusUGT72R1* et *LusUGT72R2* et aucune différence n'a été vue pour *LusUGT72N1* et *LusUGT72N2*. Enfin comme nous l'avons confirmé, *LusUGT72Q1* est plus exprimé dans les tissus externes.

Plus récemment, des analyses par RNAseq ont été réalisées sur l'apex caulinaire (dans lequel les fibres de lin commencent à devenir visibles) de plantes de 14 jours, comparé à la région basale de la tige contenant des fibres plus matures mais aussi de grandes quantités de xylème (Zhang and Deyholos, 2016). Ces deux zones représentaient ainsi les phases d'élongation et d'épaississement des fibres respectivement. Parmi les cinq UGTs, *LusUGT72Q1* était le seul à montrer une expression différente entres ces deux échantillons. Il était plus exprimé dans la partie basale, mais les plantes étant relativement jeunes et cette région très hétérogène en termes de tissus, il était difficile de conclure sur le rôle spécifique de ce gène dans le métabolisme de la paroi des fibres. De façon générale, *LusUGT72Q1* est un bon candidat pour jouer un rôle de régulateur dans le métabolisme pariétal des fibres. Une approche utilisant des gènes rapporteurs pourrait s'avérer très informatif sur la localisation précise de l'expression de ce gène mais aussi sur *LusUGT72N1* et *LusUGT72N2* dans les différents tissus de la tige.

2-2/- Les activités enzymatiques liées aux UGTs de lin

Afin de savoir si les cinq UGTs identifiées pouvaient être à l'origine de la glycosylation de monolignols et de leurs précurseurs, elles ont été produites par expression hétérologue chez *Escherichia coli* et testées vis à vis de plusieurs phénylpropanoïdes. LusUGT72R1 et R2 ne montraient aucune action de glycosylation pour l'ensemble des substrats testés. Par contre LusUGT72N1, LusUGT72N2 et LusUGT72Q1 sont actifs contre la plupart des phénylpropanoïdes vrais. Pour les 2 premiers, on retrouve globalement les mêmes profils d'activités que pour AtUGT72E2 tandis que LusUGT72Q1 a une activité plus faible.

Le fait d'avoir pu produire cinq UGTs avec des activités de glycosylation différentes constitue un bon socle pour étudier plus finement les activités catalytiques. En collaboration avec C. Bompard de l'équipe « Glycobiologie végétale » de l'UGSF, des modèles tridimensionnels de ces enzymes ont pu être établis. Ainsi, nous avons émis l'hypothèse que chez LusUGT72N1 et LusUGT72N2, qui ont une forte activité de glycosylation, deux histidines en position 84 et 183 semblent jouer un rôle important. LusUGT72R1 et R2 possèdent à la même place des acides aminés très différents d'un point de vue physico-chimique. Cette différence pourrait expliquer l'absence d'activité pour LusUGT72R1 et LusUGT72R2. Pour valider cette hypothèse, des expériences de mutagenèse dirigée ont été menées sur les deux histidines de LusUGT72N1 afin de voir si l'activité était affectée. Et en effet, même si l'on n'a pas assisté à une disparition totale des activités, globalement, elles montrent une diminution en intensité. D'autres expériences de mutagenèse avec notamment la production de doubles mutants sont actuellement en cours et des approches en cristallographie sont également prévues dans le cadre d'un stage de Master 2.

II- Relation entre la glycosylation des monolignols et la lignification

1/- Glycosylation et quantités de lignines

Comme nous l'avons déjà mentionné, un des objectifs de ce travail était de tenter d'établir un lien entre les activités enzymatiques UGTs et la lignification chez les plantes. Des travaux précédents le début de ce projet ont fourni un certain nombre d'indices laissant à penser que ce lien était possible. En ce qui concerne le modèle lin, le gène LusUGT72N1 avait déjà été identifié lors du travail d'un doctorant de notre équipe (Chantreau et al., 2014) car son expression était significativement réduite dans les tissus externes du mutant EMS *lbf1* (Chantreau et al., 2014), qui a la particularité d'avoir des fibres primaires anormalement lignifiées. Le lien entre ce gène et la lignification prend donc ici tout son sens mais dans la mesure où l'EMS entraine un grand nombre de mutations sur la plante, il semble nécessaire de croiser ce résultat avec l'étude d'autres mutants *lbf* disponibles dans le laboratoire. En effet, nous avons en notre possession des mutants ayant différents degrés de lignification au niveau des fibres et il serait donc très intéressant de pouvoir rechercher un lien éventuel avec le niveau d'expression de *LusUGT72N1* et, pourquoi pas, d'autres gènes UGTs dont *LusUGT72Q1*. Par ailleurs, dans le cadre du projet ANR PT-Flax (2010-2013), des mutants pour certains de nos gènes UGTs ont été identifiés par une approche TILLING (Targeting Induced Local Lesions in Genomes). Malheureusement, faute de temps, nous n'avons pu exploiter ces plantes mais il est fort à parier qu'elles seront d'un grand intérêt pour la suite de ce travail.

Nous avons aussi déjà fait état de la difficulté à obtenir des plantes transgéniques chez le lin. La sur-expression des gènes UGTs était malgré tout prévue dans les différents travaux programmés au début de la thèse. Malheureusement les premières expérimentations étant des échecs, nous avons rapidement abandonné car les différentes phases de culture *in vitro* étant extrêmement chronophages, nous avons préféré développer des approches expérimentales différentes. Si l'effet d'une perte partielle ou totale de fonction peut être envisagée par une approche de mutagenèse EMS, il n'en va pas de même pour la sur-expression qui elle, peut apporter des informations précieuses sur le rôle de ces gènes UGTs de lin. Ainsi, nous avons procédé à
la transformation hétérologue de plantes d'*Arabidopsis thaliana* à l'aide de constructions 35S :LusUGT72XX. Actuellement les transformants sont au stade T2 et vont être caractérisés prochainement pour leurs activités et leurs teneurs en lignines.

Ainsi, et comme c'est le cas pour un très grand nombre d'équipes de recherche de par le monde, l'utilisation du modèle Arabidopsis s'est révélée être incontournable pour nous également. Mais au delà de son utilisation pour l'expression de gènes hétérologues, cette espèce a également été étudiée pour établir un lien entre la glycosylation par des UGTs et la lignification. L'équipe de Dianna Bowles de l'Université de York a notamment été à l'origine d'un très grand nombre de travaux, souvent cités dans ce manuscrit, sur les trois UGTs de la famille 72E d'Arabidopsis. Les activités de ces enzymes avaient été déterminées après la production de l'ensemble de l'UGTome chez *E. coli* (Lim et al. 2001, 2005). Par contre, à aucun moment n'a été mentionné un impact sur la lignification notamment lorsque ces gènes avaient été réalisés par thioacidolyse et n'ont montré aucune modification (communication personnelle d'A. Lanot).

Ainsi, aucun lien entre les activités des UGTs et la lignification n'avait été montré avant le début de nos travaux. Par contre, des analyses métabolomiques faites dans différents contextes avaient déjà clairement établi un lien entre l'homéostasie des monolignols (dé)glucosylés et les UGTs chez Arabidopsis. Par exemple, sur les trois simples mutants *ugt72e1ko*, *ugt72e2kd* et *ugt72e3kd* (kd = knockdown, inactivation partielle ; ko = knockout, inactivation totale) seul *ugt72e2kd* montre une diminution de monolignols glycosylés (coniférine et syringine) dans les racines exposées à la lumière (Lanot et al., 2006). Mais cette diminution restait faible comparée au triple mutant *ugt72e1-2-3kd*, ce qui suggère une redondance fonctionnelle entre ces trois gènes (Lanot et al., 2006). Par ailleurs, les travaux de thèse d'A. Chapelle à l'Institut Jean-Pierre Bourgin (maintenant INRA Versailles Grignon), utilisant des mutants d'insertion pour ces mêmes trois gènes, n'ont pas permis de trouver de différences significatives entre les teneurs en coniférine et syringine dans les racines à la lumière chez *ugt72e1* et *ugt72e2*, alors que la quantité de ces métabolites était, de manière étonnante, plus élevée dans les tiges de ces mutants (Chapelle, 2009).

Un fait important a eu lieu en août 2016 avec la publication de travaux relatifs aux fonctions du gène *UGT72B1* chez Arabidopsis (Lin et al., 2016). Les auteurs ont

clairement démontré que ce gène, qui fait partie du même groupe phylogénétique E que les trois autres UGT72E (Li et al. 2001), a une action sur la lignification. En plus d'avoir une lignification ectopique au niveau de la hampe florale, le mutant d'insertion *ugt72b1* accumule des anthocyanes (sous-famille des flavonoïdes) au niveau de l'apex caulinaire. Les auteurs ont entrepris de produire la protéine recombinante UGT72B1 afin de tester sa spécificité vis à vis de différents substrats. Cette UGT est capable de glycosyler les monolignols et les aldéhydes in vitro mais moins efficacement que l'UGT72E2. UGT72B1 est très exprimé dans la tige jeune et spécialement dans le xylème. Les lignées surexpresseurs d'*UGT72B1* accumulent plus de coniférine que les plantes sauvages mais de façon étonnante, les mutants ugt72b1 en accumulent encore plus. Ceci pourrait éventuellement s'expliquer par une compensation de la mutation par la sur-expression détectée des gènes UGT72E2 et UGT72B3. Cette hypothèse pourrait également expliquer les résultats de Chapelle (2009) cités précédemment. Cette sur-compensation en monolignols glycosylés, retrouvés dans les hampes florales des mutants ugt72b1, *ugt72e1* et *ugt72e2* laisse donc penser que lorsqu'une UGT, qui agit sur ces métabolites est absente, une autre peut prendre le relais. Toutefois, il faut garder à l'esprit que dans le cas du mutant *ugt72b1*, le phénotype de lignification ectopique associé à la hampe florale est relativement fort. De ce fait, il peut engendrer indirectement des changements importants dans les teneurs des métabolites. Malgré tout, dans le cadre de nos travaux, des résultats préliminaires ont été obtenus en analysant l'expression d'UGT72B1 chez nos différents mutants. Ces premiers résultats suggèrent une sur-expression du gène *UGT72B1* notamment dans le mutant *ugt72e1*. Ces résultats devront être validés par des approches qRT-PCR.

Chez le mutant *ugt72b1*, de nombreux gènes de la voie des phénylpropanoïdes sont sur-exprimés comme par exemple ceux relatifs à la voie de biosynthèse des monolignols mais également certains flavonoïdes. Le mécanisme qui gouverne ce phénotype de lignification ectopique doit encore être démontré mais il semble maintenant clairement établi qu'il existe bel et bien un lien entre la glycosylation des monolignols et la lignification. Les auteurs suggèrent que *UGT72B1* joue le rôle de régulateur de la lignification. La localisation cellulaire de l'expression de *l'UGT72B1* dans les zones lignifiées (xylème) de la tige semble être essentielle à son action. En effet, il est possible de compenser le phénotype du mutant *ugt72b1* en sur-exprimant de façon ectopique (promoteur 35S) *UGT72B3* qui est, chez le sauvage, exprimé dans les feuilles

et les fleurs. La protéine recombinante UGT72B3 est active seulement sur les aldéhydes sinapyliques et coniféryliques *in vitro* contrairement à UGT72B2 qui ne possède pas d'activité contre les phénylpropanoïdes.

En termes de phénotype, sans doute à cause de la lignification ectopique et de l'épaississement global des parois des cellules corticales de la tige, le mutant ugt72b1 est de taille réduite. Il produit également moins de ramifications et de siliques comparé aux plantes sauvages. Ses grains de pollen ont la morphologie altérée probablement responsable du faible taux de germination des tubes polliniques. La raison de cette altération résulterait de la modification de la lignification au niveau des parois cellulaires des grains de pollen. Ce phénotype a déjà été observé dans des mutants liés à la biosynthèse des monolignols (Elkind et al., 1990; Preston et al., 2004). On ne peut exclure également que le dépôt de sporopollénine au niveau de l'exine des grains de pollen soit également perturbé. Celui-ci est nécessaire à la protection et à la maturation des gamètes mâles (Ariizumi and Toriyama, 2011). Il est constitué d'une partie lipidique et d'une partie phénolique (généralement des acides coumarique ou férulique). Dans le mutant *ugt72b1*, ce dépôt de sporopollenine pourrait donc être impacté par un manque d'apport en phénylpropanoïdes sous forme glycosylés et provoquer ce phénotype déjà retrouvé chez des mutants impactés pour la biosynthèse de ce polymère (de Azevedo Souza et al., 2009).

L'étude portant sur la caractérisation du mutant *ugt72b1* a été relativement surprenante car ces mutants avaient déjà été étudiés sans que ce phénotype très fort n'ait été décrit. L'UGT correspondante avait déjà été caractérisée et une faible activité *in vitro* contre l'acide 3,4-dihydroxybenzoïque (Lim et al., 2002) et des xénobiotiques tels que le 3,4-dichloroaniline (DCA) et le 2,4,5-trichlorophenol (TCP) (Loutre et al., 2003) avaient été rapportées. De façon intéressante, ces travaux montrent que *ugt72b1* semble plus résistant à de fortes concentrations de DCA dans le milieu et à l'inverse le surexpresseur est plus sensible (Brazier-Hicks and Edwards, 2005; Brazier-Hicks et al., 2007a). De plus, du DCA radiomarqué, ajouté de manière exogène dans le milieu, est retrouvé fortement lié aux plantes *ugt72b1* ce qui suggère une incorporation dans les parois cellulaires (Brazier-Hicks and Edwards, 2005). Des analyses plus approfondies ont montré que cette radioactivité était incorporée principalement dans les lignines des parois cellulaires des racines (Brazier-Hicks et al., 2007b). Le DCA s'incorporerait ainsi de façon covalente tout comme les autres monomères lorsque le gène *UGT72B1* est inactivé. Cette forme de détoxification sous forme de copolymérisation avait déjà été démontrée chez d'autres espèces (Lange et al., 1998; Sandermann et al., 1983). *UGT72B1* semble donc bien jouer un rôle dans la lignification mais son mécanisme d'action n'est pas encore bien clair.

Pour revenir à nos travaux, nous avons, à peu près au même moment que la publication des travaux sur le gène *UGT72B1*, pu à notre tour établir un lien entre les gènes UGTs et la lignification. Les mutants ugt72e1ko, ugt72e2kd, ugt72e3kd mais également uqt72e1-2-3kd utilisés dans de précédentes études ne montraient pas de changements au niveau des lignines (Lanot et al., 2006, 2008). Dans notre cas, un dosage au bromure d'acétyle sur la partie basale des hampes florales chez le triple mutant d'insertion ugt72e1-2-3 a montré une augmentation des teneurs en lignines par rapport aux lignées Col-0. D'autres analyses sont en cours pour savoir si ce phénotype est spécifique du triple mutant ou s'il est déjà présent chez un des simples mutants. Ces résultats suscitent un certain nombre de commentaires. Tout d'abord, nos expériences ont été conduites dans des conditions particulières (jours courts puis jours long) qui permettent de favoriser la lignification de la hampe florale principale (Vanholme et al., 2012). En effet, la photopériode est importante pour le développement des parois secondaires et notamment la lignification (Wang et al., 2013a). Par exemple, lorsque le mutant d'insertion *bglu45-1* (muté pour un gène de β -glucosidase) était cultivé en jours courts, il ne présentait aucune différence phénotypique avec les plantes sauvages tandis qu'en jours longs, ce mutant sur-accumule des lignines au niveau des zones interfasciculaires de la hampe florale (Chapelle et al., 2012). Il n'est donc pas possible de comparer directement les expériences faites sur des plantes cultivées en photopériodes différentes. Il semble ainsi que la photopériode soit être facteur important pouvant potentiellement faire la liaison entre la glycosylation des monolignols et la lignification. Contrairement aux études antérieures, ici, le triple mutant n'exprime plus du tout les trois UGT. Comme c'est parfois le cas pour les mutants de gènes de lignification (Wang et al., 2014b), il est possible que l'obtention d'un phénotype muté soit assujetti à l'extinction complète du gène car une petite quantité d'UGTs est peut être nécessaire et suffisante pour conserver un phénotype sauvage. On pourrait ainsi expliquer l'augmentation des lignines dans le triple mutant *ugt72e1-2-3* contrairement à *ugt72e1-2-3kd* (Lanot et al., 2008).

2/- Glycosylation et composition des lignines

La composition des lignines est très finement régulée au point de vue de son contenu en monomères et elle peut être différente selon le type cellulaire considéré (Barros et al., 2015). Peu de choses sont connues sur les mécanismes de sélection de ces monomères avant leur intégration dans le polymère. Une fois encore, les fibres primaires de lin peuvent être un bon modèle d'étude car la composition des lignines dans leur parois est relativement différente de celles retrouvées dans les tissus du xylème (Day et al., 2005; del Río et al., 2011). Des analyses par RMN montrent en effet que les lignines extraites de tissus externes de la tige ont un fort taux d'unités H comparées aux lignines des tissus internes (Day et al., 2005; del Río et al., 2011). Les activités glycosyltranférases sur l'alcool p-coumarylique sont significativement supérieures dans les tissus internes de la tige. Ceci est d'autant plus intéressant qu'il n'y a pas de différence pour l'alcool coniférylique et qu'en parallèle, l'alcool sinapylique ne semble pas être glycosylé. Dans ces conditions, il est donc possible que la glycosylation empêche l'alcool p-coumarylique d'être intégré dans les molécules de lignines au sein des parois secondaires du xylème. Plus globalement, il serait intéressant de comparer les activités glycosyltransférases sur des organes donc les compositions en lignines diffèrent, afin de confirmer ce lien possible entre la glycosylation de monomères spécifiques et la composition des lignines.

3/- Glycosylation et interactions protéines-protéines

Les approches *in silico* prédisent que les UGTs liées à la glycosylation des monolignols et leurs précurseurs sont présents dans le cytosol (Ross et al., 2001). Ceci semble relativement logique puisque les substrats y sont synthétisés (Vanholme et al., 2010). Les études métabolomiques récentes montrent bien que les monolignols sous leur forme aglycone sont peu détectés comparés aux formes glycosylées stockées elles dans la vacuole (Dima et al., 2015). A cause de leurs propriétés physico-chimiques particulières, ces molécules semblent être nocives pour les cellules (Le Roy et al., 2016). Cette toxicité supposée ainsi que les faibles teneurs en monolignols sont deux arguments qui sont en faveur d'une production par le biais de l'utilisation de métabolons (Jørgensen et al., 2005).

De plus en plus d'exemples d'interactions protéines-protéines entre certaines enzymes de la voie de biosynthèse des monolignols sont décrites. Parmi celles-ci, les interactions entre PAL et C4H (Achnine et al., 2004) C4H et C3'H (Bassard et al., 2012; Chen et al., 2011) mais aussi 4CL et HCT qui seraient associées aux membranes du réticulum endoplasmique (Bassard et al., 2012). Historiquement, des modèles de métabolons étaient proposés pour être à l'origine de la biosynthèse de différents types de phénylpropanoïdes (Jørgensen et al., 2005). L'un d'entre eux, situé sur la membrane du réticulum endoplasmique, serait constitué des trois enzymes PAL, C4H et 4CL permettant la production des molécules communes à toutes les voies, puis il existerait une partie variable composée d'enzymes spécifiques à différentes voies conduisant à la production de monolignols, de flavonoïdes... (Jørgensen et al., 2005).

Dans un deuxième temps, la glycosylation de ces molécules permettrait de les détoxifier avant qu'elles ne soient relarguées hors du métabolon. C'est par exemple le cas pour la biosynthèse de certaines molécules cyanogéniques glycosylées dont les intérmédiaires sont hautement toxiques pour les cellules (comme l'acide Nhydroxyamique ou les aldoximes) (Jørgensen et al., 2005; Nielsen et al., 2008). D'autres métabolons incluant des UGTs sont connus chez les plantes comme par exemple pour la biosynthèse de glucosinolates (Grubb and Abel, 2006) et de dhurrin (Jørgensen et al., 2005; Kannangara et al., 2011). Malgré tout, il existe encore peu d'exemples d'interactions UGTs-protéines décrits à ce jour même s'il semblerait bien qu'elles soient présentes dans des complexes. En recherchant toutes les interactions impliquant des UGTs dans la base de données de l'interactome d'Arabidopsis (Consortium, 2011), dix interactions (dont 7 non décrites actuellement) sont positives. Il est donc très vraisemblable que des associations dynamiques de ce type soit présentes chez les UGTs qui glycosylent les monolignols. Ces métabolons sont généralement constitués de plusieurs protéines solubles avec au moins une protéine qui permet l'ancrage sur une membrane (Moraes and Reithmeier, 2012). Dans cette configuration, il serait intéressant de rechercher des liens (co-expression, gène liés...) entre les UGTs et des protéines membranaires (Bassard et al., 2017).

De manière générale, il est possible d'affirmer que la glycosylation permet de réduire la toxicité de ces molécules. Dans la mesure où elles existent également sous la forme non glycosylée, on peut s'interroger sur les mécanismes permettant de les transporter jusqu'à leur destination finale comme par exemple la paroi pour les monolignols. En effet, si ces derniers sont synthétisés au niveau de la membrane du réticulum endoplasmique, on peut se demander comment ils sont transportés vers la vacuole ou l'apoplaste. Pour l'instant, il n'existe pas d'exemples de métabolons liés à des transporteurs.

Les anthocyanes sont synthétisés au niveaux du réticulum endoplasmique par un métabolon supposé (Saslowsky et al., 2005; Winkel-Shirley, 1999). Ensuite, leur transport vers la vacuole est réalisé au moins partiellement par des vésicules (Kulich and Žárský, 2014; Pourcel et al., 2010). Il pourrait en être autant pour les monolignols et les oligolignols glycosylés car de plus en plus d'exemples de métabolites spécialisés synthétisés au niveau du réticulum endoplasmique et transportés par ces vésicules vers la vacuole ou l'apoplaste ont été rapportés (Kulich and Žárský, 2014). Il est ainsi tout à fait envisageable que plusieurs types de transports de monolignols cohabitent (Pourcel et al., 2010) avec d'un coté un transport actif de type ABC ou H+ (Alejandro et al., 2012; Miao and Liu, 2010; Tsuyama et al., 2013) et de l'autre des vésicules (Kaneda et al., 2008; Pickett-Heaps, 1968).

Les interactions des UGTs avec d'autres protéines peuvent également être illustrées par leur association à des microtubules. On sait depuis longtemps que la lignification, visible sous formes d'ornementations dans les trachéides et trachées, est sous contrôle de ces microtubules (Marchant, 1979) mais c'est le cas aussi d'autres éléments de la paroi secondaire (Ménard and Pesquet, 2015). Parmi les nombreuses protéines liées aux microtubules durant la différenciation des cellules en trachéides, deux UGTs dont AtUGT72C2 ont été identifiées (Derbyshire et al., 2015). Cette enzyme n'a pas encore été bien caractérisée mais elle possède une forte homologie de séquence avec les UGTs 72E1, E2, E3 et B1 qui interagissent avec les monolignols. De façon intéressante, elle n'est pas présente tout au long de la différenciation ce qui suggère qu'elle pourrait jouer un rôle de régulateur durant la formation des parois secondaires.

4/- Glycosylation et types cellulaires

Il semble désormais communément admis qu'une cellule en cours de lignification ne soit pas elle même seule à réguler ce processus. En effet, certaines cellules peuvent ainsi être « aidées » par des cellules voisines qui vont fournir des composés pour la polymérisation (Serk et al., 2015). Cette hypothèse de « good neighbor » a été démontrée à l'aide de cultures cellulaires du modèle Zinnia elegans (Hosokawa et al., 2001; Pesquet et al., 2013; Tokunaga et al., 2005) mais également chez Arabidopsis thaliana (Smith et al., 2013). Dans cette dernière étude, un microARN artificiel ciblant le gène AtCCR1 a été exprimé sous contrôle d'un promoteur spécifique aux cellules comportant une paroi secondaire lignifiée. Par cette approche originale d'inactivation d'un gène de biosynthèse de monolignols, les auteurs se sont aperçus qu'il y avait toujours déposition de lignines au niveau des parois cellulaires du xylème. Cette observation suppose donc que d'autres cellules voisines (avec un taux de transcrits du gène AtCCR1 identique aux témoins) ont contribué à fournir les monolignols nécessaires pour la lignification. De facon intéressante, les fibres interfasciculaires ne semblent pas affectées puisqu'elles restent hypolignifiées dans ce mutant (Smith et al., 2013). Ces deux stratégies de lignification différentes, avec ou sans collaborations des cellules voisines, peuvent potentiellement inclure des étapes de (dé)-glycosylation. En effet, l'expression de AtUGT72B1 dans la hampe florale a lieu principalement dans les cellules des tissus conducteurs du xylème mais pas dans les fibres interfasciculaires (Lin et al., 2016). Sachant que la lignification des vaisseaux de xylème semble être dépendante des cellules voisines, l'action d'UGT72B1 pour glycosyler les monolignols pourrait intervenir dans un possible mécanisme de transport entre les cellules. Les monolignols synthétisés par les cellules voisines pourraient alors être glycosylés pour être transportés au niveau des parois cellulaires des cellules voisines en cours de lignification. Cette hypothèse n'est certes pas en accord avec celles concernant le transport des monolignols glycosylés qui ont été montrés comme étant préférentiellement transportés vers la vacuole (Miao and Liu, 2010; Tsuyama et al., 2013) plutôt qu'à travers la membrane plasmique, mais ne peut toutefois être exclue.

Une autre preuve possible de l'implication de monomères glycosylés dans un contexte de lignification (non-)autonome vient de la caractérisation d'un mutant d'insertion β -glucosidase-45 d'Arabidopsis thaliana (écotype WS) (Chapelle et al., 2012). Ce mutant *b*-glu45 est donc déficient pour l'un des deux gènes (avec *BGLU46*) potentiellement responsables de l'hydrolyse des monolignols glycosylés (Escamilla-Treviño et al., 2006). Cette plante ne montre pas de phénotype fort mais seulement une augmentation de lignines au niveau des fibres interfasciculaires. Cette augmentation chez *b*-glu45 reste encore difficile à expliquer, mais le fait que l'impact de la mutation soit localisé dans les types cellulaires à lignification « autonomes » (Smith et al., 2013)

pourrait être lié à des mécanismes de (dé)-glycosylations de monomères de lignines. Plus récemment, l'implication différentielle de la lignification des différentes populations cellulaires au sein du xylème a de nouveau été étudié (Smith et al., 2017). Il semble ainsi que les trois types de cellules retrouvés dans le xylème possèdent des profils de lignifications assez différents (Smith et al., 2017). Dans ce cadre, il serait intéressant d'évaluer l'action ou non des activités UGTs/ß-glucosidases.

Un autre type cellulaire lignifié, qui reste encore relativement peu étudié chez Arabidopsis, compose un type particulier de sclérenchyme appelé sclérite (Esau, 1977; Fahn, 1990). Tout comme les fibres primaires de lin, ces cellules se développent en périphérie de la tige et sont issues d'un méristème primaire (Lev-Yadun, 1997). Elles ont toutefois la particularité d'être lignifiées. Lors de la recherche d'un phénotype pour le mutant ugt72e1, nous avons observé une surabondance de ces sclérites au niveau des hypocotyles par rapport aux types sauvages. A ce jour, nous ne savons pas si cette apparition a lieu de façon plus précoce ou si ce changement est quantitatif. Le nombre de sclérites peut être considérablement augmenté en coupant la hampe florale de façon répétée (Lev-Yadun, 1994). Bien qu'il soit drastique, ce stress pourrait s'apparenter aux blessures réalisées sur les plantes de lin qui ont pour effet de moduler l'expression des cinq UGTs testées. Parmi elles, quatre sont significativement sur-exprimées au moins dans un organe analysé (tige et feuilles). Le gène LusUGT72N2 est sur-exprimé jusqu'à 85 fois lors de blessures au niveau de la tige. Bien entendu la comparaison de ces deux stress chez deux espèces différentes doit être faite de manière prudente, mais l'implication d'UGTs lié à la glycosylation des monolignols à ce niveau ne peut pas être exclue.

III- Rôle de UGT72E1 E2 et E3 dans la physiologie des plantes

1/- Développement, sénescence et stress

Afin d'éliminer d'éventuelles redondances fonctionnelles, les plantes mutantes *ugt72e1, ugt72e2* et *ugt72e3* ont été croisées successivement pour obtenir le triple mutant *ugt72e1-2-3*. Les simples mutants ainsi que le triple mutant n'ont pas fait apparaître de différences phénotypiques majeures comparées aux plantes sauvages dans des conditions de cultures optimales. Pour les nombreux caractères phénotypiques mesurés (longueur des hampes, nombre de siliques, longueur des racines, biomasse...)

seuls quelques uns étaient différents. Parmi ceux là, il a été relevé un décalage d'environ deux jours dans l'émergence de la hampe ainsi que l'anthèse pour les mutants *ugt72e1*, *ugt72e3* et *ugt72e1-2-3*. Un dosage de lignines au bromure d'acétyle a révélé que leurs teneurs étaient plus élevées dans la hampe florale de *ugt72e1-2-3*. Ce double phénotype (lignine et temps de floraison) n'est pas sans rappeler celui obtenu après la sur-expression de *PtGT1*, une UGT de *Populus tomentosa Carr.*, chez le tabac (*Nicotiana tabacum*) (Wang et al., 2012a). En plus de montrer une augmentation des lignines au niveau de la tige, ce mutant avait fleuri de façon très précoce (environ 30 jours de différence) comparé au type sauvage (Wang et al., 2012a). Lorsque PtGT1 est incluse dans une analyse phylogénétique, elle est retrouvée dans le groupe contenant les trois UGT72E d'Arabidopsis avec lesquelles elle a une bonne homologie (entre 73 et 75 % de similarité). Il semble donc que d'une espèce à l'autre, les UGTs orthologues puissent jouer un rôle sur des facteurs de développement similaires à savoir la floraison et la lignification.

En plus d'avoir une floraison retardée, le mutant *ugt72e1* connait une sénescence accélérée au niveau de ses feuilles. En effet, à l'anthèse, ce mutant possède plus de feuilles ayant commencé leur senescence que Col-0. Afin de confirmer cette observation, des expériences *in vitro* ont été conduites pour induire artificiellement (par un traitement à l'obscurité) la sénescence. Ces derniers résultats vont également prouver que la sénescence est accélérée chez le mutant *ugt72e1*. Un des symptômes de la senescence est l'augmentation de la péroxydation des lipides au niveau des membranes, ce qui a pour effet d'augmenter les quantités d'espèces réactives de l'oxygène (ROS) (Woo et al., 2013). Certains phénylpropanoïdes, sous leur forme glycosylée, ont été associés à des stress oxydatifs en tant qu'anti-oxydants comme c'est le cas pour l'esculetin et le quercetin 3-O-glycosides qui vont protéger la plante du stress oxydatif engendré lors d'un stress hydrique (Fini et al., 2012). Si l'on considère qu'en conditions *in vivo*, UGT72E1 agit sur des phénylpropanoïdes, il est alors envisageable que certains anti-oxydants ne soient plus produits chez le mutant correspondant, qui connaît alors une accélération de la sénescence.

De façon intéressante, le mutant *ugt72e1* avait déjà été caractérisé pour sa capacité à résister à de fortes concentrations en saccharose (Huang et al., 2014). En effet,

lorsqu'il est cultivé in vitro dans un milieu contenant un excès de ce sucre, ugt72e1 présentait de meilleurs taux de germination et de survie que les plantes sauvages (Huang et al., 2014). Le saccharose est une molécule impliquée dans la résistance aux stress abiotiques et biotiques (Wingler and Roitsch, 2008). Il va par exemple, en tant qu'osmolyte, jouer un rôle important dans la résistance au froid (Cook et al., 2004; Kaplan et al., 2007). Ce fait est à rapprocher de nos résultats montrant une différence de comportement du mutant ugt72e1-2-3 lors de ce stress reflété par l'accumulation d'anthocyanes. Cette molécule semble également être importante pour les interactions plante-pathogène (Wingler and Roitsch, 2008). Par exemple, lors de l'infection d'Arabidopsis thaliana par B. cinerea, la compétition pour le saccharose va être déterminante (Scharte et al., 2005; Veillet et al., 2016). Nous avions, dans une première série d'expérimentations, montré que le mutant ugt72e1 présentait une certaine résistance à l'infection par ce champignon. Il semble nécessaire d'explorer un peu plus cette interaction hôte-pathogène à travers le métabolisme du saccharose. Il pourrait être intéressant, par exemple, de tester plusieurs conditions avec des infections au niveau des racines ou sur des plantes de différents stades par exemple. Par ailleurs, l'apport ou non de saccharose lors de la croissance des mutants ugt72e1, ugt72e2 et ugt72e3 pourrait avoir un impact important sur la présence ou non de monolignols glycosylés (A. Lanot, communication personnelle). En effet, il semblerait que l'absence de saccharose dans le milieu inhiberait l'accumulation des monolignols glycosylés dans les racines à la lumière. Cette observation nécessite bien sûr des validations, mais la relation entre les UGT72E (spécialement UGT72E1) et le saccharose reste encore à explorer.

2/- Métabolisme du fer

Lorsque le mutant *ugt72e2* se développe *in vitro* en l'absence de fer, ses racines vont avoir tendance à zigzaguer de façon très importante. Nous avons vérifié que dans nos conditions expérimentales, ce phénomène était bien indépendant de la concentration en agar utilisée. Selon la littérature, ce phénomène résulte du gravitropisme (effet de la gravité), du thigmotropisme (dû au contact entre la racine et le milieu gélosé) et la circumnutation (association entre la croissance en longueur et la croissance différentielle au niveau des zones de division/élongation) (Migliaccio and Piconese, 2001). Au cours d'une étude récente, des plantes cultivées dans l'espace ont montré que ce phénotype restait visible en dehors de toute gravité (Paul et al., 2012).

Les racines de mutants déficients en certains transporteurs d'auxine ont un phénotype semblable à celui que nous observons (Luschnig et al., 1998; Rashotte et al., 2000) montrant ainsi que cette phytohormone intervient dans ce processus. Mais d'autres molécules semblent également être importantes comme les brassinostéroides (Wolf et al., 2012), les cytokinines (Kushwah et al., 2011), le saccharose et l'éthylène (Buer et al., 2003). Ce dernier semble être importante pour l'adaptation de la plante à une carence en fer (García et al., 2015). En effet, lors qu'il venait à manquer dans le milieu, le niveau d'éthylène était significativement augmenté au niveau des racines (García et al., 2010; Romera et al., 2017) et lorsqu'il était ajouté de manière exogène, il provoquait une augmentation de la fréquence des zigzags (Buer et al., 2003). Dans le cas du mutant *ugt72e2*, le fait que le milieu soit déficient en fer pourrait provoquer une production dérégulée d'étylène qui aurait pour conséquence ce phénotype. Il serait, dans ce cas, intéressant de réaliser des expériences complémentaires en faisant intervenir cette phytohormone pour confirmer ou non cette possibilité.

Une autre hypothèse impliquant le transport de fer du milieu exterieur vers la plante pourrait également être formulée. En effet, lors d'une déficience en fer, les plantes (mais pas les poacées) vont secréter des coumarines qui vont interagir avec les ions dérivant du fer, facilitant ainsi leur assimililation (Curie and Mari, 2016). Un aspect essentiel pour l'accessibilité au fer est le pH du milieu (Clemens and Weber, 2016). En effet, les ions sous la forme Fe³⁺ dans un mileu basique, sont inutilisables par la plante qui doit les reduires en Fe²⁺ avant une utilisation ultérieure. Récemment, une étude à mise en évidence un mutant d'Arabidopsis thaliana f6'h1 qui possède une plus forte sensibilité aux millieux alcalins (pH = 7,5) (Schmid et al., 2014). Le gène altéré est essentiel pour la biosynthèse de coumarines et ainsi, chez ce mutant, le groupe de molécules phénoliques sécrété par les racines est très altéré lors d'une croissance en milieu alcalin (Schmid et al., 2014). Parmi les molécules qui sont les plus impactées par la mutation, la scopoletin, la scopolin et l'esculin montrent de fortes diminutions au niveau de l'exhudat racinaire, les deux dernières ayant une forme glycosylée probablement acquise avant leur exportation. Pour faire le lien avec nos résultats, on peut noter que UGT72E2 est capable de glycosyler la scopolin et l'exculin in vitro (Lim et al., 2003). Il est donc facilement envisageable que le phénotype de zigzag racinaire provient du fait que le mutant *ugt72e2* est incapable de synthétiser correctement ces coumarines. Les conditions expérimentales ne sont certes pas les mêmes car dans un cas le fer n'est pas accessible (sous sa forme Fe^{3+}) et dans l'autre, le fer est absent mais malgré tout, des expériences supplémentaires de cultures du mutant *ugt72e2* en conditions variables de pH seront très probablement informatives.

Enfin, il est possible que d'autres phénylpropanoïdes comme les lignanes, constituées de dimères de monolignols, puissent avoir une action liée la chélation d'ions métalliques dont le fer (Fucassi et al., 2014) et que le dégré de glycosylation du lignane secoisolariciresinol (qui peut être glycosylé deux fois) soit déterminant dans cette intéraction. Chez le lin, c'est LusUGT74S1 qui catalyse ces deux glycosylations successives (Ghose et al., 2014). Celui ci est assez éloigné des trois UGT72E d'Arabidopsis qui n'auraient donc pas cette fonction mais par contre, il est donc tout à fait envisageable que l'homéostasie des precurseurs des lignanes soit dérégulée dans le mutant *ugt72e2* et que ceci ait une incidence sur la morphologie racinaire du fait de la perturbation de l'absorption du fer. En complément, il est intéressant de noter que des travaux récents (Decourtil, 2016) ont montré que plusieurs gènes en lien avec l'homéostasie du fer ont vu leur expression modifié chez un double mutant d'Arabidopsis prr1-prr2. Le rôle des PRRs (pinorésinol réductases) est de réduire le pinorésinol formé après la dimérisation des alcools coniféryliques. Chez ce mutant, la synthèse des lignanes est donc perturbée et l'imact sur le métabolisme du fer se reflète par une sur-expression de facteurs de transcription spécifiques de ce métabolisme et d'un transporteur de fer ainsi que la sous-expression de ferritines. Il reste maintenant à verifier si ces mutants ont un phénotype racinaire équivalent à celui de *ugt72e2*.

Pour conclure, cette étude a permis de dégagér des hypothèses mais aussi quelques certitudes mettant en relation les activités de glycosylation par les UGTs et la lignification. Il reste par contre encore beaucoup à faire pour cerner le(s) rôle(s) biologiques des différents acteurs potentiellement impliqués.

Matériel et méthodes

I- Matériel végétal

1/- Cultures de lin

Les plantes de lin (*Linum usitatissimum* cultivar Diane) ont été cultivées en serre avec une photo/thermopériode de 16h/20°C de jour et 8h/18°C de nuit sauf indications contraires. Elles ont été semées à raison de 16 plantes par pot de 12 x 12 cm. Tous les échantillons ont été congelés dans l'azote liquide immédiatement après leur prélèvement.

Les échantillons de racines, feuilles et tiges entières de plantes âgées de 60 jours ont été utilisés pour réaliser les essais enzymatiques et la RT-qPCR digitale. Les tissus internes et externes ont été séparés manuellement au niveau des 30 premiers centimètres à partir du collet. Les bourgeons floraux et les fleurs ont été récoltés progressivement (chaque fleur s'ouvre pendant un seul jour) sur des plantes d'environ 90 jours. Les stress hydriques ont été menés sur des plantes de 40 jours par un arrêt d'arrosage de 10 jours (l'arrosage a été maintenu pour le témoin).

Les plantes cultivées en lumière continue pendant les dernières 48h étaient âgées de 37 jours avant le prélèvement des tiges. Cette expérience a été conduite en phytotron et la photopériode du témoin correspondant a été maintenu à 16h jour/8h nuit.

Des blessures d'un centimètre réalisées au scalpel ont été occasionnées à des tiges à 40 cm du collet sur des plantes âgées de 35 jours. Après 24 h, des fragments de 3 cm encadrant la zone de blessure ont été récoltés.

Afin d'obtenir des racines exposées à la lumière, des semis ont été effectués dans des pots à coprologie contenant du milieu MS ½. La base des pots du lot témoin ainsi que la surface du milieu ont été recouvertes de papier aluminium afin d'assurer l'obscurité. Les racines ont été prélevées sur des plantes âgées de 10 jours.

2/- Cultures d'Arabidopsis thaliana

Les graines d'*Arabidopsis thaliana* Col-0 utilisées comme témoins ainsi que les lignées mutantes d'Arabidopsis ont été commandées auprès de l'institut Salk (USA). Les plantes ont été cultivées en serre dans les conditions 16h/20°C de jour et 8h/18°C de nuit. Pour le relevé de phénotype lié au développement (apparition de la hampe, anthèse, sénescence, taille de la hampe principale, nombre de siliques...), les graines avaient d'abord été placées pendant 48h dans une solution de phytagel 0,1% à l'obscurité (4°C) pour avoir une germination synchronisée, puis elles ont été semées sur

terreau. Pour la recherche de phénotypes liés aux racines ainsi que l'induction artificielle de la sénescence, les graines ont d'abord été incubées dans un microtube contenant une solution d'hypochlorite de sodium à 2% pendant 5 min puis lavées successivement pendant 1, 5, 10 et enfin 15 min avec de l'eau stérile. Elles ont ensuite été semées sur un milieu gélosé MS ½ (Murashige and Skoog, 1962).

Les boites de pétri ont été placées pendant les premières 48h à l'obscurité à 4°C pour assurer la synchronisation de la germination, puis en chambre de culture à 16h/22°C de jour et 8h/20°C de nuit.

Pour le dosage de lignines, les germinations ont d'abord été synchronisées puis transférées sur du terreau. Les cultures ont subies une photo/thermopériode de 12h/22°C de jour et 12h/20°C de nuit. Au bout de six semaines (plantes au stade rosette), celle-ci a été changée en 16h/22°C de jour et 8h/20°C de nuit. Lorsque la hampe florale mesurait 24 cm, des tronçons situés entre +1 cm du collet et +9 cm ont été collectés (Vanholme et al., 2012).

3/- Obtention du triple mutant ugt72e1-2-3

Les croisements d'*Arabidopsis thaliana* permettant d'obtenir le triple mutant *ugt72e1-2-3* ont été réalisés à l'aide de plantes cultivées à environ une semaine d'écart afin d'obtenir un décalage entre le développement des gamètes femelles (plus jeunes dans des fleurs fermées) et mâles (plus âgées dans des fleurs ouvertes). Lorsque la plante femelle était à un stade 6.00 (première fleur ouverte (Boyes et al., 2001)), les bourgeons floraux fermés ont été disséqués sous une loupe binoculaire pour laisser apparent le pistil. Les étamines d'une fleur ouverte ont ensuite été prélevées puis frottées délicatement sur le pistil. Pour maximiser les chances de réussite, cette opération a été faite sur environ deux ou trois pistils. Si le croisement a été réalisé correctement, une silique se forme quelques jours plus tard. Les graines issues de cette silique sont doubles hétérozygotes pour les deux gènes mutés considérés. Elles ont donc été semées pour donner des nouvelles graines qui vont être en ségrégation (1/9 des graines vont être homozygotes pour la double mutation). La sélection de ces plantes a été réalisée par deux amplifications de fragments spécifiques par PCR pour distinguer les hétérozygotes des homozygotes.

II- Biologie moléculaire

1/- Extraction d'ADN pour le génotypage

Une ou deux feuilles d'Arabidopsis ont été collectées puis séchées pendant une nuit dans une étuve à 40°C. Les échantillons ont alors été introduits dans un microtube de 2 mL avec une bille en verre puis broyés pendant 15 secondes dans un FastPrep®-24 à 4 m/s. La poudre obtenue a alors été incubée dans 400 μ L d'une solution d'extraction (200 mM de Tris-HCl pH=8; 200 mM NaCl; 25 mM EDTA; 0,5% de SDS). Après 10 secondes de passage au vortex, 150 μ L d'acétate de potassium (3M pH=6) ont été ajoutés dans les tubes qui ont ensuite été centrifugés à 13000g pendant 2 min. Puis 450 μ L de surnageant ont été prélevés dans un nouveau microtube contenant le même volume d'isopropanol. Le mélange a été agité puis incubé à température ambiante pendant 2 min, et enfin centrifugé à 13000g pendant 5 min pour obtenir un culot d'ADN. Le surnageant a été éliminé et le culot lavé à deux reprises avec 700 μ L d'éthanol 80%. Les culots ont été séchés pendant environ 10 à 15 min puis l'ADN solubilisé dans 200 μ L de tampon TE (10 mM Tris-HCl pH=8; 1 mM EDTA). Les tubes ont pu être conservés quelques semaines à -20°C sans que l'ADN ne soit dégradé.

2/- Extraction d'ARN, synthèse d'ADNc et quantification de l'expression

Le protocole d'extraction suivant a été appliqué aux expériences de RT-PCR semiquantitative chez les mutants d'Arabidopsis, les mesures d'expression par RT-qPCR des *UGTs* chez le lin ainsi qu'en amont du clonage des gènes *UGTs* qui contenaient un intron (LusUGT72R1 et R2). Les échantillons de plantes ont été broyés dans l'azote liquide à l'aide d'un mortier et un pilon. La poudre a alors été transvasée dans un microtube de 2 mL préalablement refroidi à l'azote liquide. Puis 1 mL de TriReagent (Euromedex) a été ajouté dans le tube qui a ensuite été vortexé pendant 10 min puis centrifugé à 12000g pendant 10 min (4°C). Le surnageant a été prélevé puis mélangé durant 15 secondes à 200 μ L de chloroforme. Les tubes ont alors été centrifugés à 12000g (4°C) pendant 15 min. La phase supérieure contenant les ARNs a alors été mélangée à 500 μ L d'isopropanol puis incubée à température ambiante pendant 10 min. Le précipité d'ARN séparé par centrifugation de 10 min à 12000g (4°C) a été lavé à l'aide d'une solution de 70% d'éthanol puis séché pendant 5 à 10 min. Les ARNs ont enfin été solubilisés dans environ 20 μ L d'eau stérile traité au DMPC. La qualité et quantité des ARNs ont été mesurées soit par électrophorèse capillaire (Experion, Bio-Rad) pour les expériences de RT-qPCR, soit à l'aide d'un Nanodrop pour la RT-PCR et les clonages. La synthèse des ADNc a été réalisée avec le kit High Capacity RNA-to-cDNA (Applied Biosystems) selon les indications du fournisseur.

L'analyse en RT-qPCR digitale a été réalisée par le système Biomark HD en utilisant des puces 96X96 Fluidigm comme cela a été décrit dans le chapitre matériel et méthodes de l'article Le Roy et al. (2017) présenté dans le chapitre I de la partie "résultats". La partie expérimentale de cette approche a été sous-traitée à la société IntegraGen. Les données ainsi obtenues ont été traitées manuellement dans le logiciel Excel (Microsoft) ou à l'aide du logiciel qBase+ (Biogazelle). Les normalisations ont été réalisées avec les gènes de référence *LusETIF5A1*, *LusUBI1* et *LusEF1A* (Huis et al., 2010).

Les analyses RT-PCR semi quantitatives chez *Arabidopsis thaliana* ont été réalisées avec l'enzyme AmpliTaq[™] (Invitrogen[™]). Une dénaturation initiale a été faite à 94°C pendant 2 min puis 25 cycles de dénaturation/hybridation/polymérisation ont été conduits (94°C - 30 sec ; 55°C - 30 sec ; 72°C - 30 sec) et enfin une polymérisation finale à 72°C a été effectuée pendant 5 min. Le gène de référence TUB1 (AT1G75780) a été utilisé pour la normalisation (Dekkers et al., 2012) (Tableau 12).

Tableau 12: Liste des amorces utilisées pour l'ensemble du travail de thèse. Les adaptateurs pour les clonages dans le système Gateway sont en rouge. Les nucléotides changés pour la mutagénèse dirigée sont également en rouge.

		1	
Aplication	Gène(s)	Nom de l'amorce	Séquence nucléotidique
Identification des mutants d'insertions homozygotes	AtUGT72E1 (At3g50740)	E1-1-F	TAACTCGTATCGATGATGGCC
		E1-1-R	TGGGTCAAGGAATGTTTCAAG
		08409	ATATTGACCATCATACTCATTGC
	AtUGT72E2 (At5g66690)	E2-F	CCCTTCCGGTAGATACTCTGG
		E2-R	TTTTGTCTCAATACTTAAATTTGCC
		Spm32	TACGAATAAGAGCGTCCATTTTAGAGTGA
	AtUGT72E3 (At5g26310)	E3-F	ACCTAGAAATCGCCTCCTTTG
		E3-R	TGATCGAGCTAGCTAAGCGTC
		LB1	GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC
RT-PCR semi quantitative	AtUGT72E1 (At3g50740)	E1-1-RT-R	TGGGTCAAGGAATGTTTCAAG
		E1-1-RT-F	CCGGACGAGTCGGTACTTTA
	AtUGT72E2 (At5g66690)	E2-RT-F	TATCTCCTTCGGGAGTGGTG
		E2-RT-R	AAACCGTTGACACTCCTTGG
	AtUGT72E3 (At5g26310)	E3-RT-F	GTGGTGACCAAGATCGGAGT
		E3-RT-R	GAACCGACTCGTTTGGTTGT
	AtUGT72B1 (At4g01070)	B1-RT-F	CCTCACCGTGACTCGTTCAA
		B1-RT-R	TGTGGAGCCCAAAAAGGGAT
	Tubuline-1 (At1g75780)	TUB-F	CTCACAGTCCCGGAGCTGACAC
		TUB-R	GCTTCAGTGAACTCCATCTCGT
Clonage des ORF complet (Adapté pour le système Gateway)	LusUGT72Q1 (Lus10003944)	attB1C_LusUGT72Q1_F	GGGGACAAGTTTGTACAAAAAGCAGGCTCCATGTCCAAATTCCATGTG
		attB2C_LusUGT72Q1_R	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCACATTAATTTCCAGCCTCGA
	LusUGT72N1 (Lus10024035)	attB1C_LusUGT72N1_F	GGGGACAAGTTTGTACAAAAAGCAGGCTCCATGCAATCCTCCGGCGCCGGAAGA
		attB2C_LusUGT72N1_R	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAAGCACCTTGAGCTTTAGCGTTGA
	LusUGT72R2 (Lus10024037)	attB1C_LusUGT72R2_F	GGGGACAAGTTTGTACAAAAAGCAGGCTCCATGGAAACGTCAACCGTCGACGTCA
		attB2C_LusUGT72R2_R	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAAAATTTGGTGGCATAACGACAAG
	LusUGT72R1 (Lus10041713)	attB1C_LusUGT72R1_F	GGGGACAAGTTTGTACAAAAAGCAGGCTCCATGGAAACATCAACCGTCGACGTC
		attB2C_LusUGT72R1_R	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAAAATTTGGTGGCATAACGACAAG
	LusUGT72N2 (Lus10041715)	attB1C_LusUGT72N2_F	GGGGACAAGTTTGTACAAAAAGCAGGCTCCATGCAATCCGCCGGCGCCGGAAGAC
		attB2C_LusUGT72N2_R	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTAAGCACCTTGAGCTTTGGCCTTGAC
	AtUGT72E3 (At5g26310)	attB1C_AtUGT72E3_F	GGGGACAAGTTTGTACAAAAAGCAGGCTCCATGCATATCACAAAACCACAC
		attB2C_AtUGT72E3_R	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTAAGCACCACGTCC
	AtUGT72E2 (At5g66690)	attB1C_AtUGT72E2_F	GGGGACAAGTTTGTACAAAAAGCAGGCTCCATGCATATCACAAAACCACAC
		attB2C_AtUGT72E2_R	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTAAGCACCACGTGA
	AtUGT72E1 (At3g50740)	attB1C_AtUGT72E1_F	GGGGACAAGTTTGTACAAAAAGCAGGCTCCATGAAGATTACAAAACCACATGT
		attB2C_AtUGT72E1_R	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTAGGCACCACGTGCCAT
Vérification de la présence du gène par PCR	LusUGT72Q1 (Lus10003944)	VLus10003944_F	ACTATACGCGGAGCAGAGGA
		VLus10003944_R	CACCAACCTCACCACCTTCT
	LusUGT72N1 (Lus10024035)	VLus10024035_F	ACATATCCGCACACGTTGAA
		VLus10024035_R	GCTTCTCGCTCCTTAGAGCA
	LusUGT72R1 et R2 (Lus10041713	VLus10041713&37_F	ACCAGACTGAACCGTTGACC
	et Lus10024037)	VLus10041713&37_R	TAGCTGGTGGGTTGATAGGG
	LusUGT72N2 (Lus10041715)	VLus10041715_F	GCTCTAAGGAGCGAGAAGCA
		VLus10041715_R	CCTTTGTCCTGGCAATGAAT
	AtUGT72E3 (At5g26310)	VAT5G26310.1_F	GTGGTGACCAAGATCGGAGT
		VAT5G26310.1_R	GAACCGACTCGTTTGGTTGT
	AtUGT72E2 (At5g66690)	VAT5G66690.1_F	TATCTCCTTCGGGAGTGGTG
		VAT5G66690.1_R	AAACCGTTGACACTCCTTGG
	AtUGT72E1 (At3g50740)	VAT3G50740.1_F	CCGGACGAGTCGGTACTTTA
		VAT3G50740.1_R	ACTCCCTCCGACGGTAGTTT
Mutagénése dirigée	LusUGT72N2 (Lus10041715)	H183D-LusUGT72N2-F	GGTTCGAGGACACGCTCGACGCGTACCTCGATTATGG
		H183D-LusUGT72N2-R	CCATAATCGAGGTACGCGTCGAGCGTGTCCTCGAACC
		no4A-LUSUG1/ZNZ-F	
		H84A-LusUGT72N2-R	JACGATGAGCTTGGTGACGAATGCGTCGGAGGGTTCGACG

3/- Clonages moléculaires

Les différents gènes étudiés dans ce travail ont été clonés dans des vecteurs adaptés à la technologie Gateway[™] (Invitrogen[™]) en utilisant les enzymes de recombinaison LR et BP clonase[™] selon le protocole du fournisseur. Dans un premier temps, ils ont été clonés dans le vecteur donneur pDONR221 (Figure 78) en utilisant des amorces spécifiques (Tableau 12) contenant les adaptateurs Gateway. Grâce au système de recombinaison, ces fragments ont été intégrés dans deux vecteurs de destination : pDEST17 pour la production de protéines recombinantes et pK2GW7 pour la surexpression des gènes dans la plante (Figure 78).



Figure 78: Les différents vecteurs utilisés pour cette étude basé sur le système Gateway. pDONR221 est le vecteur d'entrée et pK2GW7 et pDEST17 ont été utilisé respectivement pour la sur-expression in planta ainsi que pour la production des protéines recombinante.

4/- Mutagénèse dirigée

Les mutations spécifiques des nucléotiques codant les deux histidines ont été réalisées directement sur le vecteur d'expression pDEST17 à l'aide du protocole décrit par Laible and Boonrod (2009). Des amorces comportant les mutations ont été utilisées pour amplifier le vecteur entier (Tableau 12).

III- Expression hétérologue et essais enzymatiques

1/- Expression et purification des protéines recombinantes

Le vecteur pDEST17 permet de produire une protéine recombinante avec une étiquette de six histidines au niveau de la partie N-terminale. Les vecteurs contenant les gènes *UGTs* ont été introduits chez *Escherichia coli* BL21-CodonPlus-(DE3)-RIL par

électroporation. L'expression a été induite par l'ajout de 0,1 ou 0,5 mM d'isopropyl β-D-1-thiogalactopyranoside (IPTG en concentration finale) dans 200 mL de milieu LB, et les cultures ont été agitées pendant une nuit à 21°C. Les bactéries ont été récoltées par une centrifugation de 20 minutes à 4000g à 4°C et remises en suspension dans une solution d'extraction (50 mM NaH₂PO₄, 500 mM NaCl, 20 mM imidazole pH 7.4) contenant 1 mg/mL de lysozyme. La suspension a alors été incubée dans la glace pendant 30 min. Après deux passages successifs à la presse de French (1250 bars), la suspension a été centrifugée à 10000g pendant 25 min (4°C). Le surnageant a alors été chargé dans une colonne HisTrap HP (GE Healthcare) puis une première élution réalisée avec une solution contenant de l'imidazole à une concentration de 50 mM puis une seconde à 250 mM. Les différentes fractions recueillies ont été déposées sur un gel SDS-PAGE pour vérifier le niveau d'accumulation de la protéine recherchée. Les fractions d'intérêt ont alors été dessalées à l'aide d'une colonne PD-10 (Pharmacia) pour enlever l'imidazole puis les protéines concentrées à l'aide de colonnes Vivaspin (GE Healthcare) avec une coupure à 30 kDa. Les protéines ainsi purifiées ont été quantifiées par un dosage de Bradford et utilisées sans attendre pour des essais enzymatiques.

2/- Electrophorèse SDS-PAGE et Western Blot

La séparation électrophorétique a été effectuée dans le système "mini-protean III" (Bio-Rad). Les protéines préalablement dosées à l'aide du réactif de Bradford ont été dénaturées à 95°C pendant 5 min dans du tampon de charge (1/5 ; volume/volume) (40 % Glycérol ; 240 mM Tris/HCl pH 6.8 ; 8 % SDS ; 0.04 % bromophenol blue ; 5 % betamercaptoethanol). Puis ce mélange a été chargé dans un gel de concentration de composition suivante : 4 % d'acrylamide, 1 % de bisacrylamide, 0,125 M Tris-HCl (pH 6,8) et 0,1 % SDS. La polymérisation était assurée par une solution à 0,03 % de persulfate d'ammonium (APS) et 0,05 % de N, N, N', N' - tetramethylethylènediamine (TEMED) et a duré environ 45 min. Ce gel a été coulé au dessus d'un gel de résolution contenant 12,5 % d'acrylamide, 0,3 % de bisacrylamide, 0,375 M Tris- HCl (pH 8,8) et 0,1 % SDS. La polymérisation a été effectuée dans les mêmes conditions que pour le gel de concentration. Le tampon d'électrophorèse était composé de 0,025 M Tris, 0,192 M glycine, 0,1 % SDS (pH 8,2-8,3) et la migration a eu lieu à raison de 20 mA par gel pendant environ 45 min. A la suite de quoi, les gels ont subi un bain de fixation dans 15 % d'acide acétique et 20 % d'éthanol, puis de coloration dans la même solution contenant 0,25 % de bleu de Coomassie et enfin de nouveau dans la solution de fixation, chaque bain étant de 20 min.

Les protéines non colorées ont également pu être transférées sur membrane de nitrocellulose (Ø 0,45 µm, Schleicher & Schuell) à l'aide du système "Mini TRANS BLOT" (Bio Rad) dans le tampon 0,025 M Tris, 0,192 M glycine, 20 % éthanol (pH 8,5). Le transfert a été réalisé à 40 V pendant 80 minutes puis la membrane a été incubée dans un tampon bloquant (5% de lait en poudre ; TBS pH=8,5 ; 0,1% de Tween 20) à 4°C toute une nuit. Une fois ce blocage réalisé, les membranes ont ensuite été incubées dans une solution d'anticorps anti-HIS tag conjugués à une peroxydase (Sigma), dilué dans le tampon bloquant à 1/1000^{éme}, pendant 1h à température ambiante. Après 3 lavages successifs de 20 minutes, la révélation a été faite par chimioluminescence à l'aide du kit Pierce[™] ECL.

3/- Essais enzymatiques sur les protéine recombinantes

Pour déterminer les spécificités de substrat des différentes protéines recombinantes, des essais enzymatiques ont été conduits. Le mélange réactionnel de 100 µL contenait 2 ug de protéines totales, 50 mM de Tris-HCl pH=8, 5 mM d'UDPglucose (Merck; réf 670120) et 1 mM de substrat accepteur qui pouvait être l'acide cinnamique (Merck; réf 8.00235.0005), l'acide p-coumarique (Sigma; réf : C9008), l'acide férulique (Sigma ; réf F3500), l'acide sinapique (Sigma ; réf D7927), l'aldehyde pcoumarylique (TransMit ; réf C034), l'aldehyde coniférylique (Sigma ; réf 382051), l'aldehyde sinapylique (Apin chemical ; réf 09444s), l'alcool *p*-coumarylique (TransMit ; réf C019), l'alcool coniférylique (Sigma ; réf 223735), l'alcool sinapylique (Sigma ; réf 404586), l'acide salicylique (Sigma ; réf S3007), l'acide gibbérellique (Sigma ; réf G7645) ou la coumarine (Serva ; réf 17610). La réaction a été stoppée au bout d'une heure par un ajout de 50 µL de méthanol et les mélanges réactionnels stockés à -20°C. Les mélanges ont ensuite été passés sur des filtres de porosité 0,45 µm (Pall GHP, VWR) puis injectés en HPLC-UV sur une colonne C18. La séparation des molécules a été réalisée par des mélanges de solvant A (eau) puis de solvant B (acétonitrile) contenant tous les deux 0,1% d'acide ortho-phosphorique. Le débit de la phase mobile était de 1,1 mL/min et la température fixée à 45°C avec les conditions de départ de 10% de solvant B. Après l'injection, un gradient a été appliqué durant les 7 min pour arriver à 30% de solvant B. Il a été suivi d'un deuxième gradient de 3 min pour arriver à 70% et maintenu durant 1,5

min. Enfin, la concentration en acétonitrile a été abaissée à 10% en 1 min puis une rééquilibration isocratique a été appliquée pendant 8,5 min

4/- Essais enzymatiques sur des extraits protéiques de plantes

Dans le but de caractériser les activités UGTs globales de différents organes, une extraction douce de protéines a été réalisée. Pour cela, 2g de matériel frais ont été broyés avec un mortier et un pilon dans l'azote liquide et la poudre ainsi obtenue transvasée dans un tube de 15 mL contenant 3 mL de tampon d'extraction (25 mM Tris-HCl pH=6,5 ; 10% de glycérol ; 20 mM β -mercatoethanol, 1 mM PMSF, 10 % polyvinyl polypyrrolidone). Les tubes ont été agités vigoureusement puis centrifugés à 10500 g pendant 20 min à 4°C. La totalité du surnageant a de nouveau été centrifugée mais cette fois dans des microtubes à 15000g pendant 20 min (4°C) dans le but d'éliminer les derniers débris cellulaires. L'extrait protéique a été dessalé sur une colonne PD-10 (Pharmacia) puis concentré sur une colonne Vivaspin (GE Healthcare) avec une coupure à 30 kDa. Les protéines ont alors été quantifiées par un dosage de Bradford et, sans attendre, les essais enzymatiques réalisés.

Le milieu réactionnel permettant de faire les réactions enzymatiques était composé de 50 μ L d'extraits protéiques; 50 mM d'un tampon Tris-HCl pH=8; 5 mM d'UDP-glucose et 1 mM de substrat accepteur. Les mélanges ont été incubés pendant 2 h à 30°C et la réaction stoppée par l'ajout de méthanol (volume correspondant à la moitié du volume total). Les mélanges ont ensuite été passés sur des filtres de porosité 0,45 μ m (Pall GHP, VWR) puis injectés en HPLC-UV sur une colonne C18 comme décrit ci-dessus.

IV- Phénotypage

1/- Extraction des résidu pariétaux et dosage de lignines

Le protocole d'extraction des résidus pariétaux a été adapté de Mansfield et al. (2012). Les fragments de tiges ont été prélevés et séchés en étuve sous un flux d'air à 30°C pendant 48 h. Les fragments de 3 individus ont été groupés ensemble puis passés dans un broyeur à billes (Retsch PM200 ; godet 50 mL ZrO₂ et bille en ZrO₂). La totalité de la poudre a alors été reprise dans 30 mL de solution NaCl 50 mM, vortexée puis placée une nuit à 4°C sous légère agitation. Les résidus pariétaux ont été récupérés après une centrifugation de 10 min à 2800g (4°C), puis remis en suspension dans 40 mL d'éthanol 80 % et enfin traités dans un bain à ultrasons pendant 20 min. Ce lavage à

l'éthanol a été répété encore 2 fois puis suivi d'un lavage dans l'acétone 100%, d'un autre avec un mélange CHCl₃/méthanol (1/1 vol/vol) puis de nouveau dans l'acétone 100 %. Le culot a ensuite été incubé à 50°C pendant 24 h. Les résidus pariétaux ont été repris par 25 mL de tampon Tris-acétate pH=6, vortexés puis placés dans un bain marie à 90°C pendant 2h. Après un retour à 55°C, 20 et 40 unités d'amylases et d'amyloglucosidases respectivement ont été ajoutées dans le mélange qui a été incubé 2h à 55°C. Après cela, la suspension a été centrifugée à 2800g pendant 10 min (4°C) et le culot repris dans de l'eau ultrapure. Ce lavage à l'eau a été répété deux fois puis trois lavages à l'acétone 100% ont été réalisés. Les résidus pariétaux ainsi dé-protéinés et déamidonnés ont été séchés puis pesés de la manière la plus précise possible avant le dosage des lignines. Chaque échantillon a ensuite été repris dans 5 mL d'un mélange de bromure d'acétyle/acide acétique (25/75 m/m) puis placés au bain marie à 70°C pendant 30 min (avec un passage au vortex toutes les 10 min) et enfin conservés dans la glace pendant maximum 2 h. Dans un nouveau tube en verre, 2 mL de la solution ont été mélangés à 5 mL de NaOH 2 M et 13 mL d'acide acétique glacial et l'ensemble laissé au repos pendant 30 min à l'obscurité. La quantité de lignines/quantité de résidus pariétaux a été estimé après mesure au spectrophotomètre par la mesure de l'absorbance à 280 nm ($\varepsilon = 20 \text{ g}^{-1}$. l. cm⁻¹).

2/- Microscopie et préparation des échantillons

Des coupes transversales de hampes florales et de racines ont été effectuées afin de comparer les profils de lignification des mutants d'*A. thaliana*. Les échantillons ont tout d'abord été prélevés et conservés dans l'éthanol 70°. Pour inclure les fragments dans la résine Technovit 7100 (Kulzer), des bains successifs croissants en éthanol (80°, 90°, 96° et absolue) de 20 min ont été conduits. Ensuite l'inclusion a été effectuée selon le protocole du fournisseur. Des coupes de 40 µm ont été obtenues en utilisant un microtome Leica RM2065 puis elles ont été colorées par une solution contenant 0,1% de toluidine blue 0 (TBO) pour avoir une vision globale des différents types cellulaires (notamment pectines en mauve et lignines en bleu). Une coloration des lignines a été réalisée par le dépôt d'une goutte d'une solution de 2% de phloroglucinol dans 20% d'éthanol sur la préparation. Après évaporation, une goutte de l'acide chlorhydrique concentrée (12N) est ajoutée sur la coupe. Les lignines apparaissent ainsi en rouge. Pour ces deux colorations, l'observation a été faite rapidement et les lames ont été conservées à l'obscurité.

3/- Infection par Botrytis cinerea

Afin d'évaluer la résistance des plante au pathogène *B. cinerea*, 6 μl de conidies en suspension (à 10⁴ conidies/ mL) ont été déposés sur les feuilles des rosettes de plantes âgées de 5 semaines. Trois jours suivant cette infection, la taille de lésions a été mesurée.

V- Analyses in silico

1/- Analyse phylogénétique

L'arbre phylogénétique regroupant l'ensemble des UGTs de *Linum usitatissimum* et d'autres espèces a été généré par le logiciel CLC Genomics Workbench. L'alignement des séquences protéiques a été réalisé par MUSCLE (MUltiple Sequence Comparison by Log- Expectation) avec les pénalités par défaut. A partir de l'alignement, l'arbre phylogénique a été construit sur la base de l'algorithme neighbor joining et la mesure des distances effectuée par la méthode de Jukes-Cantor. Par la suite, une analyse statistique par la méthode de bootstrap a été réalisée (100 réplications).

2/- Tests statistiques

Pour savoir s'il y a des différences significatives entre deux échantillons, deux tests statistiques ont été réalisés suivant la nature de l'échantillon à l'aide du logiciel R. Lorsque l'échantillonnage était supérieur à 6 individus, le test t de Student a été réalisé. Lorsque le nombre d'échantillons était inférieur à 6 individus, le test non paramétrique de Kruskal Wallis a été préféré.

3/- Analyse des racines

La morphologie des racines de plantes cultivées *in vitro* a été évaluée après une prise de photos régulières et une analyse par le logiciel Fiji (Schindelin et al., 2012).

Bibliographie

Abrahams, S., Tanner, G.J., Larkin, P.J., and Ashton, A.R. (2002). Identification and biochemical characterization of mutants in the proanthocyanidin pathway in Arabidopsis. Plant Physiol. *130*, 561–576.

Achnine, L., Blancaflor, E.B., Rasmussen, S., and Dixon, R.A. (2004). Colocalization of L-phenylalanine ammonia-lyase and cinnamate 4-hydroxylase for metabolic channeling in phenylpropanoid biosynthesis. Plant Cell *16*, 3098–3109.

Adler, E., Björkqvist, K.J., and Häggroth, S. (1949). Über die Ursache der Farbreaktionen des Holzes (Svenska Träforskningsinstitutet).

Ageeva, M.V., Petrovská, B., Kieft, H., Sal'nikov, V.V., Snegireva, A.V., van Dam, J.E.G., van Veenendaal, W.L.H., Emons, A.M.C., Gorshkova, T.A., and van Lammeren, A. a. M. (2005). Intrusive growth of flax phloem fibers is of intercalary type. Planta *222*, 565–574.

Ageeva, M.V., Chernova, T.E., and Gorshkova, T.A. (2012). Processes of protoplast senescence and death in flax fibers: An ultrastructural analysis. Russ. J. Dev. Biol. *43*, 94–100.

Alejandro, S., Lee, Y., Tohge, T., Sudre, D., Osorio, S., Park, J., Bovet, L., Lee, Y., Geldner, N., Fernie, A.R., et al. (2012). AtABCG29 is a monolignol transporter involved in lignin biosynthesis. Curr. Biol. CB *22*, 1207–1212.

Allwood, E.G., Davies, D.R., Gerrish, C., Ellis, B.E., and Bolwell, G.P. (1999). Phosphorylation of phenylalanine ammonia-lyase: evidence for a novel protein kinase and identification of the phosphorylated residue. FEBS Lett. *457*, 47–52.

Amor, Y., Haigler, C.H., Johnson, S., Wainscott, M., and Delmer, D.P. (1995). A membrane-associated form of sucrose synthase and its potential role in synthesis of cellulose and callose in plants. Proc. Natl. Acad. Sci. U. S. A. *92*, 9353–9357.

Anterola, A.M., and Lewis, N.G. (2002). Trends in lignin modification: a comprehensive analysis of the effects of genetic manipulations/mutations on lignification and vascular integrity. Phytochemistry *61*, 221–294.

Aoki, D., Hanaya, Y., Akita, T., Matsushita, Y., Yoshida, M., Kuroda, K., Yagami, S., Takama, R., and Fukushima, K. (2016). Distribution of coniferin in freeze-fixed stem of Ginkgo biloba L. by cryo-TOF-SIMS/SEM. Sci. Rep. *6*, 31525.

Appelhagen, I., Thiedig, K., Nordholt, N., Schmidt, N., Huep, G., Sagasser, M., and Weisshaar, B. (2014). Update on transparent testa mutants from Arabidopsis thaliana: characterisation of new alleles from an isogenic collection. Planta *240*, 955–970.

Ariizumi, T., and Toriyama, K. (2011). Genetic regulation of sporopollenin synthesis and pollen exine development. Annu. Rev. Plant Biol. *62*, 437–460.

Arimura, G., and Maffei, M. (2016). Plant Specialized Metabolism: Genomics, Biochemistry, and Biological Functions (CRC Press).

Attoumbré, J., Bienaimé, C., Dubois, F., Fliniaux, M.-A., Chabbert, B., and Baltora-Rosset, S. (2010). Development of antibodies against secoisolariciresinol--application to the immunolocalization of lignans in Linum usitatissimum seeds. Phytochemistry *71*, 1979–1987.

Aung, B., Gruber, M.Y., Amyot, L., Omari, K., Bertrand, A., and Hannoufa, A. (2015). MicroRNA156 as a promising tool for alfalfa improvement. Plant Biotechnol. J. *13*, 779–790.

de Azevedo Souza, C., Kim, S.S., Koch, S., Kienow, L., Schneider, K., McKim, S.M., Haughn, G.W., Kombrink, E., and Douglas, C.J. (2009). A Novel Fatty Acyl-CoA Synthetase Is Required for Pollen Development and Sporopollenin Biosynthesis in Arabidopsis. Plant Cell *21*, 507–525.

Baley, C. (2002). Analysis of the flax fibres tensile behaviour and analysis of the tensile stiffness increase. Compos. Part Appl. Sci. Manuf. *33*, 939–948.

Barros, J., Serk, H., Granlund, I., and Pesquet, E. (2015). The cell biology of lignification in higher plants. Ann. Bot. mcv046.

Barros, J., Serrani-Yarce, J.C., Chen, F., Baxter, D., Venables, B.J., and Dixon, R.A. (2016). Role of bifunctional

ammonia-lyase in grass cell wall biosynthesis. Nat. Plants 2, 16050.

Bartley, L.E., Peck, M.L., Kim, S.-R., Ebert, B., Manisseri, C., Chiniquy, D.M., Sykes, R., Gao, L., Rautengarten, C., Vega-Sánchez, M.E., et al. (2013). Overexpression of a BAHD Acyltransferase, OsAt10, Alters Rice Cell Wall Hydroxycinnamic Acid Content and Saccharification. Plant Physiol. *161*, 1615–1633.

Barvkar, V.T., Pardeshi, V.C., Kale, S.M., Kadoo, N.Y., and Gupta, V.S. (2012). Phylogenomic analysis of UDP glycosyltransferase 1 multigene family in Linum usitatissimum identified genes with varied expression patterns. BMC Genomics *13*, 175.

Bassard, J.-E., Richert, L., Geerinck, J., Renault, H., Duval, F., Ullmann, P., Schmitt, M., Meyer, E., Mutterer, J., Boerjan, W., et al. (2012). Protein–Protein and Protein–Membrane Associations in the Lignin Pathway. Plant Cell *24*, 4465–4482.

Bassard, J.-E., Møller, B.L., and Laursen, T. (2017). Assembly of Dynamic P450-Mediated Metabolons-Order Versus Chaos. Curr. Mol. Biol. Rep. *3*, 37–51.

Baxter, H.L., and Stewart, C.N. (2013). Effects of altered lignin biosynthesis on phenylpropanoid metabolism and plant stress. Biofuels *4*, 635–650.

Bayburt, T.H., and Sligar, S.G. (2002). Single-molecule height measurements on microsomal cytochrome P450 in nanometer-scale phospholipid bilayer disks. Proc. Natl. Acad. Sci. U. S. A. 99, 6725–6730.

Bendtsen, J.D., Jensen, L.J., Blom, N., Von Heijne, G., and Brunak, S. (2004). Feature-based prediction of nonclassical and leaderless protein secretion. Protein Eng. Des. Sel. PEDS *17*, 349–356.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res. *28*, 235–242.

Bernards, M.A. (2002). Demystifying suberin. Can. J. Bot. 80, 227-240.

Berthet, S., Demont-Caulet, N., Pollet, B., Bidzinski, P., Cézard, L., Le Bris, P., Borrega, N., Hervé, J., Blondet, E.,

Balzergue, S., et al. (2011). Disruption of LACCASE4 and 17 results in tissue-specific alterations to lignification of Arabidopsis thaliana stems. Plant Cell *23*, 1124–1137.

Bhuiyan, N.H., Selvaraj, G., Wei, Y., and King, J. (2009). Role of lignification in plant defense. Plant Signal. Behav. 4, 158–159.

Bock, K.W. (2016). The UDP-glycosyltransferase (UGT) superfamily expressed in humans, insects and plants: Animal-plant arms-race and co-evolution. Biochem. Pharmacol. *99*, 11–17.

Boerjan, W., Ralph, J., and Baucher, M. (2003). Lignin Biosynthesis. Annu. Rev. Plant Biol. 54, 519–546.

Boija, E., and Johansson, G. (2006). Interactions between model membranes and lignin-related compounds studied by immobilized liposome chromatography. Biochim. Biophys. Acta BBA - Biomembr. *1758*, 620–626.

Boija, E., Lundquist, A., Edwards, K., and Johansson, G. (2007). Evaluation of bilayer disks as plant cell membrane models in partition studies. Anal. Biochem. *364*, 145–152.

Bonawitz, N.D., Soltau, W.L., Blatchley, M.R., Powers, B.L., Hurlock, A.K., Seals, L.A., Weng, J.-K., Stout, J., and Chapple, C. (2012). REF4 and RFR1, Subunits of the Transcriptional Coregulatory Complex Mediator, Are Required for Phenylpropanoid Homeostasis in Arabidopsis. J. Biol. Chem. *287*, 5434–5445.

Bonawitz, N.D., Kim, J.I., Tobimatsu, Y., Ciesielski, P.N., Anderson, N.A., Ximenes, E., Maeda, J., Ralph, J., Donohoe, B.S., Ladisch, M., et al. (2014). Disruption of Mediator rescues the stunted growth of a lignin-deficient Arabidopsis mutant. Nature *509*, 376–380.

Bourgaud, F., Gravot, A., Milesi, S., and Gontier, E. (2001). Production of plant secondary metabolites: a historical perspective. Plant Sci. *161*, 839–851.

Bourgaud, F., Hehn, A., Larbat, R., Doerper, S., Gontier, E., Kellner, S., and Matern, U. (2006). Biosynthesis of coumarins in plants: a major pathway still to be unravelled for cytochrome P450 enzymes. Phytochem. Rev. *5*, 293–308.

Bowers, J.E., Chapman, B.A., Rong, J., and Paterson, A.H. (2003). Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature *422*, 433–438.

Boyes, D.C., Zayed, A.M., Ascenzi, R., McCaskill, A.J., Hoffman, N.E., Davis, K.R., and Görlach, J. (2001). Growth Stage–Based Phenotypic Analysis of Arabidopsis A Model for High Throughput Functional Genomics in Plants. Plant Cell *13*, 1499–1510.

Brazier-Hicks, M., and Edwards, R. (2005). Functional importance of the family 1 glucosyltransferase UGT72B1 in the metabolism of xenobiotics in Arabidopsis thaliana. Plant J. Cell Mol. Biol. *42*, 556–566.

Brazier-Hicks, M., Offen, W.A., Gershater, M.C., Revett, T.J., Lim, E.-K., Bowles, D.J., Davies, G.J., and Edwards, R. (2007a). Characterization and engineering of the bifunctional N- and O-glucosyltransferase involved in xenobiotic metabolism in plants. Proc. Natl. Acad. Sci. U. S. A. *104*, 20238–20243.

Brazier-Hicks, M., Edwards, L.A., and Edwards, R. (2007b). Selection of plants for roles in phytoremediation: the importance of glucosylation. Plant Biotechnol. J. 5, 627–635.

Bronstein, J.L., Alarcón, R., and Geber, M. (2006). The evolution of plant-insect mutualisms. New Phytol. *172*, 412–428.

Brown, D.M., Zeef, L.A.H., Ellis, J., Goodacre, R., and Turner, S.R. (2005). Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. Plant Cell *17*, 2281–2295.

Buanafina, M.M. (2009). Feruloylation in grasses: current and future perspectives. Mol. Plant 2, 861–872.

Buchanan-Wollaston, V., Page, T., Harrison, E., Breeze, E., Lim, P.O., Nam, H.G., Lin, J.-F., Wu, S.-H., Swidzinski, J., Ishizaki, K., et al. (2005). Comparative transcriptome analysis reveals significant differences in gene expression and signalling pathways between developmental and dark/starvation-induced senescence in Arabidopsis. Plant J. Cell Mol. Biol. *42*, 567–585.

Buer, C.S., Wasteneys, G.O., and Masle, J. (2003). Ethylene Modulates Root-Wave Responses in Arabidopsis. Plant Physiol. *132*, 1085–1096.

Burr, S.J., and Fry, S. (2009). Feruloylated Arabinoxylans Are Oxidatively Cross-Linked by Extracellular Maize Peroxidase but Not by Horseradish Peroxidase. Mol. Plant *2*, 883–892.

Campbell, J.A., Davies, G.J., Bulone, V., and Henrissat, B. (1997). A classification of nucleotide-diphosphosugar glycosyltransferases based on amino acid sequence similarities. Biochem. J. *326*, 929–939.

Campos, L., Lisón, P., López-Gresa, M.P., Rodrigo, I., Zacarés, L., Conejero, V., and Bellés, J.M. (2014). Transgenic tomato plants overexpressing tyramine N-hydroxycinnamoyltransferase exhibit elevated hydroxycinnamic acid amide levels and enhanced resistance to Pseudomonas syringae. Mol. Plant-Microbe Interact. MPMI *27*, 1159–1169.

Cao, H., Ekstrom, A., and Yin, Y. (2015). Plant Carbohydrate Active Enzyme (CAZyme) Repertoires: A Comparative Study. In Advances in the Understanding of Biological Sciences Using Next Generation Sequencing (NGS) Approaches, G. Sablok, S. Kumar, S. Ueno, J. Kuo, and C. Varotto, eds. (Springer International Publishing), pp. 115–134.

Caputi, L., Malnoy, M., Goremykin, V., Nikiforova, S., and Martens, S. (2012). A genome-wide phylogenetic reconstruction of family 1 UDP-glycosyltransferases revealed the expansion of the family during the adaptation of plants to life on land. Plant J. Cell Mol. Biol. *69*, 1030–1042.

Carnachan, null, and Harris, null (2000). Ferulic acid is bound to the primary cell walls of all gymnosperm families. Biochem. Syst. Ecol. *28*, 865–879.

Carocha, V., Soler, M., Hefer, C., Cassan-Wang, H., Fevereiro, P., Myburg, A.A., Paiva, J.A.P., and Grima-Pettenati, J. (2015). Genome-wide analysis of the lignin toolbox of Eucalyptus grandis. New Phytol. *206*, 1297–1313.

Cartwright, A.M., Lim, E.-K., Kleanthous, C., and Bowles, D.J. (2008). A Kinetic Analysis of Regiospecific Glucosylation by Two Glycosyltransferases of Arabidopsis thaliana. J. Biol. Chem. *283*, 15724–15731.

Chantreau, M., Grec, S., Gutierrez, L., Dalmais, M., Pineau, C., Demailly, H., Paysant-Leroux, C., Tavernier, R., Trouvé, J.-P., Chatterjee, M., et al. (2013). PT-Flax (phenotyping and TILLinG of flax): development of a flax (Linum usitatissimum L.) mutant population and TILLinG platform for forward and reverse genetics. BMC Plant Biol. *13*, 159.

Chantreau, M., Portelette, A., Dauwe, R., Kiyoto, S., Crônier, D., Morreel, K., Arribat, S., Neutelings, G., Chabi, M., Boerjan, W., et al. (2014). Ectopic Lignification in the Flax lignified bast fiber1 Mutant Stem Is Associated with Tissue-Specific Modifications in Gene Expression and Cell Wall Composition[C][W]. Plant Cell *26*, 4462–4482.

Chapelle, A. (2009). Caractérisation de gènes de ß-glucisidase et d'UDP-glycosyltransférase potentiellement impliqués dans la lignification chez arabidopsis thaliana (Paris 11).

Chapelle, A., Morreel, K., Vanholme, R., Le-Bris, P., Morin, H., Lapierre, C., Boerjan, W., Jouanin, L., and Demont-Caulet, N. (2012). Impact of the Absence of Stem-Specific β -Glucosidases on Lignin and Monolignols. Plant Physiol. *160*, 1204–1217.

Chauhan, J.S., Rao, A., and Raghava, G.P.S. (2013). In silico Platform for Prediction of N-, O- and C-Glycosites in Eukaryotic Protein Sequences. PLOS ONE *8*, e67008.

Chen, F., Tobimatsu, Y., Havkin-Frenkel, D., Dixon, R.A., and Ralph, J. (2012). A polymer of caffeyl alcohol in plant seeds. Proc. Natl. Acad. Sci. *109*, 1772–1777.

Chen, F., Tobimatsu, Y., Jackson, L., Nakashima, J., Ralph, J., and Dixon, R.A. (2013). Novel seed coat lignins in the Cactaceae: structure, distribution and implications for the evolution of lignin diversity. Plant J. 73, 201–211.

Chen, G., Komatsuda, T., Pourkheirandish, M., Sameri, M., Sato, K., Krugman, T., Fahima, T., Korol, A.B., and Nevo, E. (2009a). Mapping of the eibi1 gene responsible for the drought hypersensitive cuticle in wild barley (Hordeum spontaneum). Breed. Sci. *59*, 21–26.

Chen, H.-C., Li, Q., Shuford, C.M., Liu, J., Muddiman, D.C., Sederoff, R.R., and Chiang, V.L. (2011). Membrane protein complexes catalyze both 4- and 3-hydroxylation of cinnamic acid derivatives in monolignol biosynthesis. Proc. Natl. Acad. Sci. *108*, 21253–21258.

Chen, Z., Zheng, Z., Huang, J., Lai, Z., and Fan, B. (2009b). Biosynthesis of salicylic acid in plants. Plant Signal. Behav. *4*, 493–496.

Cheng, H., Deng, W., Wang, Y., Ren, J., Liu, Z., and Xue, Y. (2014). dbPPT: a comprehensive database of protein phosphorylation in plants. Database J. Biol. Databases Curation *2014*, bau121.

Chernova, T.E., and Gorshkova, T.A. (2007). Biogenesis of plant fibers. Russ. J. Dev. Biol. 38, 221–232.

Chong, J., Poutaraud, A., and Hugueney, P. (2009). Metabolism and roles of stilbenes in plants. Plant Sci. 177, 143–155.

Clé, C., Hill, L.M., Niggeweg, R., Martin, C.R., Guisez, Y., Prinsen, E., and Jansen, M.A.K. (2008). Modulation of chlorogenic acid biosynthesis in Solanum lycopersicum; consequences for phenolic accumulation and UV-tolerance. Phytochemistry *69*, 2149–2156.

Clemens, S., and Weber, M. (2016). The essential role of coumarin secretion for Fe acquisition from alkaline soil. Plant Signal. Behav. *11*, e1114197.

Cong, F., Diehl, B.G., Hill, J.L., Brown, N.R., and Tien, M. (2013). Covalent bond formation between amino acids and lignin: Cross-coupling between proteins and lignin. Phytochemistry *96*, 449–456.

Consortium, A.I.M. (2011). Evidence for Network Evolution in an Arabidopsis Interactome Map. Science *333*, 601–607.

Cook, D., Fowler, S., Fiehn, O., and Thomashow, M.F. (2004). A prominent role for the CBF cold response pathway in configuring the low-temperature metabolome of Arabidopsis. Proc. Natl. Acad. Sci. U. S. A. *101*, 15243–15248.

Cosio, C., and Dunand, C. (2009). Specific functions of individual class III peroxidase genes. J. Exp. Bot. 60, 391–408.

Costa, M.A., Bedgar, D.L., Moinuddin, S.G.A., Kim, K.-W., Cardenas, C.L., Cochrane, F.C., Shockey, J.M., Helms, G.L., Amakura, Y., Takahashi, H., et al. (2005). Characterization in vitro and in vivo of the putative multigene 4coumarate:CoA ligase network in Arabidopsis: syringyl lignin and sinapate/sinapyl alcohol derivative formation. Phytochemistry *66*, 2072–2091.

Courtois-Moreau, C.L., Pesquet, E., Sjödin, A., Muñiz, L., Bollhöner, B., Kaneda, M., Samuels, L., Jansson, S., and Tuominen, H. (2009). A unique program for cell death in xylem fibers of Populus stem. Plant J. *58*, 260–274.

Coutinho, P.M., Stam, M., Blanc, E., and Henrissat, B. (2003). Why are there so many carbohydrate-active enzyme-related genes in plants? Trends Plant Sci. *8*, 563–565.

Curie, C., and Mari, S. (2016). New routes for plant iron mining. New Phytol. 190, 75-87.

Dam, J.E.G. van, and Gorshkova, T.A. (2003). Plant growth and development : plant fiber formation.

Dar, A.A., and Arumugam, N. (2013). Lignans of sesame: Purification methods, biological activities and biosynthesis – A review. Bioorganic Chem. *50*, 1–10.

Dastmalchi, M., Bernards, M.A., and Dhaubhadel, S. (2016). Twin anchors of the soybean isoflavonoid metabolon: evidence for tethering of the complex to the endoplasmic reticulum by IFS and C4H. Plant J. Cell Mol. Biol. *85*, 689–706.

Dastmalchi, M., Chapman, P., Yu, J., Austin, R.S., and Dhaubhadel, S. (2017). Transcriptomic evidence for the control of soybean root isoflavonoid content by regulation of overlapping phenylpropanoid pathways. BMC Genomics *18*.

D'Auria, J.C. (2006). Acyltransferases in plants: a good time to be BAHD. Curr. Opin. Plant Biol. 9, 331-340.

Davies, K.M. (2008). Modifying Anthocyanin Production in Flowers. In Anthocyanins, C. Winefield, K. Davies, and K. Gould, eds. (Springer New York), pp. 49–80.

Davin, L.B., and Lewis, N.G. (2000). Dirigent Proteins and Dirigent Sites Explain the Mystery of Specificity of Radical Precursor Coupling in Lignan and Lignin Biosynthesis. Plant Physiol. *123*, 453–462.

Davin, L.B., Wang, H.B., Crowell, A.L., Bedgar, D.L., Martin, D.M., Sarkanen, S., and Lewis, N.G. (1997). Stereoselective bimolecular phenoxy radical coupling by an auxiliary (dirigent) protein without an active center. Science *275*, 362–366.

Day, A., Ruel, K., Neutelings, G., Crônier, D., David, H., Hawkins, S., and Chabbert, B. (2005). Lignification in the flax stem: evidence for an unusual lignin in bast fibers. Planta *222*, 234–245.

Debeaujon, I., Léon-Kloosterziel, K.M., and Koornneef, M. (2000). Influence of the testa on seed dormancy, germination, and longevity in Arabidopsis. Plant Physiol. *122*, 403–414.

Decourtil, C. (2016). Etude des relations entre lignine et lignanes via la caractérisation de mutants pinorésinol réductases d'Arabidopsis thaliana.

Dekkers, B.J.W., Willems, L., Bassel, G.W., van Bolderen-Veldkamp, R.P.M., Ligterink, W., Hilhorst, H.W.M., and Bentsink, L. (2012). Identification of reference genes for RT-qPCR expression analysis in Arabidopsis and tomato seeds. Plant Cell Physiol. *53*, 28–37.

Derbyshire, P., Ménard, D., Green, P., Saalbach, G., Buschmann, H., Lloyd, C.W., and Pesquet, E. (2015). Proteomic Analysis of Microtubule Interacting Proteins over the Course of Xylem Tracheary Element Formation in Arabidopsis. Plant Cell tpc.15.00314.

Dexter, R., Qualley, A., Kish, C.M., Ma, C.J., Koeduka, T., Nagegowda, D.A., Dudareva, N., Pichersky, E., and Clark, D. (2007). Characterization of a petunia acetyltransferase involved in the biosynthesis of the floral volatile isoeugenol. Plant J. Cell Mol. Biol. *49*, 265–275.

Dhaubhadel, S., Farhangkhoee, M., and Chapman, R. (2008). Identification and characterization of isoflavonoid specific glycosyltransferase and malonyltransferase from soybean seeds. J. Exp. Bot. *59*, 981–994.

Dilokpimol, A., Poulsen, C.P., Vereb, G., Kaneko, S., Schulz, A., and Geshi, N. (2014). Galactosyltransferases

from Arabidopsis thaliana in the biosynthesis of type II arabinogalactan: molecular interaction enhances enzyme activity. BMC Plant Biol. *14*, 90.

Dima, O., Morreel, K., Vanholme, B., Kim, H., Ralph, J., and Boerjan, W. (2015). Small Glycosylated Lignin Oligomers Are Stored in Arabidopsis Leaf Vacuoles. Plant Cell *27*, 695–710.

Dixon, R.A., Liu, C., and Jun, J.H. (2013). Metabolic engineering of anthocyanins and condensed tannins in plants. Curr. Opin. Biotechnol. *24*, 329–335.

Donaldson, L.A. (2001). Lignification and lignin topochemistry — an ultrastructural view. Phytochemistry *57*, 859–873.

Dong, T., and Hwang, I. (2014). Contribution of ABA UDP-glucosyltransferases in coordination of ABA biosynthesis and catabolism for ABA homeostasis. Plant Signal. Behav. *9*.

van Doorn, W.G., and Woltering, E.J. (2008). Physiology and molecular biology of petal senescence. J. Exp. Bot. *59*, 453–480.

Dormont, L., Delle-Vedove, R., Bessière, J.-M., and Schatz, B. (2014). Floral scent emitted by white and coloured morphs in orchids. Phytochemistry *100*, 51–59.

Egertsdotter, U., van Zyl, L.M., MacKay, J., Peter, G., Kirst, M., Clark, C., Whetten, R., and Sederoff, R. (2004). Gene expression during formation of earlywood and latewood in loblolly pine: expression profiles of 350 genes. Plant Biol. Stuttg. Ger. *6*, 654–663.

Ehlting, J., Büttner, D., Wang, Q., Douglas, C.J., Somssich, I.E., and Kombrink, E. (1999). Three 4coumarate:coenzyme A ligases in Arabidopsis thaliana represent two evolutionarily divergent classes in angiosperms. Plant J. Cell Mol. Biol. *19*, 9–20.

Ek, M., Gellerstedt, G., and Henriksson, G. (2009). Volume 1 Wood Chemistry and Wood Biotechnology (Berlin, Boston: De Gruyter).

Elkind, Y., Edwards, R., Mavandad, M., Hedrick, S.A., Ribak, O., Dixon, R.A., and Lamb, C.J. (1990). Abnormal plant development and down-regulation of phenylpropanoid biosynthesis in transgenic tobacco containing a heterologous phenylalanine ammonia-lyase gene. Proc. Natl. Acad. Sci. U. S. A. *87*, 9057–9061.

Endo, H., Yamaguchi, M., Tamura, T., Nakano, Y., Nishikubo, N., Yoneda, A., Kato, K., Kubo, M., Kajita, S., Katayama, Y., et al. (2015). Multiple classes of transcription factors regulate the expression of VASCULAR-RELATED NAC-DOMAIN7, a master switch of xylem vessel differentiation. Plant Cell Physiol. *56*, 242–254.

Erickson, R.O. (1986). Symplastic growth and symplasmic transport. Plant Physiol. 82, 1153.

Esau, K. (1977). Wiley: Anatomy of Seed Plants, 2nd Edition -.

Escamilla-Treviño, L.L., Chen, W., Card, M.L., Shih, M.-C., Cheng, C.-L., and Poulton, J.E. (2006). Arabidopsis thaliana beta-Glucosidases BGLU45 and BGLU46 hydrolyse monolignol glucosides. Phytochemistry *67*, 1651–1660.

Escamilla-Treviño, L.L., Shen, H., Hernandez, T., Yin, Y., Xu, Y., and Dixon, R.A. (2014). Early lignin pathway enzymes and routes to chlorogenic acid in switchgrass (Panicum virgatum L.). Plant Mol. Biol. *84*, 565–576.

Evert, R.F. (2006). Sclerenchyma. In Esau's Plant Anatomy, (John Wiley & Sons, Inc.), pp. 191-209.

Fahn, A. (1990). Plant anatomy (Oxford; New York: Pergamon Press).

Ferrer, J.-L., Austin, M.B., Stewart, C., and Noel, J.P. (2008). Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. Plant Physiol. Biochem. PPB Société Fr. Physiol. Végétale *46*, 356–370.

Fini, A., Guidi, L., Ferrini, F., Brunetti, C., Di Ferdinando, M., Biricolti, S., Pollastri, S., Calamai, L., and Tattini, M. (2012). Drought stress has contrasting effects on antioxidant enzymes activity and phenylpropanoid biosynthesis in Fraxinus ornus leaves: an excess light stress affair? J. Plant Physiol. *169*, 929–939.

Flanagan, P.M., Kelleher, R.J., Sayre, M.H., Tschochner, H., and Kornberg, R.D. (1991). A mediator required for activation of RNA polymerase II transcription in vitro. Nature *350*, 436–438.

Fornalé, S., Lopez, E., Salazar-Henao, J.E., Fernández-Nohales, P., Rigau, J., and Caparros-Ruiz, D. (2014).

AtMYB7, a New Player in the Regulation of UV-Sunscreens in Arabidopsis thaliana. Plant Cell Physiol. *55*, 507–516. Fraenkel, G.S. (1959). The Raison d'Être of Secondary Plant Substances. Science *129*, 1466–1470.

Francoz, E., Ranocha, P., Nguyen-Kim, H., Jamet, E., Burlat, V., and Dunand, C. (2015). Roles of cell wall peroxidases in plant development. Phytochemistry *112*, 15–21.

Freudenberg, K., and Harkin, J.M. (1963). The glucosides of cambial sap of spruce. Phytochemistry 2.

Frías, I., Siverio, J.M., González, C., Trujillo, J.M., and Pérez, J.A. (1991). Purification of a new peroxidase catalysing the formation of lignan-type compounds. Biochem. J. *273*, 109–113.

Fucassi, F., Heikal, A., Mikhalovska, L.I., Standen, G., Allan, I.U., Mikhalovsky, S.V., and Cragg, P.J. (2014). Metal chelation by a plant lignan, secoisolariciresinol diglucoside. J. Incl. Phenom. Macrocycl. Chem. *80*, 345–351.

Fujita, M., and Harada, H. (1979). Autoradiographic investigations of cell wall development, (2). Mokuzai Gakkai-Shi 25, 89–94.

Fukushima, K., Taguchi, S., Matsui, N., and Yasuda, S. (1997). Distribution and seasonal changes of monolignol glucosides in Pinus thunbergii. Mokuzai Gakkaishi *43*, 254–259.

Gabaldón, C., López-Serrano, M., Pedreño, M.A., and Barceló, A.R. (2005). Cloning and Molecular Characterization of the Basic Peroxidase Isoenzyme from Zinnia elegans, an Enzyme Involved in Lignin Biosynthesis. Plant Physiol. *139*, 1138–1154.

Gachon, C.M.M., Langlois-Meurinne, M., and Saindrenan, P. (2005). Plant secondary metabolism glycosyltransferases: the emerging functional analysis. Trends Plant Sci. *10*, 542–549.

Gang, D.R., Costa, M.A., Fujita, M., Dinkova-Kostova, A.T., Wang, H.B., Burlat, V., Martin, W., Sarkanen, S., Davin, L.B., and Lewis, N.G. (1999). Regiochemical control of monolignol radical coupling: a new paradigm for lignin and lignan biosynthesis. Chem. Biol. *6*, 143–151.

Garcia, K., Delaux, P.-M., Cope, K.R., and Ané, J.-M. (2015). Molecular signals required for the establishment and maintenance of ectomycorrhizal symbioses. New Phytol. *208*, 79–87.

García, M.J., Lucena, C., Romera, F.J., Alcántara, E., and Pérez-Vicente, R. (2010). Ethylene and nitric oxide involvement in the up-regulation of key genes related to iron acquisition and homeostasis in Arabidopsis. J. Exp. Bot. *61*, 3885–3899.

García, M.J., Romera, F.J., Lucena, C., Alcántara, E., and Pérez-Vicente, R. (2015). Ethylene and the Regulation of Physiological and Morphological Responses to Nutrient Deficiencies. Plant Physiol. *169*, 51–60.

Gavnholt, B., and Larsen, K. (2002). Molecular biology of plant laccases in relation to lignin formation. Physiol. Plant. *116*, 273–280.

Geldner, N. (2013). The Endodermis. Annu. Rev. Plant Biol. 64, 531–558.

Geshi, N. (2014). Arabinogalactan Glycosyltransferases: Enzyme Assay, Protein-Protein Interaction, Subcellular Localization, and Perspectives for Application. Adv. Bot. *2014*.

Ghose, K., Selvaraj, K., McCallum, J., Kirby, C.W., Sweeney-Nixon, M., Cloutier, S.J., Deyholos, M., Datla, R., and Fofana, B. (2014). Identification and functional characterization of a flax UDP-glycosyltransferase glucosylating secoisolariciresinol (SECO) into secoisolariciresinol monoglucoside (SMG) and diglucoside (SDG). BMC Plant Biol. *14*, 82.

Gloster, T.M. (2014). Advances in understanding glycosyltransferases from a structural perspective. Curr. Opin. Struct. Biol. *28*, 131–141.

Gorshkov, O., Mokshina, N., Gorshkov, V., Chemikosova, S., Gogolev, Y., and Gorshkova, T. (2016). Transcriptome portrait of cellulose-enriched flax fibres at advanced stage of specialization. Plant Mol. Biol. 1–19.

Gorshkova, T., Brutch, N., Chabbert, B., Deyholos, M., Hayashi, T., Lev-Yadun, S., Mellerowicz, E.J., Morvan, C., Neutelings, G., and Pilate, G. (2012). Plant Fiber Formation: State of the Art, Recent and Expected Progress, and Open Questions. Crit. Rev. Plant Sci. *31*, 201–228.

Gorshkova, T.A., Sal'nikov, V.V., Chemikosova, S.B., Ageeva, M.V., Pavlencheva, N.V., and van Dam, J.E.G. (2003). The snap point: a transition point in Linum usitatissimum bast fiber development. Ind. Crops Prod. *18*, 213–221.

Gorshkova, T.A., Ageeva, M., Chemikosova, S., and Salnikov, V. (2005). Tissue-specific processes during cell wall formation in flax fiber. Plant Biosyst. - Int. J. Deal. Asp. Plant Biol. *139*, 88–92.

Gou, J.-Y., Yu, X.-H., and Liu, C.-J. (2009). A hydroxycinnamoyltransferase responsible for synthesizing suberin aromatics in Arabidopsis. Proc. Natl. Acad. Sci. U. S. A. *106*, 18855–18860.

Grubb, C.D., and Abel, S. (2006). Glucosinolate metabolism and its control. Trends Plant Sci. 11, 89–100.

Grubb, C.D., Zipp, B.J., Kopycki, J., Schubert, M., Quint, M., Lim, E.-K., Bowles, D.J., Pedras, M.S.C., and Abel, S. (2014). Comparative analysis of Arabidopsis UGT74 glucosyltransferases reveals a special role of UGT74C1 in glucosinolate biosynthesis. Plant J. *79*, 92–105.

Hagel, J.M., and Facchini, P.J. (2005). Elevated tyrosine decarboxylase and tyramine hydroxycinnamoyltransferase levels increase wound-induced tyramine-derived hydroxycinnamic acid amide accumulation in transgenic tobacco leaves. Planta *221*, 904–914.

Hakala, P., Lampi, A.-M., Ollilainen, V., Werner, U., Murkovic, M., Wähälä, K., Karkola, S., and Piironen, V. (2002). Steryl phenolic acid esters in cereals and their milling fractions. J. Agric. Food Chem. *50*, 5300–5307.

Hano, C., Martin, I., Fliniaux, O., Legrand, B., Gutierrez, L., Arroo, R.R.J., Mesnard, F., Lamblin, F., and Lainé, E. (2006). Pinoresinol-lariciresinol reductase gene expression and secoisolariciresinol diglucoside accumulation in developing flax (Linum usitatissimum) seeds. Planta *224*, 1291–1301.

Hartmann, T. (2008). The lost origin of chemical ecology in the late 19th century. Proc. Natl. Acad. Sci. *105*, 4541–4546.

Hatfield, R.D., and Chaptman, A.K. (2009). Comparing Corn Types for Differences in Cell Wall Characteristics and p-Coumaroylation of Lignin. J. Agric. Food Chem. *57*, 4243–4249.

Hatfield, R.D., Grabber, J., Ralph, J., and Brei, K. (1999). Using the acetyl bromide assay to determine lignin concentrations in herbaceous plants: some cautionary notes. J. Agric. Food Chem. *47*, 628–632.

Hayward, A.P., Moreno, M.A., Howard, T.P., Hague, J., Nelson, K., Heffelfinger, C., Romero, S., Kausch, A.P., Glauser, G., Acosta, I.F., et al. (2016). Control of sexuality by the sk1-encoded UDP-glycosyltransferase of maize. Sci. Adv. *2*.

Hemmati, S., von Heimendahl, C.B.I., Klaes, M., Alfermann, A.W., Schmidt, T.J., and Fuss, E. (2010). Pinoresinol-lariciresinol reductases with opposite enantiospecificity determine the enantiomeric composition of lignans in the different organs of Linum usitatissimum L. Planta Med. *76*, 928–934.

Henrissat, B., and Davies, G.J. (2000). Glycoside Hydrolases and Glycosyltransferases. Families, Modules, and Implications for Genomics. Plant Physiol. *124*, 1515–1519.

Heredia, A. (2003). Biophysical and biochemical characteristics of cutin, a plant barrier biopolymer. Biochim. Biophys. Acta BBA - Gen. Subj. *1620*, 1–7.

Herrero, J., Fernández-Pérez, F., Yebra, T., Novo-Uzal, E., Pomar, F., Pedreño, M.Á., Cuello, J., Guéra, A., Esteban-Carrasco, A., and Zapata, J.M. (2013). Bioinformatic and functional characterization of the basic peroxidase 72 from Arabidopsis thaliana involved in lignin biosynthesis. Planta *237*, 1599–1612.

Hiromoto, T., Honjo, E., Tamada, T., Noda, N., Kazuma, K., Suzuki, M., and Kuroki, R. (2013). Crystal structure of UDP-glucose:anthocyanidin 3-O-glucosyltransferase from Clitoria ternatea. J. Synchrotron Radiat. *20*, 894–898.

Hoffmann, L., Maury, S., Martz, F., Geoffroy, P., and Legrand, M. (2003). Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism. J. Biol. Chem. *278*, 95–103.

Hoffmann, L., Besseau, S., Geoffroy, P., Ritzenthaler, C., Meyer, D., Lapierre, C., Pollet, B., and Legrand, M.

(2004). Silencing of Hydroxycinnamoyl-Coenzyme A Shikimate/Quinate Hydroxycinnamoyltransferase Affects Phenylpropanoid Biosynthesis. Plant Cell *16*, 1446–1465.

Hong, Z., Delauney, A.J., and Verma, D.P.S. (2001a). A Cell Plate–Specific Callose Synthase and Its Interaction with Phragmoplastin. Plant Cell *13*, 755–768.

Hong, Z., Zhang, Z., Olson, J.M., and Verma, D.P.S. (2001b). A Novel UDP-Glucose Transferase Is Part of the Callose Synthase Complex and Interacts with Phragmoplastin at the Forming Cell Plate. Plant Cell *13*, 769–779.

Hosokawa, M., Suzuki, S., Umezawa, T., and Sato, Y. (2001). Progress of lignification mediated by intercellular transportation of monolignols during tracheary element differentiation of isolated Zinnia mesophyll cells. Plant Cell Physiol. *42*, 959–968.

Hsieh, L.-S., Ma, G.-J., Yang, C.-C., and Lee, P.-D. (2010). Cloning, expression, site-directed mutagenesis and immunolocalization of phenylalanine ammonia-lyase in Bambusa oldhamii. Phytochemistry *71*, 1999–2009.

Hu, J., Baker, A., Bartel, B., Linka, N., Mullen, R.T., Reumann, S., and Zolman, B.K. (2012). Plant Peroxisomes: Biogenesis and Function. Plant Cell *24*, 2279–2303.

Huang, J., Gu, M., Lai, Z., Fan, B., Shi, K., Zhou, Y.-H., Yu, J.-Q., and Chen, Z. (2010). Functional Analysis of the Arabidopsis PAL Gene Family in Plant Growth, Development, and Response to Environmental Stress. Plant Physiol. *153*, 1526–1538.

Huang, Y., Li, C.Y., Qi, Y., Park, S., and Gibson, S.I. (2014). SIS8, a putative mitogen-activated protein kinase kinase kinase, regulates sugar-resistant seedling development in Arabidopsis. Plant J. Cell Mol. Biol. *77*, 577–588.

Hughes, J., and Hughes, M.A. (1994). Multiple secondary plant product UDP-glucose glucosyltransferase genes expressed in cassava (Manihot esculenta Crantz) cotyledons. DNA Seq. J. DNA Seq. Mapp. *5*, 41–49.

Huis, R., Hawkins, S., and Neutelings, G. (2010). Selection of reference genes for quantitative gene expression normalization in flax (Linum usitatissimum L.). BMC Plant Biol. *10*, 71.

Huis, R., Morreel, K., Fliniaux, O., Lucau-Danila, A., Fénart, S., Grec, S., Neutelings, G., Chabbert, B., Mesnard, F., Boerjan, W., et al. (2012). Natural Hypolignification Is Associated with Extensive Oligolignol Accumulation in Flax Stems. Plant Physiol. *158*, 1893–1915.

Humphrey, T.V., Richman, A.S., Menassa, R., and Brandle, J.E. (2006). Spatial Organisation of Four Enzymes from Stevia rebaudiana that are Involved in Steviol Glycoside Synthesis. Plant Mol. Biol. *61*, 47–62.

Iijima, Y., Gang, D.R., Fridman, E., Lewinsohn, E., and Pichersky, E. (2004). Characterization of Geraniol Synthase from the Peltate Glands of Sweet Basil. Plant Physiol. *134*, 370–379.

Isaacson, T., Kosma, D.K., Matas, A.J., Buda, G.J., He, Y., Yu, B., Pravitasari, A., Batteas, J.D., Stark, R.E., Jenks, M.A., et al. (2009). Cutin deficiency in the tomato fruit cuticle consistently affects resistance to microbial infection and biomechanical properties, but not transpirational water loss. Plant J. Cell Mol. Biol. *60*, 363–377.

Ishihara, A., Kawata, N., Matsukawa, T., and Iwamura, H. (2000). Induction of N-hydroxycinnamoyltyramine synthesis and tyramine N-hydroxycinnamoyltransferase (THT) activity by wounding in maize leaves. Biosci. Biotechnol. Biochem. *64*, 1025–1031.

Ishihara, A., Hashimoto, Y., Tanaka, C., Dubouzet, J.G., Nakao, T., Matsuda, F., Nishioka, T., Miyagawa, H., and Wakasa, K. (2008). The tryptophan pathway is involved in the defense responses of rice against pathogenic infection via serotonin production. Plant J. Cell Mol. Biol. *54*, 481–495.

Ishihara, H., Tohge, T., Viehöver, P., Fernie, A.R., Weisshaar, B., and Stracke, R. (2015). Natural variation in flavonol accumulation in Arabidopsis is determined by the flavonol glucosyltransferase BGLU6. J. Exp. Bot. erv546.

Ishii, T. (1997). Structure and functions of feruloylated polysaccharides. Plant Sci. 127, 111–127.

Islam, S., Griffiths, C.A., Blomstedt, C.K., Le, T.-N., Gaff, D.F., Hamill, J.D., and Neale, A.D. (2013). Increased Biomass, Seed Yield and Stress Tolerance Is Conferred in Arabidopsis by a Novel Enzyme from the Resurrection Grass Sporobolus stapfianus That Glycosylates the Strigolactone Analogue GR24. PLoS ONE *8*. Jang, S.-M., Ishihara, A., and Back, K. (2004). Production of Coumaroylserotonin and Feruloylserotonin in Transgenic Rice Expressing Pepper Hydroxycinnamoyl-Coenzyme A:Serotonin N-(Hydroxycinnamoyl)transferase. Plant Physiol. *135*, 346–356.

Jeandet, P., Douillet-Breuil, A.-C., Bessis, R., Debord, S., Sbaghi, M., and Adrian, M. (2002). Phytoalexins from the Vitaceae: Biosynthesis, Phytoalexin Gene Expression in Transgenic Plants, Antifungal Activity, and Metabolism. J. Agric. Food Chem. *50*, 2731–2741.

Jeandet, P., Delaunois, B., Conreux, A., Donnez, D., Nuzzo, V., Cordelier, S., Clément, C., and Courot, E. (2010). Biosynthesis, metabolism, molecular engineering, and biological functions of stilbene phytoalexins in plants. BioFactors *36*, 331–341.

Jin, L.-F., Liu, Y.-Z., Yin, X.-X., and Peng, S.-A. (2016). Transcript analysis of citrus miRNA397 and its target LAC7 reveals a possible role in response to boron toxicity. Acta Physiol. Plant. *38*, 18.

Jones, D.H. (1984). Phenylalanine ammonia-lyase: Regulation of its induction, and its role in plant development. Phytochemistry *23*, 1349–1359.

Jones, L., Ennos, A.R., and Turner, S.R. (2001). Cloning and characterization of irregular xylem4 (irx4): a severely lignin-deficient mutant of Arabidopsis. Plant J. Cell Mol. Biol. *26*, 205–216.

Jørgensen, K., Rasmussen, A.V., Morant, M., Nielsen, A.H., Bjarnholt, N., Zagrobelny, M., Bak, S., and Møller, B.L. (2005). Metabolon formation and metabolic channeling in the biosynthesis of plant natural products. Curr. Opin. Plant Biol. *8*, 280–291.

Kai, K., Mizutani, M., Kawamura, N., Yamamoto, R., Tamai, M., Yamaguchi, H., Sakata, K., and Shimizu, B. (2008). Scopoletin is biosynthesized via ortho-hydroxylation of feruloyl CoA by a 2-oxoglutarate-dependent dioxygenase in Arabidopsis thaliana. Plant J. Cell Mol. Biol. *55*, 989–999.

Kaneda, M., Rensing, K.H., Wong, J.C.T., Banno, B., Mansfield, S.D., and Samuels, A.L. (2008). Tracking Monolignols during Wood Development in Lodgepole Pine. Plant Physiol. *147*, 1750–1760.

Kannangara, R., Motawia, M.S., Hansen, N.K.K., Paquette, S.M., Olsen, C.E., Møller, B.L., and Jørgensen, K. (2011). Characterization and expression profile of two UDP-glucosyltransferases, UGT85K4 and UGT85K5, catalyzing the last step in cyanogenic glucoside biosynthesis in cassava. Plant J. *68*, 287–301.

Kao, Y.-Y., Harding, S.A., and Tsai, C.-J. (2002). Differential expression of two distinct phenylalanine ammonialyase genes in condensed tannin-accumulating and lignifying cells of quaking aspen. Plant Physiol. *130*, 796–807.

Kaplan, F., Kopka, J., Sung, D.Y., Zhao, W., Popp, M., Porat, R., and Guy, C.L. (2007). Transcript and metabolite profiling during cold acclimation of Arabidopsis reveals an intricate relationship of cold-regulated gene expression with modifications in metabolite content. Plant J. Cell Mol. Biol. *50*, 967–981.

Karamat, F., Olry, A., Doerper, S., Vialart, G., Ullmann, P., Werck-Reichhart, D., Bourgaud, F., and Hehn, A. (2012). CYP98A22, a phenolic ester 3'-hydroxylase specialized in the synthesis of chlorogenic acid, as a new tool for enhancing the furanocoumarin concentration in Ruta graveolens. BMC Plant Biol. *12*, 152.

Karlen, S.D., Zhang, C., Peck, M.L., Smith, R.A., Padmakshan, D., Helmich, K.E., Free, H.C.A., Lee, S., Smith, B.G.,
Lu, F., et al. (2016). Monolignol ferulate conjugates are naturally incorporated into plant lignins. Sci. Adv. 2, e1600393.
Keegstra, K., and Raikhel, N. (2001). Plant glycosyltransferases. Curr. Opin. Plant Biol. 4, 219–224.

Keller, H., Hohlfeld, H., Wray, V., Hahlbrock, K., Scheel, D., and Strack, D. (1996). Changes in the accumulation of soluble and cell wall-bound phenolics in elicitor-treated cell suspension cultures and fungus-infected leaves of Solanum tuberosum. Phytochemistry *42*, 389–396.

Kenrick, P., and Strullu-Derrien, C. (2014). The Origin and Early Evolution of Roots. Plant Physiol. *166*, 570–580.

Keurentjes, J.J.B., Fu, J., de Vos, C.H.R., Lommen, A., Hall, R.D., Bino, R.J., van der Plas, L.H.W., Jansen, R.C., Vreugdenhil, D., and Koornneef, M. (2006). The genetics of plant metabolism. Nat. Genet. *38*, 842–849.
Khater, F., Fournand, D., Vialet, S., Meudec, E., Cheynier, V., and Terrier, N. (2012). Identification and functional characterization of cDNAs coding for hydroxybenzoate/hydroxycinnamate glucosyltransferases co-expressed with genes related to proanthocyanidin biosynthesis. J. Exp. Bot. *63*, 1201–1214.

Kim, S.-J., and Brandizzi, F. (2014). The Plant Secretory Pathway: An Essential Factory for Building the Plant Cell Wall. Plant Cell Physiol. pct197.

Kim, D.-Y., Scalf, M., Smith, L.M., and Vierstra, R.D. (2013a). Advanced Proteomic Analyses Yield a Deep Catalog of Ubiquitylation Targets in Arabidopsis. Plant Cell *25*, 1523–1540.

Kim, J.I., Dolan, W.L., Anderson, N.A., and Chapple, C. (2015). Indole Glucosinolate Biosynthesis Limits Phenylpropanoid Accumulation in Arabidopsis thaliana. Plant Cell *27*, 1529–1546.

Kim, W.-C., Ko, J.-H., and Han, K.-H. (2012). Identification of a cis-acting regulatory motif recognized by MYB46, a master transcriptional regulator of secondary wall biosynthesis. Plant Mol. Biol. *78*, 489–501.

Kim, W.-C., Ko, J.-H., Kim, J.-Y., Kim, J., Bae, H.-J., and Han, K.-H. (2013b). MYB46 directly regulates the gene expression of secondary wall-associated cellulose synthases in Arabidopsis. Plant J. Cell Mol. Biol. *73*, 26–36.

Kitts, D.D., Yuan, Y.V., Wijewickreme, A.N., and Thompson, L.U. (1999). Antioxidant activity of the flaxseed lignan secoisolariciresinol diglycoside and its mammalian lignan metabolites enterodiol and enterolactone. Mol. Cell. Biochem. *202*, 91–100.

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. Cell Mol. Biol. *60*, 649–665.

Koeduka, T., Fridman, E., Gang, D.R., Vassão, D.G., Jackson, B.L., Kish, C.M., Orlova, I., Spassova, S.M., Lewis, N.G., Noel, J.P., et al. (2006). Eugenol and isoeugenol, characteristic aromatic constituents of spices, are biosynthesized via reduction of a coniferyl alcohol ester. Proc. Natl. Acad. Sci. *103*, 10128–10133.

Komvongsa, J., Mahong, B., Phasai, K., Hua, Y., Jeon, J.-S., and Ketudat Cairns, J.R. (2015). Identification of Fatty Acid Glucose Esters as Os9BGlu31 Transglucosidase Substrates in Rice Flag Leaves. J. Agric. Food Chem. *63*, 9764– 9769.

Kosma, D.K., and Jenks, M.A. (2007). Eco-Physiological and Molecular-Genetic Determinants of Plant Cuticle Function in Drought and Salt Stress Tolerance. In Advances in Molecular Breeding Toward Drought and Salt Tolerant Crops, M.A. Jenks, P.M. Hasegawa, and S.M. Jain, eds. (Springer Netherlands), pp. 91–120.

Kramer, C.M., Prata, R.T.N., Willits, M.G., De Luca, V., Steffens, J.C., and Graser, G. (2003). Cloning and regiospecificity studies of two flavonoid glucosyltransferases from Allium cepa. Phytochemistry *64*, 1069–1076.

Krause, C., Richter, S., Knöll, C., and Jürgens, G. (2013). Plant secretome - from cellular process to biological activity. Biochim. Biophys. Acta *1834*, 2429–2441.

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. *305*, 567–580.

Kroon, P.A., and Williamson, G. (1999). Hydroxycinnamates in plants and food: current and future perspectives. J. Sci. Food Agric. *79*, 355–361.

Kruszka, K., Pieczynski, M., Windels, D., Bielewicz, D., Jarmolowski, A., Szweykowska-Kulinska, Z., and Vazquez, F. (2012). Role of microRNAs and other sRNAs of plants in their changing environments. J. Plant Physiol. *169*, 1664–1672.

Kubo, M., Udagawa, M., Nishikubo, N., Horiguchi, G., Yamaguchi, M., Ito, J., Mimura, T., Fukuda, H., and Demura, T. (2005). Transcription switches for protoxylem and metaxylem vessel formation. Genes Dev. *19*, 1855–1860.

Kulich, I., and Žárský, V. (2014). Autophagy-Related Direct Membrane Import from ER/Cytoplasm into the Vacuole or Apoplast: A Hidden Gateway also for Secondary Metabolites and Phytohormones? Int. J. Mol. Sci. *15*, 7462–7474.

Kumar, M., Campbell, L., and Turner, S. (2016). Secondary cell walls: biosynthesis and manipulation. J. Exp. Bot. 67, 515–531.

Kumar, S., Pandey, A.K., Kumar, S., and Pandey, A.K. (2013). Chemistry and Biological Activities of Flavonoids: An Overview, Chemistry and Biological Activities of Flavonoids: An Overview. Sci. World J. Sci. World J. 2013, 2013, e162750.

Kunieda, T., Shimada, T., Kondo, M., Nishimura, M., Nishitani, K., and Hara-Nishimura, I. (2013). Spatiotemporal Secretion of PEROXIDASE36 Is Required for Seed Coat Mucilage Extrusion in Arabidopsis. Plant Cell *25*, 1355–1367.

Kushwah, S., Jones, A.M., and Laxmi, A. (2011). Cytokinin interplay with ethylene, auxin, and glucose signaling controls Arabidopsis seedling root directional growth. Plant Physiol. *156*, 1851–1866.

Laible, M., and Boonrod, K. (2009). Homemade Site Directed Mutagenesis of Whole Plasmids. J. Vis. Exp. JoVE.

Lairson, L.L., Henrissat, B., Davies, G.J., and Withers, S.G. (2008). Glycosyltransferases: structures, functions, and mechanisms. Annu. Rev. Biochem. *77*, 521–555.

Lamb, C.J., Merritt, T.K., and Butt, V.S. (1979). Synthesis and removal of phenylalanine ammonia-lyase activity in illuminated discs of potato tuber parenchyme. Biochim. Biophys. Acta *582*, 196–212.

Lan, W., Lu, F., Regner, M., Zhu, Y., Rencoret, J., Ralph, S.A., Zakai, U.I., Morreel, K., Boerjan, W., and Ralph, J. (2015). Tricin, a Flavonoid Monomer in Monocot Lignification. Plant Physiol. *167*, 1284–1295.

Lan, W., Morreel, K., Lu, F., Rencoret, J., Río, J.C. del, Voorend, W., Vermerris, W., Boerjan, W.A., and Ralph, J. (2016). Maize Tricin-Oligolignol Metabolites and their Implications for Monocot Lignification. Plant Physiol. pp.02012.2016.

Lange, B.M., Hertkorn, N., and Sandermann, H. (1998). Chloroaniline/Lignin Conjugates as Model System for Nonextractable Pesticide Residues in Crop Plants. Environ. Sci. Technol. 32, 2113–2118.

Langlois-Meurinne, M., Gachon, C.M.M., and Saindrenan, P. (2005). Pathogen-Responsive Expression of Glycosyltransferase Genes UGT73B3 and UGT73B5 Is Necessary for Resistance to Pseudomonas syringae pv tomato in Arabidopsis. Plant Physiol. *139*, 1890–1901.

Lanot, A., Hodge, D., Jackson, R.G., George, G.L., Elias, L., Lim, E.-K., Vaistij, F.E., and Bowles, D.J. (2006). The glucosyltransferase UGT72E2 is responsible for monolignol 4-O-glucoside production in Arabidopsis thaliana. Plant J. *48*, 286–295.

Lanot, A., Hodge, D., Lim, E.-K., Vaistij, F.E., and Bowles, D.J. (2008). Redirection of flux through the phenylpropanoid pathway by increased glucosylation of soluble intermediates. Planta *228*, 609–616.

Laursen, T., Borch, J., Knudsen, C., Bavishi, K., Torta, F., Martens, H.J., Silvestro, D., Hatzakis, N.S., Wenk, M.R., Dafforn, T.R., et al. (2016). Characterization of a dynamic metabolon producing the defense compound dhurrin in sorghum. Science *354*, 890–893.

Le, T.N., Blomstedt, C.K., Kuang, J., Tenlen, J., Gaff, D.F., Hamill, J.D., and Neale, A.D. (2007). Desiccationtolerance specific gene expression in leaf tissue of the resurrection plant Sporobolus stapfianus. Funct. Plant Biol. *34*, 589–600.

Le Roy, J., Blervacq AS., Creach, A., Huss, B., Hawkins, S., and Neutelings, G. (in revision). Spatial regulation of monolignol biosynthesis and laccase genes control developmental and stress-related lignin in flax. BMC Plant Biology.

Le Roy, J., Huss, B., Creach, A., Hawkins, S., and Neutelings, G. (2016). Glycosylation Is a Major Regulator of Phenylpropanoid Availability and Biological Activity in Plants. Front. Plant Sci. *7*.

Lee, Y., Yoon, H.R., Paik, Y.S., Liu, J.R., Chung, W., and Choi, G. (2005). Reciprocal regulation of Arabidopsis UGT78D2 and BANYULS is critical for regulation of the metabolic flux of anthocyanidins to condensed tannins in developing seed coats. J. Plant Biol. *48*, 356–370.

Lee, Y., Rubio, M.C., Alassimone, J., and Geldner, N. (2013). A Mechanism for Localized Lignin Deposition in

the Endodermis. Cell 153, 402–412.

Legay, S., Sivadon, P., Blervacq, A.-S., Pavy, N., Baghdady, A., Tremblay, L., Levasseur, C., Ladouce, N., Lapierre, C., Séguin, A., 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.

Leinhos, V., and Savidge, R.A. (1993). Isolation of protoplasts from developing xylem of Pinus banksiana and Pinus strobus. Can. J. For. Res. Rev. Can. Rech. For.

Leiss, K.A., Maltese, F., Choi, Y.H., Verpoorte, R., and Klinkhamer, P.G.L. (2009). Identification of Chlorogenic Acid as a Resistance Factor for Thrips in Chrysanthemum. Plant Physiol. *150*, 1567–1575.

León, J. (2013). Role of plant peroxisomes in the production of jasmonic acid-based signals. Subcell. Biochem. *69*, 299–313.

Levigne, S.V., Ralet, M.-C.J., Quéméner, B.C., Pollet, B.N.-L., Lapierre, C., and Thibault, J.-F.J. (2004). Isolation from Sugar Beet Cell Walls of Arabinan Oligosaccharides Esterified by Two Ferulic Acid Monomers. Plant Physiol. *134*, 1173–1180.

Lev-Yadun, S. (1994). Induction of sclereid differentiation in the pith of Arabidopsis thaliana (L.) Heynh. J. Exp. Bot. 45, 1845–1849.

Lev-Yadun, S. (1997). Fibres and Fibre-sclereids in Wild-typeArabidopsis thaliana. Ann. Bot. 80, 125–129.

Li, E., Bhargava, A., Qiang, W., Friedmann, M.C., Forneris, N., Savidge, R.A., Johnson, L.A., Mansfield, S.D., Ellis, B.E., and Douglas, C.J. (2012a). The Class II KNOX gene KNAT7 negatively regulates secondary wall formation in Arabidopsis and is functionally conserved in Populus. New Phytol. *194*, 102–115.

Li, L., Modolo, L.V., Escamilla-Trevino, L.L., Achnine, L., Dixon, R.A., and Wang, X. (2007). Crystal structure of Medicago truncatula UGT85H2--insights into the structural basis of a multifunctional (iso)flavonoid glycosyltransferase. J. Mol. Biol. *370*, 951–963.

Li, Q., Lin, Y.-C., Sun, Y.-H., Song, J., Chen, H., Zhang, X.-H., Sederoff, R.R., and Chiang, V.L. (2012b). Splice variant of the SND1 transcription factor is a dominant negative of SND1 members and their regulation in Populus trichocarpa. Proc. Natl. Acad. Sci. U. S. A. *109*, 14699–14704.

Li, Y., Baldauf, S., Lim, E.-K., and Bowles, D.J. (2001). Phylogenetic Analysis of the UDP-glycosyltransferase Multigene Family of Arabidopsis thaliana. J. Biol. Chem. *276*, 4338–4343.

Li, Y., Li, P., Wang, Y., Dong, R., Yu, H., and Hou, B. (2014a). Genome-wide identification and phylogenetic analysis of Family-1 UDP glycosyltransferases in maize (Zea mays). Planta *239*, 1265–1279.

Li, Y., Kim, J.I., Pysh, L., and Chapple, C. (2015). Four Isoforms of Arabidopsis 4-Coumarate:CoA Ligase Have Overlapping yet Distinct Roles in Phenylpropanoid Metabolism. Plant Physiol. *169*, 2409–2421.

Li, Z., Zhao, Y., Liu, X., Peng, J., Guo, H., and Luo, J. (2014b). LSD 2.0: an update of the leaf senescence database. Nucleic Acids Res. *42*, D1200–D1205.

Liang, D.-M., Liu, J.-H., Wu, H., Wang, B.-B., Zhu, H.-J., and Qiao, J.-J. (2015). Glycosyltransferases: mechanisms and applications in natural product development. Chem. Soc. Rev. *44*, 8350–8374.

Liang, M., Davis, E., Gardner, D., Cai, X., and Wu, Y. (2006). Involvement of AtLAC15 in lignin synthesis in seeds and in root elongation of Arabidopsis. Planta *224*, 1185.

Lim, E.K., Li, Y., Parr, A., Jackson, R., Ashford, D.A., and Bowles, D.J. (2001). Identification of glucosyltransferase genes involved in sinapate metabolism and lignin synthesis in Arabidopsis. J. Biol. Chem. *276*, 4344–4349.

Lim, E.-K., Baldauf, S., Li, Y., Elias, L., Worrall, D., Spencer, S.P., Jackson, R.G., Taguchi, G., Ross, J., and Bowles, D.J. (2003). Evolution of substrate recognition across a multigene family of glycosyltransferases in Arabidopsis. Glycobiology *13*, 139–145.

Lim, E.-K., Doucet, C.J., Li, Y., Elias, L., Worrall, D., Spencer, S.P., Ross, J., and Bowles, D.J. (2002). The activity of

Arabidopsis glycosyltransferases toward salicylic acid, 4-hydroxybenzoic acid, and other benzoates. J. Biol. Chem. 277, 586–592.

Lim, E.-K., Jackson, R.G., and Bowles, D.J. (2005). Identification and characterisation of Arabidopsis glycosyltransferases capable of glucosylating coniferyl aldehyde and sinapyl aldehyde. FEBS Lett. *579*, 2802–2806.

Lin, J.-S., Huang, X.-X., Li, Q., Cao, Y., Bao, Y., Meng, X.-F., Li, Y.-J., Fu, C., and Hou, B.-K. (2016). UDPglycosyltransferase 72B1 catalyzes the glucose conjugation of monolignols and is essential for the normal cell wall lignification in Arabidopsis thaliana. Plant J. Cell Mol. Biol. *88*, 26–42.

Lindermayr, C., Möllers, B., Fliegmann, J., Uhlmann, A., Lottspeich, F., Meimberg, H., and Ebel, J. (2002). Divergent members of a soybean (Glycine max L.) 4-coumarate:coenzyme A ligase gene family. Eur. J. Biochem. FEBS *269*, 1304–1315.

Liu, C.-J. (2010). Biosynthesis of hydroxycinnamate conjugates: implications for sustainable biomass and biofuel production. Biofuels *1*, 745–761.

Liu, C.-J., Miao, Y.-C., and Zhang, K.-W. (2011). Sequestration and Transport of Lignin Monomeric Precursors. Molecules *16*, 710–727.

Liu, Z., Yan, J.-P., Li, D.-K., Luo, Q., Yan, Q., Liu, Z.-B., Ye, L.-M., Wang, J.-M., Li, X.-F., and Yang, Y. (2015). UDPglucosyltransferase71c5, a major glucosyltransferase, mediates abscisic acid homeostasis in Arabidopsis. Plant Physiol. *167*, 1659–1670.

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2014). The carbohydrateactive enzymes database (CAZy) in 2013. Nucleic Acids Res. *42*, D490-495.

Long, W., Zou, X., and Zhang, X. (2015). Transcriptome Analysis of Canola (Brassica napus) under Salt Stress at the Germination Stage. PLOS ONE *10*, e0116217.

Lonsdale, A., Davis, M.J., Doblin, M.S., and Bacic, A. (2016). Better Than Nothing? Limitations of the Prediction Tool SecretomeP in the Search for Leaderless Secretory Proteins (LSPs) in Plants. Front. Plant Sci. *7*.

Loutre, C., Dixon, D.P., Brazier, M., Slater, M., Cole, D.J., and Edwards, R. (2003). Isolation of a glucosyltransferase from Arabidopsis thaliana active in the metabolism of the persistent pollutant 3,4-dichloroaniline. Plant J. Cell Mol. Biol. 34, 485–493.

Lu, S., Zhou, Y., Li, L., and Chiang, V.L. (2006). Distinct Roles of Cinnamate 4-hydroxylase Genes in Populus. Plant Cell Physiol. 47, 905–914.

Lu, S., Li, Q., Wei, H., Chang, M.-J., Tunlaya-Anukit, S., Kim, H., Liu, J., Song, J., Sun, Y.-H., Yuan, L., et al. (2013). Ptr-miR397a is a negative regulator of laccase genes affecting lignin content in Populus trichocarpa. Proc. Natl. Acad. Sci. *110*, 10848–10853.

Luang, S., Cho, J.-I., Mahong, B., Opassiri, R., Akiyama, T., Phasai, K., Komvongsa, J., Sasaki, N., Hua, Y., Matsuba, Y., et al. (2013). Rice Os9BGlu31 is a transglucosidase with the capacity to equilibrate phenylpropanoid, flavonoid, and phytohormone glycoconjugates. J. Biol. Chem. *288*, 10111–10123.

Luschnig, C., Gaxiola, R.A., Grisafi, P., and Fink, G.R. (1998). EIR1, a root-specific protein involved in auxin transport, is required for gravitropism in Arabidopsis thaliana. Genes Dev. *12*, 2175–2187.

Ma, J.P., Tan, C.H., Zhu, D.Y., and Jin, L. (2011). A novel 8,4'-oxyneolignan diglycoside from Ligusticum sinensis. Chin. Chem. Lett. *22*, 1454–1456.

Mackenzie, P.I., Owens, I.S., Burchell, B., Bock, K.W., Bairoch, A., Bélanger, A., Fournel-Gigleux, S., Green, M., Hum, D.W., Iyanagi, T., et al. (1997). The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence. Pharmacogenetics *7*, 255–269.

Macoy, D.M., Kim, W.-Y., Lee, S.Y., and Kim, M.G. (2015). Biosynthesis, physiology, and functions of hydroxycinnamic acid amides in plants. Plant Biotechnol. Rep. *9*, 269–278.

Mai, H.-J., Pateyron, S., and Bauer, P. (2016). Iron homeostasis in Arabidopsis thaliana: transcriptomic

analyses reveal novel FIT-regulated genes, iron deficiency marker genes and functional gene networks. BMC Plant Biol. *16*, 211.

Mansfield, S.D., Kim, H., Lu, F., and Ralph, J. (2012). Whole plant cell wall characterization using solution-state 2D NMR. Nat. Protoc. 7, 1579–1589.

Marchant, H.J. (1979). Microtubules, cell wall deposition and the determination of plant cell shape. Nature 278, 167–168.

Marita, J.M., Ralph, J., Hatfield, R.D., Guo, D., Chen, F., and Dixon, R.A. (2003). Structural and compositional modifications in lignin of transgenic alfalfa down-regulated in caffeic acid 3-0-methyltransferase and caffeoyl coenzyme A 3-0-methyltransferase. Phytochemistry *62*, 53–65.

Matsuba, Y., Sasaki, N., Tera, M., Okamura, M., Abe, Y., Okamoto, E., Nakamura, H., Funabashi, H., Takatsu, M., Saito, M., et al. (2010). A Novel Glucosylation Reaction on Anthocyanins Catalyzed by Acyl-Glucose–Dependent Glucosyltransferase in the Petals of Carnation and Delphinium. Plant Cell *22*, 3374–3389.

Mayer, A.M., and Staples, R.C. (2002). Laccase: new functions for an old enzyme. Phytochemistry *60*, 551–565.

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.

McDougall, G.J., Morrison, I.M., Stewart, D., Weyers, J.D.B., and Hillman, J.R. (1993). Plant fibres: Botany, chemistry and processing for industrial use. J. Sci. Food Agric. *62*, 1–20.

Meagher, L.P., Beecher, G.R., Flanagan, V.P., and Li, B.W. (1999). Isolation and Characterization of the Lignans, Isolariciresinol and Pinoresinol, in Flaxseed Meal. J. Agric. Food Chem. *47*, 3173–3180.

Ménard, D., and Pesquet, E. (2015). Cellular interactions during tracheary elements formation and function. Curr. Opin. Plant Biol. *23*, 109–115.

Miao, Y.-C., and Liu, C.-J. (2010). ATP-binding cassette-like transporters are involved in the transport of lignin precursors across plasma and vacuolar membranes. Proc. Natl. Acad. Sci. U. S. A. *107*, 22728–22733.

Migliaccio, F., and Piconese, S. (2001). Spiralizations and tropisms in Arabidopsis roots. Trends Plant Sci. *6*, 561–565.

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

Modolo, L.V., Blount, J.W., Achnine, L., Naoumkina, M.A., Wang, X., and Dixon, R.A. (2007). A functional genomics approach to (iso)flavonoid glycosylation in the model legume Medicago truncatula. Plant Mol. Biol. *64*, 499–518.

Molinari, H.B.C., Pellny, T.K., Freeman, J., Shewry, P.R., and Mitchell, R.A.C. (2013). Grass cell wall feruloylation: distribution of bound ferulate and candidate gene expression in Brachypodium distachyon. Plant Biotechnol. *4*, 50.

Moraes, T.F., and Reithmeier, R.A.F. (2012). Membrane transport metabolons. Biochim. Biophys. Acta BBA - Biomembr. *1818*, 2687–2706.

Moreira-Vilar, F.C., Siqueira-Soares, R. de C., Finger-Teixeira, A., Oliveira, D.M. de, Ferro, A.P., Rocha, G.J. da, Ferrarese, M. de L.L., Santos, W.D. dos, and Ferrarese-Filho, O. (2014). The Acetyl Bromide Method Is Faster, Simpler and Presents Best Recovery of Lignin in Different Herbaceous Tissues than Klason and Thioglycolic Acid Methods. PLOS ONE *9*, e110000.

Morikawa, Y., Yoshinaga, A., Kamitakahara, H., Wada, M., and Takabe, K. (2009). Cellular distribution of coniferin in differentiating xylem of Chamaecyparis obtusa as revealed by Raman microscopy. Holzforschung *64*, 61–67.

Morreel, K., Ralph, J., Lu, F., Goeminne, G., Busson, R., Herdewijn, P., Goeman, J.L., Van der Eycken, J., Boerjan, W., and Messens, E. (2004a). Phenolic Profiling of Caffeic Acid O-Methyltransferase-Deficient Poplar Reveals Novel Benzodioxane Oligolignols. Plant Physiol. *136*, 4023–4036.

Morreel, K., Ralph, J., Kim, H., Lu, F., Goeminne, G., Ralph, S., Messens, E., and Boerjan, W. (2004b). Profiling of Oligolignols Reveals Monolignol Coupling Conditions in Lignifying Poplar Xylem. Plant Physiol. *136*, 3537–3549.

Morreel, K., Kim, H., Lu, F., Dima, O., Akiyama, T., Vanholme, R., Niculaes, C., Goeminne, G., Inzé, D., Messens, E., et al. (2010). Mass Spectrometry-Based Fragmentation as an Identification Tool in Lignomics. Anal. Chem. *82*, 8095–8105.

Morvan, C., Andème-Onzighi, C., Girault, R., Himmelsbach, D.S., Driouich, A., and Akin, D.E. (2003). Building flax fibres: more than one brick in the walls. Plant Physiol. Biochem. *41*, 935–944.

Moss, G.P. (2009). Nomenclature of Lignans and Neolignans (IUPAC Recommendations 2000). Pure Appl. Chem. 72, 1493–1523.

Mottiar, Y., Vanholme, R., Boerjan, W., Ralph, J., and Mansfield, S.D. (2016). Designer lignins: harnessing the plasticity of lignification. Curr. Opin. Biotechnol. *37*, 190–200.

Moura, J.C.M.S., Bonine, C.A.V., de Oliveira Fernandes Viana, J., Dornelas, M.C., and Mazzafera, P. (2010). Abiotic and biotic stresses and changes in the lignin content and composition in plants. J. Integr. Plant Biol. *52*, 360–376.

Muhlemann, J.K., Woodworth, B.D., Morgan, J.A., and Dudareva, N. (2014). The monolignol pathway contributes to the biosynthesis of volatile phenylpropenes in flowers. New Phytol. *204*, 661–670.

Muir, A.D., and Westcott, N.D. (2003). Flax: The genus Linum (CRC Press).

Murashige, T., and Skoog, F. (1962). A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. Physiol. Plant. *15*, 473–497.

Muravenko, O.V., Bolsheva, N.L., Yurkevich, O.Y., Nosova, I.V., Rachinskaya, O.A., Samatadze, T.E., and Zelenin, A.V. (2010). Karyogenomics of species of the genus Linum L. Russ. J. Genet. *46*, 1182–1185.

Nakano, Y., Yamaguchi, M., Endo, H., Rejab, N.A., and Ohtani, M. (2015a). NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants. Front. Plant Sci. 6.

Nakano, Y., Yamaguchi, M., Endo, H., Rejab, N.A., and Ohtani, M. (2015b). NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants. Plant Physiol. *6*, 288.

Nakatsubo, T., Mizutani, M., Suzuki, S., Hattori, T., and Umezawa, T. (2008). Characterization of Arabidopsis thaliana Pinoresinol Reductase, a New Type of Enzyme Involved in Lignan Biosynthesis. J. Biol. Chem. *283*, 15550–15557.

Naseer, S., Lee, Y., Lapierre, C., Franke, R., Nawrath, C., and Geldner, N. (2012). Casparian strip diffusion barrier in Arabidopsis is made of a lignin polymer without suberin. Proc. Natl. Acad. Sci. U. S. A. *109*, 10101–10106.

Netravali, A.N., and Pastore, C.M. (2014). Sustainable Composites: Fibers, Resins and Applications (DEStech Publications, Inc).

Nielsen, K.A., Tattersall, D.B., Jones, P.R., and Møller, B.L. (2008a). Metabolon formation in dhurrin biosynthesis. Phytochemistry *69*, 88–98.

Niggeweg, R., Michael, A.J., and Martin, C. (2004). Engineering plants with increased levels of the antioxidant chlorogenic acid. Nat. Biotechnol. *22*, 746–754.

Nishimura, Y., Tokimatsu, T., Kotera, M., Goto, S., and Kanehisa, M. (2010). Genome-wide analysis of plant UGT family based on sequence and substrate information. Genome Inform. Int. Conf. Genome Inform. *24*, 127–138.

O'Brien, T.P., Feder, N., and McCully, M.E. (1964). Polychromatic staining of plant cell walls by toluidine blue O. Protoplasma *59*, 368–373.

Oda, Y., and Fukuda, H. (2012). Secondary cell wall patterning during xylem differentiation. Curr. Opin. Plant

Biol. 15, 38-44.

Offen, W., Martinez-Fleites, C., Yang, M., Kiat-Lim, E., Davis, B.G., Tarling, C.A., Ford, C.M., Bowles, D.J., and Davies, G.J. (2006). Structure of a flavonoid glucosyltransferase reveals the basis for plant natural product modification. EMBO J. 25, 1396–1405.

Ogata, J., Itoh, Y., Ishida, M., Yoshida, H., and Ozeki, Y. (2004). Cloning and heterologous expression of cDNAs encoding flavonoid glucosyltransferases from *Dianthus caryophyllus*. Plant Biotechnol. *21*, 367–375.

Ohashi-Ito, K., Oda, Y., and Fukuda, H. (2010). Arabidopsis VASCULAR-RELATED NAC-DOMAIN6 Directly Regulates the Genes That Govern Programmed Cell Death and Secondary Wall Formation during Xylem Differentiation. Plant Cell *22*, 3461–3473.

Osmani, S.A., Bak, S., and Møller, B.L. (2009). Substrate specificity of plant UDP-dependent glycosyltransferases predicted from crystal structures and homology modeling. Phytochemistry *70*, 325–347.

Ostergaard, L., Teilum, K., Mirza, O., Mattsson, O., Petersen, M., Welinder, K.G., Mundy, J., Gajhede, M., and Henriksen, A. (2000). Arabidopsis ATP A2 peroxidase. Expression and high-resolution structure of a plant peroxidase with implications for lignification. Plant Mol. Biol. *44*, 231–243.

Ostrowski, M., and Jakubowska, A. (2014). Udp-Glycosyltransferases of Plant Hormones. Adv. Cell Biol. *4*, 43–60.

Panchy, N., Lehti-Shiu, M., and Shiu, S.-H. (2016). Evolution of Gene Duplication in Plants. Plant Physiol. *171*, 2294–2316.

Pang, Y., Peel, G.J., Sharma, S.B., Tang, Y., and Dixon, R.A. (2008). A transcript profiling approach reveals an epicatechin-specific glucosyltransferase expressed in the seed coat of Medicago truncatula. Proc. Natl. Acad. Sci. *105*, 14210–14215.

Pang, Y., Cheng, X., Huhman, D.V., Ma, J., Peel, G.J., Yonekura-Sakakibara, K., Saito, K., Shen, G., Sumner, L.W., Tang, Y., et al. (2013). Medicago glucosyltransferase UGT72L1: potential roles in proanthocyanidin biosynthesis. Planta *238*, 139–154.

Paquette, S., Møller, B.L., and Bak, S. (2003a). On the origin of family 1 plant glycosyltransferases. Phytochemistry *62*, 399–413.

Parniske, M. (2008). Arbuscular mycorrhiza: the mother of plant root endosymbioses. Nat. Rev. Microbiol. *6*, 763–775.

Passardi, F., Cosio, C., Penel, C., and Dunand, C. (2005). Peroxidases have more functions than a Swiss army knife. Plant Cell Rep. *24*, 255–265.

Patel, S. (2012). Cereal bran: the next super food with significant antioxidant and anticancer potential. Mediterr. J. Nutr. Metab. *5*, 91–104.

Paul, A.-L., Amalfitano, C.E., and Ferl, R.J. (2012). Plant growth strategies are remodeled by spaceflight. BMC Plant Biol. *12*, 232.

Paul Bolwell, G. (1992). A role for phosphorylation in the down-regulation of phenylalanine ammonia-lyase in suspension-cultured cells of french bean. Phytochemistry *31*, 4081–4086.

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.

Payyavula, R.S., Shakya, R., Sengoda, V.G., Munyaneza, J.E., Swamy, P., and Navarre, D.A. (2015a). Synthesis and regulation of chlorogenic acid in potato: Rerouting phenylpropanoid flux in HQT-silenced lines. Plant Biotechnol. J. *13*, 551–564.

Pesquet, E., Zhang, B., Gorzsás, A., Puhakainen, T., Serk, H., Escamez, S., Barbier, O., Gerber, L., Courtois-Moreau, C., Alatalo, E., et al. (2013). Non-Cell-Autonomous Postmortem Lignification of Tracheary Elements in Zinnia elegans. Plant Cell *25*, 1314–1328. Petersen, M., and Simmonds, M.S.J. (2003). Rosmarinic acid. Phytochemistry 62, 121-125.

Petersen, T.N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat. Methods *8*, 785–786.

Petrik, D.L., Karlen, S.D., Cass, C.L., Padmakshan, D., Lu, F., Liu, S., Le Bris, P., Antelme, S., Santoro, N., Wilkerson, C.G., et al. (2014). p-Coumaroyl-CoA:monolignol transferase (PMT) acts specifically in the lignin biosynthetic pathway in Brachypodium distachyon. Plant J. Cell Mol. Biol. *77*, 713–726.

Piber, M., and Koehler, P. (2005). Identification of dehydro-ferulic acid-tyrosine in rye and wheat: evidence for a covalent cross-link between arabinoxylans and proteins. J. Agric. Food Chem. *53*, 5276–5284.

Pichersky, E., Noel, J.P., and Dudareva, N. (2006). Biosynthesis of Plant Volatiles: Nature's Diversity and Ingenuity. Science *311*, 808–811.

Pickel, B., and Schaller, A. (2013). Dirigent proteins: molecular characteristics and potential biotechnological applications. Appl. Microbiol. Biotechnol. *97*, 8427–8438.

Pickel, B., Constantin, M.-A., Pfannstiel, J., Conrad, J., Beifuss, U., and Schaller, A. (2010). An Enantiocomplementary Dirigent Protein for the Enantioselective Laccase-Catalyzed Oxidative Coupling of Phenols. Angew. Chem. Int. Ed. *49*, 202–204.

Pickett-Heaps, J.D. (1968). Xylem wall deposition. Protoplasma 65, 181-205.

de Pinto, M.C., and Ros Barceló, A. (1997). Cytochemical localization of phenol-oxidizing enzymes in lignifying Coleus blumei stems. Eur. J. Histochem. EJH *41*, 17–22.

Pohjamo, S.P., Hemming, J.E., Willför, S.M., Reunanen, M.H.T., and Holmbom, B.R. (2003). Phenolic extractives in Salix caprea wood and knots. Phytochemistry *63*, 165–169.

Pollard, M., Beisson, F., Li, Y., and Ohlrogge, J.B. (2008). Building lipid barriers: biosynthesis of cutin and suberin. Trends Plant Sci. *13*, 236–246.

Pomar, F., Merino, F., and Barceló, A.R. (2002). O-4-Linked coniferyl and sinapyl aldehydes in lignifying cell walls are the main targets of the Wiesner (phloroglucinol-HCl) reaction. Protoplasma *220*, 17–28.

Pourcel, L., Irani, N.G., Lu, Y., Riedl, K., Schwartz, S., and Grotewold, E. (2010). The formation of Anthocyanic Vacuolar Inclusions in Arabidopsis thaliana and implications for the sequestration of anthocyanin pigments. Mol. Plant *3*, 78–90.

del Pozo, J.C., Boniotti, M.B., and Gutierrez, C. (2002). Arabidopsis E2Fc Functions in Cell Division and Is Degraded by the Ubiquitin-SCFAtSKP2 Pathway in Response to Light. Plant Cell *14*, 3057–3071.

Preston, J., Wheeler, J., Heazlewood, J., Li, S.F., and Parish, R.W. (2004). AtMYB32 is required for normal pollen development in Arabidopsis thaliana. Plant J. Cell Mol. Biol. *40*, 979–995.

Radominska-Pandya, A., Czernik, P.J., Little, J.M., Battaglia, E., and Mackenzie, P.I. (1999). Structural and functional studies of UDP-glucuronosyltransferases. Drug Metab. Rev. *31*, 817–899.

Raes, J., Rohde, A., Christensen, J.H., Van de Peer, Y., and Boerjan, W. (2003). Genome-Wide Characterization of the Lignification Toolbox in Arabidopsis. Plant Physiol. *133*, 1051–1071.

Ragauskas, A.J., Beckham, G.T., Biddy, M.J., Chandra, R., Chen, F., Davis, M.F., Davison, B.H., Dixon, R.A., Gilna, P., Keller, M., et al. (2014). Lignin valorization: improving lignin processing in the biorefinery. Science *344*, 1246843.

Ralph, J., Bunzel, M., Marita, J.M., Hatfield, R.D., Lu, F., Kim, H., Schatz, P.F., Grabber, J.H., and Steinhart, H. (2004a). Peroxidase-dependent cross-linking reactions of p-hydroxycinnamates in plant cell walls. Phytochem. Rev. *3*, 79–96.

Ralph, J., Lundquist, K., Brunow, G., Lu, F., Kim, H., Schatz, P.F., Marita, J.M., Hatfield, R.D., Ralph, S.A., Christensen, J.H., et al. (2004b). Lignins: Natural polymers from oxidative coupling of 4-hydroxyphenyl- propanoids. Phytochem. Rev. *3*, 29–60.

Ralph, J., Kim, H., Lu, F., Grabber, J.H., Leplé, J.-C., Berrio-Sierra, J., Derikvand, M.M., Jouanin, L., Boerjan, W.,

and Lapierre, C. (2008). Identification of the structure and origin of a thioacidolysis marker compound for ferulic acid incorporation into angiosperm lignins (and an indicator for cinnamoyl CoA reductase deficiency). Plant J. Cell Mol. Biol. *53*, 368–379.

Ranocha, P., McDougall, G., Hawkins, S., Sterjiades, R., Borderies, G., Stewart, D., Cabanes-Macheteau, M., Boudet, A.M., and Goffner, D. (1999). Biochemical characterization, molecular cloning and expression of laccases - a divergent gene family - in poplar. Eur. J. Biochem. *259*, 485–495.

Ranocha, P., Chabannes, M., Chamayou, S., Danoun, S., Jauneau, A., Boudet, A.-M., and Goffner, D. (2002). Laccase down-regulation causes alterations in phenolic metabolism and cell wall structure in poplar. Plant Physiol. *129*, 145–155.

Rashotte, A.M., Brady, S.R., Reed, R.C., Ante, S.J., and Muday, G.K. (2000). Basipetal Auxin Transport Is Required for Gravitropism in Roots of Arabidopsis. Plant Physiol. *122*, 481–490.

Rasmussen, null, and Dixon, null (1999). Transgene-mediated and elicitor-induced perturbation of metabolic channeling at the entry point into the phenylpropanoid pathway. Plant Cell *11*, 1537–1552.

Razavi, S.M. (2011). Plant Coumarins as Allelopathic Agents. Int. J. Biol. Chem. 5, 86–90.

Reisberg, E.E., Hildebrandt, U., Riederer, M., and Hentschel, U. (2013). Distinct Phyllosphere Bacterial Communities on Arabidopsis Wax Mutant Leaves. PLOS ONE *8*, e78613.

Renouard, S., Tribalatc, M.-A., Lamblin, F., Mongelard, G., Fliniaux, O., Corbin, C., Marosevic, D., Pilard, S., Demailly, H., Gutierrez, L., et al. (2014). RNAi-mediated pinoresinol lariciresinol reductase gene silencing in flax (Linum usitatissimum L.) seed coat: consequences on lignans and neolignans accumulation. J. Plant Physiol. *171*, 1372–1377.

del Río, J.C., Rencoret, J., Gutiérrez, A., Nieto, L., Jiménez-Barbero, J., and Martínez, Á.T. (2011). Structural characterization of guaiacyl-rich lignins in flax (Linum usitatissimum) fibers and shives. J. Agric. Food Chem. *59*, 11088–11099.

del Río, J.C., Rencoret, J., Prinsen, P., Martínez, Á.T., Ralph, J., and Gutiérrez, A. (2012). Structural characterization of wheat straw lignin as revealed by analytical pyrolysis, 2D-NMR, and reductive cleavage methods. J. Agric. Food Chem. *60*, 5922–5935.

Rivière, C., Pawlus, A.D., and Mérillon, J.-M. (2012). Natural stilbenoids: distribution in the plant kingdom and chemotaxonomic interest in Vitaceae. Nat. Prod. Rep. *29*, 1317–1333.

Rodríguez-Celma, J., Lin, W.-D., Fu, G.-M., Abadía, J., López-Millán, A.-F., and Schmidt, W. (2013). Mutually exclusive alterations in secondary metabolism are critical for the uptake of insoluble iron compounds by Arabidopsis and Medicago truncatula. Plant Physiol. *162*, 1473–1485.

Romera, F.J., Lucena, C., García, M.J., Alcántara, E., and Pérez-Vicente, R. (2017). The Role of Ethylene and Other Signals in the Regulation of Fe Deficiency Responses by Dicot Plants. In Stress Signaling in Plants: Genomics and Proteomics Perspective, Volume 2, M. Sarwat, A. Ahmad, M.Z. Abdin, and M.M. Ibrahim, eds. (Springer International Publishing), pp. 277–300.

Rosler, J., Krekel, F., Amrhein, N., and Schmid, J. (1997). Maize Phenylalanine Ammonia-Lyase Has Tyrosine Ammonia-Lyase Activity. Plant Physiol. *113*, 175–179.

Ross, J., Li, Y., Lim, E.-K., and Bowles, D.J. (2001). Higher plant glycosyltransferases. Genome Biol. 2.

Rubio-Somoza, I., and Weigel, D. (2011). MicroRNA networks and developmental plasticity in plants. Trends Plant Sci. *16*, 258–264.

Ruegger, M., and Chapple, C. (2001). Mutations That Reduce Sinapoylmalate Accumulation in Arabidopsis thaliana Define Loci With Diverse Roles in Phenylpropanoid Metabolism. Genetics *159*, 1741–1749.

Saint von Paul, V., Zhang, W., Kanawati, B., Geist, B., Faus-Keßler, T., Schmitt-Kopplin, P., and Schäffner, A.R. (2011). The Arabidopsis Glucosyltransferase UGT76B1 Conjugates Isoleucic Acid and Modulates Plant Defense and

Senescence. Plant Cell 23, 4124–4145.

Sánchez-Campillo, M., Gabaldon, J.A., Castillo, J., Benavente-García, O., Del Baño, M.J., Alcaraz, M., Vicente, V., Alvarez, N., and Lozano, J.A. (2009). Rosmarinic acid, a photo-protective agent against UV and other ionizing radiations. Food Chem. Toxicol. Int. J. Publ. Br. Ind. Biol. Res. Assoc. 47, 386–392.

Sandermann, H.J., Scheel, D., and Trenck, T. v d (1983). Metabolism of Environmental Chemicals by Plants -Copolymerization into Lignin. J Appl Polym Sci Appl Polym Symp U. S.

Saslowsky, D.E., Warek, U., and Winkel, B.S.J. (2005). Nuclear localization of flavonoid enzymes in Arabidopsis. J. Biol. Chem. *280*, 23735–23740.

Satake, H., Ono, E., and Murata, J. (2013). Recent Advances in the Metabolic Engineering of Lignan Biosynthesis Pathways for the Production of Transgenic Plant-Based Foods and Supplements. J. Agric. Food Chem. *61*, 11721–11729.

Savidge, R.A. (1991). Notes: Seasonal Cambial Activity in Larix laricina Saplings in Relation to Endogenous Indol-3-ylacetic Acid, Sucrose, and Coniferin. For. Sci. *37*, 953–958.

Scharte, J., Schön, H., and Weis, E. (2005). Photosynthesis and carbohydrate metabolism in tobacco leaves during an incompatible interaction with Phytophthora nicotianae. Plant Cell Environ. *28*, 1421–1435.

Scheible, W.-R., and Pauly, M. (2004). Glycosyltransferases and cell wall biosynthesis: novel players and insights. Curr. Opin. Plant Biol. *7*, 285–295.

Schiestl, F.P., and Ayasse, M. (2001). Post-Pollination Emission of a Repellent Compound in a Sexually Deceptive Orchid: A New Mechanism for Maximising Reproductive Success? Oecologia *126*, 531–534.

Schiestl, F.P., Ayasse, M., Paulus, H.F., Löfstedt, C., Hansson, B.S., Ibarra, F., and Francke, W. (1999). Orchid pollination by sexual swindle. Nature *399*, 421–421.

Schilmiller, A.L., Last, R.L., and Pichersky, E. (2008). Harnessing plant trichome biochemistry for the production of useful compounds. Plant J. *54*, 702–711.

Schilmiller, A.L., Stout, J., Weng, J.-K., Humphreys, J., Ruegger, M.O., and Chapple, C. (2009). Mutations in the cinnamate 4-hydroxylase gene impact metabolism, growth and development in Arabidopsis. Plant J. Cell Mol. Biol. *60*, 771–782.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.

Schmid, G., Hammer, D.K., Ritterbusch, A., and Grisebach, H. (1982). Appearance and immunohistochemical localization of UDP-glucose: coniferyl alcohol glucosyltransferase in spruce (Picea abies (L.) Karst.) seedlings. Planta *156*, 207–212.

Schmid, N.B., Giehl, R.F.H., Döll, S., Mock, H.-P., Strehmel, N., Scheel, D., Kong, X., Hider, R.C., and von Wirén, N. (2014). Feruloyl-CoA 6'-Hydroxylase1-dependent coumarins mediate iron acquisition from alkaline substrates in Arabidopsis. Plant Physiol. *164*, 160–172.

Schmidt, A., Scheel, D., and Strack, D. (1998). Elicitor-stimulated biosynthesis of hydroxycinnamoyltyramines in cell suspension cultures of Solanum tuberosum. Planta *205*, 51–55.

Schmidt, T.J., Hemmati, S., Klaes, M., Konuklugil, B., Mohagheghzadeh, A., Ionkova, I., Fuss, E., and Wilhelm Alfermann, A. (2010). Lignans in flowering aerial parts of Linum species – Chemodiversity in the light of systematics and phylogeny. Phytochemistry *71*, 1714–1728.

Schneider, K., Hövel, K., Witzel, K., Hamberger, B., Schomburg, D., Kombrink, E., and Stuible, H.-P. (2003). The substrate specificity-determining amino acid code of 4-coumarate:CoA ligase. Proc. Natl. Acad. Sci. U. S. A. *100*, 8601–8606.

Schröder, G., Brown, J.W.S., and Schröder, J. (1988). Molecular analysis of resveratrol synthase. Eur. J.

Biochem. 172, 161–169.

Schuetz, M., Benske, A., Smith, R.A., Watanabe, Y., Tobimatsu, Y., Ralph, J., Demura, T., Ellis, B., and Samuels, A.L. (2014). Laccases direct lignification in the discrete secondary cell wall domains of protoxylem. Plant Physiol. *166*, 798–807.

Seca, A.M.L., and Silva, A.M.S. (2010). A new 4',7-epoxy-8,3'-oxyneolignan from the acetone extract of Juniperus brevifolia leaves. Phytochem. Lett. *3*, 126–128.

Seigler, D.S. (1998). Plant Secondary Metabolism (Springer Science & Business Media).

Serk, H., Gorzsás, A., Tuominen, H., and Pesquet, E. (2015). Cooperative lignification of xylem tracheary elements. Plant Signal. Behav. *10*, e1003753.

Shadle, G., Chen, F., Srinivasa Reddy, M.S., Jackson, L., Nakashima, J., and Dixon, R.A. (2007). Down-regulation of hydroxycinnamoyl CoA: Shikimate hydroxycinnamoyl transferase in transgenic alfalfa affects lignification, development and forage quality. Phytochemistry *68*, 1521–1529.

Shao, H., He, X., Achnine, L., Blount, J.W., Dixon, R.A., and Wang, X. (2005). Crystal structures of a multifunctional triterpene/flavonoid glycosyltransferase from Medicago truncatula. Plant Cell *17*, 3141–3154.

Sharma, R., Rawat, V., and Suresh, C.G. (2014). Genome-Wide Identification and Tissue-Specific Expression Analysis of UDP-Glycosyltransferases Genes Confirm Their Abundance in Cicer arietinum (Chickpea) Genome. PLOS ONE 9, e109715.

Shirley, B.W., Kubasek, W.L., Storz, G., Bruggemann, E., Koornneef, M., Ausubel, F.M., and Goodman, H.M. (1995). Analysis of Arabidopsis mutants deficient in flavonoid biosynthesis. Plant J. Cell Mol. Biol. *8*, 659–671.

Sisó-Terraza, P., Luis-Villarroya, A., Fourcroy, P., Briat, J.-F., Abadía, A., Gaymard, F., Abadía, J., and Álvarez-Fernández, A. (2016). Accumulation and Secretion of Coumarinolignans and other Coumarins in Arabidopsis thaliana Roots in Response to Iron Deficiency at High pH. Front. Plant Sci. *7*.

Šmehilová, M., Dobrůšková, J., Novák, O., Takáč, T., and Galuszka, P. (2016). Cytokinin-Specific Glycosyltransferases Possess Different Roles in Cytokinin Homeostasis Maintenance. Front. Plant Sci. *7*.

Smith, R.A., Schuetz, M., Roach, M., Mansfield, S.D., Ellis, B., and Samuels, L. (2013). Neighboring parenchyma cells contribute to Arabidopsis xylem lignification, while lignification of interfascicular fibers is cell autonomous. Plant Cell *25*, 3988–3999.

Smith, R.A., Schuetz, M., Karlen, S.D., Bird, D.A., Tokunaga, N., Sato, Y., Mansfield, S.D., Ralph, J., and Samuels, A.L. (2017). Defining the Diverse Cell Populations Contributing to Lignification in Arabidopsis thaliana Stems. Plant Physiol. pp.00434.2017.

Snegireva, A., Chernova, T., Ageeva, M., Lev-Yadun, S., and Gorshkova, T. (2015). Intrusive growth of primary and secondary phloem fibres in hemp stem determines fibre-bundle formation and structure. AoB PLANTS *7*.

Soler, M., Plasencia, A., Larbat, R., Pouzet, C., Jauneau, A., Rivas, S., Pesquet, E., Lapierre, C., Truchet, I., and Grima-Pettenati, J. (2016). The Eucalyptus linker histone variant EgH1.3 cooperates with the transcription factor EgMYB1 to control lignin biosynthesis during wood formation. New Phytol. *190*, 75–87

Sterjiades, R., Dean, J.F., and Eriksson, K.E. (1992). Laccase from Sycamore Maple (Acer pseudoplatanus) Polymerizes Monolignols. Plant Physiol. *99*, 1162–1168.

Sterjiades, R., Dean, J.F.D., Gamble, G., Himmelsbach, D.S., and Eriksson, K.-E.L. (1993). Extracellular laccases and peroxidases from sycamore maple (Acer pseudoplatanus) cell-suspension cultures. Planta *190*, 75–87.

Stracke, R., De Vos, R.C.H., Bartelniewoehner, L., Ishihara, H., Sagasser, M., Martens, S., and Weisshaar, B. (2009). Metabolomic and genetic analyses of flavonol synthesis in Arabidopsis thaliana support the in vivo involvement of leucoanthocyanidin dioxygenase. Planta *229*, 427–445.

Sun, H., Li, Y., Feng, S., Zou, W., Guo, K., Fan, C., Si, S., and Peng, L. (2013a). Analysis of five rice 4coumarate:coenzyme A ligase enzyme activity and stress response for potential roles in lignin and flavonoid biosynthesis in rice. Biochem. Biophys. Res. Commun. 430, 1151–1156.

Sun, Y.-G., Wang, B., Jin, S.-H., Qu, X.-X., Li, Y.-J., and Hou, B.-K. (2013b). Ectopic expression of Arabidopsis glycosyltransferase UGT85A5 enhances salt stress tolerance in tobacco. PloS One *8*, e59924.

Sunkar, R., Kapoor, A., and Zhu, J.-K. (2006). Posttranscriptional Induction of Two Cu/Zn Superoxide Dismutase Genes in Arabidopsis Is Mediated by Downregulation of miR398 and Important for Oxidative Stress Tolerance. Plant Cell *18*, 2051–2065.

Suzuki, S., and Umezawa, T. (2007). Biosynthesis of lignans and norlignans. J. Wood Sci. 53, 273–284.

Takabe, K., Fujita, M., Harada, H., and Saiki, H. (1985). Autoradiographic investigation of lignification in the cell walls of cryptomeria (Cryptomeria japanica D. Don). Mokuzai Gakkai-Shi *31*, 613–619.

Tattersall, D.B., Bak, S., Jones, P.R., Olsen, C.E., Nielsen, J.K., Hansen, M.L., Høj, P.B., and Møller, B.L. (2001). Resistance to an Herbivore Through Engineered Cyanogenic Glucoside Synthesis. Science *293*, 1826–1828.

Taylor-Teeples, M., Lin, L., de Lucas, M., Turco, G., Toal, T.W., Gaudinier, A., Young, N.F., Trabucco, G.M., Veling, M.T., Lamothe, R., et al. (2015). An Arabidopsis gene regulatory network for secondary cell wall synthesis. Nature *517*, 571–575.

Terashima, N., and Fukushima, K. (1988). Heterogeneity in formation of lignin—XI: An autoradiographic study of the heterogeneous formation and structure of pine lignin. Wood Sci. Technol. *22*, 259–270.

Terrier, N., Torregrosa, L., Ageorges, A., Vialet, S., Verriès, C., Cheynier, V., and Romieu, C. (2009). Ectopic expression of VvMybPA2 promotes proanthocyanidin biosynthesis in grapevine and suggests additional targets in the pathway. Plant Physiol. *149*, 1028–1041.

Thakur, V.K., and Thakur, M.K. (2015). Recent advances in green hydrogels from lignin: a review. Int. J. Biol. Macromol. *72*, 834–847.

Tobimatsu, Y., Chen, F., Nakashima, J., Escamilla-Treviño, L.L., Jackson, L., Dixon, R.A., and Ralph, J. (2013). Coexistence but Independent Biosynthesis of Catechyl and Guaiacyl/Syringyl Lignin Polymers in Seed Coats. Plant Cell *25*, 2587–2600.

Tokunaga, N., Sakakibara, N., Umezawa, T., Ito, Y., Fukuda, H., and Sato, Y. (2005). Involvement of extracellular dilignols in lignification during tracheary element differentiation of isolated Zinnia mesophyll cells. Plant Cell Physiol. *46*, 224–232.

Tsutsumi, Y., Matsui, K., and Sakai, K. (1998). Substrate-Specific Peroxidases in Woody Angiosperms and Gymnosperms Participate in Regulating the Dehydrogenative Polymerization of Syringyl and Guaiacyl Type Lignins. Holzforschung *52*, 275–281.

Tsuyama, T., and Takabe, K. (2014). Distribution of lignin and lignin precursors in differentiating xylem of Japanese cypress and poplar. J. Wood Sci. *60*, 353–361.

Tsuyama, T., Kawai, R., Shitan, N., Matoh, T., Sugiyama, J., Yoshinaga, A., Takabe, K., Fujita, M., and Yazaki, K. (2013). Proton-Dependent Coniferin Transport, a Common Major Transport Event in Differentiating Xylem Tissue of Woody Plants1[W]. Plant Physiol. *162*, 918–926.

Turlapati, P.V., Kim, K.-W., Davin, L.B., and Lewis, N.G. (2010). The laccase multigene family in Arabidopsis thaliana: towards addressing the mystery of their gene function(s). Planta *233*, 439–470.

Tyree, M.T., and Zimmermann, M.H. (2002). Xylem Structure and the Ascent of Sap (Berlin, Heidelberg: Springer Berlin Heidelberg).

Tzin, V., and Galili, G. (2010). The Biosynthetic Pathways for Shikimate and Aromatic Amino Acids in Arabidopsis thaliana. Arab. Book Am. Soc. Plant Biol. 8.

Umezawa, T. (2003). Phylogenetic Distribution of Lignan Producing Plants - Version details.

Vanholme, R., Demedts, B., Morreel, K., Ralph, J., and Boerjan, W. (2010). Lignin Biosynthesis and Structure. Plant Physiol. *153*, 895–905.

Vanholme, R., Storme, V., Vanholme, B., Sundin, L., Christensen, J.H., Goeminne, G., Halpin, C., Rohde, A., Morreel, K., and Boerjan, W. (2012). A Systems Biology View of Responses to Lignin Biosynthesis Perturbations in Arabidopsis[W]. Plant Cell *24*, 3506–3529.

Vannozzi, A., Dry, I.B., Fasoli, M., Zenoni, S., and Lucchin, M. (2012). Genome-wide analysis of the grapevine stilbene synthase multigenic family: genomic organization and expression profiles upon biotic and abiotic stresses. BMC Plant Biol. *12*, 130.

Vassão, D.G., Kim, S.-J., Milhollan, J.K., Eichinger, D., Davin, L.B., and Lewis, N.G. (2007). A pinoresinollariciresinol reductase homologue from the creosote bush (Larrea tridentata) catalyzes the efficient in vitro conversion of p-coumaryl/coniferyl alcohol esters into the allylphenols chavicol/eugenol, but not the propenylphenols p-anol/isoeugenol. Arch. Biochem. Biophys. *465*, 209–218.

Veillet, F., Gaillard, C., Coutos-Thévenot, P., and La Camera, S. (2016). Targeting the AtCWIN1 Gene to Explore the Role of Invertases in Sucrose Transport in Roots and during Botrytis cinerea Infection. Front. Plant Sci. *7*.

Venugopala, K.N., Rashmi, V., Odhav, B., Venugopala, K.N., Rashmi, V., and Odhav, B. (2013). Review on Natural Coumarin Lead Compounds for Their Pharmacological Activity, Review on Natural Coumarin Lead Compounds for Their Pharmacological Activity. BioMed Res. Int. BioMed Res. Int. *2013*, 2013, e963248.

Verdonk, J.C., Ric de Vos, C.H., Verhoeven, H.A., Haring, M.A., van Tunen, A.J., and Schuurink, R.C. (2003). Regulation of floral scent production in petunia revealed by targeted metabolomics. Phytochemistry *62*, 997–1008.

Vicente, M.R.-S., and Plasencia, J. (2011). Salicylic acid beyond defence: its role in plant growth and development. J. Exp. Bot. *62*, 3321–3338.

Villegas, M., Brodelius, P.E., and Kylin, A. (1990). Elicitor-induced hydroxycinnamoyl-CoA:tyramine hydroxycinnamoyltransferase in plant cell suspension cultures. Physiol. Plant. *78*, 414–420.

Vishwanath, S.J., Delude, C., Domergue, F., and Rowland, O. (2014). Suberin: biosynthesis, regulation, and polymer assembly of a protective extracellular barrier. Plant Cell Rep. *34*, 573–586.

Vogt, T. (2002). Substrate specificity and sequence analysis define a polyphyletic origin of betanidin 5- and 6-O-glucosyltransferase from Dorotheanthus bellidiformis. Planta *214*, 492–495.

Vogt, T. (2010). Phenylpropanoid Biosynthesis. Mol. Plant 3, 2–20.

Voinnet, O. (2009). Origin, Biogenesis, and Activity of Plant MicroRNAs. Cell 136, 669-687.

Voxeur, A., Wang, Y., and Sibout, R. (2015). Lignification: different mechanisms for a versatile polymer. Curr. Opin. Plant Biol. *23*, 83–90.

Waese, J., Pasha, A., Wang, T.T., van Weringh, A., Guttman, D.S., and Provart, N.J. (2016). Gene Slider: sequence logo interactive data-visualization for education and research. Bioinforma. Oxf. Engl. *32*, 3670–3672.

Wagner, S.C. (2011). Biological Nitrogen Fixation. Nat. Educ. Knowl. 3.

Wagner, A., Tobimatsu, Y., Phillips, L., Flint, H., Torr, K., Donaldson, L., Pears, L., and Ralph, J. (2011). CCoAOMT suppression modifies lignin composition in Pinus radiata. Plant J. Cell Mol. Biol. *67*, 119–129.

Wang, B., and Qiu, Y.-L. (2006). Phylogenetic distribution and evolution of mycorrhizas in land plants. Mycorrhiza 16, 299–363.

Wang, C.-H., Yu, J., Cai, Y.-X., Zhu, P.-P., Liu, C.-Y., Zhao, A.-C., Lü, R.-H., Li, M.-J., Xu, F.-X., and Yu, M.-D. (2016). Characterization and Functional Analysis of 4-Coumarate:CoA Ligase Genes in Mulberry. PLOS ONE *11*, e0155814.

Wang, C.-Y., Zhang, S., Yu, Y., Luo, Y.-C., Liu, Q., Ju, C., Zhang, Y.-C., Qu, L.-H., Lucas, W.J., Wang, X., et al. (2014a). MiR397b regulates both lignin content and seed number in Arabidopsis via modulating a laccase involved in lignin biosynthesis. Plant Biotechnol. J. *12*, 1132–1142.

Wang, J., Ma, X.-M., Kojima, M., Sakakibara, H., and Hou, B.-K. (2011). N-glucosyltransferase UGT76C2 is involved in cytokinin homeostasis and cytokinin response in Arabidopsis thaliana. Plant Cell Physiol. *52*, 2200–2213.

Wang, J., Kucukoglu, M., Zhang, L., Chen, P., Decker, D., Nilsson, O., Jones, B., Sandberg, G., and Zheng, B.

(2013b). The Arabidopsis LRR-RLK, PXC1, is a regulator of secondary wall formation correlated with the TDIF-PXY/TDR-WOX4 signaling pathway. BMC Plant Biol. *13*, 94.

Wang, J.P., Naik, P.P., Chen, H.-C., Shi, R., Lin, C.-Y., Liu, J., Shuford, C.M., Li, Q., Sun, Y.-H., Tunlaya-Anukit, S., et al. (2014b). Complete Proteomic-Based Enzyme Reaction and Inhibition Kinetics Reveal How Monolignol Biosynthetic Enzyme Families Affect Metabolic Flux and Lignin in Populus trichocarpa. Plant Cell *26*, 894–914.

Wang, J.P., Chuang, L., Loziuk, P.L., Chen, H., Lin, Y.-C., Shi, R., Qu, G.-Z., Muddiman, D.C., Sederoff, R.R., and Chiang, V.L. (2015). Phosphorylation is an on/off switch for 5-hydroxyconiferaldehyde O-methyltransferase activity in poplar monolignol biosynthesis. Proc. Natl. Acad. Sci. *112*, 8481–8486.

Wang, Y., Chantreau, M., Sibout, R., and Hawkins, S. (2013b). Plant cell wall lignification and monolignol metabolism. Plant Biotechnol. 4, 220.

Wang, Y.-W., Wang, W.-C., Jin, S.-H., Wang, J., Wang, B., and Hou, B.-K. (2012a). Over-expression of a putative poplar glycosyltransferase gene, PtGT1, in tobacco increases lignin content and causes early flowering. J. Exp. Bot. *63*, 2799–2808.

Wang, Z., Hobson, N., Galindo, L., Zhu, S., Shi, D., McDill, J., Yang, L., Hawkins, S., Neutelings, G., Datla, R., et al. (2012b). The genome of flax (Linum usitatissimum) assembled de novo from short shotgun sequence reads. Plant J. 72, 461–473.

Wang, Z., Hobson, N., Galindo, L., Zhu, S., Shi, D., McDill, J., Yang, L., Hawkins, S., Neutelings, G., Datla, R., et al. (2012c). The genome of flax (Linum usitatissimum) assembled de novo from short shotgun sequence reads. Plant J. Cell Mol. Biol. *72*, 461–473.

Wani, S.H., Kumar, V., Shriram, V., and Sah, S.K. (2016). Phytohormones and their metabolic engineering for abiotic stress tolerance in crop plants. Crop J. *4*, 162–176.

Weng, J.-K., and Chapple, C. (2010). The origin and evolution of lignin biosynthesis. New Phytol. *187*, 273–285.

Weng, J.-K., Akiyama, T., Bonawitz, N.D., Li, X., Ralph, J., and Chapple, C. (2010). Convergent Evolution of Syringyl Lignin Biosynthesis via Distinct Pathways in the Lycophyte Selaginella and Flowering Plants. Plant Cell *22*, 1033–1045.

Wetterhorn, K.M., Newmister, S.A., Caniza, R.K., Busman, M., McCormick, S.P., Berthiller, F., Adam, G., and Rayment, I. (2016). Crystal Structure of 0s79 (0s04g0206600) from Oryza sativa: A UDP-glucosyltransferase Involved in the Detoxification of Deoxynivalenol. Biochemistry (Mosc.) *55*, 6175–6186.

Whetten, R., Sun, Y.H., Zhang, Y., and Sederoff, R. (2001). Functional genomics and cell wall biosynthesis in loblolly pine. Plant Mol. Biol. *47*, 275–291.

Wilkerson, C.G., Mansfield, S.D., Lu, F., Withers, S., Park, J.-Y., Karlen, S.D., Gonzales-Vigil, E., Padmakshan, D., Unda, F., Rencoret, J., et al. (2014). Monolignol ferulate transferase introduces chemically labile linkages into the lignin backbone. Science *344*, 90–93.

Wingler, A., and Roitsch, T. (2008). Metabolic regulation of leaf senescence: interactions of sugar signalling with biotic and abiotic stress responses. Plant Biol. Stuttg. Ger. *10 Suppl 1*, 50–62.

Winkel-Shirley, B. (1999). Evidence for enzyme complexes in the phenylpropanoid and flavonoid pathways. Physiol. Plant. *107*, 142–149.

Winkel-Shirley, B. (2001). Flavonoid Biosynthesis. A Colorful Model for Genetics, Biochemistry, Cell Biology, and Biotechnology. Plant Physiol. *126*, 485–493.

Wolf, S., Mravec, J., Greiner, S., Mouille, G., and Höfte, H. (2012). Plant Cell Wall Homeostasis Is Mediated by Brassinosteroid Feedback Signaling. Curr. Biol. *22*, 1732–1737.

Woo, H.-H., Jeong, B.R., Hirsch, A.M., and Hawes, M.C. (2007). Characterization of Arabidopsis AtUGT85A and AtGUS gene families and their expression in rapidly dividing tissues. Genomics *90*, 143–153.

Woo, H.R., Kim, H.J., Nam, H.G., and Lim, P.O. (2013). Plant leaf senescence and death – regulation by multiple layers of control and implications for aging in general. J Cell Sci *126*, 4823–4833.

Wu, B., Gao, L., Gao, J., Xu, Y., Liu, H., Cao, X., Zhang, B., and Chen, K. (2017). Genome-Wide Identification, Expression Patterns, and Functional Analysis of UDP Glycosyltransferase Family in Peach (Prunus persica L. Batsch). Front. Plant Sci. *8*.

Xu, W., Lepiniec, L., and Dubos, C. (2014). New insights toward the transcriptional engineering of proanthocyanidin biosynthesis. Plant Signal. Behav. *9*, e28736.

Yamaguchi, M., and Demura, T. (2010). Transcriptional regulation of secondary wall formation controlled by NAC domain proteins. Plant Biotechnol. *27*, 237–242.

Yamauchi, K., Yasuda, S., Hamada, K., Tsutsumi, Y., and Fukushima, K. (2003). Multiform biosynthetic pathway of syringyl lignin in angiosperms. Planta *216*, 496–501.

Yang, Z., Zhang, T., Lang, T., Li, G., Chen, G., and Nevo, E. (2013). Transcriptome Comparative Profiling of Barley eibi1 Mutant Reveals Pleiotropic Effects of HvABCG31 Gene on Cuticle Biogenesis and Stress Responsive Pathways. Int. J. Mol. Sci. *14*, 20478–20491.

Yap, V.A., Loong, B.-J., Ting, K.-N., Loh, S.H.-S., Yong, K.-T., Low, Y.-Y., Kam, T.-S., and Lim, K.-H. (2015). Hispidacine, an unusual 8,4'-oxyneolignan-alkaloid with vasorelaxant activity, and hispiloscine, an antiproliferative phenanthroindolizidine alkaloid, from Ficus hispida Linn. Phytochemistry *109*, 96–102.

Yeats, T.H., and Rose, J.K.C. (2013). The Formation and Function of Plant Cuticles. Plant Physiol. 163, 5-20.

Yonekura-Sakakibara, K., and Hanada, K. (2011). An evolutionary view of functional diversity in family 1 glycosyltransferases. Plant J. Cell Mol. Biol. *66*, 182–193.

Yonekura-Sakakibara, K., Tohge, T., Niida, R., and Saito, K. (2007). Identification of a Flavonol 7-0-Rhamnosyltransferase Gene Determining Flavonoid Pattern in Arabidopsis by Transcriptome Coexpression Analysis and Reverse Genetics. J. Biol. Chem. *282*, 14932–14941.

Yoon, J., Choi, H., and An, G. (2015). Roles of lignin biosynthesis and regulatory genes in plant development. J. Integr. Plant Biol. *57*, 902–912.

Zamioudis, C., Hanson, J., and Pieterse, C.M.J. (2014). β-Glucosidase BGLU42 is a MYB72-dependent key regulator of rhizobacteria-induced systemic resistance and modulates iron deficiency responses in Arabidopsis roots. New Phytol. *204*, 368–379.

Zeng, Y., Zhao, S., Yang, S., and Ding, S.-Y. (2014). Lignin plays a negative role in the biochemical process for producing lignocellulosic biofuels. Curr. Opin. Biotechnol. *27*, 38–45.

Zhang, N., and Deyholos, M.K. (2016). RNASeq Analysis of the Shoot Apex of Flax (Linum usitatissimum) to Identify Phloem Fiber Specification Genes. Front. Plant Sci. 7.

Zhang, G.-Z., Jin, S.-H., Jiang, X.-Y., Dong, R.-R., Li, P., Li, Y.-J., and Hou, B.-K. (2016). Ectopic expression of UGT75D1, a glycosyltransferase preferring indole-3-butyric acid, modulates cotyledon development and stress tolerance in seed germination of Arabidopsis thaliana. Plant Mol. Biol. *90*, 77–93.

Zhang, X., Gou, M., and Liu, C.-J. (2013a). Arabidopsis Kelch repeat F-box proteins regulate phenylpropanoid biosynthesis via controlling the turnover of phenylalanine ammonia-lyase. Plant Cell *25*, 4994–5010.

Zhang, Y.-C., Yu, Y., Wang, C.-Y., Li, Z.-Y., Liu, Q., Xu, J., Liao, J.-Y., Wang, X.-J., Qu, L.-H., Chen, F., et al. (2013b). Overexpression of microRNA OsmiR397 improves rice yield by increasing grain size and promoting panicle branching. Nat. Biotechnol. *31*, 848–852.

Zhao, Q. (2016). Lignification: Flexibility, Biosynthesis and Regulation. Trends Plant Sci.

Zhao, J., Huhman, D., Shadle, G., He, X.-Z., Sumner, L.W., Tang, Y., and Dixon, R.A. (2011). MATE2 Mediates Vacuolar Sequestration of Flavonoid Glycosides and Glycoside Malonates in Medicago truncatula. Plant Cell *23*, 1536– 1555. Zhao, Q., Nakashima, J., Chen, F., Yin, Y., Fu, C., Yun, J., Shao, H., Wang, X., Wang, Z.-Y., and Dixon, R.A. (2013). Laccase is necessary and nonredundant with peroxidase for lignin polymerization during vascular development in Arabidopsis. Plant Cell *25*, 3976–3987.

Zhao, Y., Sun, J., Xu, P., Zhang, R., and Li, L. (2014). Intron-mediated alternative splicing of WOOD-ASSOCIATED NAC TRANSCRIPTION FACTOR1B regulates cell wall thickening during fiber development in Populus species. Plant Physiol. *164*, 765–776.

Zhong, R., and Ye, Z.-H. (2009). Secondary Cell Walls. In eLS, (John Wiley & Sons, Ltd).

Zhong, R., and Ye, Z.-H. (2012). MYB46 and MYB83 bind to the SMRE sites and directly activate a suite of transcription factors and secondary wall biosynthetic genes. Plant Cell Physiol. *53*, 368–380.

Zhong, R., 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.

Zhong, R., Richardson, E.A., and Ye, Z.-H. (2007a). Two NAC domain transcription factors, SND1 and NST1, function redundantly in regulation of secondary wall synthesis in fibers of Arabidopsis. Planta *225*, 1603–1611.

Zhong, R., Richardson, E.A., and Ye, Z.-H. (2007b). The MYB46 Transcription Factor Is a Direct Target of SND1 and Regulates Secondary Wall Biosynthesis in Arabidopsis. Plant Cell *19*, 2776–2792.

Zhong, R., Lee, C., Zhou, J., McCarthy, R.L., and Ye, Z.-H. (2008). A battery of transcription factors involved in the regulation of secondary cell wall biosynthesis in Arabidopsis. Plant Cell *20*, 2763–2782.

Zhong, R., Lee, C., and Ye, Z.-H. (2010). Global analysis of direct targets of secondary wall NAC master switches in Arabidopsis. Mol. Plant *3*, 1087–1103.

Zhou, J., Lee, C., Zhong, R., and Ye, Z.-H. (2009). MYB58 and MYB63 are transcriptional activators of the lignin biosynthetic pathway during secondary cell wall formation in Arabidopsis. Plant Cell *21*, 248–266.

Zhou, J., Zhong, R., and Ye, Z.-H. (2014). Arabidopsis NAC Domain Proteins, VND1 to VND5, Are Transcriptional Regulators of Secondary Wall Biosynthesis in Vessels. PLOS ONE *9*, e105726.

Annexes

<u>Annexe 1</u>

La glycosylation est un régulateur majeur de la disponibilité et de l'activité biologique des phénylpropanoïdes chez les plantes

Avant propos

Parmi tous les métabolites spécialisés, les phénylpropanoïdes constituent un ensemble de molécules d'une grande diversité et essentiels pour la plante. Il est connu depuis longtemps que les phénylpropanoïdes interviennent aussi bien lors de stress biotiques et abiotiques que pour contribuer au développement de la plante. Leur grande diversité s'explique par le grand nombre d'enzymes "secondaires" qui vont venir décorer les métabolites à la suite de leur biosynthèse. En effet, des modifications comme des méthylations, acylations, prenylations, oxidations, réductions ou glycosylations vont considérablement impacter les propriétés physico-chimiques des molécules et de ce fait, la physiologie de la plante (Vogt 2010). La glycosylation des phénylpropanoïdes est dans certains cas une des dernières étapes de biosynthèse comme pour quelques flavonoïdes (Shirley et al. 1995), et dans d'autre cas, elle intervient en réponse à l'infection de virus en glycosylant d'autres phénylpropanoïdes comme des coumarines (Chong et al. 2002). Mais dans la majorité des cas, la raison pour laquelle la glycosylation de phénylpropanoïdes a lieu reste encore inexpliquée.

Dans un premier temps, cet article permet de faire un état des lieux des fonctions connues de glycosylation des phénylpropanoïdes. Il est fait état du rôle de la glycosylation dans le développement des plantes avec une attention particulière pour le lien entre la glycosylation des monolignols et leur impact sur la lignification. Cet article de synthèse s'intéresse également aux liens entres la glycosylation des phénylpropanoïdes et la tolérance/sensibilité des plantes face aux pathogènes ou les UVs.

Dans un second temps, des analyses de co-expression entre les UGTs d'*Arabidopsis thaliana* et les gènes de biosynthèse des phénylpropanoïdes ont été analysées pour permettre de mieux comprendre la régulation des UGTs et, *in fine*, leurs fonctions.

L'article est publié ici sous sa forme soumise et acceptée à Frontiers in Plant Science en 2016. Sa forme définitive est disponible en ligne sur le site du journal (DOI: <u>10.3389/fpls.2016.00735</u>)

Article: Glycosylation is a major regulator of phenylpropanoid availability and biological activity in plants.

Julien Le Roy, Brigitte Huss, Anne Creach, Simon Hawkins, Godfrey Neutelings*

Université de Lille, CNRS, UMR 8576 - UGSF - Unité de Glycobiologie Structurale et Fonctionnelle, Lille, F-59000, France

Correspondence:

Godfrey Neutelings godfrey.neutelings@univ-lille1.fr

Abstract:

The phenylpropanoid pathway in plants is responsible for the biosynthesis of a huge amount of secondary metabolites derived from phenylalanine and tyrosine. Both flavonoids and lignins are synthesized at the end of this very diverse metabolic pathway, as well as many intermediate molecules whose precise biological functions remain largely unknown. The diversity of these molecules can be further increased under the action of UDP-glycosyltransferases (UGTs) leading to the production of glycosylated hydroxycinnamates and related aldehydes, alcohols and esters. Glycosylation can change phenylpropanoid solubility, stability and toxic potential, as well as influencing compartmentalization and biological activity. (De)-glycosylation therefore represents an extremely important regulation point in phenylpropanoid homeostasis. In this article we review recent knowledge on the enzymes involved in regulating phenylpropanoid glycosylation status and availability in different subcellular compartments. We also examine the potential link between monolignol glycosylation and lignification by exploring co-expression of lignin biosynthesis genes and phenolic (de)glycosylation genes. Of the different biological roles linked with their particular chemical properties, phenylpropanoids are often correlated with the plant's stress management strategies that are also regulated by glycosylation. UGTs can for instance influence the resistance of plants during infection by microorganisms and be involved in the mechanisms related to environmental changes. The impact of flavonoid glycosylation on the color of flowers, leaves, seeds and fruits will also be discussed. Altogether this paper underlies the fact that glycosylation and deglycosylation are powerful mechanisms allowing plants to regulate phenylpropanoid localisation, availability and biological activity.

Keywords: phenylpropanoids, glycosylation, UDP-glycosyltransferase, betaglucosidase, lignin, flavonoids, compartmentalization

1. Introduction

Terrestrialization was probably one of the most important evolutionary steps in the history of life. After the transition from water to land, a part of the green lineage was anchored in the soil and became sessile. Unable to move, plants are continuously subjected to UV radiations, large temperature variations and other extreme growing conditions currently accentuated by climate change such as drought, waterlogging and cold stresses as well as increased ozone levels and other industrial pollutants. Furthermore, plants are attacked by a wide range of pathogens including viruses, fungi, nematodes and bacteria, and are also targeted by herbivores that use plants for food. Finally, since the development of agriculture, plants must increasingly deal with different types of pesticides. Both biotic and abiotic stresses represent unfavourable growth conditions negatively affecting plant growth and survival and leading to dramatic crop losses. On the other hand, plants also depend on abiotic and biotic agents for pollen and seed dispersal, on symbiotic microorganisms such as mycorrhizal fungi for water and mineral absorption, as well as rhizobia nitrogen-fixing bacteria in leguminous species. To manage both negative and positive interactions with the environment and other forms of life, individual plants species continually synthesize a complex and changing mixture of different secondary metabolites including phenylpropanoids, phenylpropanoid-derived compounds, terpenoids and alkaloids to respond to specific environmental situations (Timell, 1986; Cheynier et al., 2013; Mierziak et al., 2014; Pusztahelyi et al., 2015; Holbein et al., 2016; Ishihara et al., 2016). Although a recent estimate put the number of secondary metabolites at more than 100,000 (Gachon et al., 2005) across plant species, it is likely that the real number is far higher. The increasing interest for metabolomics during the past decade has accelerated improvements in the sensitivity and specificity of spectrometric- and chromatographic-based technics as well as nuclear magnetic resonance (NMR) approaches.

Phenylpropanoids are a large group of secondary metabolites containing a phenyl group linked to a 3-C propane side chain. Both the position of the propenyl double bond and the substituents on the benzene ring result in different compounds with diverse biological activities (Koeduka et al., 2014). In many different species, the genes involved in the phenylpropanoid pathway belong to multigenic families (Ehlting et al., 1999) and are mainly regulated by MYB and NAC transcription factors (Sablowski et al., 1994; Zhong et al., 2006). The first enzymatic step is catalysed by phenylalanine ammonia lyase (PAL), which converts phenylalanine issued from the shikimate pathway into cinnamic acid (figure 1), which in turn is transformed to *p*-coumaric acid by cinnamate 4-hydrolase (C4H). Hydroxycinnamates (HCAs) derived from p-coumaric acid, CoA esters or hydroxycinnamaldehydes constitute the simplest phenylpropanoids (Strack, 2001). They can accumulate as ester- or amide-conjugates with monosaccharides, organic acids, lipids and amines and, when activated with CoenzymeA (CoA) or glucose, become acyl donors for modifications of secondary metabolites. HCAs are involved in the plant's response to pathogens and interactions with symbionts (Babst et al., 2014). Grass cell walls are characterized by large amounts of HCAs such as ferulic and pcoumaric acids. In these species, ferulic acid binds to lignin, polysaccharides and structural proteins and leads to their cross-linking but they are also present in smaller quantities in the cell walls of most dicots (for review see de Oliveira et al. (2015)). In sugarcane, sinapic, *p*-coumaric and ferulic acids are deposited during the early stages of lignification, and levels of hydroxycinnamic acids were higher in parenchyma walls than in the vascular bundles (He and Terashima, 1990; 1991). They can also be integrated into lignins being involved in ligno-polysaccharide cross-links (Tobimatsu et al., 2012). The subsequent production of *p*-coumaroyl CoA by 4-coumaroyl CoA-ligase (4CL) constitutes an important branch point (Ehlting et al., 1999) leading to the biosynthesis of the different subfamilies of phenylpropanoid-derived molecules (Vogt, 2010), the most important being undoubtedly lignin, the second most abundant plant polymer on the Earth's surface.

Lignin heteropolymers are derived mainly from three different monolignols: *p*coumaryl alcohol, coniferyl alcohol, and sinapyl alcohol, synthesized in the cytoplasm through at least ten different enzymatic reactions and are oxidized by laccases and/or peroxidases prior to polymerization into the lignin polymer (Freudenberg, 1968; Sterjiades et al., 1992; Boerjan et al., 2003). Instead of being exported to the apoplast, monolignols can also be stored as glycosylated conjugates in the vacuole (Dima et al., 2015). Lignins play a crucial role during the plant's life by providing cell integrity, mechanical support, water transport, assimilation through the vascular system and are also involved in defense mechanisms. (Neo)lignans are found in a wide range of species and represent a large group of diverse phenolics derived from the oxidative dimerization of two monolignols. Lignans are dimerized on their C3 side chains in a tail to tail structure whereas neolignans result from a head to tail dimerization (Cheynier et al., 2013). Flaxseed lignans have gained particular interest since some of them (e.g. SDG: secoisolariciresinol diglucoside) have potential antimutagenic, anti-inflammatory, antioxydant, antimicrobial and antiobesity effects (Imran et al., 2015).

Flavonoids are ubiquitously present across the plant kingdom. They are synthesized from the condensation of *p*-coumaroyl-CoA with three malonyl-CoA molecules by a chalcone synthase (CHS), which in turn produces a flavanone containing a 2-phenylchroman backbone. The same 2-phenylchroman backbone also gives rise to flavanols, isoflavonoids, flavonols, flavones and anthocyanidins. These compounds do not have a typical phenylpropanoid structure and are therefore considered as phenylpropanoid-derived compounds (Stafford, 1991). The different carbons of the flavonoid skeleton show numerous substitutions catalysed by isomerases, reductases, hydroxylases, methylases and glycosyltransferases leading to high structural diversity (Harborne and Williams, 2001; Ferrer et al., 2008).

Coumarins play roles in plant defense against fungi and possess both antioxidative and antimicrobial activities (Kai et al., 2006) as they can scavenge reactive oxygen intermediates produced during the hypersensitive response (Chong et al., 2002a). The core structure derives from cinnamates via *ortho*-hydroxylation of the aromatic ring, *trans/cis* isomerization and lactonization. They are widely distributed in

the plant kingdom and are classified into several categories depending on whether they harbour simple hydroxyl, alkoxy and/or alkyl groups in the benzene ring, or if additional benzene rings are present (Shimizu, 2014). Only 12 plant families have been listed to synthesize stilbenes from flavonoids precursors through stilbene synthases (Holl et al., 2013). There is a particular interest for *trans*-resveratrol synthesized in grapevine, cranberry, blueberry and peanut because of its pharmacological properties associated with protection towards cardiovascular, neurodegenerative diseases, cancer and diabetes. Finally phenylpropenes are volatiles emitted to attract both pollinators - through fragrances and seed dispersers through fruit scents. When they are produced in vegetative tissues, they also act as deterrents, toxicants and antifeedants to repel herbivores and bacterial pathogens.

The basic phenylpropanoid molecular backbone is synthesised in one of the numerous core metabolic pathways invented by plants during evolution, but the massive expansion of these different subfamilies of molecules is mainly due to the grafting of side chains on the basal structures. Among these modifications, glycosylation appears as a major reaction leading not only to an increase in structural diversity but also to an increase in the diversity of their properties. The glycosylation status of phenylpropanoids is regulated by the combination of regioselective glycosyltransferases (GTs) and glycoside hydrolases (GHs) with relatively broad substrate spectrum. Glycosyltransferases catalyse the transfer of sugar moieties from activated donor molecules to specific acceptor molecules such as sugars, lipids, proteins or small molecules including phenylpropanoids. Information on GHs and GTs, as well as polysaccharide lyases and carbohydrate esterases, can be found in the CarbohydrateActive enZymes (CAZy) database (Davies et al., 2005) and a specific PlantCAZyme database is now also available (Ekstrom et al., 2014).

CAZy takes into account similarities in amino acid sequences, 3D-structures, sugar donors, transferred sugars, acceptors, catalytic mechanisms (inverting or retaining mechanisms) and more recently the combination of modules which can be catalytic or not (Lombard et al., 2014). Among the 98 GT families present in the CAZy database in may 2016, the GTs catalysing secondary metabolite conjugation all belong to the GT1 family and are commonly named UDP-glycosyltransferases (UGTs) as UDP-

sugars are used as the sugar donor. When they act on the conjugation of secondary metabolites, most UGTs perform O-glycosylation, but on some xenobiotics, UDP-glucose can act as sugar donor for S-, C- or N-glycosylations. Pollutants such as 2,4,5-trichlorophenyl and 3,4-dichloroaniline can potentially be metabolized by 44 different UGTs in *Arabidopsis* through O- and N-glucosylation, respectively, whereas one UGT (UGT72B1) was shown to act as a bifunctional enzyme carrying both activities (Brazier-Hicks et al., 2007). Conjugated xenobiotics may then be exported from the cell to the apoplast or imported into the vacuole. C-glucosylation of 2-hydroxyflavanone precursors of flavonoids was also reported for the rice OsCGT (Brazier-Hicks et al., 2009). S-glycosylation takes place during the biosynthesis of glucosinolates involved in plant defense against pathogen and herbivores (Grubb et al., 2004).

When an acceptor presents multiple possible binding sites for a sugar, UGTs exhibit regioselectivity by transferring the sugar to a specific position. This was demonstrated mainly for different coumarins (Lim et al., 2003) and flavonoids for which prediction of regioselectivity was assessed (Jackson et al., 2011). The glycosylation status is also regulated by glycoside hydrolases (GHs) involved in hydrolysis and/or rearrangement of glycosidic bonds. Among these, several β -glucosidase subfamilies can remove the non-reducing terminal β -D-glucosyl from glucoconjugates. Most β -glucosidases are included in the CAZy GH1 family (Rouyi et al., 2014) but others can be found in GH3, GH5, GH9, GH30 and GH116 families. The different GH families contain enzymes showing different substrate specificities (Lombard et al., 2014) and most likely broad range specificities because of the high number of glycoconjugates within the plant compared to the number of β -glucosidases (Ketudat Cairns et al., 2015).

In this present review we will first summarize the current knowledge on UGTs and β -glucosidases related to the glycosylation status of phenylpropanoids and then we will evaluate the impact of sugar conjugation on their compartmentation. The function of these glucoconjugates will also be discussed in the context of lignification, development and the stress response.

2. Catalytic activities regulating phenylpropanoid glycosylation

2.1 UDP-glycosyltransferases

UGTs are usually named with a specific nomenclature including family numbers 1-50 used for animals, 51-70 for yeasts, 71-100 for plants and 101-200 for bacteria followed by a letter for the sub-family defined by proteins sharing at least 60% homology, and finally an arabic number to describe the gene (Mackenzie et al., 1997; Ross et al., 2001). According to the nomenclature committee of the international union of biochemistry and molecular biology (IUBMB), they are also named as EC 2.4.1.x, EC 2.4.2.x and 2.4.99.x. UGTs include enzymes able to transfer a sugar linked to uridine diphosphate (UDP) to a large range of acceptors. In plants, the nucleotidic sugar donor is mostly glucose but can also be galactose, xylose, rhamnose, arabinose or glucuronic acid, the latter also being the main sugar derivative donor for UGTs in mammalians (Ross et al., 2001; Osmani et al., 2008; Yonekura-Sakakibara et al., 2008). The conjugation of acceptor molecules such as hormones, xenobiotics or secondary metabolites by UGTs allows the plant cell to modulate their biochemical proprieties and thus has a strong influence on their biological activity and compartment storage. Although the primary sequences of the UGT proteins show only low similarities (Vogt and Jones, 2000), a highly conserved 44-amino acid sequence named plant secondary product glycosyltransferase (PSPG)-box believed to be involved in binding to the UDP moiety, is present in the C-terminal region (Mackenzie et al., 1997). This sequence is therefore commonly used as a query to identify the UGT genes in plant genomes by using the Blastp program in addition to position-specific weight matrix (PSWM) guided search and a hidden Markov model (HMMER) program (figure 2). The sequences can then be confirmed by using the MEME suite or the Pfam (http://pf am.sanger.ac.uk/search), **SMART** (http://smart.embl-heidelberg.de/) and Interpro (https://www.ebi.ac.uk/interpro/) databases. The very large spectrum of molecules potentially glycosylated by UGTs is in accordance with the high number of sequences identified at a genome-wide scale in plants as compared to other kingdoms including animals. However there does not seem to be any relationship between the number of predicted UGT models and the taxonomic families such as liliopsida and eudicotyledons.

Among the numerous phenylpropanoids, flavonoids are probably among the best-characterized molecules in terms of glycosylation due to their numerous medical and commercial benefits. The basic structure of flavonoids consists of two phenyl groups joined by three carbons cyclized with oxygen (figure 1). On this double ring, the carbons are annotated from 2 to 8 and on the third ring anchored on C2 they are annotated from 1' to 6'. Successive glycosylation by UGTs is possible on ring positions bearing hydroxyl groups mostly on C3 and C7. In Arabidopsis, three homologous enzymes belonging to the UGT78D family catalyse the first 3-O-glycosylation steps on flavonols. These three enzymes share the same aglycone substrates kaempferol, quercetin and isorhamnetin but differ by their specificity towards the donor sugar, with UGT78D1 using UDPrhamnose, UGT78D2 using UDP-glucose and UGT78D3 using UDP-arabinose (Yonekura-Sakakibara et al., 2008). The following step consists of a 7-0-glycosylation during which rhamnose is transferred by UGT89C1 or glucose by UGT73C6. Whether this transfer takes place or not, a secondary glycosylation step is also possible on the conjugated C3 by addition of a glucose on the flavonoid 3-O-glucoside by UGT79B6 (Yonekura-Sakakibara et al., 2014), or a rhamnose via a 1'-2" link forming the triglycoside kaempferol 3-O-beta-[alpha-rhamnopyranosyl(1-->2)-glucopyranoside]-7-O-alpharhamnopyranoside. In the same metabolic pathway, recombinant protein or mutagenesis approaches showed that the production of anthocyanins also depended on the transfer of sugar moieties onto the carbon rings. Both UGT78D2 and UGT75C1 were identified in Arabidopsis as glucosyltransferases specific to the C3 and C5 positions, respectively (Tohge et al., 2005; Kubo et al., 2007) and UGT79B1 was shown to be involved in anthocyanin biosynthesis by converting cyanidin 3-O-glucoside to cyanidin 3-0-xylosyl(1 \rightarrow 2)glucoside (Yonekura-Sakakibara et al., 2012). To a lesser extent the same enzyme is also capable of recognizing cyanidin 3-O-rhamnosyl $(1\rightarrow 2)$ glucoside as substrate. Beyond the Arabidopsis model, numerous different UGT activities on flavonoids were identified in medicinal and aromatic plants: a flavonoid 7-0glucuronosyltransferase in Perilla frutescens and Scutellaria laeteviolacea (Noguchi et al., 2009), a flavonoid glucoside: 6"-O-glycosyltransferase in Catharanthus roseus (Masada et al., 2009) or a flavonol 3-O-glucoside: 2"-O-glucosyltransferase in Crocus sativus (Trapero et al., 2012). The identification of the complete plant set of flavonoidassociated UGTs is still ongoing and major advances will undoubtedly be made with the growing number of full genome sequences available.

The identification and characterization of UGTs used to construct final complex molecular structures such as flavonoids or anthocyanins can be quite straightforward but is less obvious for those acting on more upstream intermediates such as hydroxycinnamic acids, aldehydes or alcohols. The role of these UGTs may indeed be difficult to reveal because of the proximity of related UGT genes in the Arabidopsis genome preventing generation of multiple mutants (Meißner et al., 2008). The use of techniques allowing fine gene targeting such as CRISPR-Cas9 may allow circumventing this constraint in the future. In addition the presence of overlapping substrates between several enzymes and the broad range of substrates recognized under *in vitro* conditions by each enzyme may also be a source of difficulty.. One example is the TOGT1 recombinant enzyme derived from a tobacco gene and active towards ferulic, caffeic, cinnamic, p-coumaric and o-coumaric acids as well as conifervl alcohol (Fraissinet-Tachet et al., 1998) although it is important to be cautious with respect to the results obtained by in vitro activity approaches. In Arabidopsis, the UGT72E subfamily was extensively studied by the group of D. Bowles in the context of their possible implication in regulating lignin precursor availability (Lim et al., 2005). A series of substrates was tested on the recombinant proteins and results showed that UGT72E1 and UGT72E2 displayed catalytic activities on coniferyl and sinapyl aldehydes and that the monolignols coniferyl and sinapyl alcohols were glycosylated by UGT72E2 and UGT72E3 to the respective 4-O-ß -D-glucosides forms (coniferin and syringin) (Lim et al., 2001; Lim et al., 2005). It should be noted that the enzymatic tests were performed on HCAs and on most of the intermediates of the lignification biosynthetic pathway but in general the CoA esters were omitted most likely because they have only recently become commercially available.

When used as a substrate by glucosyltransferases, caffeic acid can form caffeoyl-3-*O*-glucoside and caffeoyl-4-*O*-glucoside but also 1-*O*-caffeoylglucose, a high-energy glucose ester potentially used as a donor molecule during the formation of various hydroxycinnamic acid *O*- esters in plants (Mock and Strack, 1993; Lim et al., 2003). In the same way, 1-*O*-sinapoylglucose is formed by an UDP-glucose:sinapate glucosyltransferase (SGT) enzymatic activity and serves as a sinapoyl donor in various acyltransfer reactions leading to the accumulation of sinapate esters (Milkowski and Strack, 2010). During seed development in *Brassicaceae*, it is converted to sinapoylcholine (sinapine), the major seed phenylpropanoid that accumulates to levels of 1–2% of seed dry weight (Baumert et al., 2005), and is then further hydrolyzed in germinating seeds (Milkowski and Strack, 2010). In the seedling, sinapoylglucose is also channelled towards sinapoylmalate (2-O-sinapoyl-L-malate), a UV-B sunscreen protectant that accumulates in the vacuoles of the leaf epidermis and causes the adaxial surface to fluoresce in blue under UV light (Sinlapadech et al., 2007). In Arabidopsis, recombinant proteins of the four members of the UGT84A clade can convert HCA to 1-0ß-glucose esters (Milkowski et al., 2000) and the overexpression of UGT84A1, UGT84A2, and UGT84A3 leads to an increased level of sinapoylglucose in seeds and seedlings (Meißner et al., 2008). In *Brassica napus*, the Sinapoylglycosyltransferase (SGT) activity is due to UGT84A9, an enzyme with glucosylating capacities on sinapate, cinnamate, ferulate, 4-coumarate and caffeate (Milkowski et al., 2000). The closely related UGT84A10 can also catalyze the formation of HCA glucose esters with preference for ferulate and sinapate when produced in E. coli (Mittasch et al., 2007). Recent work showed that the suppression of UGT84A9 in RNAi lines induced strong metabolic modifications in late stages of seed development confirming the essential upstream role of the SGT activity on metabolites whose synthesis depends on 1-O-sinapoylglucose (Hettwer et al., 2016). A similar SGT activity in *Populus* has been assigned to GT1-316. Interestingly, the corresponding gene has a higher expression in plants grown under Nlimiting conditions. Similarly, a stress response of the *Fragaria* x ananassa FaGT2 gene was also observed after oxidative stress treatments in both strawberry fruit and cell cultures confirming that the final acceptors probably play an important role in the stress resistance mechanisms (Lunkenbein et al., 2006).

To reveal potential relations between genes involved in the phenylpropanoid backbone construction and those used by the plant for regulating the glycosylation status, we performed a mutual ranking classification using ATTED-II (Obayashi et al., 2009). This coexpression database (v8.0; March 2016) contains 58 experiments, 1388 GeneChips collected by AtGenExpress and about 20,000 publicly available files to provide gene-to-gene mutual ranks. We focused on a specific comparison between the known phenylpropanoid synthesis genes and the UGT genes found to be expressed in *Arabidopsis*. The complete results are presented as a heat map in supp. data figure 1 and a focus on five zones determined by hierarchical clustering containing a high number of coregulated genes is presented in figure 3. The UGT genes known to be involved in the glycosylation of the major phenylpropanoids are included in groups 1, 2 and 5 When the

flavonoid biosynthesis pathway is considered (figure 3, group 1), UGT78D1, UGT78D2 and UGT89C1 are all highly co-regulated with the four successive genes involved in the transformation of *p*-coumaroyl CoA to kaempferol/quercetin *ie* chalcone synthase (TT4), chalcone isomerase (TT5), flavanone 3-hydroxylase (F3H) and flavonol synthase (FLS). Only the three latest genes are coexpressed with UGT78D3. In the same group, the sinapoylglucose specific genes *UGT84A1* and *UGT84A2* are coexpressed with the flavonoid pathway genes, which is in accordance with the fact that this component is also a substrate for anthocyanin sinapoyltransferase (Fraser et al., 2007). UGT75C1 and UGT79B1 are present in group 2 but are not coregulated with the corresponding phenylpropanoid genes.

2.2 ß-glucosidases

ß-glucosidases (EC 3.2.1.21) have been described in Archaea, Eubacteria, and Eukaryotes (Ketudat Cairns and Esen, 2010). In plants, they show many different quaternary structures and can form dimers (Dharmawardhana et al., 1995), tetramers (Hosel et al., 1978) or decamers (Fan and Conn, 1985). In *Cassava*, linamarases can also be present as homooligomers with different numbers of monomers forming very large aggregates with sizes up to 2000 kDa (Sermsuvityawong et al., 1995). They catalyse the hydrolysis of terminal, non-reducing β -D-glucosyl residues with release of β -D-glucose and the corresponding free aglycone from various substrates including glucosides, 1-*O*glucosyl esters and oligosaccharides. ß-D-glucosidase activity is among the oldest enzymatic activity known as it was reported in the 19th century that almond emulsin was able to cleave the cyanogenic glucoside amygdalin (Wöhler and Liebig, 1837).

Plant glycosylated phenylpropanoids are not just substrates for endogenous (plant) enzymes but are also potential targets for ß-glucosidases from many different organisms. In yeast, hydrolytic activity towards 7- or 4'-O-glucosides of isoflavones, flavonols, flavones, and flavanones can be attributed to 3 genes and is probably related to metabolism of flavonoids from plant material within the culture broth (Schmidt et al., 2011). During human digestion, monoglycosylated and aglycone forms of flavonoids are more easily absorbed than the original multi-glycosylated compound (Vila-Real et al., 2011). The absorption of dietary flavonoid glycosides includes a critical deglycosylation step which is though to occur in the oral cavity by microbiota showing β -glucosidase

activity necessary for the delivery of the biologically active aglycones to epithelial cell surfaces (Walle et al., 2005). The plants themselves also show &-glucosidase activity towards a large range of their own phenylpropanoids underling the fact that the glycosylated forms of these molecules are unlikely to represent dead-end products. Several examples of the implication of deglycosylated phenylpropanoids in defense have been clearly illustrated (Konig et al., 2014). Nevertheless, this enzymatic activity has also been proposed to mediate the turnover of flavonol bisglycosides and may therefore have important consequences on plant growth and development (Roepke and Bozzo, 2015).

On a genome-wide scale, the number of ß-glucosidase genes has only clearly been determined in a small number of species. In Arabidopsis, 47 GH1 ß-glucosidase genes annotated from *BGLU1* to *BGLU47* were shown to have a potential common evolutionary origin according to their position within the phylogenetic tree (Xu et al., 2004). The first global analysis of the rice GH1 family identified at least 31 genes (of 34 apparently functional genes) expressed in different organs and growth stages (Opassiri et al., 2006). In this species, duplication events occurring after the divergence with the Arabidopsis thaliana lineage gave rise to most of the identified genes. Although the global number of predicted gene models is higher in maize compared to Arabidopsis, only 26 ßglucosidase paralogs were identified in the inbred line B73. Gomez-Anduro et al. (2011) explained this difference by the fact that a glucosinolate-mediated biotic stress response is absent in maize, reducing the amount of glycosylated substrates. The number of GH1 genes compared to the global number of gene models was also low in *Sorghum bicolor* (Sekhwal et al., 2013) suggesting that this feature may be characteristic of liliopsida genomes, however this should be confirmed by an in depth analysis of a much larger number of plant genomes based on the CAZy annotation. In agreement with the large number and diversity of glycoconjugates present in a given plant and the limited number of ß-glucosidases, these enzymes present overlapping specificities. For example the recombinant BGLU45 and BGLU46 from Arabidopsis, hydrolyze the glycosylated monolignols syringin, coniferin, and *p*-coumaryl alcohol 4-O-ß-D-glucoside (Escamilla-Trevino et al., 2006) but BGLU45 preferentially hydrolyses syringin whereas BGLU46 presents a broader specificity to other phenolic glucosides This promiscuity is exploited during in vitro assays since many ß-glucosidases are first tested with p-nitrophenylsugars before a more in depth characterization of the preferred cleaved sugar (Escamilla-Trevino et al., 2006; Baiya et al., 2014). On the contrary other ß-glucosidases seem to show a more strict specificity for their natural glycosylated compounds (Ketudat Cairns et al., 2015). For instance the *Arabidopsis* BGLU21, BGLU22 and BGLU23 exclusively cleave the coumarin scopolin into the aglycone scopoletin but do not hydrolyse *p*- or *o*-nitrophenyl-sugars (Ahn et al., 2010).

During lignin biosynthesis plant cells produce the monolignols *p*-coumaryl alcohol, coniferyl alcohol and sinapyl alcohol. Their corresponding glucoconjugates *p*coumaryl alcohol glucoside, coniferin and syringin can be used as substrates by a specific group of coniferin β -glucosidases. These coniferin β -glucosidases (EC 3.2.1.126) belong to the GH3 family and were first described in Picea abies seedlings (Marcinowski and Grisebach, 1978) and in *Cicer arietinum* suspension cells (Hosel et al., 1978). In both species the enzymes deglycosylated coniferin as well as syringin. It was previously speculated that the role of these specific enzymes could be to hydrolyse monolignol glucosides prior to incorporation of the aglycones in the lignin macromolecule by oxidative radical coupling thereby regulating the amount of precursors for lignification. Lignins from gymnosperms are rich in G units (Weng and Chapple, 2010) and support for a possible role of β -glucosidases in lignification was provided by the observation that high amounts of coniferin are accumulated in the cambium together with the colocalization of coniferin ß-glucosidase and peroxidase activity in the differentiating xylem of *Pinus contorta* (Dharmawardhana et al., 1995). The presence of both glucosylated monolignols and ß-glucosidase activity was also previously reported in Pinus banksiana (Leinhos and Savidge, 1993). Despite such observations, the use of radioactive precursors in *Pinus contorta* indicated that coniferin was probably not the main source of the coniferyl alcohol incorporated in lignins (Kaneda et al., 2008). The presence of coniferin ß-glucosidase is not exclusive to gymnosperms and these enzymes have also been found in vegetative and reproductive tissues of Oryza sativa (Baiya et al., 2014), in the xylem of Betula pendula (Marjamaa et al., 2003) and in the interfascicular fibers and protoxylem of Arabidopsis (Chapelle et al., 2012). The relation between lignification and monolignol 4-O-ß -D-glucosides is further discussed below.

3. The role of glycosylation in the cellular compartmentation

One of the most predominant questions about phenylpropanoid biosynthesis concerns the spatial distribution of both precursors and final products within the cells. Many phenylpropanoid molecules necessary for plant development and/or defense are also toxic and an important challenge for plants is how to successfully mange the production/storage of these potentially dangerous molecules (Väisänen et al., 2015). While the association between PAL and C4H enzymes on microsomal membranes is believed to be responsible for the rapid transformation of phenylalanine to 4-coumaric acid in tobacco (Rasmussen and Dixon, 1999), much less evidence exists to support the existence of any large scale, tight metabolic channelling that could prevent accumulation of intermediates synthesized in the cytosol. The vacuole is one major compartment that has often been identified as an important site of storage for phenylpropanoid-derived products. Coniferin was shown to be accumulated in protoplasts of pine cambial cells, most likely in the vacuoles (Leinhos and Savidge, 1993) and recent phenolic profiling of vacuoles from Arabidopsis leaves identified coniferin, coniferin hexoside, and 34 other diverse glycosylated phenolics including derivatives of coniferyl alcohol, sinapyl alcohol, ferulic acid and sinapic acid (Dima et al., 2015).

The cell wall is another major compartment where the oxidation machinery including peroxidases and laccases form monolignol free-radicals that are subsequently cross-coupled in a non-stereospecific reaction to form lignin (Boerjan et al., 2003). In the apoplast, two monolignols can also undergo a enantioselective radical coupling reaction controlled by dirigent proteins to form lignans (Davin et al., 1997). Several lines of evidence support the biosynthesis of lignans in this compartment including the presence of an N-terminal signal peptide in dirigent proteins (Ralph et al., 2006), immunolocalization of these proteins in the cell wall (Burlat et al., 2001), and their presence in apoplastic protein extracts (Uzal et al., 2009). A previous report assuming that (neo)lignans were also located in vacuoles of *Linum* cell cultures (Kuhlmann et al., 2002) was confirmed recently by their detection in *Arabidopsis* leaf vacuoles (Dima et al., 2015). This paper also reported for the first time that in *Arabidopsis*, the combinatorial radical coupling reactions leading to (neo)lignans synthesis usually detected in the apoplast could also be detected in the cytoplasm. This observation greatly contributes to resolving the apparent paradox of how oligolignol hexosides arise

in plant cells. Logically their formation requires an oxidative coupling reaction - usually thought to occur exclusively in the cell wall, followed by a glycosyl conjugation occurring in the cytoplasm. Nevertheless, this explanation does not totally exclude the possibility that a proportion of the cytoplasmic oligolignols may be formed in the apoplasm and subsequently transferred to the cytoplasm via specific (neo)lignin transporters as in mammalian systems (Miguel et al., 2014).

The glycosylation step is essential for the synthesis, transport and sequestration of some phenylpropanoids. Because of their abundance and their role in flower coloration, the transport of glycosylated anthocyanins has been extensively studied over the past years (Yazaki, 2006). The involvement of Multidrug Resistance-associated Protein (MRP) also called ATP Binding Cassette subfamily C (ABCC) transporters in the vacuolar transport of such phenolic glucosides was first proposed in maize (Marrs et al., 1995) and later in carnation (Larsen et al., 2003) and Arabidopsis (Kitamura et al., 2004). However, these transporters are known to translocate a wide range of glutathione (GSH)-conjugated substrates through the membranes and anthocyanin-GSH conjugates have never been detected in plants (Ohkama-Ohtsu et al., 2011). Nevertheless a potential co-transportation of GSH and glycosylated anthocyanidins was suggested by the heterologous expression of grapevine ABCC1 in yeast (Francisco et al., 2013). A similar transportation system must still be validated in other species. In the lignification process the glycosylation of hydroxycinnamic aldehydes and alcohols has been proposed to form part of an overall regulatory network of lignin precursor homoeostasis. In this model ABC-like transporters located in the plasma and vacuolar membranes show differential affinities towards glycosylated and non-glycosylated molecules thereby controlling subcellular compartmentation (Miao and Liu, 2010). The 4-O-glycosylated forms are preferentially transported (in the presence of ATP) across the tonoplast into the vacuole whereas the specificity of the plasma membrane transporters is more relaxed with respect to ATP-dependence and favours the passage of aglycone forms. Given the existence of such preferential transport it is tempting to speculate that unglycosylated lignin precursors are transported across the plasma membrane by specific ABC-like transporters and undergo free-radical polymerization while a surplus of monolignols could be glycosylated and transferred to the vacuole for storage by a transporter specific to glycoconjugates. ABC genes constitute large multigene families (approximately 130 members in Arabidopsis) and both structural and biochemical approaches are needed to fully characterize these transporters and evaluate the importance of glycosyl groups for substrate specificity. For example the transporters characterized in this work are located in Arabidopsis rosette leaves that contain only small amounts of lignins, and were shown to be mainly associated with the transport of *p*-coumaryl alcohol, a minor lignin component. More recently, Tsuyama et al. (2013) showed that a light membrane fraction prepared from differentiating xylem of hybrid and wild poplar, Japanese cypress and pine had clear ATP-dependent transport activity specific for coniferin suggesting the existence of a common transport mechanism dependant on a proton gradient created by a vacuolar-type H+ ATPase (V-ATPase) in the lignifying tissues of both woody angiosperms and gymnosperms. Considering the fact that these V-ATPases were also located in the trans-Golgi network, it could be argued that coniferin may, in fact, be transported through the cell in small vesicles as suggested nearly 50 years ago by Pickett-Heaps (1968). It still remains necessary to determine whether the transport systems are universal or are dependent on molecules and/or species.

When considering the compartmentation of phenylpropanoids within the plant cell, the location of the enzymes implicated in their glycosylation status should also be taken into account. It seems clear that the absence of predicted signal peptides in plant UGT sequences supports the belief that they are located in the cytoplasm (Ross et al., 2001). The transfer of the glycosyl group on acceptor molecules can then be considered as exclusive to this compartment. On the other hand, the removal of sugar moieties by ß-glucosidase can potentially occur in a different compartment. Historically, the first ß-glucosidase capable of hydrolysing phenylpropanoid glycoconjugates was found in the cell wall fractions of roots and hypocotyls of *Picea abies* seedlings (Marcinowski and Grisebach, 1978) and cell suspensions of *Cicer arietinum* (Hosel et al., 1978). Since then, different &-glucosidases have been identified in many plant species - generally by – omics approaches - but very few have been characterized at the biochemical level. It has been assumed for almost 40 years that the enzyme and the glycosylated substrate can be present in different locations and that contact between them occurs when the tissue or cells are damaged compromising cell integrity. This physical separation exists at the

cellular level in sorghum seedlings where the cyanogenic glycoside dhurrin and a specific ß-glucosidase are located in different cells, the substrate in the epidermis and the enzyme in the mesophyll tissues (Kojima et al., 1979). However, it should be noted that due to the difficulty of physically separating these tissues, the results were obtained using protoplasts prepared from leaves and may therefore not represent the real situation in planta. In the case of phenylpropanoids, this separation generally occurs at the subcellular level. For example in *Melilota alba* leaves, coumarin (glycosylated form of 2-hydroxycinnamic acid) is located in the vacuole (also prepared from protoplasts) whereas the specific ß-glucosidase is found in the cell wall (Oba et al., 1981). The optimum pH determined for purified or recombinant ß-glucosidases exhibiting coniferin specificity in gymnosperms, monocots and dicots (Hosel et al., 1978; Marcinowski and Grisebach, 1978; Escamilla-Trevino et al., 2006; Baiya et al., 2014) generally ranges between 4.5 and 5.5 which is consistent with the acidic environment of the cell wall. The role of ß-glucosidases in cell wall lignification and plant defense responses is discussed further in the following section of this paper. This separation is a critical need for molecules that are toxic under their aglycone form but may not be necessarily widespread. In rice, a ß-glucosidase showing *in vitro* activity against a *p*-coumaryl alcohol glucoside but not coniferin with an optimum pH of 6.5 was identified in the cytoplasm (Rouyi et al., 2014) suggesting that both glucosylation (by cytosolic UGTs) and deglucosylation (by cytosolic ß-glucosidases) may occur within the same cellular compartment (but not necessarily the same cell) at least in this species. Transglucosidases are GH1 enzymes that can transfer a sugar from a donor other than a nucleotide phosphate or phospholipid phosphate to an acceptor. Among these, the rice Os9BGlu31 enzyme is located in the vacuole of senescing flag leaves and developing seeds and is induced in seedlings in response to drought stress and different phytohormones (Luang et al., 2013). Subsequent functional characterization indicated that the enzyme showed the highest activity with an optimum pH of 4.5 towards 1-0feruloyl-ß-D-glucose, as well as towards flavonoid glucosides but with lower efficiency. The presence of such an activity inside the vacuole would suggest that transported glycosylated phenylpropanoids can undergo further enzymatic modifications and so should not be considered as 'dead-end' storage products. The detection of six GH1 glucosidases in Arabidopsis leaf vacuoles by proteomics also supports this idea (Carter et al., 2004).
In conclusion, a thorough understanding of the biological roles of (glycosylated) phenylpropanoids not only depends on detailed knowledge about their biosynthesis, but also a better comprehension of their localisation and transport within, and between different tissues, cells and compartments.

4. At the crossroads of monolignol glycosylation and lignin biosynthesis

There are still many uncertainties about whether pooled glucoconjugated monolignols - and by extension hydroxycinnamic acids and aldehydes – really constitute a stock for subsequent lignin construction. Terashima and Fukushima (1988) followed the incorporation of radiolabeled glycosylated monolignols in *Pinus thumbergii* by microautoradiography in softwood differentiating xylem. When incubated at concentrations of up to 1mg glucoside in 100 μ l, the labelled monolignols were incorporated in the cell wall leading the authors to conclude that glycosylated forms could contribute to the lignification process. Although it is tempting to think that incorporation occurs after a simple hydrolysis of the sugar moieties, an alternative pathway involving the oxidization of coniferin to coniferaldehyde glucoside prior to the deglucosylation step, followed by the transformation to coniferyl alcohol was suggested by the use of radiolabeled coniferin in *Gingko biloba* shoots (Tsuji et al., 2005). Either way, these experiments show that exogenous coniferin can serve as a lignin precursor.

Lignin composition is modified when plants are submitted to a mechanical stress. Compression wood (CW) is formed on the lower part of stems and branches of inclined softwoods (Timell, 1986). When compared to normal wood (NW), CW tracheids contain less cellulose but more lignin with (in addition to G-units) a higher proportion of *p*hydroxyphenyl units (H-units) that are almost absent in NW and opposite woods (Terashima and Fukushima, 1988; Donaldson, 2001). In addition to increased lignin content, Yoshinaga et al. (2015) also reported that coniferin was less abundant in compression wood than in normal wood and that the *p*-coumaryl alcohol glucoside was undetectable in the two Japanese softwoods *Chamaecyparis obtusa* and *Cryptomeria japonica*. In the light of these results the authors suggested that if glycosylated monolignols do indeed function as lignin precursors, then their incorporation as G- and H- subunits must occur immediately after their synthesis without being pooled in the cell. A possible relationship between the accumulation of glycoconjugated phenylpropanoids and lignin status has also been reported in flax. In this species, bast fibers located in the outer part of the stem possess very thick secondary cell walls containing low amounts (4 % cell wall residue) of lignin compared to inner stem xylem cell walls containing high amounts of lignin (25% cell wall residue) (Day et al., 2005). An ultra-high-performance liquid chromatography-Fourier transform ion cyclotron resonance-mass spectrometry (UHPLC-FT-ICR-MS) phenolic profiling approach showed that flax inner- and outer-stem tissues had different phenolic profiles (Huis et al., 2012). In total 81 different phenolic compounds including monolignols, oligolignols and (neo)lignans were identified and among these, 39 were glycosylated and present (except for one compound) in much higher amounts (up to 92-fold higher) in the outer stem tissues. Whether the observed differences in glycosylated phenolics content can be related to the differences in cell wall lignin content in the two stem tissues remains an interesting question for future research.

It is now well known that transgenic approaches targeting general phenylpropanoid or more specific monolignol genes can modify both lignin composition and quantity (Chabannes et al., 2001; Voelker et al., 2010). For example, inactivation of a caffeic acid O-methyltransferase (COMT) gene in Panicum virgatum decreased lignin content concomitantly with metabolites related to trans-sinapyl alcohol including the conjugation-product syringin (Fu et al., 2011; Tschaplinski et al., 2012). In Arabidopsis, a global view of the metabolic shifts in the lignin-related biosynthesis was obtained by a systems biology approach (Vanholme et al., 2012). These published data revealed that mutations in phenylpropanoid genes have different impacts on glycosylated monolignol production. In the *pal1, c4h, 4cl, ccoaomt1* and *ccr1* mutants, the amounts of syringin and coniferin were significantly lower when compared to the wild type. All these mutants except for *pal1* also show reduced lignification. An increase in the syringin and coniferin levels was also reported in the 4cl2 and f5h1 mutants respectively without any change in lignin contents. The overall effect depends on the mutated gene and the phenylpropanoid considered. A rather general tendency shows that the glycosylated metabolites beyond coniferaldehyde in the biosynthetic pathway tend to decrease in *pal1, c4h, 4cl1, ccoaomt1* and *ccr1* and that the glycosylated metabolites produced before the mutated gene increase in *4cl1*, *ccoaomt1*, *ccr1* and *comt*. The results obtained in this species clearly support the hypothesis that monolignols whose destiny is to become either glycosylated or polymerized as lignins are produced in a common metabolic pathway.

Comparison of lignin biosynthetic and UGT gene expression patterns reveal interesting co-expressions hinting at links between lignification and monolignol glycosylation. In figure 3, the group 5 shows that two (UGT72E1 and UGT72E2) of the three monolignol-associated UGT72E subfamily genes are coexpressed with the peroxidase *PRX49* and that all three members are coexpressed with *PRX72*. These two peroxidases were recently described as major actors of lignin polymerization in the cell wall (Herrero et al., 2013; Fernandez-Perez et al., 2015) and since both peroxidases and UGTs use the same substrates the observed coexpression would strongly suggest the existence of a functional link between these enzymes. Further evidence for such an idea is provided by the effects of mutations in the triple mutant of 3 laccase genes (LAC4, LAC11 and LAC17) in Arabidopsis (Zhao et al., 2013). The growth of these plants is blocked in the early stages of development with (among other) phenotypic characters, a narrow root diameter and vascular development arrest. The authors attribute these severe growth perturbations to the lack of lignin and secondary cell walls in the mutants and to an overaccumulation of free forms of monolignols. Indeed levels of coniferin and syringin were both increased in the mutant along with those of sinapoylglucose and three kaempferol glycosides. Interestingly, the expression of most monolignol biosynthesis genes was not affected showing the lack of a negative feedback regulation on the transcription. On the other hand, the accumulation of monolignol glucosides in this triple mutant could be correlated to significant up-regulation of UGT72E2 and UGT72E3 potentially related to monolignol detoxification mechanisms in the absence of lignin polymerization.

The inversely proportional expression of *UGTs* and genes involved in lignin polymerization was also demonstrated in the *lbf1* flax mutant isolated from an ethyl methanesulfonate (EMS) population (Chantreau et al., 2014). In these plants, the bast fibers located in the external parts of the stems show ectopic lignification of their secondary cell walls. This additional production of lignin is accompanied by a higher *cinnamoyl-CoA reductase* (*CCR*), *caffeic acid O-methyltransferase* (*COMT*), *cinnamyl*

alcohol dehydrogenase (*CAD*) and PRX71 ortholog transcript accumulation indicating a possible increase in monolignol production and oxidation. In parallel, the reduction in the abundance of transcripts corresponding to a *UGT72E* ortholog observed in the outer tissues was explained by the reduced necessity for detoxification due to a higher incorporation of monolignols into the lignin polymer.

Altogether, these several examples illustrate the fact that variations in lignin amounts can have a potential effect on monolignol glucoside production but the opposite is not necessarily true. The overexpression of UGT72E2 and UGT72E3 in Arabidopsis (Lanot et al., 2006; Lanot et al., 2008) mostly led to increased accumulation of coniferin in light grown roots but had little impact on syringin levels. Functional validation of both genes also came from the significant reduction in coniferin and syringin in down-regulated transgenic plants. Perhaps the most interesting result coming from these transgenic approaches is that regardless of the effects on monolignol glucoside levels, no impact on lignin was ever observed, even when the silencing of the three UGT genes by an RNAi approach led to a 90% reduction (Wang et al., 2013). These results strongly suggest that changes in monolignol glucosides do not affect lignin levels. Nevertheless, it is interesting to note that the heterologous overexpression of a poplar UGT gene in tobacco (Wang et al., 2012) led to an increase of lignin in stems. However, in this case the authors indicate that the corresponding recombinant enzyme was unable to glycosylate phenylpropanoid components suggesting that the observed effect was indirect.

Overexpression of *UGT72E2* and *UGT72E3* also led to the accumulation of coniferin and syringin in the leaves associated to a significant reduction in the amounts of sinapoyl malate. A link between monolignol glucosides and hydroxycinnamic acid esters was also suggested by the results obtained by Hemm et al. (2004) showing that the light-induced production of coniferin and syringin in roots is accompanied by upregulation of the *4CL3* (At1g65060) gene involved in the production of sinapate esters (Nair et al., 2004). Interestingly, the analysis of the gene expression heatmap shown in figure 3 (group 1) indicates that *4CL3* is coexpressed with the two glycosyltransferases *UGT84A1* and *UGT84A2* implicated in the production of sinapoyl glucose (Milkowski et al., 2000).

As stated above any link between lignin and phenylpropanoid glucoconjugates should not only take into account UGT activities and gene expression but also ß-glucosidase behaviour. In the systems biology approach developed by Vanholme et al. (2012), the decrease in the amounts of coniferin and syringin were nicely related to the lower *UGT72E2* and *UGT72E3* expression levels, but also to the increase in both *BGLU45* and *BGLU46* transcript accumulation. These 2 genes were shown to be specific for monolignol glycosides (Escamilla-Trevino et al., 2006) and the corresponding knockout mutants displayed a significant increase in stem coniferin content but, as for UGT overexpressors, no changes in lignin content were observed (Chapelle et al., 2012). A general overview of the effects of mutations of genes involved in monolignol production, polymerization and glycosylation is shown in figure 4.

In conclusion, although monolignol glucosides have long been considered as participants in the lignification process (Freudenberg and Harkin, 1963; Freudenberg and Torres-Serres, 1967) it seems unlikely that the plant actually uses these molecules stocked in the vacuole to synthesize lignin. In contrast, the perturbation of genes involved in lignification has a clear impact on the production of glycoconjugated monolignols. However, it would be important to check whether the results obtained in *Arabidopsis* are also valid for woody plants. Although secondary cell wall formation and especially lignin formation have been extensively studied in poplar, few studies (Wang et al., 2012) have addressed the potential links between phenylpropanoid glycosylation and lignification in trees.

5. The role of glycosylation in phenylpropanoid function

5.1 Glycosylation and phenylpropanoid stability

Many phenylpropanoids are toxic and unstable molecules and so rarely accumulate in their aglycone form in plant cells (Whetten and Sederoff, 1995; Alejandro et al., 2012; Väisänen et al., 2015). The mechanism of their toxicity is not clear but it has been suggested Väisänen et al. (2015) that: (*i*) high concentrations of coniferyl alcohol could react with lipids or proteins in cellular environments in the presence of reactive oxygen species, (*ii*) since coniferyl alcohol could be easily degraded (e.g. to vanillin or

ferulic acid) it is possible that its toxicity results from the accumulation of degradation products (*iii*) the availability of coniferyl-alcohol leads to the formation of a lignin-like compound and reduced growth of plant cells. More generally, aglycone forms of phenylpropanoids are nucleophilic molecules with high damage potentialities on diverse cellular components. The attachment of carbohydrate moieties will then reduce their reactivity and improve their stability (Jones and Vogt, 2001). In the case of monolignols, it is known that the conversion into their glycone forms by UGT activity reduces their toxicity (Bock, 2016). This is also the case of anthocyanidins for example, which only accumulate in cells in their glycosylated forms (Boss et al., 1996).

Glycosylation may not only prevent the toxicity of aglycone phenylpropanoids but can also contribute to the production of protectant molecules against reactive oxygen species (ROS). Flax seed coats contain the lignan secoisolariciresinol diglucoside (SDG) whereas their aglycone precursors pinoresinol, lariciresinol and secoisolariciresinol are undetectable whatever the development stage considered (Hano et al., 2006). SDG was previously shown to prevent lipid peroxidation in a concentrationdependent manner (Prasad, 1997) and it is believed that it may play a similar role in the plant. This hypothesis is consistent with the concomitant presence of very high amounts of SDG and highly reactive polyunsaturated fats in the flax seeds (Lorenc-Kukula et al., 2005). Nevertheless, a peroxyl radical scavenging capacity comparable to that of glutathione has been described for the aglycone form of the isoflavone daidzein and the non-glycosylated lignan honokiol (Kim et al., 2010) suggesting that glycosylation is not essential for ROS protection capacity. Glycosylation can therefore contribute to reducing phenylpropanoid toxicity and increasing stability thereby explaining the widespread occurence of phenylpropanoids in plant development and resistance/tolerance to major biotic and abiotic stresses (Lim and Bowles, 2004; Bowles et al., 2005)

5.2 Phenylpropanoid glycosylation and plant defense

The phenylpropanoid biosynthetic pathway is stimulated during the hypersensitive response (HR) induced by pathogen infection and is associated with the production of phenylpropanoid-derived phytoalexins such as coumarins and isoflavonoids (Hammond-Kosack and Jones, 1996; Dorey et al., 1997; Langlois-Meurinne et al., 2005). Chong et al. (2002b) reported that the glycosylation of phenylpropanoids in

tobacco played a significant role during Tobacco Mosaic Virus (TMV) infection and that the down-regulation of a tobacco glycosyltransferase gene (TOGT1) led to reduced TMV resistance. The authors suggested that scopolin (a glycoconjugated coumarin) is a storage form of the aglycone form scopoletin that functions as an antiviral agent potentially involved in scavenging ROS accumulated during the infection. Overexpression of the TOGT gene also enhanced resistance against the potatovirus Y (PVY) in tobacco (Matros and Mock, 2004). Recent work also shows that sinapate esters, coniferin and the glycosylated pinoresinol and lariciresinol lignans accumulate when Arabidopsis leaves are infected by the soil-borne ascomycota Verticillium longisporum (Konig et al., 2014). When tested on agar plates, only the sinapovl glucose and conifervl alcohol, but not sinapyl alcohol, managed to inhibit the growth of the fungi, whereas coniferin specifically inhibited melanisation. The authors proposed that the growthinhibiting differences observed between the two aglycone monolignols was solely due to the oxidization of coniferyl alcohol to ferulic acid, which was also shown to have an inhibitory effect on V. longisporum. Coniferin-accumulating transgenic plants overexpressing UGT72E2 were less susceptible to the fungi. In this case it was suggested that coniferin would operate as a storage form of coniferyl alcohol, which could further be transformed into ferulic acid. Plant species can also display glycosylation-related cultivar- and pathogen-dependent responses to biotic stresses. In wheat, genes encoding cytochrome P450 (CYP709C1) and a UGT were more strongly upregulated during the infection by Fusarium graminearum compared to Magnaporthe grisea and cultivar resistance to *F. graminearum* led to a stronger expression in incompatible (resistant) interactions with the chinese spring wheat cultivar Sumai 3 (Ha et al., 2016). Based on these observations the authors propose the existence of different resistance genes related to both fungi and confirm the necessity to combine genes in breeding programs to obtain multi-resistance plants (Schafer et al., 1963). Deoxynivalenol produced by Fusarium is glycosylated by DOGT1 (UGT73C5) and the constitutive overexpression of the corresponding gene in *Arabidopsis* led to enhanced tolerance against this mycotoxin (Poppenberger et al., 2003). Other studies have underlined the importance of some UGT genes implicated in pathogen response and redox status during the pathogen infection. For example, it was reported that UGT73B3 and UGT73B5 are involved in the hypersensitive response to the bacteria *Pseudomonas syringae* pv tomato in *Arabidopsis* and participate in the regulation of the redox status of plant cells. However, the exact role of these UGT *in planta* remains unknown (Langlois-Meurinne et al., 2005; Simon et al., 2014).

There is now clear evidence that UGTs play an important role in disease resistance even though their precise contribution still remains unclear. However it appears that the addition of a glycosyl group does not necessarily play a major role in determining the biological activity of phenylpropanoids and it may even block the biological activity of a particular molecule towards the pathogen. In this case, it is possible that glycosylation may play a role in maintaining the homeostasis of an ensemble of pathogen- specific molecules. Langenbach et al. (2013) suggested that disruption of the *UGT84A2/BRT1* gene can lead to an enhanced accumulation of certain sinapate derivatives or new metabolites absent from wild-type plants, which would confirm the important role of UGTs in maintaining a specific pool of phenylpropanoid-related metabolites.

5.3 Phenylpropanoid glycosylation and abiotic stress

The Arabidopsis UGT84A2 and its Brassica napus ortholog UGT84A9 are both involved in the production of sinapoyl malate. This HCA glucose ester accumulates in vacuoles of sub-epidermal tissues and is believed to play a role in the protection against the harmful effects of UV-B radiation (Landry et al., 1995; Meißner et al., 2008) (Dean et al., 2014). Li et al. (2010) isolated the Arabidopsis Pna-10 accession containing a 13-kb deletion that eliminates the *SNG1* gene responsible for the conversion of sinapoylglucose to sinapoylmalate, and the gene encoding sinapoylglucose:anthocyanin sinapoyltransferase (SAT). As a result Pna-10 is unable to accumulate both sinapoylmalate and sinapoylated anthocyanins. Under field conditions, this mutation is not lethal suggesting that the presence of sinapoylmalate in the leaves may not be absolutely essential for plant protection. More recently, (Hectors et al., 2014) highlighted the role of the rhamnosylated kaempferol and quercetin glycosides during UV acclimation in Arabidopsis. The concentration of these compounds increased as a result of UV-stress (and related oxidative stress) and could be correlated with the upregulation of the flavonol-7-O-rhamnosyltransferase UGT89C1. The biological role of these flavonoid glycosides is not clear as they are less effective anti-oxidants that the corresponding aglycone forms (Vogt and Jones, 2000; Gachon et al., 2005). The authors suggest that the accumulation of flavonol glycosides constitutes a reserve of flavonols which could be easily mobilized at any given time and especially under UV stress conditions.

A relationship between aglycone/glycone flavonoids and oxidative stress in Arabidopsis was also shown by Kim et al. (2010). In this case, the loss of function of three genes UGT73B1, UGT73B2 and UGT73B3 involved in the glycosylation of flavonoids led to a greater tolerance to oxidative stress whereas overexpression of *UGT73B2* increased the sensitivity to ROS. The significant role of HCA glycosylation was also highlighted in senescent and water stressed plants (Fini et al., 2012; Torras-Claveria et al., 2012). An accumulation of the coumarin esculetin and quercetin 3-0-glucosides was observed in water stressed Fraxinus ornus leaves simultaneously with a decrease in antioxidant enzyme activities (Fini et al., 2012). The authors suggested that these glycosides could act as H₂O₂ scavengers during water stress. These results are also consistent with those of Kylli et al. (2008) who showed that sinapic and ferulic acid glycoside esters were efficient antioxidants. More recently, Babst et al. (2014) showed an up-regulation of UGT84A17 expression under environmental stress (in particular under N limitation) in transgenic *Populus*, leading to the accumulation of hydroxycinnamate glucose esters (especially caffeoyl-4-coumaroyl- and cinnamoyl-glucose esters). This accumulation was glucoside synthesis, to the detriment of flavonoid which also requires hydroxycinnamates in their free form as intermediates. These authors highlighted a metabolic trade-off associated with stress dependent phenylpropanoid glycosylation (Dauwe et al., 2007). Finally, a novel Arabidopsis glycosyltransferase gene UGT85A5 (Sun et al. 2013) and more recently UGT85U1 (Ahrazem et al., 2015) are both significantly induced by salt stress. Nevertheless, their connection with the phenylpropanoid pathway and their substrate affinity require more detailed characterization.

Although the antioxidant proprieties of phenolic molecules have been known for a long time, the precise role(s) of glycosylation and, by extension the role of UGTs and ßglucosidases, in abiotic stresses is difficult to understand because of the existence of numerous signal pathways that can positively or negatively interact with each other. Such cross-talk is sometimes illustrated by the production of common glycosylated phenylpropanoids under different stresses. For example roots of hydroponically-grown *Arabidopsis* plants produce the glycosylated coumarin scopolin and monolignol coniferin when they are submitted to oxidative stress, root wounding, or nitrate deprivation (Ward et al., 2011). A major challenge is therefore to deepen our knowledge on the response mechanisms of plants to environmental changes in order to understand how these glycocojugates can contribute to plant protection.

5.4 Phenylpropanoid glycosylation: aroma, taste and color of plant products

Flavonoid glycosylation plays a significant role in determining flower, leaf, seed and fruit color necessary for attracting animals for flower pollination and/or seed dispersion (Tanaka et al., 2008). For example, strawberry fruits contain anthocyanins (glycosides of anthocyanidins) that give them the attractive red color in the ripe fruit (Griesser et al., 2008). Six anthocyanidins are generally present and their glycosylation is usually driven by UDP-glucose:flavonoid-O-glycosyl-transferase (UFGT) (Jaakola, 2013). However, the mechanism is quite complex and takes place in a species-specific manner. These anthocyanidins are most frequently O-glycosylated at the C3- followed by the C5-and/or C7-position (Tanaka et al., 2008). The glycosyl moieties of anthocyanins are themselves usually later modified by aromatic and/or aliphatic acyl moities. After these modifications anthocyanins become stable compounds and accumulate in vacuoles. The range of different mechanisms causes a variety of flower and fruit colors (Vogt and Jones, 2000; Grotewold, 2006; He et al., 2015; Li et al., 2016). In the case of strawberry fruits, Griesser et al. (2008) demonstrated that FaGT1, an anthocyanidin-3-O-glucosyltransferase, is a key enzyme in the phenylpropanoid biosynthesis because it channels the flavonoid pathway into anthocyanins and not into the glycosylated flavonols. Anthocyanidin 3-O-glucosyltransferases have also been characterized in many ornamental commercial plants (Tanaka et al., 1996; Yamazaki et al., 2002; Bowles et al., 2005; Sasaki and Nakayama, 2015). Recently, Matsuba et al. (2010) have shown that a novel glucosylation reaction on anthocyanins occurs with 1-0-b-D-vanilly-glucose as the sugar donor molecule in the petals of carnation and *Delphinium*.

The glycosylation of flavonoids may also be responsible for fruit flavour and aroma (Frydman et al., 2013; Tikunov et al., 2013; Hjelmeland and Ebeler, 2015). In Citrus, bitter species (e.g. pummelo, grapefruit) accumulate bitter flavonone 7-Oneohesperidosides (neohesperidin and naringin) which are synthesized through 1,2 rhamnosyltransferase activity and non-bitter species (e.g. oranges, mandarins)

flavanone-7-0-rutinosides accumulate (hesperidin and narirutin) with 1.6 rhamnosyltransferases (Frydman et al., 2013). The roles of these different glycosylated flavonoids have not yet been determined but Del Rio et al. (2004) suggested that they are involved in plant defense against pathogens. Phenylpropanoid volatiles or volatile organic compounds (VOCs) are responsible for fruit aroma (Hjelmeland and Ebeler, 2015). In the case of green tomatoes, the aroma is termed "smoky" (Tikunov et al., 2013) and results from the presence of volatile phenylpropanoid diglycosides (with a hexose pentose sugar moiety), which can be easily cleaved. In mature fruits glycosylation is more complex and triglycosides are synthetized (with a dihexose-pentose moiety) that are less easily cleaved. Tikunov et al. (2013) have characterized a glycosyltransferase named NON-SMOKY GLYCOSYLTRANSFERASE1 (NSGT1) in ripe fruits unable to produce and release the 'smoky' aroma. NSGT1 is unable to glycosylate the aglycone form of phenylpropanoids and is only involved in further elongation of the glycosidic moiety of glycosides. The physiological and ecological roles of these aromas have not yet been elucidated. Some authors propose that they may repel (in the case of green fruits) or attract (in ripe fruits) animals (Borges et al., 2008). Glycosyltransferase activities would therefore indirectly affect seed dispersion.

6. Conclusions

In response to a wide range of developmental and environmental cues, plants continuously produce (and modify) an extremely large pool of different phenylpropanoid-based secondary metabolites thereby contributing to the success of these autotrophic and, in many cases, sessile organisms. Many plant phenylpropanoids are also of economic interest for humans and a better understanding of the different factors controlling their biosynthesis and compartmentation should facilitate production and plant improvement. It is now becoming clear that the synthesis, localisation and biological activity of plant phenylpropanoids is not only regulated via the transcriptional and post-transcriptional regulation of different biosynthetic genes/enzymes, but also involves the reversible glycosylation of these molecules by regioselective glycosyltransferases (GTs) and glycosyl hydrolases (GHs). A major future research challenge will be to obtain a better understanding of how plants regulate the individual members of these multigenic GT/GH families during development and in response to abiotic and biotic stress. Given the wide diversity of phenylpropanoid structure across

the plant kingdom it will be important to obtain information on a number of different species. In addition it will also be necessary to explore the links between GT/GH regulation and the regulation of other genes/enzymes involved in determining phenylpropanoid availability and activity (e.g. phenylpropanoid biosynthesis, peroxidases/laccases, transporters, UDP-sugar production). Such knowledge will be particularly valuable for maintaining/improving cultivated plant production in a context of increased atmospheric CO₂ levels, global warming and modified pathogen biodiversity.

Figure legends

Figure 1: Schematic view of the phenylpropanoid biosynthesis. The general phenylpropanoid pathway begins by successive reactions resulting in the transformation of phenylalanine into *p*-coumaryl CoA, which is the common precursor of stilbenes, coumarins, phenylpropenes, 2-phenylchroman-containing flavonoids and monolignols.. The pathways leading to the production of hydroxycinnamic acids were obtained based on phenylpropanoid analysis in *Arabidopsis* mutants.

Figure 2: Predicted UGT enzymes in different higher plants. Species are arranged in an order based on taxonomic families. [1] Barvkar et al. (2012); [2] Caputi et al. (2012); [3] Cao et al. (2008); [4] Ekstrom et al. (2014); [5] Huang et al. (2015); [6] Li et al. (2014); [7] Ross et al. (2001); [8] Sharma et al. (2014); [9] Song et al. (2015); [10] Yonekura-Sakakibara and Hanada (2011).

Figure 3: Co-expression of UGT and phenylpropanoid genes. Five relevant groups of coexpressed genes delimited by hierarchical clustering indicated on the left of the heatmap were extracted from the global heat map presented in supp data figure 1. Groups 1, 2 and 5 contain genes known to be involved in the glycosylation of the major phenylpropanoids and groups 3 and 4 contain genes with high coregulation values. When the mutual rank (MR) value < 50, co-expression between the genes is considered as strong; if 51<MR<1000, there is co-expression and if MR>1001, there is no co-expression.

Figure 4: Effects of phenylpropanoid gene mutations or UGT gene overexpression. A schematic view of lignin, glucoconjugated monolignols, hydroxycinnamate glucosides or hydroxycinnamoyl esters are presented in *ccr-1* and *lac4,11,17* mutants and *35S:UGT72E2-3* overexpressors. For metabolites and transcripts, a colour code indicates an accumulation or a reduction compared to the wild-type. Data were extracted from Vanholme et al., 2012, Zhao et al., 2013 and Lanot et al., 2006, 2008 respectively.

Supp data figure 1: Co-expression of UGT and phenylpropanoid genes. The heat map was constructed using ATTED-II (Obayashi et al., 2009). When the mutual rank (MR) value < 50, co-expression between the genes is considered as strong; if 51<MR<1000, there is co-expression and if MR>1001, there is no co-expression.

Figure 1



Figure 2

			Plant species	Number of predicted UGTs	Identification method	References
			Picea abies	219	HMMER	[4]
			Oryza sativa	180/191/224/213	Blastp/HMMER	[2]/[4]/[3]/[10]
	П		Brachypodium distachyon	143/163	HMMER	[10]/[4]
_	-	Г	Zea mays	147/167/168	Blastp-Pfam-SMART/HMMER	[6]/[4]/[10]
			Setaria italica	211	HMMER	[4]
		1 –	Sorghum bicolor	180/188/201	Blastp/HMMER	[2]/[4]/[10]
		4	Panicum virgatum	359	HMMER	[4]
l r			Mimulus guttatus	100/147	Blastp/HMMER	[2]/[4]
1	Г		Cucumis sativus	85/105/100	Blastp/HMMER	[2]/[10]/[4]
			Cicer arietinum	96	Blastp- PSWM-HMMER	[8]
			Glycine max	182/233/242	Blastp/HMMER	[2]/[4]/[10]
	Γ	-	Medicago truncatula	182/187	HMMER	[4]/[10]
			Phaseolus vulgaris	174	HMMER	[4]
	L		Manihot esculenta	137/153	HMMER	[10]/[4]
			Ricinus communis	109/117	HMMER	[4]/[10]
	+		Linum usitatissimum	137/192	Blastp/HMMER	[1]/[4]
	\parallel		Populus trichocarpa	178/208/236	Blastp/HMMER	[2]/[4] [10]
l	١L		Fragaria vesca	133	HMMER	[4]
			Fragaria x ananassa	199	NS	[9]
			Malus domestica	241/284	Blastp/HMMER	[2]/[4]
	1-		Prunus persica	162	HMMER	[4]
			Brassica rapa	144	HMMER	[4]
			Arabidopsis lyrata	116/118	HMMER	[4]/[10]
	L	L_[Arabidopsis thaliana	109	Blastp/PSI-BLAST	[7]
		ון נ	Capsella rubella	110	HMMER	[4]
	Ш		Thellungiella halophila	87	HMMER	[4]
	11		Carica papaya	56/68	HMMER	[10]/[4]
			Gossypium arboreum	149	Blastp-PFAM-MEME	[5]
	\parallel		Gossypium hirsutum	196	Blastp-PFAM-MEME	[5]
			Gossypium raimondii	142/152	Blastp-PFAM-MEME	[5]/[4]
	\parallel		Eucalyptus grandis	370	HMMER	[4]
	IL		Citrus clementina	152	HMMER	[4]
	[Citrus sinensis	126	HMMER	[4]
	\vdash		Vitis vinifera	74/102/181	Blastp/HMMER	[4]/[2]/[10]
	L		Aquilegia coerulea	146	HMMER	[4]

Figure 3



Figure 4



Supp data 1



References

Ahn, Y.O., Shimizu, B., Sakata, K., Gantulga, D., Zhou, C., Bevan, D.R., et al. (2010). Scopolin-hydrolyzing beta-glucosidases in roots of Arabidopsis. Plant Cell Physiol 51(1), 132-143. doi: 10.1093/pcp/pcp174.

Ahrazem, O., Rubio-Moraga, A., Trapero-Mozos, A., Climent, M.F., Gomez-Cadenas, A., and Gomez-Gomez, L. (2015). Ectopic expression of a stress-inducible glycosyltransferase from saffron enhances salt and oxidative stress tolerance in Arabidopsis while alters anchor root formation. Plant Science 234, 60-73. doi: 10.1016/j.plantsci.2015.02.004.

Alejandro, S., Lee, Y., Tohge, T., Sudre, D., Osorio, S., Park, J., et al. (2012). AtABCG29 Is a Monolignol Transporter Involved in Lignin Biosynthesis. Current Biology 22(13), 1207-1212. doi: http://dx.doi.org/10.1016/j.cub.2012.04.064.

Babst, B.A., Chen, H.Y., Wang, H.Q., Payyavula, R.S., Thomas, T.P., Harding, S.A., et al. (2014). Stress-responsive hydroxycinnamate glycosyltransferase modulates phenylpropanoid metabolism in Populus. Journal of Experimental Botany 65(15), 4191-4200. doi: 10.1093/jxb/eru192 eru192 [pii].

Baiya, S., Hua, Y., Ekkhara, W., and Ketudat Cairns, J.R. (2014). Expression and enzymatic properties of rice (Oryza sativa L.) monolignol beta-glucosidases. Plant Science 227, 101-109. doi: http://dx.doi.org/10.1016/j.plantsci.2014.07.009.

Barvkar, V.T., Pardeshi, V.C., Kale, S.M., Kadoo, N.Y., and Gupta, V.S. (2012). Phylogenomic analysis of UDP glycosyltransferase 1 multigene family in Linum usitatissimum identified genes with varied expression patterns. BMC Genomics 13, 175. doi: 10.1186/1471-2164-13-175.

Baumert, A., Milkowski, C., Schmidt, J., Nimtz, M., Wray, V., and Strack, D. (2005). Formation of a complex pattern of sinapate esters in Brassica napus seeds, catalyzed by enzymes of a serine carboxypeptidase-like acyltransferase family? Phytochemistry 66(11), 1334-1345. doi: 10.1016/j.phytochem.2005.02.031.

Bock, K.W. (2016). The UDP-glycosyltransferase (UGT) superfamily expressed in humans, insects and plants: Animal-plant arms-race and co-evolution. Biochemical Pharmacology 99, 11-17. doi: 10.1016/j.bcp.2015.10.001.

Boerjan, W., Ralph, J., and Baucher, M. (2003). Lignin biosynthesis. Annual Review of Plant Biology 54, 519-546. doi: 10.1146/annurev.arplant.54.031902.134938.

Borges, R.M., Bessiere, J.M., and Hossaert-McKey, M. (2008). The chemical ecology of seed dispersal in monoecious and dioecious figs. Functional Ecology 22, 484–493.

Boss, P.K., Davies, C., and Robinson, S.P. (1996). Expression of anthocyanin biosynthesis pathway genes in red and white grapes. Plant Molecular Biology 32(3), 565-569.

Bowles, D., Isayenkova, J., Lim, E.-K., and Poppenberger, B. (2005). Glycosyltransferases: managers of small molecules. Current Opinion in Plant Biology 8(3), 254-263. doi: http://dx.doi.org/10.1016/j.pbi.2005.03.007.

Brazier-Hicks, M., Evans, K.M., Gershater, M.C., Puschmann, H., Steel, P.G., and Edwards, R. (2009). The C-glycosylation of flavonoids in cereals. Journal of Biological Chemistry 284(27), 17926-17934. doi: 10.1074/jbc.M109.009258.

Brazier-Hicks, M., Offen, W.A., Gershater, M.C., Revett, T.J., Lim, E.K., Bowles, D.J., et al. (2007). Characterization and engineering of the bifunctional N- and O-glucosyltransferase involved in xenobiotic metabolism in plants. Proceedings of the National Academy of Sciences 104(51), 20238-20243. doi: 10.1073/pnas.0706421104.

Burlat, V., Kwon, M., Davin, L.B., and Lewis, N.G. (2001). Dirigent proteins and dirigent sites in lignifying tissues. Phytochemistry 57(6), 883-897. doi: S0031-9422(01)00117-0 [pii].

Cao, P.J., Bartley, L.E., Jung, K.H., and Ronald, P.C. (2008). Construction of a rice glycosyltransferase phylogenomic database and identification of rice-diverged glycosyltransferases. Mol Plant 1(5), 858-877. doi: 10.1093/mp/ssn052.

Caputi, L., Malnoy, M., Goremykin, V., Nikiforova, S., and Martens, S. (2012). A genome-wide phylogenetic reconstruction of family 1 UDP-glycosyltransferases revealed the expansion of the family during the adaptation of plants to life on land. Plant J 69(6), 1030-1042. doi: 10.1111/j.1365-313X.2011.04853.x.

Carter, C., Pan, S., Zouhar, J., Avila, E.L., Girke, T., and Raikhel, N.V. (2004). The vegetative vacuole proteome of Arabidopsis thaliana reveals predicted and unexpected proteins. The Plant Cell 16(12), 3285-3303. doi: tpc.104.027078 [pii] 10.1105/tpc.104.027078.

Chabannes, M., Barakate, A., Lapierre, C., Marita, J.M., Ralph, J., Pean, M., et al. (2001). Strong decrease in lignin content without significant alteration of plant development is induced by simultaneous down-regulation of cinnamoyl CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD) in tobacco plants. The Plant Journal 28(3), 257-270.

Chantreau, M., Portelette, A., Dauwe, R., Kiyoto, S., Cronier, D., Morreel, K., et al. (2014). Ectopic lignification in the flax lignified bast fiber1 mutant stem is associated with tissue-specific modifications in gene expression and cell wall composition. The Plant Cell 26(11), 4462-4482. doi: 10.1105/tpc.114.130443.

Chapelle, A., Morreel, K., Vanholme, R., Le-Bris, P., Morin, H., Lapierre, C., et al. (2012). Impact of the absence of stemspecific beta-glucosidases on lignin and monolignols. Plant Physiology 160(3), 1204-1217. doi: 10.1104/pp.112.203364.

Cheynier, V., Comte, G., Davies, K.M., Lattanzio, V., and Martens, S. (2013). Plant phenolics: Recent advances on their biosynthesis, genetics, and ecophysiology. Plant Physiology and Biochemistry 72, 1-20. doi: http://dx.doi.org/10.1016/j.plaphy.2013.05.009.

Chong, J., Baltz, R., Schmitt, C., Beffa, R., Fritig, B., and Saindrenan, P. (2002a). Downregulation of a Pathogen-Responsive Tobacco UDP-Glc:Phenylpropanoid Glucosyltransferase Reduces Scopoletin Glucoside Accumulation, Enhances Oxidative Stress, and Weakens Virus Resistance. The Plant Cell 14(5), 1093-1107. doi: 10.1105/tpc.010436. Chong, J., Baltz, R., Schmitt, C., Beffa, R., Fritig, B., and Saindrenan, P. (2002b). Downregulation of a pathogen-responsive tobacco UDP-Glc:phenylpropanoid glucosyltransferase reduces scopoletin glucoside accumulation, enhances oxidative stress, and weakens virus resistance. The Plant Cell 14(5), 1093-1107.

Dauwe, R., Morreel, K., Goeminne, G., Gielen, B., Rohde, A., Van Beeumen, J., et al. (2007). Molecular phenotyping of ligninmodified tobacco reveals associated changes in cell-wall metabolism, primary metabolism, stress metabolism and photorespiration. The Plant Journal 52(2), 263-285. doi: 10.1111/j.1365-313X.2007.03233.x.

Davies, G.J., Gloster, T.M., and Henrissat, B. (2005). Recent structural insights into the expanding world of carbohydrateactive enzymes. Current Opinion in Structural Biology 15(6), 637-645. doi: http://dx.doi.org/10.1016/j.sbi.2005.10.008.

Davin, L.B., Wang, H.-B., Crowell, A.L., Bedgar, D.L., Martin, D.M., Sarkanen, S., et al. (1997). Stereoselective Bimolecular Phenoxy Radical Coupling by an Auxiliary (Dirigent) Protein Without an Active Center. Science 275(5298), 362-367.

Day, A., Ruel, K., Neutelings, G., Cronier, D., David, H., Hawkins, S., et al. (2005). Lignification in the flax stem: evidence for an unusual lignin in bast fibers. Planta 222(2), 234-245. doi: 10.1007/s00425-005-1537-1.

de Oliveira, D.M., Finger-Teixeira, A., Rodrigues Mota, T., Salvador, V.H., Moreira-Vilar, F.C., Correa Molinari, H.B., et al. (2015). Ferulic acid: a key component in grass lignocellulose recalcitrance to hydrolysis. Plant Biotechnol J 13(9), 1224-1232. doi: 10.1111/pbi.12292.

Dean, J.C., Kusaka, R., Walsh, P.S., Allais, F., and Zwier, T.S. (2014). Plant Sunscreens in the UV-B: Ultraviolet Spectroscopy of Jet-Cooled Sinapoyl Malate, Sinapic Acid, and Sinapate Ester Derivatives. Journal of the American Chemical Society 136(42), 14780-14795. doi: 10.1021/ja5059026.

Del Rio, D., Stewart, A.J., Mullen, W., Burns, J., Lean, M.E., Brighenti, F., et al. (2004). HPLC-MSn analysis of phenolic compounds and purine alkaloids in green and black tea. Journal of Agriculture and Food Chemistry 52(10), 2807-2815. doi: 10.1021/jf0354848.

Dharmawardhana, D.P., Ellis, B.E., and Carlson, J.E. (1995). A beta-glucosidase from lodgepole pine xylem specific for the lignin precursor coniferin. Plant Physiology 107(2), 331-339.

Dima, O., Morreel, K., Vanholme, B., Kim, H., Ralph, J., and Boerjan, W. (2015). Small glycosylated lignin oligomers are stored in Arabidopsis leaf vacuoles. The Plant Cell 27(3), 695-710. doi: 10.1105/tpc.114.134643 tpc.114.134643 [pii].

Donaldson, L.A. (2001). Lignification and lignin topochemistry - an ultrastructural view. Phytochemistry 57(6), 859-873.

Dorey, S., Baillieul, F., Pierrel, M.-A., Saindrenan, P., Fritig, B., and Kauffmann, S. (1997). Spatial and Temporal Induction of Cell Death, Defense Genes, and Accumulation of Salicylic Acid in Tobacco Leaves Reacting Hypersensitively to a Fungal Glycoprotein Elicitor. Molecular Plant-Microbe Interactions 10(5), 646-655. doi: 10.1094/MPMI.1997.10.5.646.

Ehlting, J., Buttner, D., Wang, Q., Douglas, C.J., Somssich, I.E., and Kombrink, E. (1999). Three 4-coumarate:coenzyme A ligases in Arabidopsis thaliana represent two evolutionarily divergent classes in angiosperms. The Plant Journal 19(1), 9-20.

Ekstrom, A., Taujale, R., McGinn, N., and Yin, Y. (2014). PlantCAZyme: a database for plant carbohydrate-active enzymes. Database (Oxford) 2014. doi: 10.1093/database/bau079.

Escamilla-Trevino, L.L., Chen, W., Card, M.L., Shih, M.C., Cheng, C.L., and Poulton, J.E. (2006). Arabidopsis thaliana beta-Glucosidases BGLU45 and BGLU46 hydrolyse monolignol glucosides. Phytochemistry 67(15), 1651-1660. doi: 10.1016/j.phytochem.2006.05.022.

Fan, T.W., and Conn, E.E. (1985). Isolation and characterization of two cyanogenic beta-glucosidases from flax seeds. Archives of Biochemistry and Biophysics 243(2), 361-373.

Fernandez-Perez, F., Pomar, F., Pedreno, M.A., and Novo-Uzal, E. (2015). Suppression of Arabidopsis peroxidase 72 alters cell wall and phenylpropanoid metabolism. Plant Science 239, 192-199. doi: 10.1016/j.plantsci.2015.08.001.

Ferrer, J.L., Austin, M.B., Stewart, C., Jr., and Noel, J.P. (2008). Structure and function of enzymes involved in the biosynthesis of phenylpropanoids. Plant Physiology and Biochemistry 46(3), 356-370. doi: 10.1016/j.plaphy.2007.12.009.

Fini, A., Guidi, L., Ferrini, F., Brunetti, C., Di Ferdinando, M., Biricolti, S., et al. (2012). Drought stress has contrasting effects on antioxidant enzymes activity and phenylpropanoid biosynthesis in Fraxinus ornus leaves: an excess light stress affair? Journal of Plant Physiology 169(10), 929-939. doi: 10.1016/j.jplph.2012.02.014.

Fraissinet-Tachet, L., Baltz, R., Chong, J., Kauffmann, S., Fritig, B., and Saindrenan, P. (1998). Two tobacco genes induced by infection, elicitor and salicylic acid encode glucosyltransferases acting on phenylpropanoids and benzoic acid derivatives, including salicylic acid. FEBS Letters 437(3), 319-323. doi: 10.1016/s0014-5793(98)01257-5.

Francisco, R.M., Regalado, A., Ageorges, A.s., Burla, B.J., Bassin, B., Eisenach, C., et al. (2013). ABCC1, an ATP Binding Cassette Protein from Grape Berry, Transports Anthocyanidin 3-O-Glucosides. The Plant Cell 25(5), 1840-1854.

Fraser, C.M., Thompson, M.G., Shirley, A.M., Ralph, J., Schoenherr, J.A., Sinlapadech, T., et al. (2007). Related Arabidopsis serine carboxypeptidase-like sinapoylglucose acyltransferases display distinct but overlapping substrate specificities. Plant Physiology 144(4), 1986-1999. doi: 10.1104/pp.107.098970.

Freudenberg, K. (1968). Lignin biosynthesis. Berlin, Heidelberg, New York: Springer.

Freudenberg, K., and Harkin, J.M. (1963). The glucosides of cambial sap of spruce. Phytochemistry 2(2), 189-193. doi: http://dx.doi.org/10.1016/S0031-9422(00)82980-5.

Freudenberg, K., and Torres-Serres, J. (1967). [Conversion of the phenylalanines in lignin-component glucosides]. Justus Liebigs Annalen der Chemie 703(0075-4617 (Print)), 225-230.

Frydman, A., Liberman, R., Huhman, D.V., Carmeli-Weissberg, M., Sapir-Mir, M., Ophir, R., et al. (2013). The molecular and enzymatic basis of bitter/non-bitter flavor of citrus fruit: evolution of branch-forming rhamnosyltransferases under domestication. The Plant Journal 73(1), 166-178. doi: 10.1111/tpj.12030.

Fu, C., Mielenz, J.R., Xiao, X., Ge, Y., Hamilton, C.Y., Rodriguez, M., Jr., et al. (2011). Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. Proceedings of the National Academy of Sciences 108(9), 3803-3808. doi: 10.1073/pnas.1100310108.

Gachon, C.M., Langlois-Meurinne, M., and Saindrenan, P. (2005). Plant secondary metabolism glycosyltransferases: the emerging functional analysis. Trends in Plant Science 10(11), 542-549. doi: 10.1016/j.tplants.2005.09.007.

Gomez-Anduro, G., Ceniceros-Ojeda, E.A., Casados-Vazquez, L.E., Bencivenni, C., Sierra-Beltran, A., Murillo-Amador, B., et al. (2011). Genome-wide analysis of the beta-glucosidase gene family in maize (Zea mays L. var B73). Plant Molecular Biology 77(1-2), 159-183. doi: 10.1007/s11103-011-9800-2.

Griesser, M., Hoffmann, T., Bellido, M.L., Rosati, C., Fink, B., Kurtzer, R., et al. (2008). Redirection of flavonoid biosynthesis through the down-regulation of an anthocyanidin glucosyltransferase in ripening strawberry fruit. Plant Physiology 146(4), 1528-1539. doi: 10.1104/pp.107.114280.

Grotewold, E. (2006). The genetics and biochemistry of floral pigments. Annual Review of Plant Biology 57, 761-780. doi: 10.1146/annurev.arplant.57.032905.105248.

Grubb, C.D., Zipp, B.J., Ludwig-Muller, J., Masuno, M.N., Molinski, T.F., and Abel, S. (2004). Arabidopsis glucosyltransferase UGT74B1 functions in glucosinolate biosynthesis and auxin homeostasis. The Plant Journal 40(6), 893-908. doi: 10.1111/j.1365-313X.2004.02261.x.

Ha, X., Koopmann, B., and von Tiedemann, A. (2016). Wheat Blast and Fusarium Head Blight Display Contrasting Interaction Patterns on Ears of Wheat Genotypes Differing in Resistance. Phytopathology 106(3), 270-281. doi: 10.1094/PHYTO-09-15-0202-R.

Hammond-Kosack, K.E., and Jones, J.D. (1996). Resistance gene-dependent plant defense responses. The Plant Cell 8(10), 1773-1791.

Hano, C., Martin, I., Fliniaux, O., Legrand, B., Gutierrez, L., Arroo, R.R., et al. (2006). Pinoresinol-lariciresinol reductase gene expression and secoisolariciresinol diglucoside accumulation in developing flax (Linum usitatissimum) seeds. Planta 224(6), 1291-1301. doi: 10.1007/s00425-006-0308-y.

Harborne, J.B., and Williams, C.A. (2001). Anthocyanins and other flavonoids. Natural Products Reports 18(3), 310-333.

He, F., Chen, W.K., Yu, K.J., Ji, X.N., Duan, C.Q., Reeves, M.J., et al. (2015). Molecular and biochemical characterization of the UDP-glucose: Anthocyanin 5-O-glucosyltransferase from Vitis amurensis. Phytochemistry 117, 363-372. doi: 10.1016/j.phytochem.2015.06.023.

He, L., and Terashima, N. (1990). Formation and Structure of Lignin in Monocotyledons. III. Heterogeneity of Sugarcane (Saccharum officinarum L.) Lignin with Respect to the Composition of Structural Units in Different Morphological Regions. Journal of Wood Chemistry and Technology 10(4), 435-459. doi: 10.1080/02773819008050251.

He, L., and Terashima, N. (1991). "Formation and Structure of Lignin in Monocotyledons IV. Deposition Process and Structural Diversity of the Lignin in the CellWall of Sugarcane and Rice Plant Studied by Ultraviolet Microscopic Spectroscopy", in: Holzforschung - International Journal of the Biology, Chemistry, Physics and Technology of Wood.).

Hectors, K., Van Oevelen, S., Geuns, J., Guisez, Y., Jansen, M.A., and Prinsen, E. (2014). Dynamic changes in plant secondary metabolites during UV acclimation in Arabidopsis thaliana. Physiologia Plantarum 152(2), 219-230. doi: 10.1111/ppl.12168.

Hemm, M.R., Rider, S.D., Ogas, J., Murry, D.J., and Chapple, C. (2004). Light induces phenylpropanoid metabolism in Arabidopsis roots. The Plant Journal 38(5), 765-778. doi: 10.1111/j.1365-313X.2004.02089.x.

Herrero, J., Esteban-Carrasco, A., and Zapata, J.M. (2013). Looking for Arabidopsis thaliana peroxidases involved in lignin biosynthesis. Plant Physiology and Biochemistry 67, 77-86. doi: 10.1016/j.plaphy.2013.02.019.

Hettwer, K., Böttcher, C., Frolov, A., Mittasch, J., Albert, A., von Roepenack-Lahaye, E., et al. (2016). Dynamic metabolic changes in seeds and seedlings of Brassica napus (oilseed rape) suppressing UGT84A9 reveal plasticity and molecular regulation of the phenylpropanoid pathway. Phytochemistry 124, 46-57. doi: http://dx.doi.org/10.1016/j.phytochem.2016.01.014.

Hjelmeland, A., and Ebeler, S. (2015). Glycosidically Bound Volatile Aroma Compounds in Grapes and Wine: A Review. American Journal of Enology and Viticulture 66(1), 1-11.

Holbein, J., Grundler, F.M., and Siddique, S. (2016). Plant basal resistance to nematodes: an update. Journal of Experimental Botany 67(7), 2049-2061. doi: 10.1093/jxb/erw005.

Holl, J., Vannozzi, A., Czemmel, S., D'Onofrio, C., Walker, A.R., Rausch, T., et al. (2013). The R2R3-MYB transcription factors MYB14 and MYB15 regulate stilbene biosynthesis in Vitis vinifera. The Plant Cell 25(10), 4135-4149. doi: 10.1105/tpc.113.117127.

Hosel, W., Surholt, E., and Borgmann, E. (1978). Characterization of beta-Glucosidase Isoenzymes Possibly Involved in Lignification from Chick Pea (Cicer arietinum L.) Cell Suspension Cultures. European Journal of Biochemistry 84(2), 487-492. doi: 10.1111/j.1432-1033.1978.tb12190.x.

Huang, J., Pang, C., Fan, S., Song, M., Yu, J., Wei, H., et al. (2015). Genome-wide analysis of the family 1 glycosyltransferases in cotton. Mol Genet Genomics 290(5), 1805-1818. doi: 10.1007/s00438-015-1040-8.

Huis, R., Morreel, K., Fliniaux, O., Lucau-Danila, A., Fenart, S., Grec, S., et al. (2012). Natural hypolignification is associated with extensive oligolignol accumulation in flax stems. Plant Physiology 158(4), 1893-1915. doi: 10.1104/pp.111.192328.

Imran, M., Ahmad, N., Anjum, F.M., Khan, M.K., Mushtaq, Z., Nadeem, M., et al. (2015). Potential protective properties of flax lignan secoisolariciresinol diglucoside. Nutrition Journal 14(1), 1-7. doi: 10.1186/s12937-015-0059-3.

Ishihara, H., Tohge, T., Viehover, P., Fernie, A.R., Weisshaar, B., and Stracke, R. (2016). Natural variation in flavonol accumulation in Arabidopsis is determined by the flavonol glucosyltransferase BGLU6. Journal of Experimental Botany 67(5), 1505-1517. doi: 10.1093/jxb/erv546.

Jaakola, L. (2013). New insights into the regulation of anthocyanin biosynthesis in fruits. Trends in Plant Science 18(9), 477-483. doi: 10.1016/j.tplants.2013.06.003.

Jackson, R., Knisley, D., McIntosh, C., and Pfeiffer, P. (2011). Predicting Flavonoid UGT Regioselectivity. Advances in Bioinformatics 2011, 506583. doi: 10.1155/2011/506583.

Jones, P., and Vogt, T. (2001). Glycosyltransferases in secondary plant metabolism: tranquilizers and stimulant controllers. Planta 213(2), 164-174. doi: 10.1007/s004250000492.

Kai, K., Shimizu, B.-i., Mizutani, M., Watanabe, K., and Sakata, K. (2006). Accumulation of coumarins in Arabidopsis thaliana. Phytochemistry 67(4), 379-386. doi: http://dx.doi.org/10.1016/j.phytochem.2005.11.006.

Kaneda, M., Rensing, K.H., Wong, J.C., Banno, B., Mansfield, S.D., and Samuels, A.L. (2008). Tracking monolignols during wood development in lodgepole pine. Plant Physiology 147(4), 1750-1760. doi: 10.1104/pp.108.121533 pp.108.121533 [pii].

Ketudat Cairns, J.R., and Esen, A. (2010). beta-Glucosidases. Cell Mol Life Sci 67(20), 3389-3405. doi: 10.1007/s00018-010-0399-2.

Ketudat Cairns, J.R., Mahong, B., Baiya, S., and Jeon, J.S. (2015). beta-Glucosidases: Multitasking, moonlighting or simply misunderstood? Plant Science 241, 246-259. doi: 10.1016/j.plantsci.2015.10.014 S0168-9452(15)30098-4 [pii].

Kim, S.J., Kwon do, Y., Kim, Y.S., and Kim, Y.C. (2010). Peroxyl radical scavenging capacity of extracts and isolated components from selected medicinal plants. Archives of Pharmacal Research 33(6), 867-873. doi: 10.1007/s12272-010-0609-3.

Kitamura, S., Shikazono, N., and Tanaka, A. (2004). TRANSPARENT TESTA 19 is involved in the accumulation of both anthocyanins and proanthocyanidins in Arabidopsis. The Plant Journal 37(1), 104-114. doi: 1943 [pii].

Koeduka, T., Sugimoto, K., Watanabe, B., Someya, N., Kawanishi, D., Gotoh, T., et al. (2014). Bioactivity of natural 0prenylated phenylpropenes from Illicium anisatum leaves and their derivatives against spider mites and fungal pathogens. Plant Biology 16(2), 451-456. doi: 10.1111/plb.12054.

Kojima, M., Poulton, J.E., Thayer, S.S., and Conn, E.E. (1979). Tissue Distributions of Dhurrin and of Enzymes Involved in Its Metabolism in Leaves of Sorghum bicolor. Plant Physiology 63(6), 1022-1028.

Konig, S., Feussner, K., Kaever, A., Landesfeind, M., Thurow, C., Karlovsky, P., et al. (2014). Soluble phenylpropanoids are involved in the defense response of Arabidopsis against Verticillium longisporum. New Phytologist 202(3), 823-837. doi: 10.1111/nph.12709.

Kubo, H., Nawa, N., and Lupsea, S.A. (2007). Anthocyaninless1 gene of Arabidopsis thaliana encodes a UDP-glucose:flavonoid-3-O-glucosyltransferase. Journal of Plant Research 120(3), 445-449. doi: 10.1007/s10265-006-0067-7.

Kuhlmann, S., Kranz, K., Löcking, B., Alfermann, A.W., and Petersen, M. (2002). Aspects of cytotoxic lignan biosynthesis in suspension cultures of Linum nodiflorum. Phytochemistry Reviews 1(1), 37-43. doi: 10.1023/a:1015876001812.

Kylli, P., Nousiainen, P., Biely, P., Sipila, J., Tenkanen, M., and Heinonen, M. (2008). Antioxidant potential of hydroxycinnamic acid glycoside esters. Journal of Agriculture and Food Chemistry 56(12), 4797-4805. doi: 10.1021/jf800317v.

Landry, L.G., Chapple, C.C., and Last, R.L. (1995). Arabidopsis mutants lacking phenolic sunscreens exhibit enhanced ultraviolet-B injury and oxidative damage. Plant Physiology 109(4), 1159-1166.

Langenbach, C., Campe, R., Schaffrath, U., Goellner, K., and Conrath, U. (2013). UDP-glucosyltransferase UGT84A2/BRT1 is required for Arabidopsis nonhost resistance to the Asian soybean rust pathogen Phakopsora pachyrhizi. New Phytologist 198(2), 536-545. doi: 10.1111/nph.12155.

Langlois-Meurinne, M., Gachon, C.M., and Saindrenan, P. (2005). Pathogen-responsive expression of glycosyltransferase genes UGT73B3 and UGT73B5 is necessary for resistance to Pseudomonas syringae pv tomato in Arabidopsis. Plant Physiology 139(4), 1890-1901. doi: 10.1104/pp.105.067223.

Lanot, A., Hodge, D., Jackson, R.G., George, G.L., Elias, L., Lim, E.K., et al. (2006). The glucosyltransferase UGT72E2 is responsible for monolignol 4-O-glucoside production in Arabidopsis thaliana. The Plant Journal 48(2), 286-295. doi: 10.1111/j.1365-313X.2006.02872.x.

Lanot, A., Hodge, D., Lim, E.K., Vaistij, F.E., and Bowles, D.J. (2008). Redirection of flux through the phenylpropanoid pathway by increased glucosylation of soluble intermediates. Planta 228(4), 609-616. doi: 10.1007/s00425-008-0763-8.

Larsen, E.S., Alfenito, M.R., Briggs, W.R., and Walbot, V. (2003). A carnation anthocyanin mutant is complemented by the glutathione S-transferases encoded by maize Bz2 and petunia An9. Plant Cell Reports 21(9), 900-904. doi: 10.1007/s00299-002-0545-x.

Leinhos, V., and Savidge, R.A. (1993). Isolation of protoplasts from developing xylem of Pinus banksiana and Pinus strobus. Canadian Journal of Forest Research 23(3), 343-348. doi: 10.1139/x93-050.

Li, X., Bergelson, J., and Chapple, C. (2010). The ARABIDOPSIS Accession Pna-10 Is a Naturally Occurring sng1 Deletion Mutant. Molecular Plant 3(1), 91-100. doi: 10.1093/mp/ssp090.

Li, X.J., Zhang, J.Q., Wu, Z.C., Lai, B., Huang, X.M., Qin, Y.H., et al. (2016). Functional characterization of a glucosyltransferase gene, LcUFGT1, involved in the formation of cyanidin glucoside in the pericarp of Litchi chinensis. Physiologia Plantarum. doi: 10.1111/ppl.12391.

Li, Y., Li, P., Wang, Y., Dong, R., Yu, H., and Hou, B. (2014). Genome-wide identification and phylogenetic analysis of Family-1 UDP glycosyltransferases in maize (Zea mays). Planta 239(6), 1265-1279. doi: 10.1007/s00425-014-2050-1.

Lim, E.K., and Bowles, D.J. (2004). A class of plant glycosyltransferases involved in cellular homeostasis. EMBO Journal 23(15), 2915-2922. doi: 10.1038/sj.emboj.7600295.

Lim, E.K., Higgins, G.S., Li, Y., and Bowles, D.J. (2003). Regioselectivity of glucosylation of caffeic acid by a UDP-glucose: glucosyltransferase is maintained in planta. Biochemical Journal 373(3), 987-992. doi: 10.1042/bj20021453.

Lim, E.K., Jackson, R.G., and Bowles, D.J. (2005). Identification and characterisation of Arabidopsis glycosyltransferases capable of glucosylating coniferyl aldehyde and sinapyl aldehyde. FEBS Lett 579(13), 2802-2806. doi: 10.1016/j.febslet.2005.04.016.

Lim, E.K., Li, Y., Parr, A., Jackson, R., Ashford, D.A., and Bowles, D.J. (2001). Identification of glucosyltransferase genes involved in sinapate metabolism and lignin synthesis in Arabidopsis. Journal of Biological Chemistry 276(6), 4344-4349. doi: 10.1074/jbc.M007263200

Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Research 42(Database issue), D490-495. doi: 10.1093/nar/gkt1178gkt1178 [pii].

Lorenc-Kukula, K., Amarowicz, R., Oszmianski, J., Doermann, P., Starzycki, M., Skala, J., et al. (2005). Pleiotropic effect of phenolic compounds content increases in transgenic flax plant. Journal of Agriculture and Food Chemistry 53(9), 3685-3692. doi: 10.1021/jf047987z.

Luang, S., Cho, J.I., Mahong, B., Opassiri, R., Akiyama, T., Phasai, K., et al. (2013). Rice Os9BGlu31 is a transglucosidase with the capacity to equilibrate phenylpropanoid, flavonoid, and phytohormone glycoconjugates. Journal of Biological Chemistry 288(14), 10111-10123. doi: 10.1074/jbc.M112.423533.

Lunkenbein, S., Bellido, M., Aharoni, A., Salentijn, E.M., Kaldenhoff, R., Coiner, H.A., et al. (2006). Cinnamate metabolism in ripening fruit. Characterization of a UDP-glucose:cinnamate glucosyltransferase from strawberry. Plant Physiology 140(3), 1047-1058. doi: pp.105.074955 [pii] 10.1104/pp.105.074955.

Mackenzie, P.I., Owens, I.S., Burchell, B., Bock, K.W., Bairoch, A., Belanger, A., et al. (1997). The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence. Pharmacogenetics 7(4), 255-269.

Marcinowski, S., and Grisebach, H. (1978). Enzymology of lignification: Cell wall beta-glucosidase for coniferin from spruce (Picea abies) seedlings. European Journal of Biochemistry 87(1), 37-44.

Marjamaa, K., Lehtonen, M., Lundell, T., Toikka, M., Saranpaa, P., and Fagerstedt, K.V. (2003). Developmental lignification and seasonal variation in beta-glucosidase and peroxidase activities in xylem of Scots pine, Norway spruce and silver birch. Tree Physiology 23(14), 977-986.

Marrs, K.A., Alfenito, M.R., Lloyd, A.M., and Walbot, V. (1995). A glutathione S-transferase involved in vacuolar transfer encoded by the maize gene Bronze-2. Nature 375(6530), 397-400.

Masada, S., Terasaka, K., Oguchi, Y., Okazaki, S., Mizushima, T., and Mizukami, H. (2009). Functional and structural characterization of a flavonoid glucoside 1,6-glucosyltransferase from Catharanthus roseus. Plant and Cell Physiology 50(8), 1401-1415. doi: 10.1093/pcp/pcp088.

Matros, A., and Mock, H.P. (2004). Ectopic expression of a UDP-glucose:phenylpropanoid glucosyltransferase leads to increased resistance of transgenic tobacco plants against infection with Potato Virus Y. Plant and Cell Physiology 45(9), 1185-1193. doi: 10.1093/pcp/pch140.

Matsuba, Y., Sasaki, N., Tera, M., Okamura, M., Abe, Y., Okamoto, E., et al. (2010). A novel glucosylation reaction on anthocyanins catalyzed by acyl-glucose-dependent glucosyltransferase in the petals of carnation and delphinium. The Plant Cell 22(10), 3374-3389. doi: 10.1105/tpc.110.077487.

Meißner, D., Albert, A., Böttcher, C., Strack, D., and Milkowski, C. (2008). The role of UDP-glucose:hydroxycinnamate glucosyltransferases in phenylpropanoid metabolism and the response to UV-B radiation in Arabidopsis thaliana. Planta 228(4), 663-674. doi: 10.1007/s00425-008-0768-3.

Miao, Y.-C., and Liu, C.-J. (2010). ATP-binding cassette-like transporters are involved in the transport of lignin precursors across plasma and vacuolar membranes. Proceedings of the National Academy of Sciences 107(52), 22728-22733. doi: 10.1073/pnas.1007747108.

Mierziak, J., Kostyn, K., and Kulma, A. (2014). Flavonoids as Important Molecules of Plant Interactions with the Environment. Molecules 19(10). doi: 10.3390/molecules191016240.

Miguel, V.n., Otero, J.A., $Garc\sqrt{\neq}a$ -Villalba, R.o., $Tom\sqrt{\circ}s$ -Barber $\sqrt{\circ}n$, F., $Esp\sqrt{\neq}n$, J.C., Merino, G., et al. (2014). Role of ABCG2 in Transport of the Mammalian Lignan Enterolactone and its Secretion into Milk in Abcg2 Knockout Mice. Drug Metabolism and Disposition 42(5), 943-946. doi: 10.1124/dmd.113.055970.

Milkowski, C., Baumert, A., and Strack, D. (2000). Identification of four Arabidopsis genes encoding hydroxycinnamate glucosyltransferases. FEBS Letters 486(2), 183-184. doi: http://dx.doi.org/10.1016/S0014-5793(00)02270-5.

Milkowski, C., and Strack, D. (2010). Sinapate esters in brassicaceous plants: biochemistry, molecular biology, evolution and metabolic engineering. Planta 232(1), 19-35. doi: 10.1007/s00425-010-1168-z.

Mittasch, J., Strack, D., and Milkowski, C. (2007). Secondary product glycosyltransferases in seeds of Brassica napus. Planta 225(2), 515-522. doi: 10.1007/s00425-006-0360-7.

Mock, H.-P., and Strack, D. (1993). The International Journal of Plant BiochemistryEnergetics of the uridine 5'diphosphoglucose: Hydroxycinnamic acid acyl-glucosyltransferase reaction. Phytochemistry 32(3), 575-579. doi: http://dx.doi.org/10.1016/S0031-9422(00)95139-2.

Nair, R.B., Bastress, K.L., Ruegger, M.O., Denault, J.W., and Chapple, C. (2004). The Arabidopsis thaliana REDUCED EPIDERMAL FLUORESCENCE1 Gene Encodes an Aldehyde Dehydrogenase Involved in Ferulic Acid and Sinapic Acid Biosynthesis. The Plant Cell 16(2), 544-554. doi: 10.1105/tpc.017509.

Noguchi, A., Horikawa, M., Fukui, Y., Fukuchi-Mizutani, M., Iuchi-Okada, A., Ishiguro, M., et al. (2009). Local differentiation of sugar donor specificity of flavonoid glycosyltransferase in Lamiales. The Plant Cell 21(5), 1556-1572. doi: 10.1105/tpc.108.063826 tpc.108.063826 [pii].

Oba, K., Conn, E.E., Canut, H., and Boudet, A.M. (1981). Subcellular Localization of 2-(beta-d-Glucosyloxy)-Cinnamic Acids and the Related beta-glucosidase in Leaves of Melilotus alba Desr. Plant Physiology 68(6), 1359-1363.

Obayashi, T., Hayashi, S., Saeki, M., Ohta, H., and Kinoshita, K. (2009). ATTED-II provides coexpressed gene networks for Arabidopsis. Nucleic Acids Research 37(Database issue), D987-991. doi: 10.1093/nar/gkn807.

Ohkama-Ohtsu, N., Sasaki-Sekimoto, Y., Oikawa, A., Jikumaru, Y., Shinoda, S., Inoue, E., et al. (2011). 12-Oxo-Phytodienoic Acid–Glutathione Conjugate is Transported into the Vacuole in Arabidopsis. Plant and Cell Physiology 52(1), 205-209.

Opassiri, R., Pomthong, B., Onkoksoong, T., Akiyama, T., Esen, A., and Ketudat Cairns, J.R. (2006). Analysis of rice glycosyl hydrolase family 1 and expression of 0s4bglu12 beta-glucosidase. BMC Plant Biology 6, 33. doi: 1471-2229-6-33 [pii] 10.1186/1471-2229-6-33.

Osmani, S.A., Bak, S., Imberty, A., Olsen, C.E., and Moller, B.L. (2008). Catalytic key amino acids and UDP-sugar donor specificity of a plant glucuronosyltransferase, UGT94B1: molecular modeling substantiated by site-specific mutagenesis and biochemical analyses. Plant Physiology 148(3), 1295-1308. doi: 10.1104/pp.108.128256.

Pickett-Heaps, J.D. (1968). Xylem Wall Deposition. Radioautographic Investigations Using Lignin Precursors. Protoplasma 65, 181-205.

Poppenberger, B., Berthiller, F., Lucyshyn, D., Sieberer, T., Schuhmacher, R., Krska, R., et al. (2003). Detoxification of the Fusarium mycotoxin deoxynivalenol by a UDP-glucosyltransferase from Arabidopsis thaliana. J Biol Chem 278(48), 47905-47914. doi: 10.1074/jbc.M307552200.

Prasad, K. (1997). Hydroxyl radical-scavenging property of secoisolariciresinol diglucoside (SDG) isolated from flax-seed. Molecular and Cell Biochemistry 168(1-2), 117-123.

Pusztahelyi, T., Holb, I.J., and Pocsi, I. (2015). Secondary metabolites in fungus-plant interactions. Frontiers in Plant Science 6, 573. doi: 10.3389/fpls.2015.00573.

Ralph, S., Park, J.-Y., Bohlmann, J., and Mansfield, S.D. (2006). Dirigent Proteins in Conifer Defense: Gene Discovery, Phylogeny, and Differential Wound- and Insect-induced Expression of a Family of DIR and DIR-like Genes in Spruce (Picea spp.). Plant Molecular Biology 60(1), 21-40. doi: 10.1007/s11103-005-2226-y.

Rasmussen, S., and Dixon, R.A. (1999). Transgene-mediated and elicitor-induced perturbation of metabolic channeling at the entry point into the phenylpropanoid pathway. The Plant Cell 11(8), 1537-1552.

Roepke, J., and Bozzo, G.G. (2015). Arabidopsis thaliana beta-glucosidase BGLU15 attacks flavonol 3-O-beta-glucoside-7-O-alpha-rhamnosides. Phytochemistry 109, 14-24. doi: 10.1016/j.phytochem.2014.10.028.

Ross, J., Li, Y., Lim, E., and Bowles, D.J. (2001). Higher plant glycosyltransferases. Genome Biology 2(2), REVIEWS3004.

Rouyi, C., Baiya, S., Lee, S.-K., Mahong, B., Jeon, J.-S., Ketudat-Cairns, J.R., et al. (2014). Recombinant Expression and Characterization of the Cytoplasmic Rice β-Glucosidase 0s1BGlu4. PLoS One 9(5), e96712. doi: 10.1371/journal.pone.0096712.

Sablowski, R.W., Moyano, E., Culianez-Macia, F.A., Schuch, W., Martin, C., and Bevan, M. (1994). A flower-specific Myb protein activates transcription of phenylpropanoid biosynthetic genes. The EMBO Journal 13(1), 128-137.

Sasaki, N., and Nakayama, T. (2015). Achievements and perspectives in biochemistry concerning anthocyanin modification for blue flower coloration. Plant and Cell Physiology 56(1), 28-40. doi: 10.1093/pcp/pcu097.

Schafer, J., Caldwell, R., Patterson, F., and Compton, L. (1963). Wheat leaf rust combinations. Phytopathology 53, 569-573.

Schmidt, S., Rainieri, S., Witte, S., Matern, U., and Martens, S. (2011). Identification of a Saccharomyces cerevisiae Glucosidase That Hydrolyzes Flavonoid Glucosides. Applied and Environmental Microbiology 77(5), 1751-1757. doi: 10.1128/AEM.01125-10.

Sekhwal, M.K., Sharma, V., and Sarin, R. (2013). Annotation of glycoside hydrolases in Sorghum bicolor using proteins interaction approach. Journal of Proteome Science and Computational Biology 2(1). doi: 10.7243/2050-2273-2-2.

Sermsuvityawong, K., Svasti, M.R.J., Sawang-Areetrakul, P., Kisamanonta, P., and Chulavatnatol, M. (1995). Aggregation of Cassava linamarase. Journal of the Science Society of Thailand 21, 283-292.

Sharma, R., Rawat, V., and Suresh, C.G. (2014). Genome-wide identification and tissue-specific expression analysis of UDPglycosyltransferases genes confirm their abundance in Cicer arietinum (Chickpea) genome. PLoS One 9(10), e109715. doi: 10.1371/journal.pone.0109715.

Shimizu, B. (2014). 2-Oxoglutarate-dependent dioxygenases in the biosynthesis of simple coumarins. Frontiers in Plant Science 5, 549. doi: 10.3389/fpls.2014.00549.

Simon, C., Langlois-Meurinne, M., Didierlaurent, L., Chaouch, S., Bellvert, F., Massoud, K., et al. (2014). The secondary metabolism glycosyltransferases UGT73B3 and UGT73B5 are components of redox status in resistance of Arabidopsis to Pseudomonas syringae pv. tomato. Plant Cell and Environment 37(5), 1114-1129. doi: 10.1111/pce.12221.

Sinlapadech, T., Stout, J., Ruegger, M.O., Deak, M., and Chapple, C. (2007). The hyper-fluorescent trichome phenotype of the brt1 mutant of Arabidopsis is the result of a defect in a sinapic acid: UDPG glucosyltransferase. The Plant Journal 49(4), 655-668. doi: TPJ2984 [pii]10.1111/j.1365-313X.2006.02984.x.

Song, C., Gu, L., Liu, J., Zhao, S., Hong, X., Schulenburg, K., et al. (2015). Functional Characterization and Substrate Promiscuity of UGT71 Glycosyltransferases from Strawberry (Fragaria x ananassa). Plant Cell Physiol 56(12), 2478-2493. doi: 10.1093/pcp/pcv151.

Stafford, H.A. (1991). Flavonoid evolution: an enzymic approach. Plant Physiology 96(3), 680-685.

Sterjiades, R., Dean, J.F., and Eriksson, K.E. (1992). Laccase from Sycamore Maple (Acer pseudoplatanus) Polymerizes Monolignols. Plant Physiology 99(3), 1162-1168.

Strack, D. (2001). "Enzymes involved in hydroxycinnamate metabolism," in Methods in Enzymology. Academic Press), 70-81.

Tanaka, Y., Sasaki, N., and Ohmiya, A. (2008). Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. The Plant Journal 54(4), 733-749. doi: 10.1111/j.1365-313X.2008.03447.x.

Tanaka, Y., Yonekura, K., Fukuchi-Mizutani, M., Fukui, Y., Fujiwara, H., Ashikari, T., et al. (1996). Molecular and biochemical characterization of three anthocyanin synthetic enzymes from Gentiana triflora. Plant and Cell Physiology 37(5), 711-716.

Terashima, N., and Fukushima, K. (1988). Heterogeneity in formation of lignin—XI: An autoradiographic study of the heterogeneous formation and structure of pine lignin. Wood Science and Technology 22(3), 259-270. doi: 10.1007/bf00386021.

Tikunov, Y.M., Molthoff, J., de Vos, R.C., Beekwilder, J., van Houwelingen, A., van der Hooft, J.J., et al. (2013). Non-smoky glycosyltransferase1 prevents the release of smoky aroma from tomato fruit. The Plant Cell 25(8), 3067-3078. doi: 10.1105/tpc.113.114231.

Timell, T. (1986). Compression wood in gymnosperms, vol 1. Berlin-Heidelberg.

Tobimatsu, Y., Elumalai, S., Grabber, J.H., Davidson, C.L., Pan, X., and Ralph, J. (2012). Hydroxycinnamate conjugates as potential monolignol replacements: in vitro lignification and cell wall studies with rosmarinic acid. ChemSusChem 5(4), 676-686. doi: 10.1002/cssc.201100573.

Tohge, T., Nishiyama, Y., Hirai, M.Y., Yano, M., Nakajima, J.-i., Awazuhara, M., et al. (2005). Functional genomics by integrated analysis of metabolome and transcriptome of Arabidopsis plants over-expressing an MYB transcription factor. The Plant Journal 42(2), 218-235. doi: 10.1111/j.1365-313X.2005.02371.x.

Torras-Claveria, L., Jauregui, O., Codina, C., Tiburcio, A.F., Bastida, J., and Viladomat, F. (2012). Analysis of phenolic compounds by high-performance liquid chromatography coupled to electrospray ionization tandem mass spectrometry in senescent and water-stressed tobacco. Plant Science 182, 71-78. doi: 10.1016/j.plantsci.2011.02.009.

Trapero, A., Ahrazem, O., Rubio-Moraga, A., Jimeno, M.L., Gomez, M.D., and Gomez-Gomez, L. (2012). Characterization of a glucosyltransferase enzyme involved in the formation of kaempferol and quercetin sophorosides in Crocus sativus. Plant Physiology 159(4), 1335-1354. doi: 10.1104/pp.112.198069.

Tschaplinski, T.J., Standaert, R.F., Engle, N.L., Martin, M.Z., Sangha, A.K., Parks, J.M., et al. (2012). Down-regulation of the caffeic acid O-methyltransferase gene in switchgrass reveals a novel monolignol analog. Biotechnology for Biofuels 5(1), 71. doi: 10.1186/1754-6834-5-71.

Tsuji, Y., Chen, F., Yasuda, S., and Fukushima, K. (2005). Unexpected behavior of coniferin in lignin biosynthesis of Ginkgo biloba L. Planta 222(1), 58-69. doi: 10.1007/s00425-005-1517-5.

Tsuyama, T., Kawai, R., Shitan, N., Matoh, T., Sugiyama, J., Yoshinaga, A., et al. (2013). Proton-Dependent Coniferin Transport, a Common Major Transport Event in Differentiating Xylem Tissue of Woody Plants. Plant Physiology 162(2), 918-926. doi: 10.1104/pp.113.214957.

Uzal, E.N., Gomez-Ros, L.V., Hernandez, J.A., Pedreno, M.A., Cuello, J., and Ros Barcelo, A. (2009). Analysis of the soluble cell wall proteome of gymnosperms. Journal of Plant Physiology 166(8), 831-843. doi: 10.1016/j.jplph.2008.11.009 S0176-1617(08)00358-1 [pii].

Väisänen, E.E., Smeds, A.I., Fagerstedt, K.V., Teeri, T.H., Willför, S.M., and Kärkönen, A. (2015). Coniferyl alcohol hinders the growth of tobacco BY-2 cells and Nicotiana benthamiana seedlings. Planta 242(3), 747-760. doi: 10.1007/s00425-015-2348-7.

Vanholme, R., Storme, V., Vanholme, B., Sundin, L., Christensen, J.H., Goeminne, G., et al. (2012). A systems biology view of responses to lignin biosynthesis perturbations in Arabidopsis. The Plant Cell 24(9), 3506-3529. doi: 10.1105/tpc.112.102574.

Vila-Real, H., Alfaia, A.J., Bronze, M.R., Calado, A.R., and Ribeiro, M.H. (2011). Enzymatic Synthesis of the Flavone Glucosides, Prunin and Isoquercetin, and the Aglycones, Naringenin and Quercetin, with Selective alpha-L-Rhamnosidase and beta-D-Glucosidase Activities of Naringinase. Enzyme Research 2011, 692618. doi: 10.4061/2011/692618.

Voelker, S.L., Lachenbruch, B., Meinzer, F.C., Jourdes, M., Ki, C., Patten, A.M., et al. (2010). Antisense down-regulation of 4CL expression alters lignification, tree growth, and saccharification potential of field-grown poplar. Plant Physiology 154(2), 874-886. doi: 10.1104/pp.110.159269 pp.110.159269 [pii].

Vogt, T. (2010). Phenylpropanoid Biosynthesis. Molecular Plant 3(1), 2-20. doi: http://dx.doi.org/10.1093/mp/ssp106.

Vogt, T., and Jones, P. (2000). Glycosyltransferases in plant natural product synthesis: characterization of a supergene family. Trends in Plant Science 5(9), 380-386. doi: 10.1016/s1360-1385(00)01720-9.

Walle, T., Browning, A.M., Steed, L.L., Reed, S.G., and Walle, U.K. (2005). Flavonoid Glucosides Are Hydrolyzed and Thus Activated in the Oral Cavity in Humans. The Journal of Nutrition 135(1), 48-52.

Wang, Y., Chantreau, M., Sibout, R., and Hawkins, S. (2013). Plant cell wall lignification and monolignol metabolism. Front Plant Sci 4, 220. doi: 10.3389/fpls.2013.00220.

Wang, Y.W., Wang, W.C., Jin, S.H., Wang, J., Wang, B., and Hou, B.K. (2012). Over-expression of a putative poplar glycosyltransferase gene, PtGT1, in tobacco increases lignin content and causes early flowering. Journal of Experimental Botany 63(7), 2799-2808. doi: 10.1093/jxb/ers001.

Ward, J.L., Baker, J.M., Llewellyn, A.M., Hawkins, N.D., and Beale, M.H. (2011). Metabolomic analysis of Arabidopsis reveals hemiterpenoid glycosides as products of a nitrate ion-regulated, carbon flux overflow. Proceedings of the National Academy of Sciences 108(26), 10762-10767. doi: 10.1073/pnas.1018875108.

Weng, J.K., and Chapple, C. (2010). The origin and evolution of lignin biosynthesis. New Phytologist 187(2), 273-285. doi: 10.1111/j.1469-8137.2010.03327.x.

Whetten, R., and Sederoff, R. (1995). Lignin Biosynthesis. The Plant Cell 7(7), 1001-1013. doi: 10.1105/tpc.7.7.1001.

Wöhler, F., and Liebig, J. (1837). Ueber die Bildung des Bittermandelöls. Annalen der Pharmacie 22, 1-24.

Xu, Z., Escamilla-Trevino, L., Zeng, L., Lalgondar, M., Bevan, D., Winkel, B., et al. (2004). Functional genomic analysis of Arabidopsis thaliana glycoside hydrolase family 1. Plant Molecular Biology 55(3), 343-367. doi: D000000790 [pii]

10.1007/s11103-004-0790-1.

Yamazaki, M., Yamagishi, E., Gong, Z., Fukuchi-Mizutani, M., Fukui, Y., Tanaka, Y., et al. (2002). Two flavonoid glucosyltransferases from Petunia hybrida: molecular cloning, biochemical properties and developmentally regulated expression. Plant Molecular Biology 48(4), 401-411.

Yazaki, K. (2006). ABC transporters involved in the transport of plant secondary metabolites. FEBS Letters 580(4), 1183-1191. doi: http://dx.doi.org/10.1016/j.febslet.2005.12.009.

Yonekura-Sakakibara, K., Fukushima, A., Nakabayashi, R., Hanada, K., Matsuda, F., Sugawara, S., et al. (2012). Two glycosyltransferases involved in anthocyanin modification delineated by transcriptome independent component analysis in Arabidopsis thaliana. The Plant Journal 69(1), 154-167. doi: 10.1111/j.1365-313X.2011.04779.x.

Yonekura-Sakakibara, K., and Hanada, K. (2011). An evolutionary view of functional diversity in family 1 glycosyltransferases. Plant J 66(1), 182-193. doi: 10.1111/j.1365-313X.2011.04493.x.

Yonekura-Sakakibara, K., Nakabayashi, R., Sugawara, S., Tohge, T., Ito, T., Koyanagi, M., et al. (2014). A flavonoid 3-0-glucoside:2"-O-glucosyltransferase responsible for terminal modification of pollen-specific flavonols in Arabidopsis thaliana. The Plant Journal 79(5), 769-782. doi: 10.1111/tpj.12580.

Yonekura-Sakakibara, K., Tohge, T., Matsuda, F., Nakabayashi, R., Takayama, H., Niida, R., et al. (2008). Comprehensive flavonol profiling and transcriptome coexpression analysis leading to decoding gene-metabolite correlations in Arabidopsis. The Plant Cell 20(8), 2160-2176. doi: 10.1105/tpc.108.058040.

Yoshinaga, A., Kamitakahara, H., and Takabe, K. (2015). Distribution of coniferin in differentiating normal and compression woods using MALDI mass spectrometric imaging coupled with osmium tetroxide vapor treatment. Tree Physiology.

Zhao, Q., Nakashima, J., Chen, F., Yin, Y., Fu, C., Yun, J., et al. (2013). Laccase is necessary and nonredundant with peroxidase for lignin polymerization during vascular development in Arabidopsis. The Plant Cell 25(10), 3976-3987. doi: 10.1105/tpc.113.117770.

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(11), 3158-3170. doi: 10.1105/tpc.106.047399.