Université des Sciences et Technologies Lille 1

Ecole doctorale des Sciences de la matière, du rayonnement et de l'environnement

Unité de Glycobiologie Structurale et Fonctionnelle, UMR CNRS-Lille1 8576 UFR de Biologie, USTL

Thèse de doctorat
Présentée à l'Université des Sciences et Technologies de Lille

Le 10 décembre 2015 par :

Malika CHABI

Titre :

Approches transcriptomique et protéomique pour étudier les rôles de l'environnement et du génotype sur le métabolisme pariétal chez le lin

En vue de l'obtention du titre de :
Docteur de l'Université des Sciences et Technologies Lille 1

Spécialité :
Ingénierie des fonctions biologiques

Jury :

Elisabeth Jamet - DR CNRS Toulouse - Rapporteur
Gilles Pilate - DR INRA Orléans - Rapporteur
Eric Lainé - Professeur Université d'Orléans - examinateur
Brigitte Chabbert - CR1 INRA - Reims - examinateur
Anca Lucau - MCF (HDR) - Université de Lille - directrice de thèse
Simon Hawkins - Professeur Université de Lille - directeur de thèse

Remerciements

Lorsque je commence ces lignes je pense à ceux que je souhaite remercier. Et il y a comme une évidence qui se dessine alors à l'horizon. A toutes les personnes que j'ai croisées dédie cette thèse.

Au commencement il y a eu la thèse, et son financement, le sujet était vaste : il y avait des projets et des encadrants au-dessus des idées.
Pour les années passées au labo je remercie tous mes encadrants.
A Simon Hawkins : Ma plus grande reconnaissance

Les étudiants dirent : qu'il y est une pause-café. Et les pauses furent. Et ils virent que la pause était bonne.
Pour ces pauses je remercie mes co-llègues/détenues/pains.

Ma famille dit : « Il y a une étendue d'eau entre deux continents, cependant il y a une même façon de penser sur ces deux terres. » Sur ces deux terres, j'ai grandi et appris. Pour l'éducation, les valeurs et la force que les membres de ma famille m'ont inculquées, je ne les remercierai jamais assez.

Mon mari dit : « Je t'aime » et il m'aida à trouver ma voie même au cœur des ténèbres.
Pour cela je le remercie : قـلـبي

Ma belle-famille dit : que l'amour vous porte et multiplie vos rêves. Pour leurs encouragements et leur soutien, je les remercie.

Les amis dirent : « Buvons du Champ…omy et faisons la fête. » Pour ces soirées, merci !

Aux membres du jury qui donnent de leur temps pour échanger avec moi, et pour corriger ma thèse, en disant « oui » pour venir dans la grisaille, pour m'aider à apprendre encore :
mes plus vifs remerciements.

Si, au final, sans être le Messi(e) ou le Ronaldo de ma science, je suis en mesure de soutenir cette thèse avec le sourire, c'est grâce à toutes ces personnes.

Encore merci à tous ceux que j'ai croisé durant ma thèse.

Mille mercis.

Table des matières

Liste des figures

Liste des tableaux

Liste des abréviations

4CL : 4-Coumarate CoA ligase

AA : auxiliary activites

AAA : Acide aminé aromatique

ABC : ATP binding cassette

ADN : Acide desoxyribonucléique

ADNc : ADN complémentaire

AG : Arabinogalactane

AGI/ II : arabinogalactanes de type I et de type II

AGP : Arabinogalactan-rich protein

ALA : Alpha linolenic acid

Api : Apiose

Ara : Arabinose

AraT : Arabinosyltransferase

ARN : Acide ribonucléique

AX : Arabinoxylan

C3H : Coumarate 3-hydroxylase

C4H : Cinnamate-4-hydroxylase

CAD : Cinnamyl alcohol dehydrogenase

CAZy : Carbohydrate active enzyme

CBM : Carbohydrate binding modules

CC: Cell corner

CCoAOMT : Caffeoyl CoA

CCR : Cinnamoyl-CoA reductase

CESA : Cellulose Synthase A

CHS : Chalcone Synthase

CML: compound middle lamella (middle lamella + primary cell wall)

CNX1 : calnexin 1

Co: Cortex

COMT : Caffeic acid 3-O-methyltransferase

CRT3 : Calreticulin 3

CSE : Caffeoyl shikimate esterase

CSL : Cellulose synthase like

CTL : Chitinase like

CWP : Cell wall protein

DCL : DICER-like

DGL1 : defective glycosylation

DTT : dithiothréitol

EBS1 : EMS-mutagenized bri1 suppressor 1

EDTA : Éthylène Diamine Tétra-Acétique

En : Endoderm

Ep : Epiderme

EST : Expressed sequence tag

Expn ou EXP : Expansin

EXT : Extensine

F5H : Ferulate 5-hydroxylase

FAOSTAT : Food and agriculture organization statistics

FR: Région fibreuse

FSV = Fiber spring variety

FT-IR : Fourier transform infrared

FUT : Fucosyltansferase

FWV = Fiber winter variety

G : Guaiacyl

GalA : Galacturonic acid

GalAT : Galacturonosyltransferase

GAX : Glucuronoarabinoxylane

GEVES : Groupe d'Etude et de contrôle des Variétés Et des Semences

GPI : Glycosylphosphatidylinositol

GRP : Glycine-rich protein

GT : Glycosyl-transferase

Gtons : Gigatonnes

H : p-hydroxyphenyl

HCT : Hydroxycinnamoyl-CoA

HG : Homogalacturonan

HRGP : Hydroxyproline Rich glycoprotein

IPCC : Intergovernmental Panel on Climate Change

KBBE : Knowledge-Based Bio-Economy

LAC : Laccase

LC-MS/MS : liquid chromatography tandem mass spectrometry

LM : lipid metabolism

LRX : Leucine-rich-repeat extensin

ManS : Mannan, synthase

miRNA : microARN

Misc : miscellaneous

MLG: Multi linkage glucanes

NCP : Polysaccharides non cellulosiques

NCP : Non-cellulosic polysaccharides

OR : Oxidoreductases

OWV :Oil winter variety

PACW : Proteins acting on cell wall polysaccharides

PAL : Phenylalanine ammonia-lyase

Pases : Proteases

PCR : Polymerase chain reaction

Phe : Phénylalanine

PID : Protein interacting domains

PME : Pectin methylesterases

PMSF : Fluorure de phénylméthylsulfonyle

PRP : Proline-rich protein

qRT-PCR : Quantitative real-time polymerase chain reaction

RG : Rhamno-galacturonan

RGI : Rhamnogalacturonane I

RGII : Rhamnogalacturonane II

Rha : Rhamnose

RhaT : Rhamnosyltransferase

RNAi : ARN interférant

RT-PCR : Real-time polymerase chain reaction

S : Syringyl

S2: Couche S2 de la paroi secondaire

SAD : Sinapyl alcohol dehydrogenase

SDS : dodécylsulfate de sodium

SHD : SHEPHERD

Sig : Signaling

SP : Structural proteins

SuSy : Sucrose synthase

TBO : Toluidine-Blue O

TCA : Thiocarbonic acid

Trp : Tryptophane

Tyr : Tyrosine

UDP : Uridine diphosphate

UGT : UDP-glycosyl-transférase

UN : unknown function

UV : Ultraviolet

VIGS : Virus-induced Gene Silencing

XGA : Xylogalacturonan.

XTH : Xyloglucan endo-transglycosylase/hydrolase

XyG : Xyloglucane

Chapitre 1 : Synthèse bibliographique

La paroi cellulaire végétale

Introduction

Il y a environ 470 millions d'années, a commencé la colonisation du milieu terrestre par les algues multicellulaires, cet événement unique a donné naissance à toute la diversité du monde végétale terrestre (Becker and Marin, 2009; Wodniok et *al.*, 2011), passant par les bryophytes et les plantes vasculaires ; ces dernières étant le groupe le plus abondant sur la surface terrestre (Sørensen et *al.*, 2010). Le passage du milieu aquatique au milieu terrestre a engendré de nombreuses contraintes non présentes dans le premier environnement. Ainsi les modifications de la paroi et la mise en place des tissus vasculaires ont été des étapes pivots qui ont permis aux plantes vasculaires de conquérir les habitats terrestres, et de former ainsi le sous-règne des embryophytes (constitué des Bryophytes, Lycopodiophytes, Pteridophytes, Gymnospermes, Poales et angiospermes non Poales) (Figure 1) (Kenrick and Crane, 1997; McCourt et *al.*, 2004; Becker and Marin, 2009; Niklas and Kutschera, 2010)

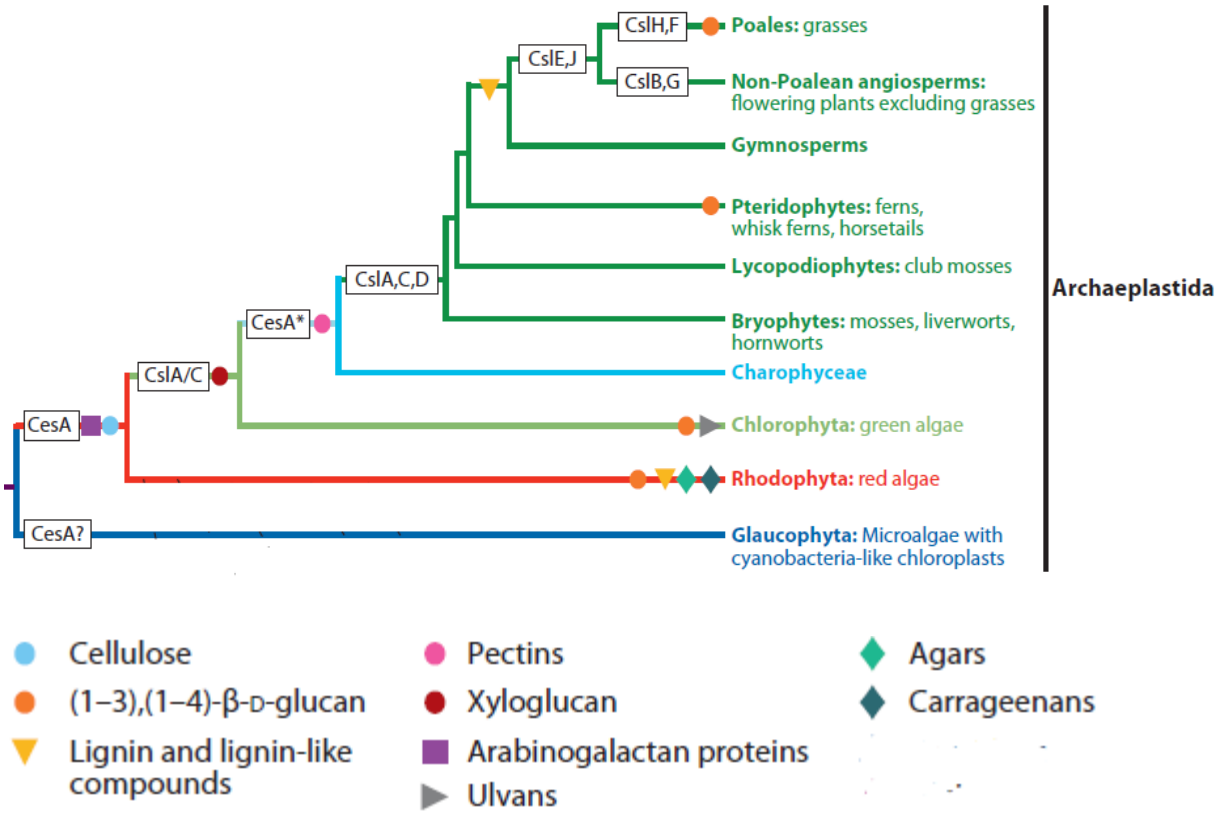


Figure 1 : Arbre évolutif représentant l'apparition des différents composants de la paroi cellulaire chez les Archaeplastida. Adapté de (Popper et *al.*, 2011).

Les cellules végétales sont caractérisées par une paroi cellulaire, qui fut longtemps considérée comme statique, cependant de nombreuses études ont révélé qu'elle est d'un grand dynamisme puisqu'elle est capable d'évoluer au cours de la croissance et de la différenciation des cellules, en fonction des besoins de celles-ci (Popper, 2008). La paroi cellulaire assure de nombreuses fonctions vitales telle que la détermination de la forme des innombrables types cellulaires existants, qui donnent naissance aux différents tissus qui à leur tour forment les organes de la plante.o !m La paroi cellulaire sert d'interface entre cellules adjacentes jouant ainsi un rôle crucial dans la communication intercellulaire (Pilling and Höfte, 2003), mais aussi dans les interactions plantes-microorganismes qu'elles soient des interactions symbiotiques (Rich et *al.*, 2014)ou défensives (Vorwerk et *al.*, 2004). Ainsi, la paroi, de par sa participation active à la physiologie de la cellule, peut être considérée comme un organite à part entière.

La paroi cellulaire végétale est d'une grande complexité structurale. Elle est majoritairement constituée de polymères polysaccharidiques tels que la cellulose, les hémicelluloses ou les pectines, et de glycoprotéines et d'enzymes, mais aussi dans certaines paroi de polymère phénolique lignine (Lewis and Yamamoto, 1990; Burton et *al.*, 2010). Ces différents polymères sont bâtis et assemblés grâce à des liaisons covalentes (ex : liaison glycosidique) et non covalentes (ex : liaison d'hydrogène), en une structure hétérogène, représentant ainsi la biomasse la plus abondante sur terre. En dépit de la grande diversité des structures pariétales retrouvées au sein du monde végétal, la paroi s'organise généralement sous forme de deux ou trois couches distinctes : une lamelle moyenne, une paroi primaire et, si présente, une paroi secondaire. Toutes les cellules végétales sont généralement caractérisées par la présence d'une lamelle moyenne et d'une paroi primaire. Certains types cellulaires (sclérenchyme, xylème) possèdent en plus une paroi secondaire lignifiée.

Structure et organisation de la paroi végétale

La lamelle moyenne est la première couche pariétale synthétisée, elle se met en place lors de la division cellulaire, et est commune à deux cellules adjacentes. Sa composition diffère de celle des parois primaire et secondaire, en effet elle est riche en pectine et de ce fait constitue le ciment assurant la cohésion inter-cellulaire. Cependant cette structure peut devenir lignifiée au moment de la formation de la paroi secondaire (Figure 2) (Gierlinger and Schwanninger, 2006).

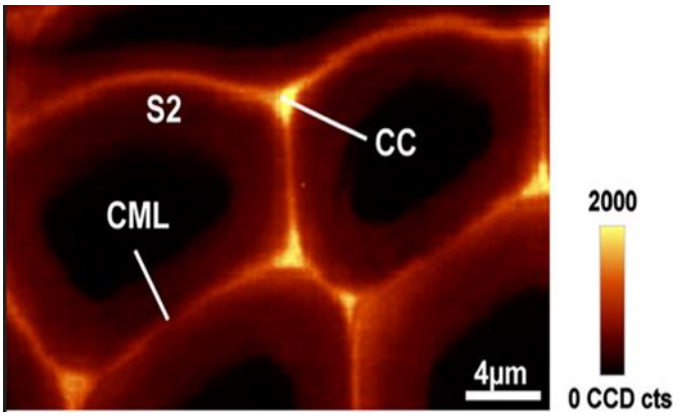


Figure 2 : Coupe transversale du bois du peuplier. Image montrant la présence de lignine dans les différentes couches de la paroi cellulaire en utilisant l'imagerie Raman. CC: Cell corner; S2: Couche S2 de la paroi secondaire; CML: Compound middle lamella (lamelle moyenne + paroi cellulaire primaire). (Gierlinger and Schwanninger, 2006).

Après la lamelle moyenne, en se rapprochant du centre de la cellule, se situe la paroi primaire constituée de micro-fibrilles de cellulose moins cristallines que celles des parois secondaires. Cette cellulose est enchâssée dans une matrice polysaccharidique amorphe riche en pectines, hémicelluloses et glycoprotéines. Elle est flexible permettant ainsi la croissance et l'expansion cellulaire.

Deux types de parois primaires sont distingués : les parois de type I (Carpita et Gibeaut 1993), retrouvées principalement chez les dicotylédones, mais également chez certaines monocotylédones (ex : famille Liliaceae), contiennent des hémicelluloses de type xyloglucanes, tandis que les parois de type II, caractérisant les monocotylédones (ex. Poaceae), contiennent une quantité plus importante d'hétéroxylanes (ex. les glucuronoarabinoxylanes). Une fois la croissance cellulaire terminée au niveau de la taille et de la forme, la synthèse de la paroi secondaire est enclenchée au cours de la différentiation cellulaire comme c'est le cas des cellules xylèmiennes.

Cette paroi secondaire se met en place dans les tissus de soutien, tels que le sclérenchyme mais également dans les tissus conducteurs comme le xylème. Comme la paroi primaire, elle est constituée de cellulose, d'hémicelluloses et dans la majorité des cas, de lignines. La paroi secondaire est elle-même composée de trois couches différenciées par l'orientation des micro-fibrilles de cellulose (Figure 3A). Cette orientation joue un rôle primordial dans la définition des propriétés physiques de ces parois (Keckes et *al.*, 2003). Les 3 couches sont nommées de l'extérieur vers l'intérieur S1, S2 et S3. La couche S2 est la plus épaisse des trois couches et ses

micro-fibrilles de cellulose sont disposées suivant un schéma hélicoïdal presque parallèle à l'axe de la cellule (Figure 3B) (Barnett and Bonham, 2004).

Cependant, cette composition de base de la paroi cellulaire végétale peut être modifiée en cas de stress abiotique ou biotique, selon les espèces, induisant ainsi des modifications dans les propriétés physicochimiques et mécaniques de la paroi.

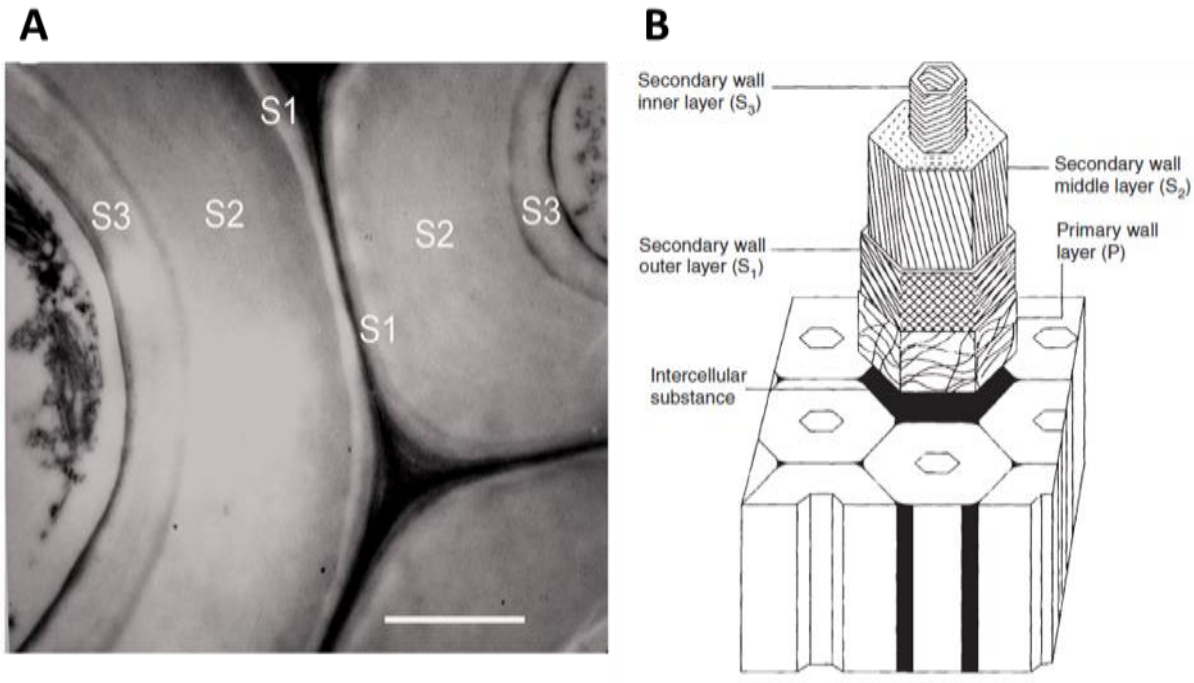


Figure 3 : Les couches composant la paroi cellulaire végétale. (A) image en microscopie électronique à transmission montrant les 3 couches S1, S2 et S3 retrouvées au sein de la paroi secondaire des fibres interfasciculaires de la tige d'*Arabidopsis thaliana* (Zhong and Ye, 2014). (B) schéma représentatif de l'orientation des microfibrilles de cellulose dans la paroi primaire et les trois couches de la paroi secondaire (Doblin et *al.*, 2010).

Composition chimique des parois cellulaires végétales

Les polysaccharides

Les polysaccharides représentent les principaux composés des parois des cellules végétales. Ils se présentent sous forme de polymères glycosidiques homogènes ou hétérogènes, linéaires ou ramifiés. En fonction de la composition osidique et des liaisons qui relient les différentes unités structurales, trois polysaccharides différents sont majoritairement distingués : la cellulose, les hémicelluloses, et les pectines (Carpita, 1996; Lerouxel et *al.*, 2006; Caffall and Mohnen, 2009).

La cellulose

La cellulose est le premier carbohydrate constituant les parois des cellules végétales, commun à toutes les plantes et le plus abondant polymère sur la surface terrestre. Celle-ci forme le principal réseau de support de charge des parois qui contribuent à la protection des cellules ainsi qu'à leur croissance. La cellulose est un homopolymère simple non-ramifié constitué de résidus D-glucose liés en β-1,4, de conformation en chaise. Les résidus glycopyranosidiques qui se succèdent pour former la chaîne linéaire sont orientés de 180°C l'un par rapport à l'autre, définissant ainsi un motif répété, appelé cellobiose (Figure 4) (Taylor, 2008). Cette architecture spécifique confère une polarisation aux polymères saccharidiques, ainsi on distingue une extrémité « réductrice », due à la présence d'une fonction hémiacétylique, et une extrémité « non réductrice ». Les polymères cellulosiques sont agencés de façon parallèle et liés par des ponts hydrogènes formant ainsi des microfibrilles para-cristallines insolubles et bien structurées, renfermant un nombre toujours débattu entre 18 et 36 chaînes linéaires avec un diamètre de 2 à 5nm (Somerville, 2006, Thomas et *al.* 2013). Ces microfibrilles peuvent atteindre une longueur de plusieurs micromètres et confèrent aux parois cellulaires une très grande résistance aux pressions de turgescence que la cellule peut subir aux cours de différents processus physiologiques.

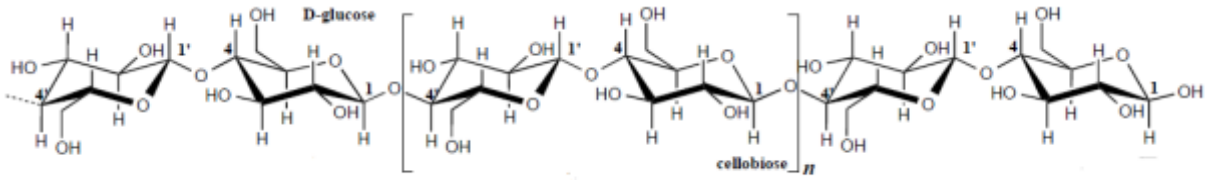


Figure 4 : Illustration de la cellulose linéaire et du motif répété la constituant appelé cellobiose.D'après (Taylor, 2008).

Les hémicelluloses

Hémicellulose est un terme qui regroupe différentes classes de polysaccharides hémicellulosiques qui varient considérablement entre les espèces et les tissus (Tableau 1). Elles regroupent les xyloglucanes, les xylanes, les hétéromannanes et les β-glucanes. Les hémicelluloses sont liées aux microfibrilles de cellulose par des liaisons hydrogènes.

Tableau 1 : Les hémicelluloses chez les plantes et leurs répartitions tissulaires (% w/m). (Scheller and Ulvskov, 2010)

Polysaccharide	Dicot walls		Grass walls		Conifer walls	
	Primary	Secondary	Primary	Secondary	Primary	Secondary
Xyloglucan	20–25	Minor	2–5	Minor	10	–
Glucuronoxylan	–	20–30	–	–	–	–
Glucuronoarabinoxylan	5	–	20–40	40–50	2	5–15
(Gluco)mannan	3–5	2–5	2	0–5	–	–
Galactoglucomannan	–	0–3	–	–	+	10–30
β-(1→3,1→4)-glucan	Absent	Absent	2–15	Minor	Absent	Absent

Les xyloglucanes

Les xyloglucanes sont les hémicelluloses les plus abondantes chez les dicotyledons angiospermes et les monocotylédones non-commelinoides où ils représentent environ 30% des hémicelluloses. Chez les gymnospermes le pourcentage tombe à 20 % et seulement 2% chez les Poaceae.

Les xyloglucanes comme leur nom l'indique sont composés de glucose et de xylose. Les résidus de glucose sont liés entre eux par des liaisons β (1-4) formant de la sorte un squelette semblable à celui de la cellulose. Généralement, sur ce squelette, des α-D-xyloses viennent se greffer en positon O6 sur trois résidus de glucose sur quatre, et couvriront ainsi environ 75% des résidus glucose. D'autres oses tels que le galactose, le fructose et l'arabinose, peuvent être retrouvés au sein d'une molécule de xyloglucane. En effet, en fonction de ces oses, deux configurations de xyloglucane peuvent être retrouvées : les fucogalactoxyloglucanes et les arabinogalactoxyloglucanes. Dans le premier, du galactose et du fructose liés entre eux par une liaison α (1-2) se greffent au résidu xylose en position β (1-2) (Figure 5). Cette configuration est principalement retrouvée dans la paroi des dicotylédones.

Pour la deuxième configuration (arabinogalactoxyloglucanes) les substitutions sur le squelette de glucose sont soit des résidus xylose seuls, soit un résidu xylose lié en β (1-2) avec un galactose, soit lié en α (1-2) avec un arabinose.

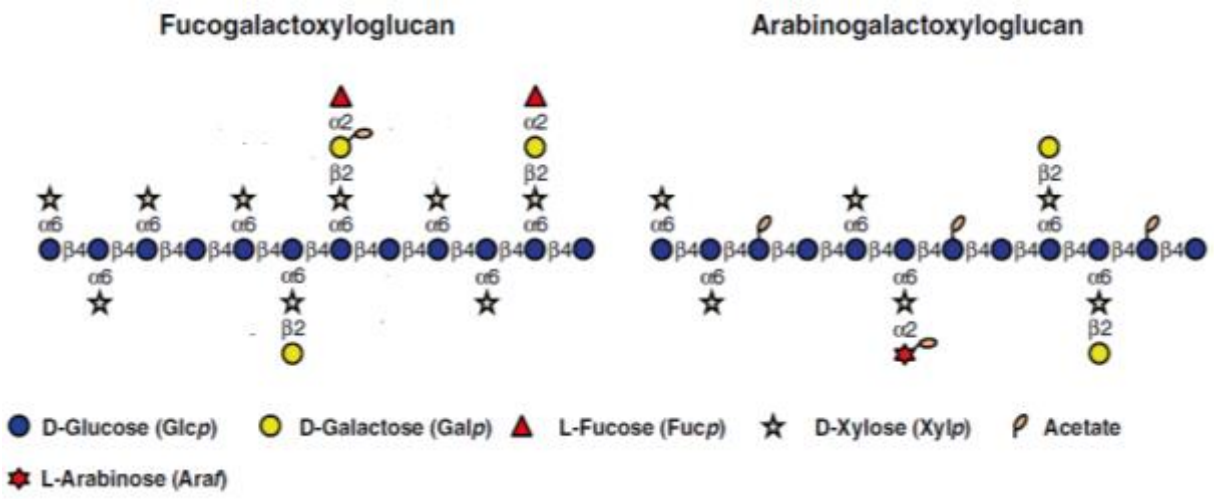


Figure 5 : Représentation schématique des deux formes de xyloglucanes les plus abondantes chez les Eudicots (fucogalactoxyloglucane) et les monocots (arabinogalactoxyloglucane). Adaptée de (Pauly et *al.*, 2013).

Les xylanes

Constituant majeur des parois secondaires des angiospermes, les hémicelluloses de type xylane sont constituées d'une chaine linéaire de résidus D-xylose liées en β (1-4) (Figure 6). Les xyloses sont substituées différemment suivant les espèces et les tissus. En effet, chez les dicotylédones les xyloses sont substituées pas des acides glucuroniques liés en α (1-2), par des O-4-methyl-acides glucuroniques, et peuvent être acétylés (Gille and Pauly, 2012). De plus, même si cela n'a pas été rapporté chez *Arabidopsis thaliana*, ni chez le peuplier, il semblerait que des arabinoses peuvent être greffés au squelette xylosidique chez d'autres espèces.

Les arabinoses, rarement présents chez les dicotylédones, sont beaucoup plus abondants comme substituant sur les xyloses des glucoronoarabinoxylanes caractéristiques des monocotylédones. Outre l'abondance des arabinoses, deux types de liaison glucosidique sont retrouvés au sein de ce type de xylane : des α (1-2) ou α (1-3).

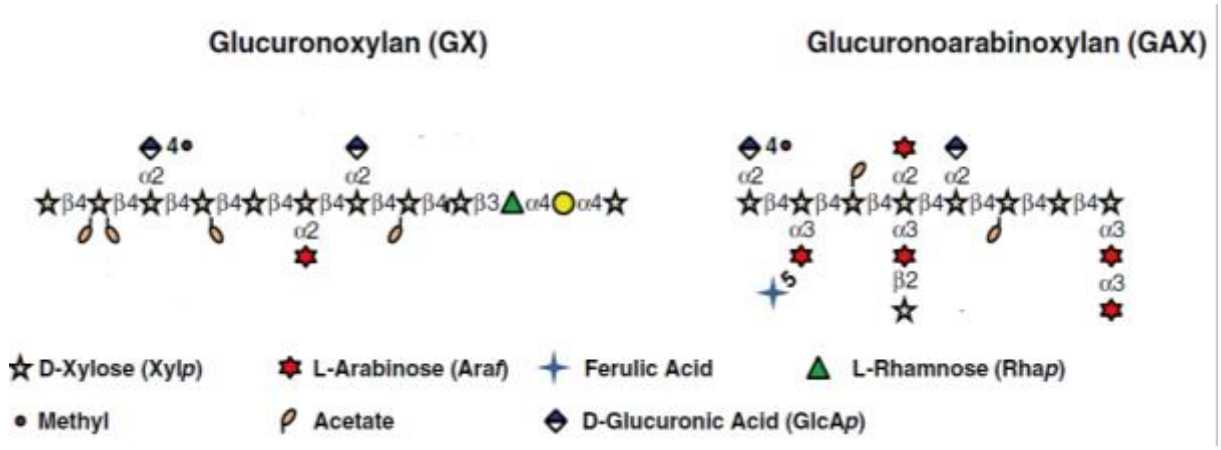


Figure 6 : Représentation schématique des deux formes de xylane(Adaptée de Pauly et *al.*, 2013).

Les hétéromannanes

Les hétéromannanes peuvent être répartis en quatre types : les mannanes, les glucomannanes, les galactomannanes et les galactoglucomannanes (Scheller and Ulvskov, 2010). Les mannanes et glucomannanes sont constitués de squelettes linéaires, composés uniquement de mannose liée en β (1-4) dans le cas des mannanes, et d'une succession de résidus de mannose et de résidus de glucose avec une fréquence 2:1 dans le cas des glucomannanes. Les résidus mannoses dans ces deux composés peuvent être substitués avec des résidus de galactose en position α (1-6), ce qui donne les galactomannanes et les galactoglucomannanes.

Les galactoglucomannanes sont les composants majeurs de la paroi secondaire des gymnospermes (Pauly et *al.*, 2013).

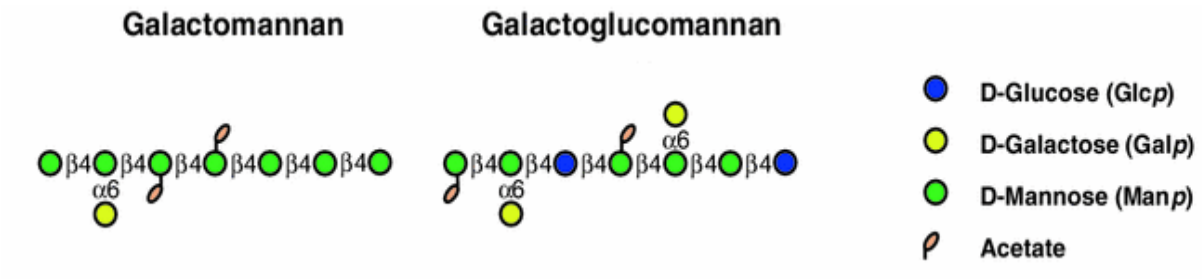


Figure 7 : représentation schématique des deux formes des hétéromannanes. Adaptée de (Pauly et *al.*, 2013).

Les (1-3), (1-4) β glucanes

Les β glucanes sont une succession linéaire et non ramifiée de D-glucose, liés en β (1-4) mais aussi en β (1-3) (Figure 8).

Les β glucanes sont des hémicelluloses, qui caractérisent la paroi de la famille des Poaceae, mais ils sembleraient présents aussi chez les algues vertes (Popper et *al.*, 2011). Ils sont également régulés durant le développement et sont modifiés durant l'élongation cellulaire.

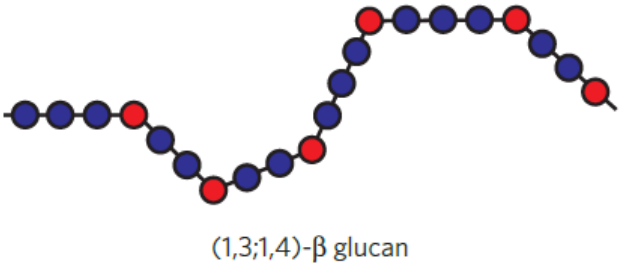


Figure 8 : illustration de la structure des (1-3 ; 1-4) β glucanes. : D-Glucose liaison (1-4); : D-Glucose liaison (1-3). (Adapté de Burton et *al.*, 2010)

Les pectines

Les pectines jouent un rôle important de « ciment » entre cellules adjacentes et les défauts dans leur biosynthèse sont souvent associés aux défauts d'intégrité tissulaire. À titre d'exemple, le mutant haploïde de *Nicotiana plumbaginifolia* présentant une déficience dans une glycosyltranférase nécessaire à la synthèse des pectines est caractérisé par une faible cohésion cellulaire (Iwai et *al.*, 2002).

Les pectines possèdent une composition chimique plus complexe, qui se remanie durant la croissance ou en réponse aux stresses (Willats et *al.*, 2001). Elles sont riches en acide galacturonique et forment des chaînes polysaccharidiques acides de haute masse moléculaire. Elles s'organisent généralement sous forme de trois domaines, le plus simple est linéaire - l'homogalacturonane. Les deux autres structures (rhamnogalacturonanes I et II) sont ramifiés avec une composition beaucoup plus complexe (Figure 9).

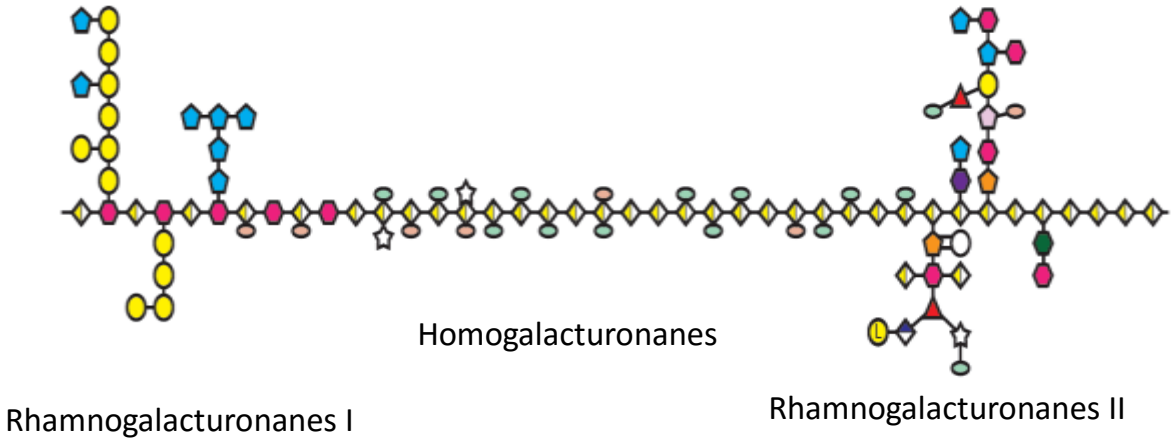


Figure 9 : Représentation des trois domaines constituant les pectines. (Adapté de Burton et *al.*, 2010).Il est a noté que différents modèles coexistent (Schols et Voragen 1996 ; Vincken 2003).

L'homogalacturonane

L'homogalacturonane (HG) est constitué d'une succession linéaire de résidus d'acide galacturonique, liés en position α (1-4). Cette structure est la plus abondante dans la paroi primaire et dans la lamelle moyenne des dicotylédones constituant environ 65% de la totalité des pectines retrouvées dans la paroi. Les résidus d'acide galacturonique peuvent être methylesterifié en C6, ainsi que O-acetylé sur le O2 ou O3. Ces modifications peuvent concerner 80% pour la methylesterification et plus de 10% pour l'acetylation (Wolf et *al.*, 2009; Gou et *al.*, 2012). D'autres modifications ont été rapportées telles que la substitution de l'acide galacturonique par un β-D-xylose en position O3 pour former ainsi le xylogalacturonan (XGA).

Les rhamnogalacturonanes I

Le rhamnogalacturonane I (RGI) est le second domaine des pectines et représente entre 20 à 30% des pectines de la paroi. Il est retrouvé principalement dans la paroi primaire et la lamelle moyenne (Goubet et *al.*, 1995; Ridley et *al.*, 2001), mais aussi dans la paroi secondaire de quelque types cellulaires telles que les fibres cellulosiques chez le lin (Gorshkova and Morvan, 2006; Gurjanov et *al.*, 2008). La chaîne principale de ce polysaccharide est composée d'une succession de résidus de L-rhamnose et de l'acide galacturonique, ce qui forme une chaîne linéaire de 2-α-L rhamnopyranose-(1,4)-α-D où viennent se greffer des chaînes latérales de sucres neutres telles que les arabinanes et les arabinogalactanes de type I (pectines) et de type II (AGP) (AGI et II) (Voragen et *al.*, 2009). Les arabinanes sont constitués d'un squelette de α (1,5) L-arabinofuranose qui sont substitués parfois en O2 et/ou O3 par d'autres résidus α-L-

arabinofuranose. Les AGI sont composés de chaînes de D-Gal liées en β (1-4), sur lesquelles s'attachent, en position O3, des chaînes courtes de L-arabinofuranose liées en α (1-5). Tandis que les AG de type II sont constitués d'une chaîne principale de galactane liée en β (1,3) et de chaînes latérales composées de mono- ou d'oligo-β(1,6) galactane qui portent à leur tour différents sucres tels que L-arabinofuranose, L-arabinopyranose, acide D-glucuronique, 4,O-méthyl-D-glucuronique, L-rhamnose et le L-fucose (Willats and Knox, 1996). Les groupes galacturonyl du RGI, quant à eux, peuvent être acétylés sur le O2 ou le O3.

L'abondance et la structure des différentes ramifications retrouvées dans le RGI varient entre les plantes (Vincken, 2003; Harholt et *al.*, 2010) et impactent les propriétés physicochimiques des pectines.

Les rhamnogalacturonanes II

Le rhamnogalacturonane II (RGII) (Figure 10) possède une structure hautement conservée (Matsunaga et *al.*, 2004), constituée de 12 glycosyles liés entres eux par une vingtaine de liaisons glycosidiques différentes. Sa biosynthèse nécessite l'implication de pas moins de 22 glycosyltransférases spécifiques (Bar-Peled et *al.*, 2012), ce qui génère une structure polysaccharidique des plus complexe avec une masse moléculaire estimée entre 5-10 kDa (Whitcombe et *al.*, 1995; Vidal et *al.*, 2000). Au niveau architectural les RGII se caractérisent par une chaîne linéaire courte de (Gal*p*A) (continuité du HG) sur laquelle viennent se greffer 5 chaînes latérales nommées de A à E (Whitcombe et *al.*, 1995; Pérez et *al.*, 2003; Buffetto et *al.*, 2014). Les chaînes latérales de A à D sont les plus complexes et elles sont les plus connues dans la littérature. La présence d'une 5ème chaine nommée E qui n'est composée que d'un seul résidu α-L-Ara*f* liée en positon 3 au Gal*p*A de la chaîne principale, n'est pas toujours admise, étant donné que ce résidu n'est pas souvent détecté faute de méthodes d'extraction de ce polysaccharide.

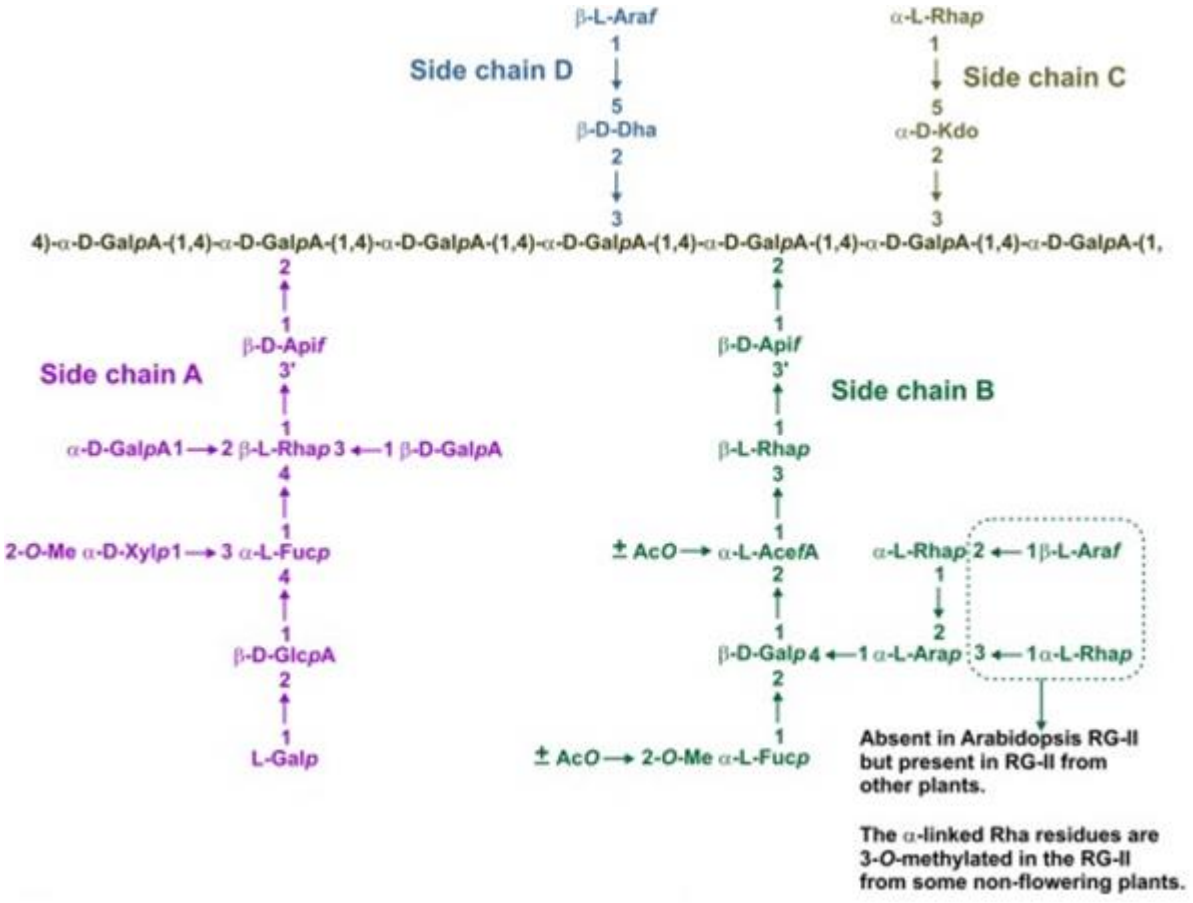


Figure 10 : Structure d'un RGII détaillant la structure des différentes chaines.(D'après Bar-Peled et *al.*, 2012).

Les RGII représentent environ 10% des pectines retrouvées dans la paroi des plantes supérieures (O'Neill et *al.*, 2004), dont 90% sont sous forme dimérique (Ridley et *al.*, 2001). Ces dimères se forment grâce aux ions de borate (BO_2^-) qui relient deux unités du RGII au niveau des Araf*s* des chaînes latérales A via des liaisons ester-borate (Figure 11). La forme dimérique est d'une importance capitale pour l'intégrité de la paroi, sa mécanique et dans la germination du pollen(Dumont et *al.*, 2014).

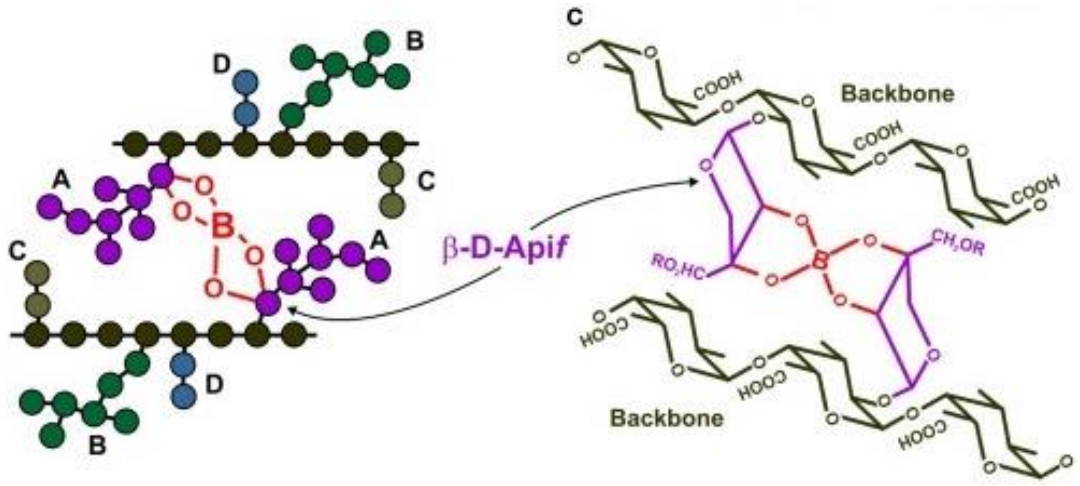


Figure 11: liaisons ester-borate mettant en liaison deux RGII. D'après (Bar-Peled et *al.*, 2012). Le code couleur des figure A, B, C et D se référé à la figure 10.

Les protéines

Les protéines sont les macromolécules le moins représentées dans la paroi cellulaire avec une teneur entre 5 et 10 % (Cassab and Varner, 1988). Elles peuvent être structurales ou bien des enzymes agissant au sein même de cette paroi, en étant ainsi les acteurs qui orchestrent toutes les modifications pariétales nécessaires à la croissance et au développement de la plante y compris en réponse aux stress (a)biotiques (Fry, 2004; Passardi et *al.*, 2004). Différentes manières sont proposées pour classer ces protéines : suivant leurs caractéristiques physicochimiques, fonctionnelles ou structurales. La première se base sur les interactions que ces protéines peuvent avoir avec leur environnement : protéines fortement liées à la paroi, protéines liées à la paroi ou protéines faiblement liées à la paroi. La seconde classification reste incomplète vu qu'elle ne couvre actuellement que les protéines extractibles (celles qui sont labiles ou faiblement liées à la paroi). Cette classification fonctionnelle regroupe les protéines pariétales en 9 classes qui sont des protéines ayant une action envers les polysaccharides (glycosylhydrolases, estérases, expansines, lyases), des oxydoréductases, des protéines de signalisation, des protéases, des lectines, des protéines contenant des carbohydrate binding modules (CBMs) ou des protéines de fonction inconnue (Bosch and Hepler, 2005; Sampedro and Cosgrove, 2005; Jamet et *al.*, 2006; Minic et *al.*, 2007; Obembe et *al.*, 2007). Enfin, il y a la troisième classification basée sur la structure de ces protéines pariétales et qui est décrite ci-dessous.

Les protéines structurales de la paroi

Les protéines structurales de la paroi sont divisées en 2 groupes majeurs les HRGPs et les GRPs en fonction de leur structure chimique (composition en acides aminés et profils de glycosylation).

Les Hydroxyproline-rich glycoproteins (HRGPs)

Les (Hyp)-riche O-glycoprotéines sont des macromolécules complexes présentant une importante diversité structurelle et fonctionnelle. Suivant le degré de O-glycosylation, elles sont divisées en 3 classes : les HRGPs modérément glycosylés représentés par les extensines (EXTs) ; les arabinogalactan proteines (AGPs) hyperglycosylés et les hyp-/pro-rich proteine(H/PRP) (Kieliszewski, 2001; Tan et *al.*, 2004; Estévez et *al.*, 2006).

Les extensines (EXTs)

Les extensines sont des glycoprotéines, généralement décrites comme des protéines capables de former des liaisons covalentes entre elles (Cannon et *al.*, 2008; Velasquez et *al.*, 2012). Elles subissent des O-glycosylations post-transcriptionnelles sur le motif Ser-(Hyp). Ce motif Ser-(Hyp), peut être en nombre de 4, 5 ou 6 (Smith et *al.*, 1984; Qi et *al.*, 1995; Campargue et *al.*, 1998). Des chaînes linéaires de résidus arabinoses, viennent se greffer sur la partie (Hyp) du motif (Ogawa-Ohnishi et *al.*, 2013), tandis qu'une peptidyl-Ser galactosyltransferase (SGT1) est responsable de l'ajout d'un seul résidu α Gal sur le résidu Ser du motif Ser-(Hyp) (Saito et *al.*, 2014) (Figure 12).

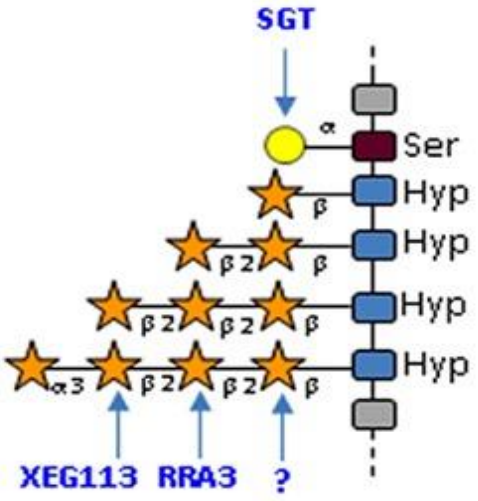


Figure 12 : représentation schématique de la structure des extensines. Ser : Sérine, Hyp : Hydroxyproline ; SGT : peptidyl-Ser-galactosyltranferase ; RRA3 : Reduced Residual Arabinose 3 ; XEG113 : xyloglucanase 113; : galactose; : arabonose. Adapté de (Nguema-Ona et *al.*, 2014).

Chez *Arabidopsis thaliana*, les EXTs sont représentées par 59 gènes, certaines étant des EXT classiques, tandis que d'autres sont des « EXT-like ». Malgré le nombre important de protéines pariétales possédant une domaine EXT (Lamport et *al.*, 2011), peu de choses sont connues au

sujet de leurs fonctions précises. Cependant l'importance de ces protéines est parfaitement illustrée par les problèmes d'assemblage de la paroi, observés chez le mutant létal *d'Arabidopsis thaliana AtEXT3* (codant une extensine) (Hall and Cannon, 2002; Cannon et *al.*, 2008). De même des défauts au niveau de la paroi ont été également constatés chez les mutants *ext6*, *ext7*, *ext10*, *ext11* et *ext12* d'Arabaidopsis (Velasquez et *al.*, 2011).

Les arabinogalactanes protéines (AGP)

Les arabinogalactane proteins (AGPs) sont les protéines les plus glycosylées de la famille des HRGPs. Ces protéines qui sont spécifiques aux plantes, sont caractérisées par une partie saccharidique représentant plus de 90% de la masse totale et très probablement associée aux rôles biologiques des AGPs *in planta* (Seifert and Roberts, 2007; Kitazawa et *al.*, 2013). Elles contiennent en général un motif dipeptidique répété de Ala-hyp, Ser-hyp, Thr-hyp, et Val-hyp. Un nombre d'AGPs sont attachés à la membrane plasmique par des ancres GPI (glycosylphosphatidylinositol) qui peuvent être clivées par des phospholipases (Schultz, 2004).

La caractérisation des O-glycanes covalemment liée aux deux motifs peptidiques [Ala(Hyp)]n, $[Ser(Hyp)]_n$ a été réalisée par Tan et *al.* en 2010 permettant de définir précisément la structure d'un AG de type II (Figure 13 B). Celui-ci diffère de l'AG de type I connu comme étant une des chaines latérales de RGI (Voragen et *al.*, 2009), mais aussi de l'AG de type III retrouvé chez les allergènes Artv1 et Amba4 décrit respectivement chez *Artemisia vulgaris*(Leonard et *al.*, 2005) et *Ambrosia artemisiifolia*(Léonard et *al.*, 2010) (Figure 13).

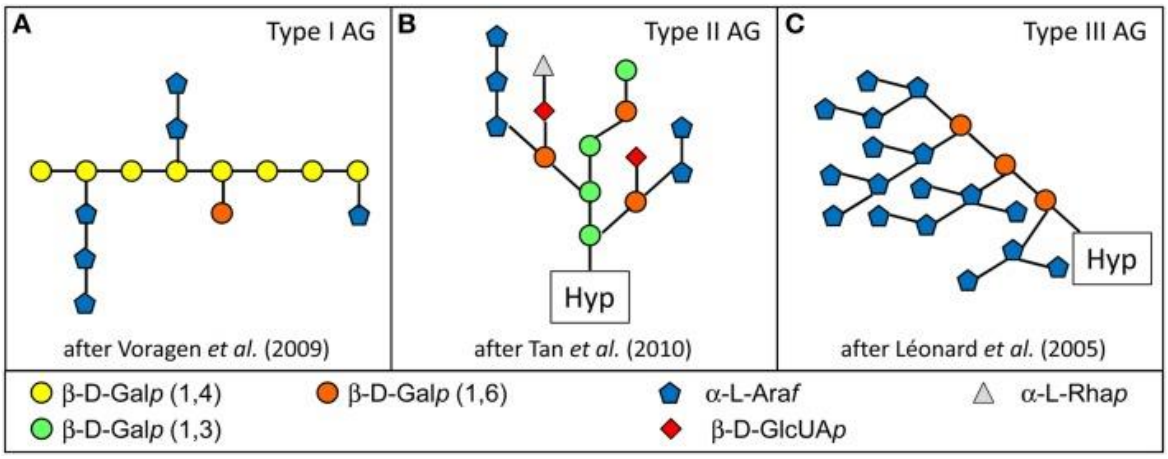


Figure 13 : Structure des trois types d'arabinogalactanes. D'après (Hijazi et *al.*, 2014).

De nombreuses études suggèrent que les AGPs sont covalemment liées aux pectines et plus spécifiquement au RGI chez différentes espèces et dans différents tissues (Keegstra et *al.*, 1973; Jauh and Lord, 1996; Immerzeel et *al.*, 2006). Un lien entre ces AGP et les hémicelluloses a été

mis en évidence par les travaux de Tan et collaborateurs réalisés chez Arabidopsis où ils ont démontrés que deux isoformes d'un AGP (At3g45230) sont covalement liées non seulement aux pectines, mais aussi aux hémicelluloses (Tan et *al.*, 2010; Tan et *al.*, 2012)

Les AGPs sont impliquées dans de nombreux processus physiologiques tels que l'expansion cellulaire, le développement, la reproduction et la différenciation du xylème (Motose et *al.*, 2004), l'embryogénèse somatique (van Hengel and Roberts, 2003), la signalisation et la défense (Seifert and Roberts, 2007) et les interactions plante micro-organismes (Nguema-Ona et *al.*, 2013)

Les protéines riches en proline (H/PRPs)

Les PRPs sont riches en proline et peuvent être faiblement ou fortement-glycosylées (Hijazi et *al.*, 2014). Ces protéines sont peu documentées, cependant elles auraient un rôle dans le développement et la défense contre les stresses biotique et abiotique (Bernhardt and Tierney, 2000; Battaglia et *al.*, 2006), ainsi que dans l'assemblage et le maintien de la paroi (Gothandam et *al.*, 2009; Xu et *al.*, 2013).

Les protéines riches en glycine (GRP)

Les glycine-rich proteins (GRPs) constituent le deuxième groupe majeur de protéines pariétales. (Keller et *al.*, 1989; Sachetto-Martins et *al.*, 2000). Elles sont retrouvées chez de nombreux organismes (Ringli et *al.*, 2001) et forment un groupe de protéines très hétérogènes mais qui se caractérisent toutes par leur teneur élevée en glycine estimé à 40 à 70% (Sachetto-Martins et *al.*, 2000; Mousavi and Hotta, 2005). Les résidus de glycine ne sont pas glycosylées (Showalter, 1993; Mangeon et *al.*, 2010). Les GRPs sont divisées en cinq classes en fonction de l'organisation des motifs glycine riche, ainsi que de la présence de domaines conservés (Figure 14).

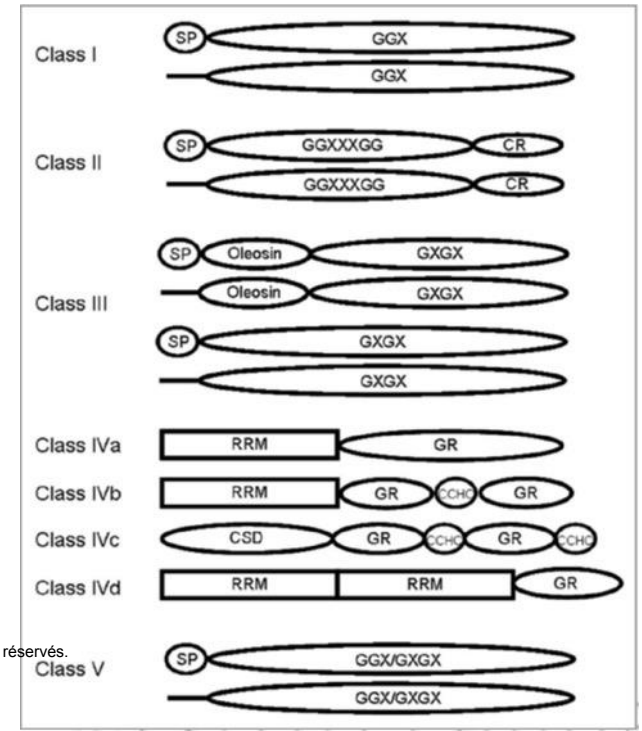


Figure 14 : représentation schématique des 5 classes des GRPs chez les plantes. SP : peptide signale, CR : domaine riche en cystéine Oleosin : domaine Oleosin-conservé ; RRM : RNA-recognition motif ; GR, Glycine-riche domaine ; CCHC, zinc-finger; CSD, Cold-shock domaine. Les motifs

répétitifs sont Glycine-rich et ils sont indiqués par : GGX, GGXXXGG, GXGX et GGX/GXGX, où G représente la glycine et X n'importe quel autre acide aminé. D'après (Mangeon et *al.*, 2010).

Les enzymes de modification de la paroi cellulaire

Le deuxième groupe majeur de protéines pariétales est composé des enzymes impliquées soit dans les étapes finales de la biosynthèse (ex. laccases, peroxidases pour la lignine), soit dans le remodelage des différents polymères polysaccharidiques/protéines lors de la différenciation de la cellule, ou en réponse aux stress (a)biotiques. En effet, la nature complexe de la paroi cellulaire et ses nombreuses fonctions biologiques, lui imposent la nécessité de mobiliser un nombre considérable d'enzymes. Suivant ses besoins physiologiques, les enzymes qui interviennent sont différentes. Les enzymes assurent la structuration, restructuration ou déstructuration de la paroi cellulaire. L'expansion cellulaire, par exemple, fera appel aux enzymes capables d'hydrolyser les composants de la paroi (Cosgrove, 1999; Minic, 2008). Tandis que l'assemblage et/ ou le renforcement de la paroi utiliseront des transglycanases, (Thompson and Fry, 2001). Ces enzymes sont recensées dans la base de donnés CAZy pour Carbohydrate Active enZymes, cette classification est subdivisée en 6 classes : glycoside hydrolase (GH), glycosyltransférase (GT), polysaccharide lyases (PL), les carbohydrate esterase (CE), les Carbohydrate binding module (CBMs), et la classe des ActivitésAuxiliaires (AA) qui est une nouvelle classe, dans laquelle sont regroupés par exemple les peroxydases et les laccases. Selon la spécificité de leur substrat, chacune des 6 classes est subdivisée en sous famille, et certaines d'entre elles sont liées à la paroi cellulaire (Lombard et *al.*, 2014). Le tableau 2 regroupe un certain nombre d'enzymes qui sont décrites dans la littérature comme agissant sur la paroi cellulaire.

Tableau 2 : Les enzymes des familles cazy impliqués dans le métabolisme pariét*al.* Adapté de Pauly et *al.*, 2013.

Family	Catalytic domain	Enzyme activity	Reference on transglycosylation
GH1	$(\beta/\alpha)_8$	β-Glucosidase	Crombie *et al.* (1998); Opassiri *et al.* (2003)
		β-Mannosidase	
GH2	$(\beta/\alpha)_8$	β-Mannosidase	–
		Mannosylglycoprotein endo-β-mannosidase	–
		β-Galactosidase	–
GH3		α-Arabinofuranosidase /β-1,4-xylosidase	–
		β-1,4-Xylosidase	–
		β-Glucosidase	–
		β-Glucosidase (preferred	–

		substrates are polysaccharides, thus 'exo-β-glucanase')	
GH5	$(\beta/\alpha)_8$	β-Mannanase, trans-β-mannanase	Hrmova *et al.* (2006); Schröder *et al.* (2006)
GH9	$(\alpha/\alpha)_6$	β-1,4-Glucanase (cellulase)	–
GH10	$(\beta/\alpha)_8$	β-1,4-Xylanase, trans-β-xylanase	Johnston *et al.* (2013)
GH16	β-Jelly roll	Xyloglucan endotransglucosylase	Xu *et al.* (1995); Campbell and Braam, 1999)
		Xyloglucan endohydrolase	De Silva *et al.* (1993); Fanutti *et al.* (1993); Tabuchi *et al.* (2001); Baumann *et al.* (2007); Zhu *et al.* (2012)
GH17	$(\beta/\alpha)_8$	β-1,3-Glucanase	–
		'Lichenase' (MLG-specific β-1,4-glucanase)	–
GH 27	$(\beta/\alpha)_8$	α-Galactosidase	–
GH28	β-helix	α-Galacturonidase ('exo-polygalacturonase')	–
		Galacturonanase ('endo-polygalacturonase', pectinase)	–
GH29		α-1,3-Fucosidase, α-1,4-Fucosidase	–
GH31	$(\beta/\alpha)_8$	α-Glucosidase, α-Xylosidase	Sampedro *et al.* (2010)
		α-Xylosidase	–
GH35	$(\beta/\alpha)_8$	β-Galactosidase	–
GH36	$(\beta/\alpha)_8$	α-Galactosidase	–
GH38	$(\beta/\alpha)_7$	α-Mannosidase	–
GH51	$(\beta/\alpha)_8$	α-Arabinofuranosidase	–
		α-Arabinofuranosidase, β-xylosidase	–
GH81	ND	β-1,3-Glucanase	–
GH95	$(\alpha/\alpha)_6$	α-1,2-Fucosidase	–
CE6	(α/β/α)-Sandwich	Xylan acetylesterase	–
CE8	β-Helix	Pectin methylesterase	–
CE13	(α/β/α)-Sandwich	Pectin acetylesterase	–
PL1	Parallel β-helix	Pectate lyase	–

La lignine

L'apparition de l'hétéropolymère lignine définit un moment clef de la colonisation et de la dominance des végétaux en milieux terrestres. Au niveau évolutif, le passage du milieu aquatique vers le milieu terrestre a nécessité de réelles modifications métaboliques permettant l'adaptation à ce nouvel environnement. En effet le premier ancêtre des végétaux devait faire face à de nombreuses contraintes dans son nouvel environnement, à savoir l'exposition aux rayons UV-B, l'absence de soutien, la sécheresse et enfin la présence des herbivores et des pathogènes. Ainsi ces différents éléments ont convergé vers l'édification de nouveaux métabolismes, y compris celui responsable de la biosynthèse de la lignine, le métabolisme des phénylpropanoïdes (Lowry et *al.*, 1980).

La lignine est un polymère aromatique, fortement ramifié et retrouvé chez toutes les plantes vasculaires. Elle représente environ 30% de la biomasse produite dans la biosphère et la lignine est donc, après la cellulose, le polymère végétal le plus largement répandu sur le globe terrestre (Boerjan et *al.*, 2003). Elle est le polymère le plus difficilement dégradable, une propriété très avantageuse permettant la préservation de l'intégrité de la paroi cellulaire, ainsi que l'acheminement de l'eau et des solutions minérales sur de grandes distances. La lignine joue aussi un rôle central dans la défense contre les pathogènes (Boerjan et *al.*, 2003; Ralph et *al.*, 2004) pas uniquement en tant que barrière physique (Buendgen et *al.*, 1990; Bonello and Blodgett, 2003), mais également dans l'immunité de la plante (Miedes et *al.*, 2014).

La lignine est constituée principalement de trois alcools hydroxycinnamyliques appelés monolignols : l'alcool *p*-coumarylique, l'alcool conifériylique et l'alcool sinapylique qui sont, respectivement non méthoxylé, monométhoxylé et bimethoxylé (Figure 15). Liés entre eux par plusieurs liaisons éther (C-O-C) et/ou des liaisons C-C, ils peuvent également former des liaisons covalentes avec les hémicelluloses via des acides phénoliques contribuant à l'intégrité de la paroi cellulaire. La lignine est formée par polymérisation radicalaire des monolignols formant ainsi les unités *p*-hydroxyphenyl (H) dérivées du précurseur alcool para-coumarylique, les unités guaiacyl (G) dérivées du précurseur alcool conifériylique et les unités syringyl (S) dérivées du précurseur alcool sinapylique.

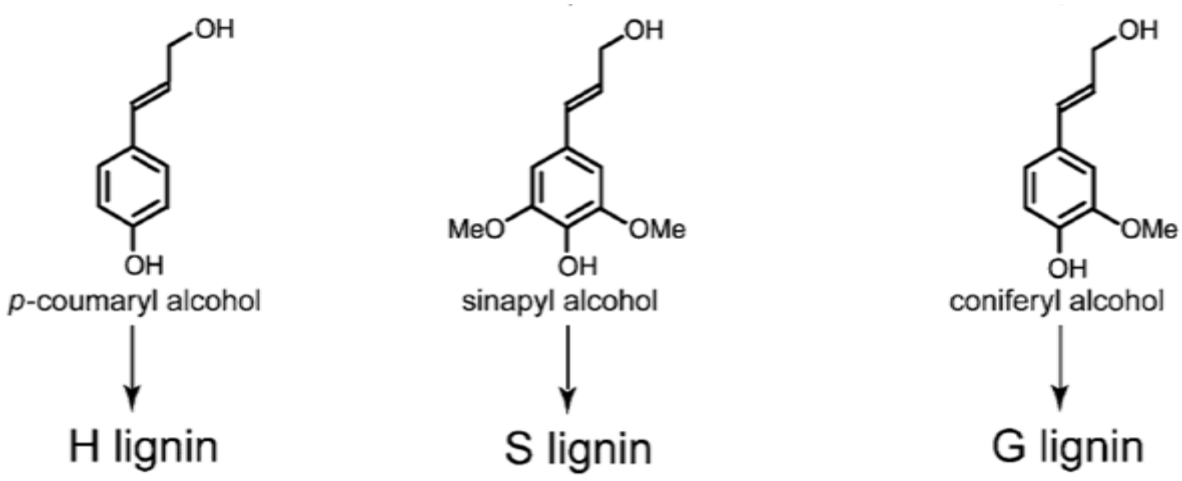


Figure 15 : Structure des trois alcools hydroxycinnamyliques qui forment les unités monomériques H, S et G de la lignine. D'après (Weng et Chapple, 2010).

La biosynthèse des polymères pariétaux

Les différents composants de la paroi cellulaire sont produits via l'action de différentes enzymes de biosynthèse et de polymérisation localisées dans plusieurs compartiments cellulaires

Les cellulose synthases

La synthèse des microfibrilles de cellulose est réalisée via des complexes enzymatiques multimériques de cellulose synthase (CESA) avec une taille de plus de 500kDa, localisés sur la membrane plasmique des cellules. Ces complexes enzymatiques sont visualisés par microscopie électronique et cela, au sein de nombreux organismes telles que : les mousses, les fougères, les algues et les plantes vasculaires (Delmer, 1999; Kimura, 1999; Tsekos, 1999).

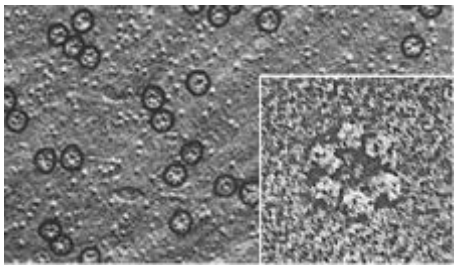


Figure 16 : Observation des rosettes de cellulose synthase par microscopie électronique chez *Lepidium*. D'après Herth 1985.

Les complexes enzymatiques de celluloses synthases agencées en structure hexamèrique forment une rosette avec un diamètre de 20-30 nm, et qui comporte six domaines globulaires (Figure 16) (Kimura, 1999). Chaque domaine globulaire composant la rosette est constitué de six polypeptides CesA ayant une masse moléculaire de 110kDa renfermant l'activité catalytique de l'enzyme. Ces protéines CesA identifiées comme appartenant à la superfamille des glycosyltransférases (GT2), utilisent l'uridine diphospho-glucose (UDP-Glu) comme substrat afin de synthétiser les longues chaînes glucosidiques, et cela, au niveau de la face interne de la membrane plasmique. Trois types de polypeptides CesA, α_1, α_2 et β, sont connus à ce jour, intervenant dans la formation de la rosette, en mettant en œuvre des interactions spécifiques entre les 3 sous unités. Dans ce modèle, il est suggéré que chaque rosette synthétise simultanément 36 (6x6) chaînes polysaccharidiques dans le cytoplasme. Ce modèle, longtemps accepté a été récemment remis en cause par McFarlan et *al.* (2014), qui propose un nouveau modèle avec la production de 18 chaines (Figure 17). Cette nouvelle configuration est basée sur les caractéristiques physicochimiques de la cellulose et ses possibles interactions avec les éléments qui l'entourent (Mcfarlane et Anett, 2014)

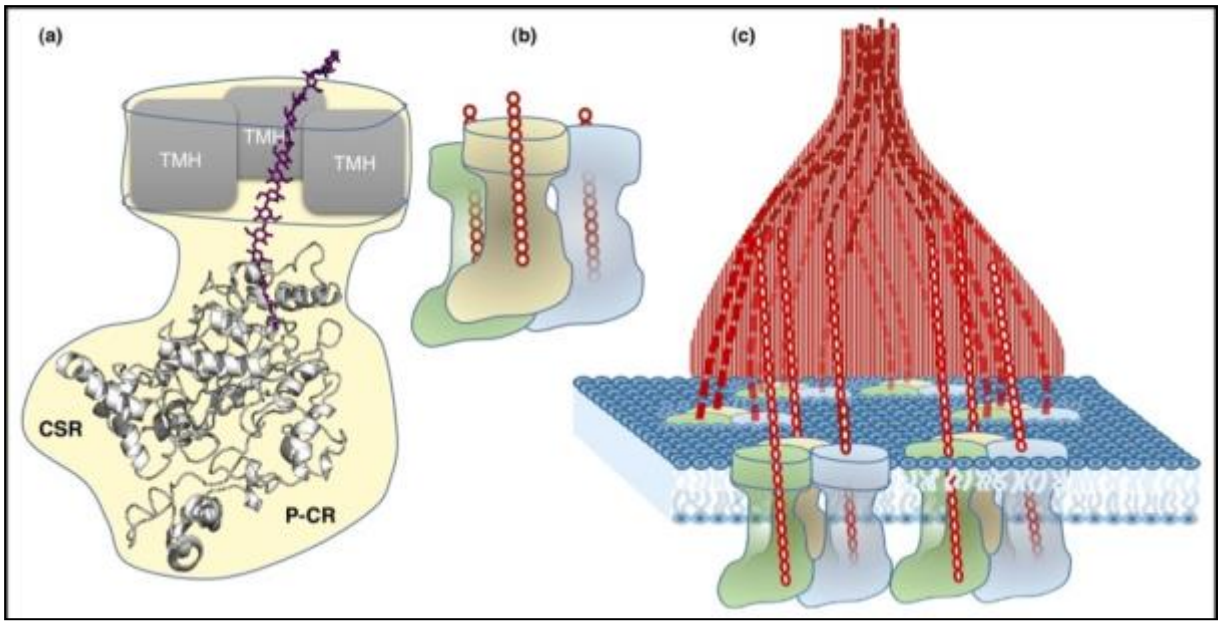


Figure 17 : Miss à jour de la structure des CESAs. a) Représentation de la structure du domaine catalytique d'une CESA. Trois CESAs, codées par trois gènes différents peuvent s'associer pour former un trimère (b), qui à son tour s'associe pour former une rosette hexamèrique. (c) Ainsi,ici sont représentées, 18 chaines de cellulose représentées en rouge sont simultanément synthétisées. D'après (Mcfarlane et Anett, 2014).

L'identification de gènes codant les celluloses synthases chez les plantes a été réalisée pour la première fois par alignement de séquences d'ADNc de coton avec les celluloses synthases bactériennes. Depuis, plusieurs séquences géniques codent des protéines CesA, présentant une structure très conservée, ont été identifiées chez de nombreuses plantes terrestres.

Effectivement, chez Arabidopsison on trouve 10 gènes codant des *CesA*, le maïs et le riz comptent 12 et 9 gènes respectivement (Holland, 2000; Richmond, 2000; Appenzeller et *al.*, 2004). Les différents gènes *CesA* chez les plantes permettent l'assemblage de nombreux complexes catalytiques spécifiques qui interviennent dans la synthèse des celluloses de la paroi primaire ou de la paroi secondaire (Desprez et *al.*, 2007; Taylor, 2008; Carroll et *al.*, 2012).

La synthèse des hémicelluloses et pectines

Contrairement à la cellulose, la synthèse des hémicelluloses et pectines se réalise au niveau de l'appareil de Golgi, puis les différents polysaccharides sont acheminés par transport vésiculaire vers la paroi cellulaire, où ils sont incorporés à la matrice pariétale.

La synthèse du xyloglucane fait intervenir deux types d'enzymes : des UDP-Glc-dépendante glucane synthases telles que les celluloses synthases like (CSL) (Cocuron et *al.*, 2007), et des glycosyl-transférases (GT). Les CSLs sont responsables de la synthèse du squelette, tandis que les GTs interviennent dans les ramifications (Perrin et *al.*, 1999; Keegstra and Raikhel, 2001; Schmid et *al.*, 2005; Zabotina, 2012). En revanche, la synthèse du xylane n'implique qu'un ensemble de glycosyl-transférases (Lao et *al.*, 2003; Brown et *al.*, 2005; Zhong et *al.*, 2005; Brown et *al.*, 2007; Lee et *al.*, 2010; Zeng et *al.*, 2010; Lee et *al.*, 2011; Anders et *al.*, 2012).La synthèse des mannanes, se fait grâce à l'activité de β-mannane synthases (ManS) qui sont des CSLs identifiées pour la première fois dans les graines de *Cyamopsis tetragonoloba* (Dhugga et *al.*, 2004), puis d'*Arabidopsis* (Liepman et *al.*, 2005; Verhertbruggen et *al.*, 2011).

En ce qui concerne les pectines, la synthèse *in vitro* des HGs réalisée par Villemez et collalborateurs en 1965 a fourni des informations concernant leur biosynthèse. L'enzyme responsable de la biosynthèse fait partie de la famille 8 des GTs selon la classification CAZy (Carbohydrate-Active enZymes). Cette famille, grâce à son activité HG : α-1,4-D-galacturonosyltransferase (HG : GalAT) transfère le D-GalA de l'UDP-D-GalA sur l'extrémité non réductrice d'un accepteur HG (Mohnen, 2008; Cantarel et *al.*, 2009). Afin de synthétiser le squelette du RG-I deux activités sont nécessaires : l'activité GalAT et l'activité de la rhamnosyltransferase (RhaT), tandis que l'initiation, l'allongement et le branchement des chaînes latérales font appel à de multiples galactosyltransferases (GalTs) et arabinosyltransferases (AraTs) (Liwanag et *al.*, 2012; Atmodjo et *al.*, 2013).

La biosynthèse du RG-II ne nécessiterait pas moins de 22 GTs (Bar-Peled et *al.*, 2012). Actuellement quatre enzymes appartenant au sous-groupe B de la famille GT77 ont été identifiées et clairement impliquées dans la synthèse de ce polymère, il s'agit des xylosyltransférases 1; 2; 3 et 4 (RGXT1; 2; 3 et 4) qui sont capables de transférer le xylose du donneur UDP-xylose sur un résidu fucose. (Egelund et *al.*, 2006; Petersen et *al.*, 2009; Liu et *al.*, 2011) grâce à leur activité α-1,3-xylosyltransférase.

La biosynthèse de la lignine

Les unités monomériques de la lignine, les monolignols sont généralement synthétisées via deux voies successives : la voie générale des phénylpropanoïdes, et la voie spécifique des monolignols (Figure 18). Ensuite ces derniers sont acheminés vers la paroi, où ils seront activés par oxydation déhydrogénative, puis intégrés dans le polymère de lignine déjà existant (Boerjan et *al.*, 2003; Ralph et *al.*, 2004; Vanholme et *al.*, 2010).

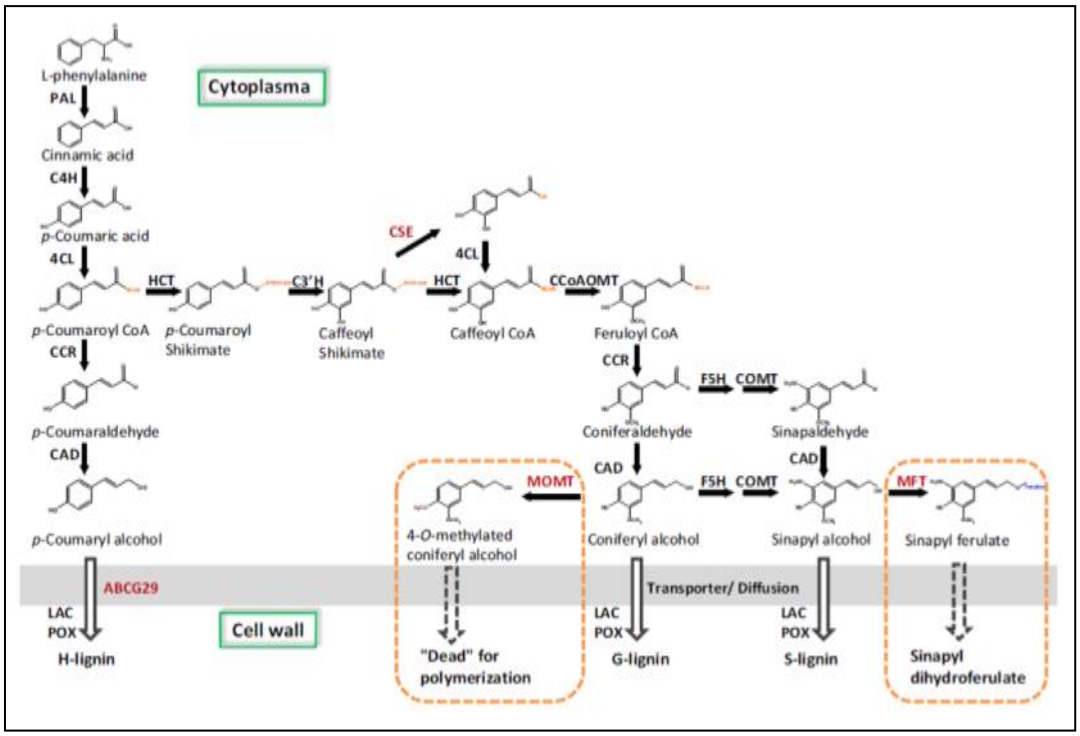


Figure 18 : La voie de biosynthèse de la lignine d'après Liu et *al.*, 2014 impliquant la CSE, la nouvelle enzyme décrite dans cette voie. CSE : Caffeoyl Shikimate Esterase, PAL : Phenylalanine Ammonia Lyase ; C4H : Cinnamic acid 4-hydroxylase ; 4CL : 4-Coumarate CoA ligase ; CCR : Cinnamoyl Coenzyme Reductase; CAD : Cinnamyl Alcohol Desydrogenase ; HCT : Hydroxycinnamoyl CoA Transferase ; C3'H : p-Coumaroylshikimate 3'-Hydroxylase ; F5H : Ferulic acide 5-hydroxylase ; COMT : Cafeic acide /5-hydroxyferulic 3-,5-O-methyltransferase ; CCoAOMT : Caffeoyl Coenzyme A 3-,5-O-methyltransferase. LAC : laccase ; POX : peroxidase ; ABCG29 : transporteur ABC

Le dépôt des sous-unités constituant la lignine se fait d'une manière hétérogène, en effet, elle est, non seulement, différente entre gymnosperme et angiosperme, mais elle affiche aussi une grande diversité qualitative et quantitative au sein des espèces, des tissus et cellules de la même plante (Figure 19) (Nakashima et *al.*, 2008; Tobimatsu et *al.*, 2013). À l'échelle d'une cellule, la disposition de la lignine se fait de l'extérieur vers l'intérieur, débutant par les jonctions tri-cellulaires, puis se propage dans la lamelle moyenne, la paroi primaire, pour finir dans toute la paroi secondaire (Donaldson et *al.*, 2001; Tobimatsu et *al.*, 2013).

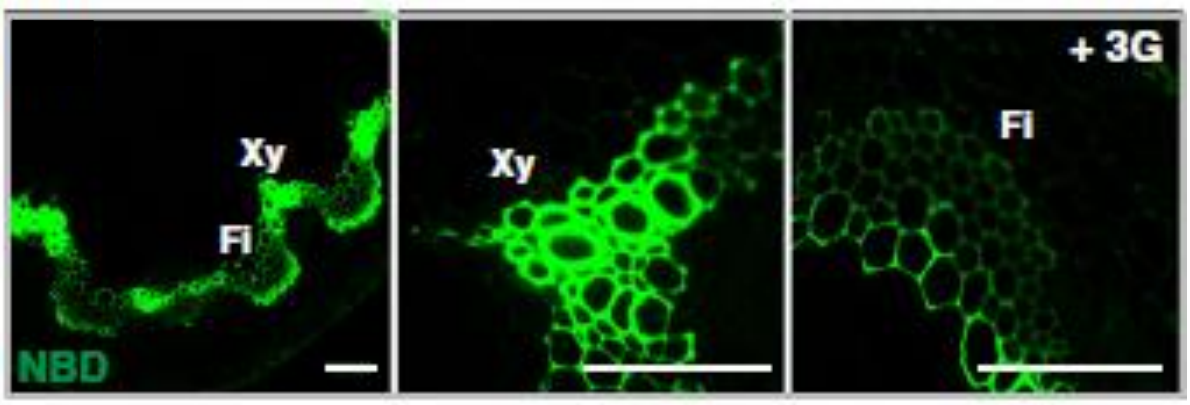


Figure 19 : Imagerie de microscopie confocale montrant l'incorporation dans le polymère de la lignine de monoligols marqués, dans différents tissus provenant d'*Arabidopsis thalina.* Xy : xylème, Fi : fibres, hampe florale d'*Arabidopsis*. D'après Tobimatsu et *al.*, 2013.

La voie générale des phénylpropanoïdes (Figure 18) commence avec l'acide aminé aromatique, la phénylalanine, qui subit trois séries de réactions. La première réaction implique la désamination de la phénylalanine par la première enzyme de la voie : la phénylalanine ammonia lyase (PAL). Puis suit une série de modifications du cycle aromatique via des hydroxylations du cycle en position *para*, puis une thioestérification du groupe carboxyl par le coenzyme A (formation du *p*-coumaroyl-CoA). Ensuite des méthylations et des estérifications auront lieu et la réduction de la fonction thioester en aldéhyde complètera cette voie générale. La voie des phénylpropanoïdes servira ensuite à d'autres synthèses comme celle de la lignine/des lignanes, ou bien des flavonoides. La voie spécifique des monolignols génère les trois hydroxycinnamoyl alcools, *p*-coumaryl, coniferyl et sinapyl alcool (Vanholme et *al.*, 2010).

Plus de dix enzymes successives (Figure 18) catalysent cette voie de synthèse. Trois d'entre elles nommées Cinnamique acid 4-hydroxylase (C4H), *p*-Coumaroylshikimate 3'-hydroxylase (C3'H), et coniferaldehyde/ Ferulic acide 5-hydroxylase (F5H) catalysent les étapes d'hydroxylation. Les enzymes Cafeique acide /5-hydroxyferulic 3-,5-O-methyltransferase (COMT) ou aldéhyde OMT, la Caffeoyl Coenzyme A 3-,5-O-methyltransferase (CCoAOMT) sont responsables des étapes de méthylation. Enfin, la 4-coumarate CoA ligase (4CL) catalyse l'étape d'estérification. Pour aboutir au précurseur de la lignine, il manque encore des réductases qui sont la Cinnamoyl Coenzyme Reductase (CCR) et la Cinnamyl Alcohol Desydrogenase (CAD).

Cette voie dans laquelle l'hydroxycinnamoyl-CoA : shikimate/quinate hydroxycinnamoyl transferase (HCT) est définie comme la première enzyme qui engage le flux de carbone de la voie des phenylpropanoides vers la voie spécifique des monolignols, a été largement étudiée.

Valorisation de la paroi végétale

Les cellules végétales sont caractérisées par la présence d'une paroi cellulaire considérée comme le squelette des plantes. Comme la (Figure 20) l'illustre, à l'échelle macroscopique, la « biomasse » des plantes est composée d'organes, ces organes telle que la tige sont un ensemble de tissus. Ces tissus sont eux même composés de cellules avec parois, constituées majoritairement d'un assemblage de polymères polysaccharidiques, et phénoliques dans certain cas (ex. le bois). Ainsi, d'une certaine façon, la biomasse des plantes correspond à l'ensemble des parois cellulaires. Le bois a été l'une des premières matières utilisées par l'homme que ce soit comme outil de chasse, source de chaleur, ou pour la construction de son habitat. Avec l'âge industriel, la paroi végétale et ses dérivés ont trouvé leur place dans l'industrie notamment dans le textile et la papeterie. Par la suite, de nouveaux débouchés ont été ouverts et cette paroi a été utilisée dans la production de matériaux d'isolation, et dans la production de filtres et de membranes. Au cours de la dernière décennie, la prise de conscience des limites des énergies fossiles, des problèmes liés à leur consommation : approvisionnement, coût, et implication dans le réchauffement climatique a engendré l'orientation de la recherche et de l'industrie vers la production d'énergies propres et renouvelables. La biomasse végétale, de par sa composition, et la diversité des polymères la constituant, est perçue comme étant une des meilleures alternatives, aux produits fossiles, car, outre la possibilité d'obtenir une énergie propre, il est possible de générer des bio-polymères pouvant servir de matière première.

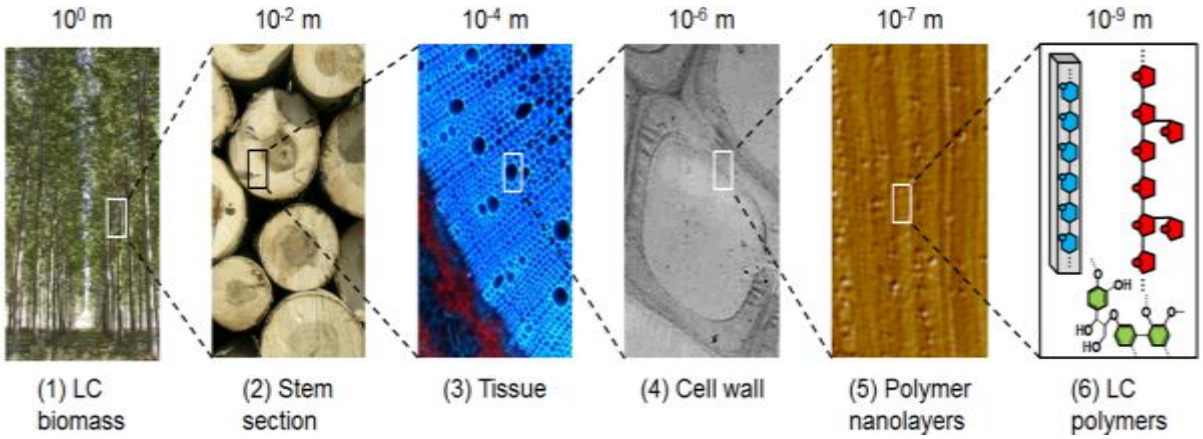


Figure 20 : De l'arbre (biomasse lignocellulosique) aux polymères lignocellulosiques. (D'après Paës, 2014).

Cet intérêt pour la biomasse végétale est conforté par le fait que la cellulose est considérée comme le polymère le plus abondant sur la surface terrestre avec une production annuelle

globale estimée à 180 Gtons (Festucci-Buselli et *al.*, 2007). Pour créer de nouveaux produits à partir de ces matières premières, les industriels utilisent la biomasse lignocellulosique constituée de différents composés chimiques : la lignine, la cellulose et l'hémicellulose. Ces deux dernières sont principalement utilisées pour la production du bioéthanol et d'autres bio-fuels (Figure 21). La lignine, quant à elle, est le deuxième élément d'origine végétale le plus abondant dans le milieu terrestre, et est la seule à avoir des composants aromatiques, qui sont particulièrement intéressants dans la production de plastique.

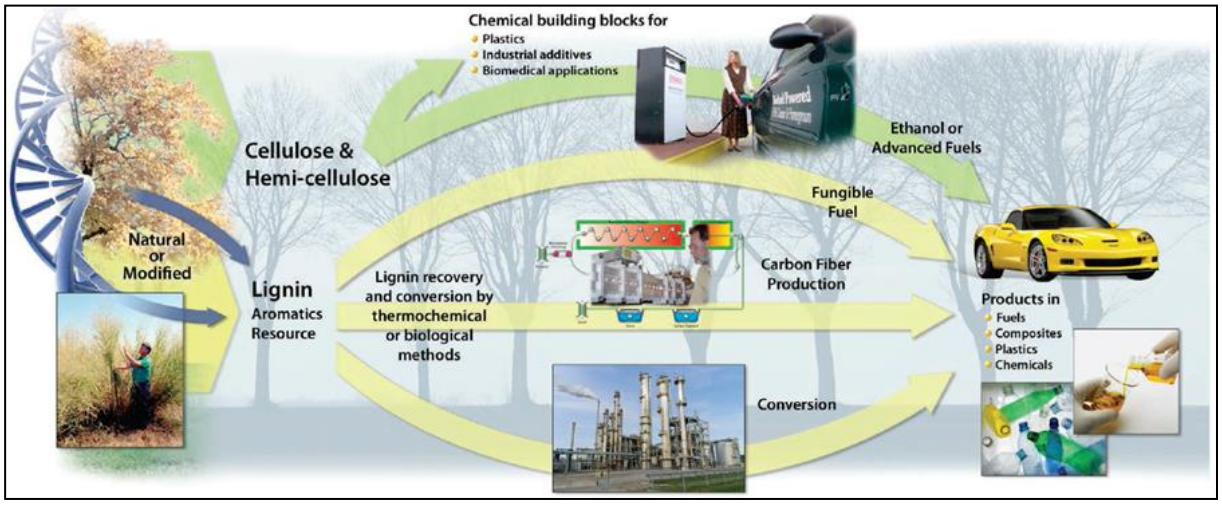


Figure 21: Valorisation industrielle de la biomasse lignocellulosique. D'après (Ragauskas et *al.*, 2014).

Afin de valoriser au mieux ses composants, il faut dans un premier temps déstructurer la paroi pour pouvoir séparer ses différents polymères : par exemple, la séparation de la lignine des autres polymères polysaccharidiques. Or ceci est l'un des problèmes les plus importants pour les industriels. En effet, bien que la lignine soit un hétéro-polymère nécessaire et indispensable au bon développement des plantes vasculaires, elle constitue un obstacle majeur pour l'exploitation des polysaccharidiques cellulosiques et non-cellulosiques en industrie (Grabber, 2005; Lygin et *al.*, 2011; Vanholme et *al.*, 2012). Ces dernières années, la structure et l'organisation des composants polysaccharidique eux-mêmes ont été mis en cause dans ce qui a été appelé la « récalcitrance » de la paroi cellulaire (Pauly and Keegstra, 2008; Le Ngoc Huyen et *al.*, 2010; DeMartini et *al.*, 2013). Ainsi l'opportunité de l'utilisation de cette ressource naturelle et renouvelable, est tributaire de la composition de la paroi cellulaire, c'est à dire la structure de ses polymères, leur abondance relative au sein de la matrice pariétale, et les interactions entre eux. Or tous ces paramètres ne demeurent pas fixes, mais ils sont continuellement remodelés, et

ce, en fonction de plusieurs facteurs. Alors quels sont ces facteurs qui peuvent impacter la composition et la structure de la paroi cellulaire ?

Les facteurs majeurs qui peuvent impacter la paroi

La paroi cellulaire, est un élément indispensable aux plantes, sa composition est complexe, variable. En effet, les polysaccharides présents dans la paroi varient non seulement d'une espèce à une autre mais aussi au sein même d'une espèce ; ceci est dû à la variabilité génétique du monde vivant. Il est aussi admis que la paroi cellulaire est d'une grande plasticité, et qu'elle réagit à son environnement par des remodelages de sa structure, et un réajustement dans sa composition (Carpita et *al.*, 2001; Trethewey et *al.*, 2005).

Variabilité génotypique

La vie est caractérisée par une variabilité incroyable de formes. L'étude de cette abondance est reconnue dans l'utilisation du mot « biodiversité ». Celle-ci a été définie en 1992 par la Convention sur la Diversité Biologique émise par les Nations Unis, comme « variabilité des organismes vivants de toute origine y compris, entres autres, les écosystèmes terrestres, marins et autres écosystèmes aquatiques et les complexes écologiques dont ils font partie : cela comprend la diversité au sein des espèces et entres les espèces ainsi que celle des écosystèmes ». (https://www.cbd.int/doc/legal)

La biodiversité est complexe et s'organise en trois niveaux distincts : biodiversité écosystémique, biodiversité interspécifique et biodiversité intra-spécifique.

La diversité interspécifique réside dans les différences morphologiques observées entre les espèces. Ces différences sont en partie orchestrées par la paroi cellulaire vu qu'elle détermine la forme des cellules. Généralement, la paroi cellulaire s'organise en deux/trois couches : lamelle moyenne, paroi primaire et paroi secondaire (si présent). Si globalement l'organisation de la paroi est semblable dans la plupart des plantes terrestres, des différences résident néanmoins dans la structure infime de la paroi, aux niveaux de l'abondance relative des différents polysaccharides et de leurs interactions. Afin de mettre en lumière ces différences, des études chimiques de la composition de la paroi primaire et secondaire ont été menées sur différents espèces appartenant à des groupes phylogénétiquement différents. Ainsi des différences considérables et significatives, ont été démontrées entre les parois primaires et secondaires des monocotylédones et des dicotylédones. Ces différences ont été illustrées par deux modèles architecturaux de la paroi primaire, définie comme paroi de type I qui regroupe la paroi des

dicotylédones, des gymnospermes et des monocotylédones non-Commelinidées (Figure 22), et la paroi de type II qui est typique des monocotylédones graminées.

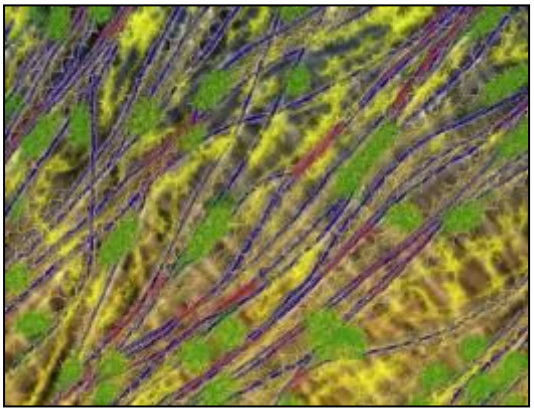


Figure 22 : Un modèle récent proposé pour illustrer la paroi primaire de type I, les microfibrilles de cellulose sont représentés en bleu, les pectines en jaune, les hémicelluloses en vert. D'après Cosgrove, 2014.

Au sein de ces deux groupes la cellulose ne semble pas montrer de différences majeures, alors qu'il existe une plus grande variabilité au niveau des polysaccharides non-cellulosiques, de la lignine, ainsi qu'au niveau des protéines.

En effet, au sein de la paroi primaire des monocotylédones de type Poacées (Carpita et Gibeaut 1993), les microfibrilles de cellulose sont noyés dans une matrice polysaccharidique composée de glucuronoarabinoxylans, et de (1-3, 1-4) β glucanes. Au sein de cette matrice polysaccharidique, les pectines et les xyloglucanes sont minoritaires (Smith and Harris, 1999), tandis qu'ils sont les principaux éléments retrouvés au sein de la paroi primaire des dicotylédones (en plus de la cellulose). La paroi secondaire est plus épaisse que la paroi primaire et se caractérise par une importante présence des xylanes et de lignine que ce soit chez les monocotylédones ou les dicotylédones (Tableau 3). Cependant, des différences majeures existent. En effet, chez les dicotylédones, la lignine est composée principalement de sous unités G et S, tandis que chez les monocotylédones, la sous-unité H est aussi significativement représentée (Grabber, 2005). L'autre différence retrouvée concerne les protéines structurales qui sont différentes dans la paroi primaire des monocotylédones en comparison avec les dicotylédones (Vogel, 2008).

Tableau 3 : Composition approximative en (% de la masse sèche) de la paroi primaire et secondaire typique des monocotyledones (Poacaea) et des dicotyledones. MLG: Multi linkage glucanes. XyG: xyloglucanes. (D'après Vogel, 2008).

	Primary wall		Secondary wall	
	Grass	**Dicot**	**Grass**	**Dicot**
Cellulose	20–30	15–30	35–45	45–50
Hemicelluloses				
Xylans	20–40	5	40–50	20–30
MLG	10–30	Absent	Minor	Absent
XyG	1–5	20–25	Minor	Minor
Mannans and glucomannans	Minor	5–10	Minor	3–5
Pectins	5	20–35	0.1	0.1
Structural proteins	1	10	Minor	Minor
Phenolics				
Ferulic acid and *p*-coumaric acid	1–5	Minor except order Caryophyllales	0.5–1.5	Minor except order Caryophyllales
Lignin	Minor	Minor	20	7–10
Silica			5–15	Variable

Les différences observées au niveau de la paroi cellulaire ne sont pas uniquement interspécifiques, mais ces différences sont aussi retrouvées au sein même d'une espèce. Une récente étude de De Souza et *al.*2015 a mis en évidence la diversité intra-spécifique retrouvée au sein de la paroi cellulaire de l'espèce *Miscanthus*. En effet l'analyse de sept génotypes différents de cette espèce a révélé des différences significatives au niveau de la lignine entre trois génotypes. Puis une analyse plus approfondie de la composition en polysaccharides de ces trois génotypes a permis de mettre en évidence des différences dans la quantité en cellulose, des hémicelluloses et dans le ratio hexose/pentose des monosaccharides pariétaux (Souza et *al.*, 2015).

Variabilité environnementale

Si certaines différences de structure pariétale peuvent être mises en rapport avec la génétique, il est également important de noter que la composition de la paroi n'est pas figée dans le temps, puisque celle-ci varie en fonction de l'environnement dans lequel l'individu se trouve. En effet, les plantes réagissent continuellement à leur environnement, si les conditions environnementales ne sont pas favorables à une croissance optimale, on parle de stress. Deux types de stress sont

connus, les stress abiotiques et les stress biotiques. Ces derniers regroupent toutes les réactions engendrées par les organismes vivants, que ce soit des pathogènes, ou juste des compétitions inter-spécifiques pour des nutriments au sein d'un même espace de vie. En ce qui concerne les stress abiotiques, ce sont toutes les conditions environnementales défavorables à une croissance optimale telles que le manque (ou excès) d'éléments nutritifs, ainsi que toutes les réactions engendrées par le climat telles que les variations de température et/ou radiation (UV), des ressources hydriques (sécheresse, inondation, salinité…etc).

Le changement climatique et les stress abiotiques chez les végétaux

Le changement climatique correspond à une modification durable des paramètres statistiques du climat global de la Terre ou des différents climats régionaux. Ces changements peuvent être dus à des processus intrinsèques à la Terre, à des influences extérieures ou, plus récemment, aux activités humaines. Actuellement, on assiste à un changement climatique anthropique à cause des émissions de gaz à effet de serre engendrées par les activités humaines et modifiant la composition de l'atmosphère de la planète, auxquelles viennent s'ajouter les variations naturelles du climat (Scarwell and Roussel, 2010 et IPCC de 2007).

Au sein des plantes, la paroi cellulaire est une des premières lignes de défense qui se dresse contre ces changements climatiques. La paroi cellulaire réagit aux différents stress abiotiques par différents mécanismes, contribuant ainsi à la défense de la plante, lui permettant de trouver un rééquilibre de survie, temporaire ou durable, en fonction des espèces.

Les réactions de la paroi aux stress abiotiques impliquent des modifications de sa composition. En effet, diverses études comparatives réalisées sur différentes variétés de blé, montrant des tolérances variables aux stress salin (Uddin et *al.*, 2013; An et *al.*, 2014), ou hydrique (Leucci et *al.*, 2008), ont révélé à chaque fois, une augmentation dans la quantité de pectines présentes dans les parois des plantes tolérantes à ces stress. Cette tolérance est probablement expliquée par la présence d'un gel hydratant constitué de pectine, qui permet de limiter les dommages cellulaires (Leucci et *al.*, 2008). L'augmentation de la proportion de pectine, ainsi que d'autres polysaccharides, a été aussi rapportée chez *Arabidopsis halleri*, sujette à un stress au cadmium. L'augmentation de la quantité des polysaccharides tels que les arabinoglucanes chez le blé (Rakszegi et *al.*, 2014), des β glucanes chez Miscanthus (Domon et *al.*, 2013), ainsi que de la lignine, ont été mis en évidence, au sein de la paroi cellulaire, suite à des études comparatives réalisées sur des génotypes tolérantsaux différents stress.

Le stress hydrique

En tant que phénomène naturel, le stress hydrique est connu sous le terme de « sécheresse », et représente des longues périodes de disponibilité d'eau en quantités insuffisantes pour une croissance optimale. De nos jours, les sécheresses augmentent en ampleur et en fréquence. Ce phénomène provoque l'impact le plus marquant parmi les différents types de stress sur tous les types d'organismes vivants.

Chez les plantes, l'eau est essentielle pour la photosynthèse, la croissance, le maintien du port érigé (turgescence), les différents mouvements, le refroidissement par évapotranspiration, ainsi que pour l'absorption et le transport de solutés en tant que solvant. L'exposition à un stress hydrique mène à un déséquilibre entre la quantité d'eau perdue par la transpiration et la quantité d'eau absorbée, ce qui crée un déficit hydrique dans la plante. La déshydratation au niveau cellulaire cause la diminution du volume cytosolique et vacuolaire, ce qui se traduit par un stress osmotique. Ce déficit survient aussi bien suite à un asséchement du sol, qu'à un remaniement des forces osmotiques limitant la prise d'eau, ou bien suite à un réchauffement atmosphérique qui accroit l'évaporation d'eau (Neumann, 2011).

Ainsi, comme l'eau est indispensable à la vie et est essentielle pour le métabolisme, son manque impact négativement la croissance des plantes et leur productivité.Cette situation empirant progressivement en raison du changement climatique global (http://www.ipcc.ch) aura des effets majeurs sur les écosystèmes, avec un risque d'extinction de certaines espèces ainsi qu'une modification des aires de répartition de celles-ci (Nicotra et *al.*, 2010). Il est donc important d'améliorer non seulement notre compréhension sur les réponses des plantes à ce stress, mais aussi sur les modifications quantitatives et qualitatives des productions végétales en condition de stress hydrique.

Les effets d'un stress hydrique s'étendent de la structure moléculaire des plantes à leurs morphologies globales, et impactent toutes leurs étapes de croissance. On compte parmi ces effets le ralentissement de la croissance, une diminution du contenu en eau, la diminution de la photosynthèse, des modifications dans l'allocation des réserves en favorisant les racines aux parties aériennes de la plante (Wu and Cosgrove, 2000), une intensification de la respiration, et des dommages oxydatifs. Ces effets ont été décrits en détail chez de nombreuses espèces (Ingram and Bartels, 1996; Penna, 2003; Agarwal, 2006; Farooq et *al.*, 2009; Kantar et *al.*, 2011)

En conditions de stress hydrique, le contenu en eau des tissus est réduit, par conséquent la turgescence diminue drastiquement et l'élongation cellulaire s'arrête. La diminution de la photosynthèse associée au stress hydrique réduit l'approvisionnement en métabolites nécessaires pour la division cellulaire. Par conséquence, les divisions mitotiques, l'élongation et l'expansion cellulaire sont diminuées avec un effet négatif sur la croissance générale de la plante. Cet effet impact la germination des graines et le développement des plantules (Kaya et *al.*, 2006; Zeid and Shedeed, 2006), mais aussi la croissance de la plante entière et son développement (Tripathy et *al.*, 2000; Manickavelu et *al.*, 2006). Ces évènements (division, élongation et différenciation cellulaire) sont étroitement liés et orchestrés par la paroi cellulaire.

En condition de stress hydrique on peut observer des modifications compositionnelles et architecturales de la paroi des cellules des feuilles(Moore et *al.*, 2008).Toutefois, des études menées sur l'effet du même stress sur les racines donnent des conclusions variables suivant l'espèce étudiée. À titre d'exemple, Alvarez et *al.* en 2008 rapportent que la lignification diminue chez le maïs, alors que chez d'autres espèces telles que *Trifolium repens*(Lee et *al.*, 2007), et le riz (Yang et *al.*, 2006) la lignification augmente. La lignification est associée à l'expression des gènes des peroxydases, or plusieurs études ont rapporté l'augmentation des peroxydases suite à un stress hydrique. D'autres études ont aussi rapportées l'augmentation de l'expression de différents gènes impliqués dans la synthèse des monolignols c'est le cas du *PAL* chez le riz et du *CCR* dans la racine du maïs(Fan et *al.*, 2006).

Le stress hydrique engendre aussi un déséquilibre entre la production des espèces réactives d'oxygène (ROS) et les mécanismes de défense antioxydant, ceci fait que les ROS s'accumulent et induisent un stress oxydatif au niveau des protéines, des membranes et des divers compartiments cellulaires (Fu and Huang, 2001; Ramachandra Reddy et *al.*, 2004). Leur accumulation au niveau de la paroi en condition de stress, favorisera la lignification car elles sont utilisées par les peroxydases pour activer les monolignols, permettant ainsi leur polymérisation, ces ROS provoqueront aussi des changements de paroi du à leur action sur les protéines structurales. La réunion de ces trois facteurs permet de proposer un modèle reliant la lignification et la croissance cellulaire en condition de stress (Figure 23).

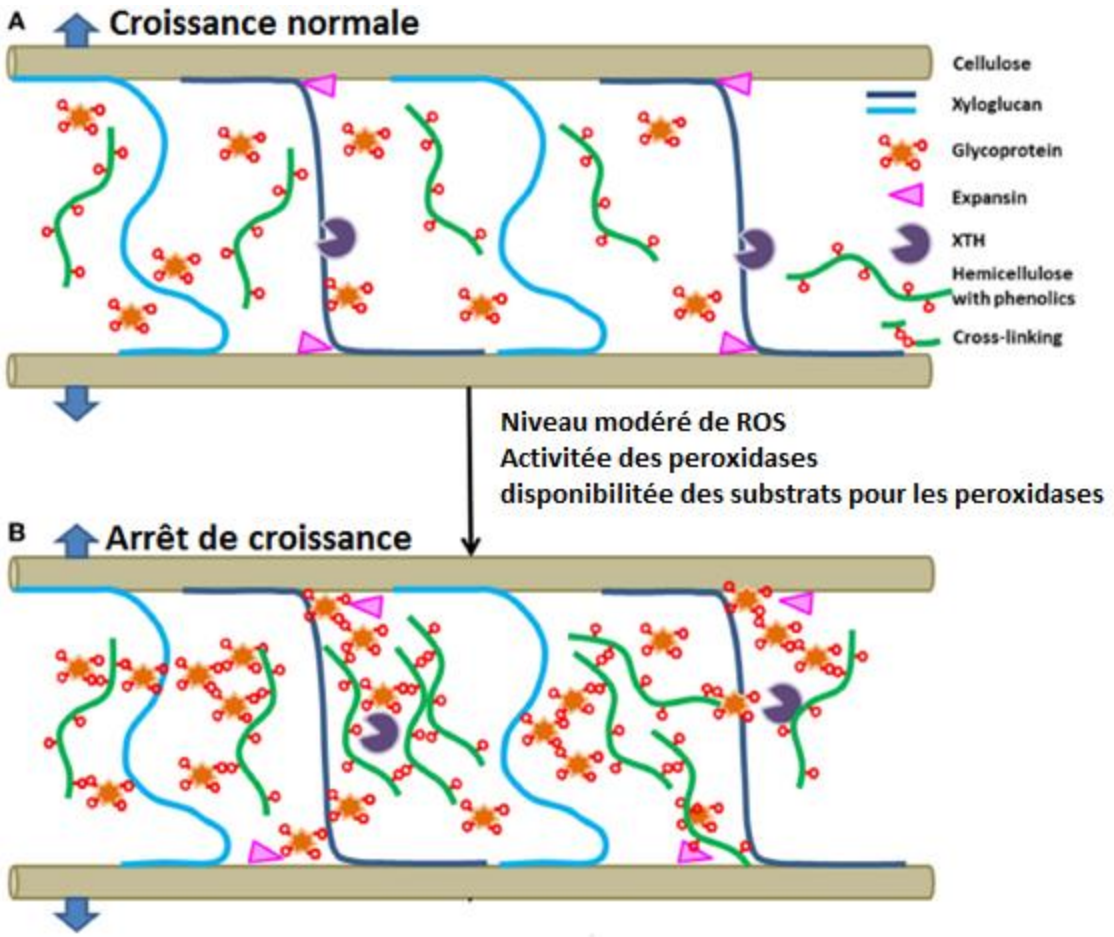


Figure 23 : Modèle représentant de la structure pariétale dans des conditions normales (croissance cellulaire) (A), et en réponse à un stress abiotique (arrêt de croissance) (B). Adaptés de (Tenhaken, 2015).

Les peroxydases ne sont pas les seules enzymes de la paroi cellulaire dont le niveau d'expression change en réponse au stress hydrique. En effet, l'expression de gènes codant autres enzymes telles que les expansines et les XTHs (Dong et *al.*, 2011; Han et *al.*, 2012)est également modifiée. En modifiant l'expression de ces gènes, le stress hydrique induit des modifications architecturales de la paroi cellulaire. Divers travaux rapportent aussi des modifications de la composition de la paroi en réponse à un stress hydrique : une augmentation de la quantité des pectines (Piro et *al.*, 2003; Konno et *al.*, 2008; Leucci et *al.*, 2008), une diminution de la quantité de la cellulose chez de nombreuses espèces (Piro et *al.*, 2003; Bray, 2004), tandis que chez le coton, une étude suggère une augmentation dans la synthèse de la cellulose (Zhong and Ye, 2014).

Le stress hydrique a donc un effet sur la structure et la composition de la paroi cellulaire, mais son effet est tissu- et espèce-dépendant.

Le lin comme modèle d'étude de la paroi végétale

Le lin cultivé est une angiosperme appartenant à la famille des Linacées qui regroupe 22 genres répartis sur environ 300 espèces (McDill et *al.*, 2009). Le genre *Linum* est le plus largement répandu au sein de cette famille, avec environ 200 espèces recensées.

Morphologiquement, la plante de lin se dresse à l'aide d'une tige principale cylindrique, susceptible d'avoir à sa base des ramifications, dont le nombre peut varier en fonction du climat régional, de la richesse du sol (Dillman and Brinsmade, 1938) et de la densité de semi. La tige de lin est fine et haute. Elle a en effet, quelques millimètres de diamètre (1 à 3mm) pour une hauteur qui varie de 20 à 150 cm. Cette tige porte des feuilles simples disposées sous forme spiralée due à la superposition de trois hélices foliaires. Les feuilles, sessiles et tri-nervurées, sont d'une longueur variable de 15 à 33 mm et d'une largeur également variable de 3 à 13 mm. Elles deviennent sénescentes à maturité, puis tombent.

L'apparition d'une inflorescence en forme de cyme sur l'axe primaire caractérise le début du cycle de reproduction. À partir de cet axe, se développent des axes secondaires qui peuvent se ramifier à leur tour. Sur chacune des branches, une fleur apparaîtra. Les fleurs, dont il existe une panoplie de couleurs pour les pétales, sont généralement bleues, et sont hermaphrodites de type 5 (5 pétales) (Figure 24). La plante du lin est généralement autogame, donc généralement avant l'ouverture de la fleur, la fécondation aura déjà eu lieu, et donnera un fruit de type capsule. Al'intérieur de cette dernière se développeront les graines qui peuvent être en nombre de 10, logées dans les 10 lodicules dont la capsule est divisée (Diederichsen and Richards, 2003).

Figure 24 : Les deux couleurs principalement retrouvées chez les fleurs du lin, le bleu étant représentatif des lignées originelles (M.Chabi).

Utilisation et valorisation du lin « Tout est bon dans le lin »

Le lin est cultivé depuis longtemps pour ses fibres cellulosiques et pour ses graines riches en huile.

Utilisation des variétés à fibres

Les variétés à fibres sont traditionnellement cultivées pour l'industrie textile. Néanmoins, elles sont de plus en plus utilisées dans d'autres secteurs industriels, et ce grâce aux propriétés physico-chimiques et mécaniques dont ces fibres naturelles sont dotées. En effet, ces cellules fibreuses sont très longues et se caractérisent par une paroi secondaire épaisse, constituée principalement de cellulose (65 à 85%) (Baley, 2002) et d'hémicelluloses. Les fibres de lin contiennent un taux très faible (entre 1,5 et 4,2%) de lignine (Day et *al.*, 2005b).

L'association de ces trois caractéristiques : (i) la faible teneur en lignine de la paroi secondaire, (ii) la richesse en cellulose et (iii) la longueur exceptionnelle des fibres élémentaires, permet d'avoir des propriétés mécaniques très intéressantes pour l'industrie .La souplesse et la flexibilité sont des caractéristiques très recherchées dans les industries textiles. Les fibres sont aussi utilisées pour la production de matériaux isolants du fait de leurs propriétés acoustiques et de leur faible conductance thermique. De plus, la faible densité, et le module spécifique (tenue à l'allongement à masse identique) de ses fibres cellulosiques comparables à celle de la laine de verre (Baley, 2002; Joshi et *al.*, 2004), leur ont ouvert de nouvelles voies d'utilisation comme l'industrie des matériaux bio-composites, où les fibres courtes du xylème sont déjà utilisées dans la fabrication du plastique modifié et comme charge pour le ciment (Zini et *al.*, 2003; Kymäläinen and Sjöberg, 2008).

Récemment, une nouvelle génération des composites entièrement biodégradables et bioactif a été décrite. Elle est basée sur des fibres de lin bioplastiques (lin transgénique contenant du polyhydroxybutyrate) et des polymères biodégradables qui offrent des propriétés remarquables de par leur non-cytotoxicité, bactériostatique et plaquettes anti agrégation. (Wróbel-Kwiatkowska et *al.*, 2012). De nouvelles opportunités d'application peuvent aussi être envisagées pour les fibres de lin dans le domaine médic*al*. En effet des travaux de recherche menés par Michel et *al.* (2014), ont révélé des propriétés anti-inflammatoires des fibres de lin de par la présence d'une molécule active similaire au canabinoïde : ceci a été confirmé par Michel et *al.*, (2014) qui ont rapporté la possibilité d'une utilisation biomédicale des fibres de lin dans la chirurgie.

Utilisation des variétés à graines

Le lin à graines est cultivé pour sa teneur élevée en huile qui représente 35 à 45% de la graine. Cette huile a été longtemps le centre d'intérêt de nombreuses industries oléochimiques, principalement dans la production du linoléum, d'encre et de peinture. La trituration des graines de lin, génère un tourteau qui est incorporé dans l'alimentation animale. Plus récemment, leurs teneurs élevées en acide gras insaturé, spécialement l'acide linolénique représentant 45 à 52% de l'huile extraite, ont été mise en évidence Cette acide linolénique est classé comme un acide gras de type oméga-3, bénéfique pour la santé. Un autre point fort des graines de lin pour la santé, a trait à leurs teneurs élevées en lignanes biologiquement actifs (anti-inflammatoire, anti-carcinogène, anti-microbien, protecteur cardio-vasculaire) (Prasad, 1997; Kitts et *al.*, 1999; Adolphe et *al.*, 2009).Ces propriétés ont étendu leur utilisation dans l'engraissement animalier, de manière à produire par exemple de la viande avec une composition en graisse améliorée, plus bénéfique pour la santé humaine (Touré and Xueming, 2010; Yi et *al.*, 2014).

En plus des innombrables intérêts que présentent les graines de lin, la partie végétative de la plante dite « paille » a ses propres débouchés. En effet, elle est utilisée à différentes fins, que ce soit comme matériaux d'isolation, de brûlage, ou autres (Figure 25).

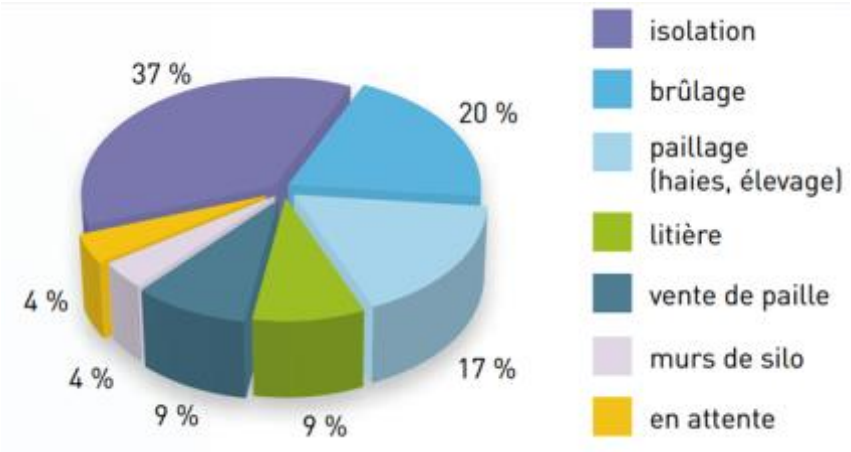


Figure 25 : Valorisation des pailles de lin à huile. (Enquête CETIOM 2009).

La structure de la fibre de lin

Chez les plantes, les « fibres » sont classées en trois types, en fonction de leur localisation. En effet, il existe des fibres dont l'origine sont les fruits ou les graines : c'est l'exemple des fibres du coton ou du kapkier, des fibres dont l'origine est la feuille, c'est le cas chez le Sisal et le

Abaca et enfin les fibres originaires des tiges, c'est l'exemple du lin, du jute, et du chanvre (M. Sfiligoj Smole, 2013). Une analyse histologique d'une coupe transversale de la tige du lin, permet de distinguer différents tissus (Figure 26 A). Les fibres se situent entre le cortex et le phloème, à la périphérie du protophloème (Esau, 1943), et elles s'organisent sous forme de faisceaux de 12 à 36 cellules reliées entre elle par un ciment pectique (Pritchard et *al.*, 2000; Morvan et *al.*, 2003a), le nombre de faisceaux pouvant varier de 20 à 50 par tige.

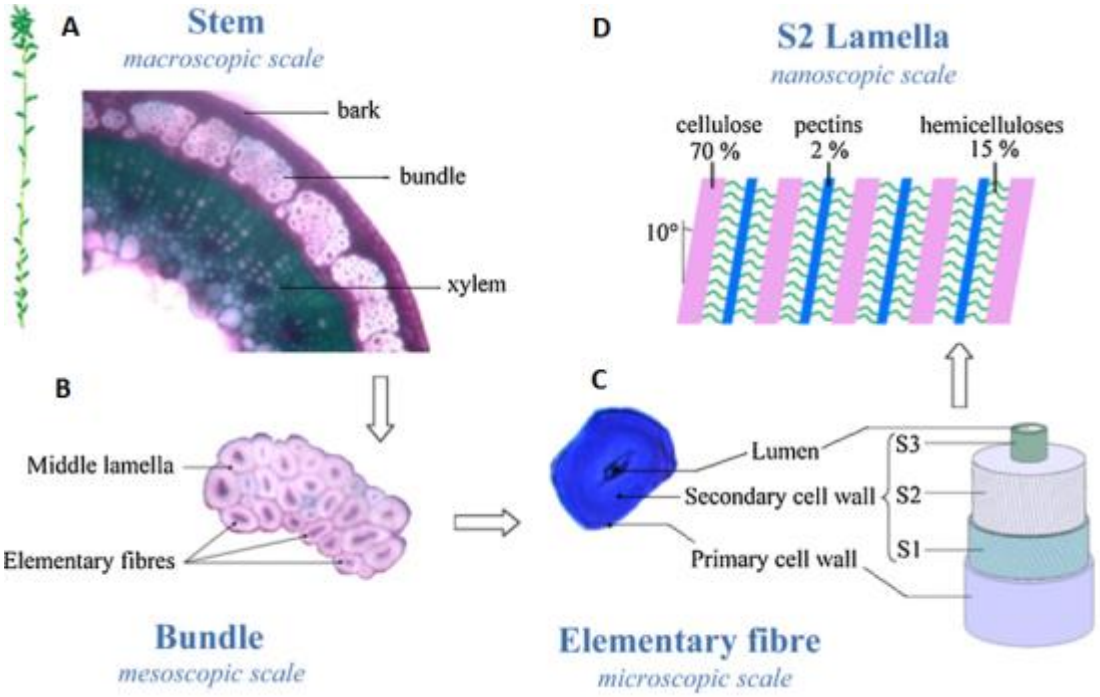


Figure 26 : Représentation schématique des fibres de lin à différentes échelles. (A) Coupe transversale d'une tige de lin qui représente les différents tissus et la localisation des fibres de lin au sein de la tige ; (B) Focalisation sur un faisceau formé de plusieurs fibres élémentaires reliée par le ciment pectique de la lamelle moyenne ; (C) Cellule fibreuse et l'organisation de sa paroi cellulaire ; (D) La composition de la couche S2 de la paroi secondaire des fibres élémentaires du lin. D'après (Charlet et *al.*, 2010)

Les cellules fibreuses, dite fibres élémentaires, diffèrentes des fibres techniques, terme utilisé pour désigner le faisceau en entier (ensemble de cellules). Tandis que les fibres élémentaires représentent une cellule fusiforme unique d'une longueur moyenne qui va de 6 à 65 mm avec un diamètre de 20 µm en moyenne (McDougall, 1993; M. Sfiligoj Smole, 2013). Afin d'ateindre ces dimensions, ces cellules fibreuses trouvent leur origine dans l'aisselle d'une feuille et se

terminent en s'insérant dans la feuille qui suit, croissant rapidement dans un premier temps en longueur d'une manière intrusive de type intercalaire entre les cellules existantes. Ceci se déroule dans la partie apicale de la tige formant ainsi des cellules coenocytiques qui ne possèdent pas de plasmodesmes, les rendant de la sorte complètement isolées au sein de la tige (Ageeva et *al.*, 2005).

Au sein de la tige de lin, la croissance en longueur et l'épaississement de la paroi sont séparés dans l'espace, et ce point de séparation est dit « snap-point »(Figure 27A). En effet, la croissance en longueur se déroule au-dessus de ce point tandis que la croissance en largeur (épaississement de la paroi) se déroule au-dessous de ce point-là. Et celle-ci (croissance en largeur) est centripète, c'est-à-dire que le remplissage des fibres se fait de l'extérieur vers l'intérieur de la tige, et peut durer de trois à cinq semaines (Figure 27B).

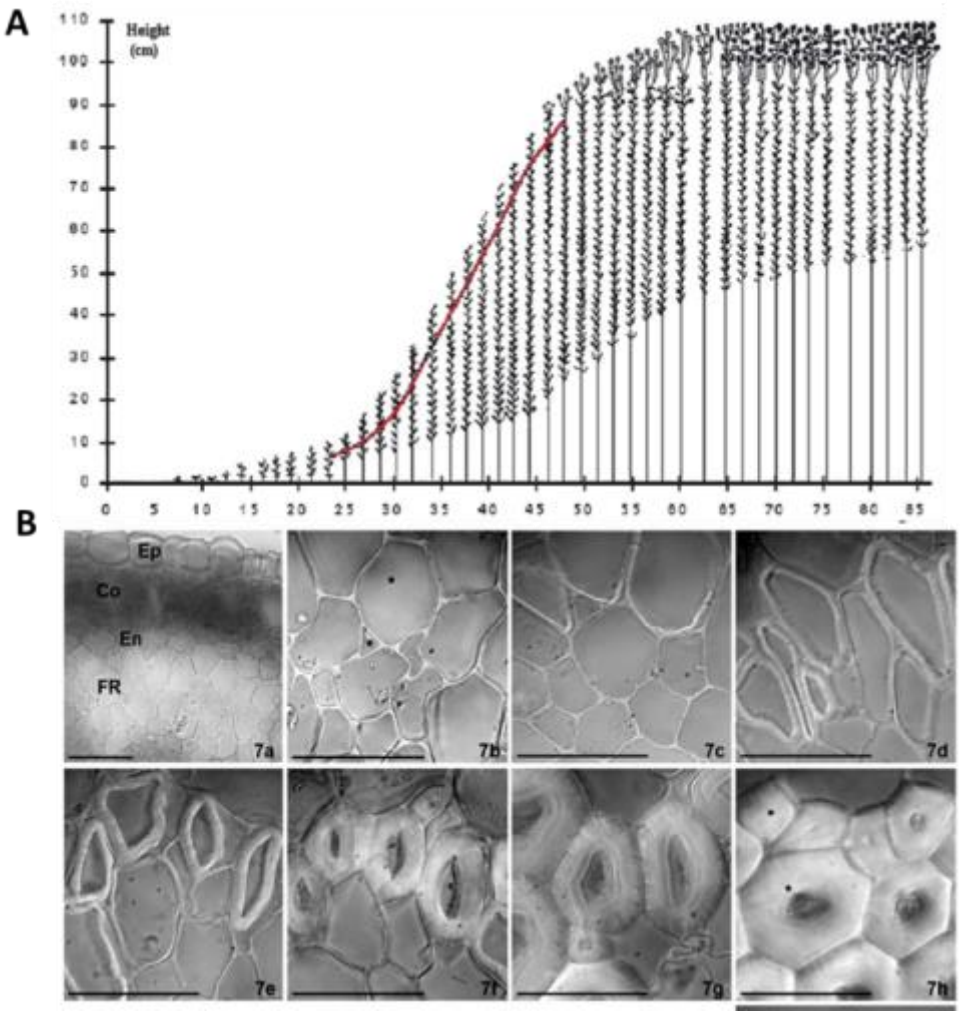


Figure 27 : position du snap-point et le remplissage des fibres chez le lin. (A) Suivie de l'évolution de la position du snap-point durant la croissance de la tige de lin représentée par la ligne rouge (la taille de la tige en cm en fonction de l'âge de la plante en jour. D'après

Gorshkova et *al.*, 2003. (B) 7a. Coupe transversale dans une tige de lin, la position des fibres dans la tige est indiquée par FR. Ensuite les fibres à différentes distances de l'extrémité supérieur de la tige sont représentées de 7b à 7h. Les photographies de 7b à 7h montre le développement des fibres phloèmiennes à 3 ; 5.5 ; 7 ; 8 ; 15 ; 25 et de 70 cm. Ep : Epiderme ; Co : cortex, En : Endoderm; FR: région fibreuse. Barre d'échelle = 50µm. D'après (Ageeva et *al.*, 2005).

Le lin comme modèle d'étude de la paroi cellulaire

Le lin peut être considéré comme un modèle d'étude pour la paroi cellulaire végétale, pas uniquement comme exemple d'une fibre, mais aussi dans le contexte plus large des recherches portant sur l'exploitation de la biomasse végétale. En effet, l'utilisation optimale de ces ressources végétales, nécessite une compréhension plus approfondie de la paroi cellulaire, des interactions existants entre ses polymères, et des mécanismes contrôlant les teneurs relatives des différents polymères. Dans ce contexte, les mécanismes régulant le rapport entre les polymères polysaccharidiques et la lignine sont d'un intérêt particulier car la quantité de lignine est corrélée de façon négative avec la production de bio-éthanol. La tige de lin contient des cellules avec des parois secondaires lignifiées (xylème : entre 23 et 31 %) et hypolignifiées (fibres cellulosiques) (Day et *al.*, 2005c; Del Río et *al.*, 2011). Cette observation laisse supposer que les quantités relatives de lignine et autres polymères polysaccharidiques ne soient pas régulées de la même façon dans les deux tissus. De ce fait, le lin représente un modèle intéressant pour étudier ces régulations pariétales. La disparité de la quantité de lignine n'est pas la seule différence existant entre ces deux tissus, puisque la paroi secondaire des fibres est beaucoup plus épaisse que celle des cellules xylémiennes, et peut atteindre à maturité jusqu'à 15µm d'épaisseur(Gorshkova et *al.*, 2003; Morvan et *al.*, 2003a).

La paroi des cellules fibreuses présente aussi une composition atypique en polysaccharides. Elle est très riche (65 – 85 %) en cellulose, qui présente un degré de cristallinité estimé à 74% (Focher et *al.*, 2001; Baley, 2002), ce qui est très élevé en comparaison à d'autres espèces. De plus cette cellulose s'organise en microfibrilles avec orientation axiale, (Roland et al, 1995) (figure 26 D). Le rhamnogalacturonane I est aussi retrouvé au sein de cette paroi secondaire. Ce RGI présente un ratio GalA/Rha proche de 1, avec un degré de substitution en α-l-Rha*p* de 96% mais qui diminue jusqu'à 72% une fois incorporé à la paroi cellulaire. Ainsi une partie des

propriétés mécaniques des fibres de lin semble être attribuée à cette composition (Gorshkova et *al.*, 1996; Gurjanov et *al.*, 2008; Mellerowicz and Gorshkova, 2012).

La culture du lin

L'espèce *Linum usitatissimum* L., même si moins cultivée aujourd'hui a joué un rôle très important dans l'économie mondiale. Sa domestication, d'après Vavilov en 1926, remonterait à environ 9000 ans. Cette espèce serait originaire du Moyen-Orient, même si des études plus récentes (Zohary and Hopf, 2000) ont démontré une origine qui prête à discussion entre le bassin méditerranéen, l'Asie centrale et l'Inde. Quoiqu'il en soit, le lin fait partie des premières espèces cultivées par l'homme, en étant l'objet d'intérêt à la fois pour ses graines oléagineuses et pour ses fibres. L'origine exacte du lin cultivé reste inconnue, mais il existe deux espèces qui sont décrites dans la littérature, et qui sont supposées être les ancêtres du lin cultivé qui sont *L. bienne*, ou *L. angustifolium* (Muravenko et *al.*, 2001; Muravenko et *al.*, 2003; Muravenko et *al.*, 2004). Une nouvelle étude réalisée dernièrement sur 11 espèces du genre *Linum*, a permis de les situer entre elle, et donc *L. bienne* ne peux être que la proche cousine du *L. usitatisimum* (Sveinsson et *al.*, 2014) (Figure 28).

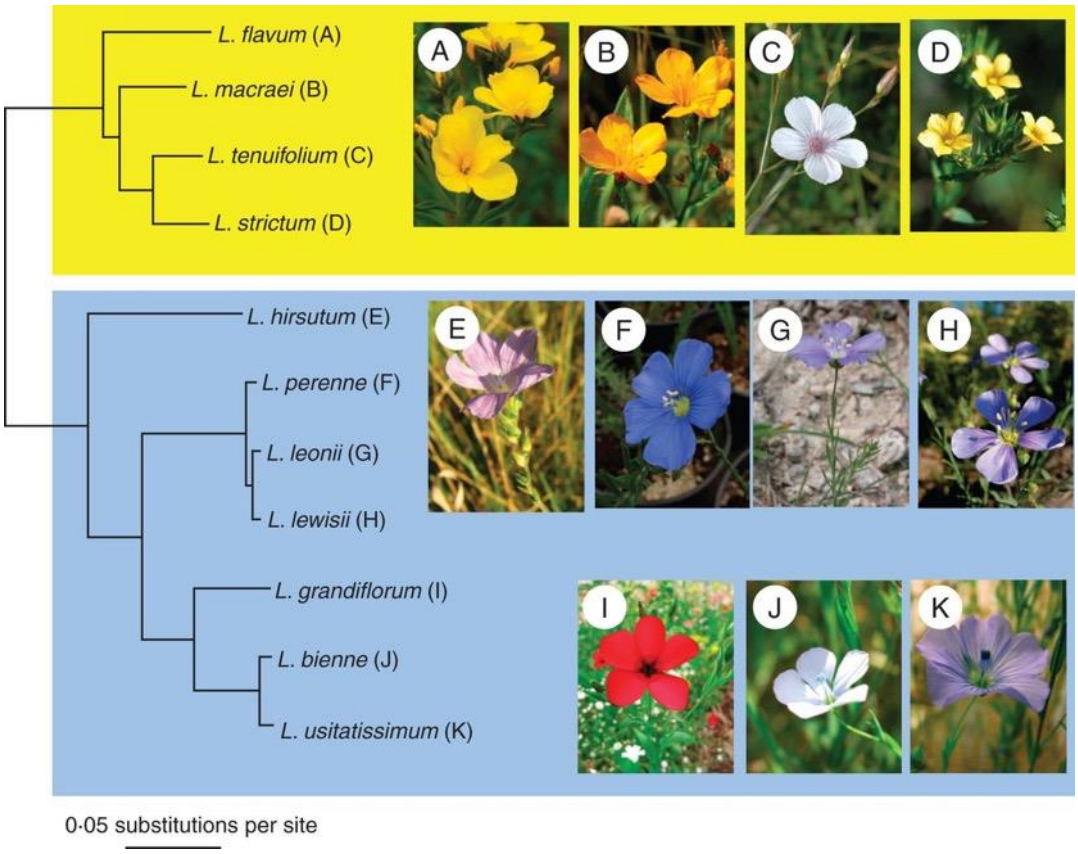


Figure 28 : Arbre phylogénique des 11 espèces du genre *Linum* et la séparation des deux principaux groupes retrouvés en fonction de la couleur de leurs fleurs.D'après Sveinsson et *al.*, 2014.

Le lin, comme d'autres plantes très tôt domestiquées, présente une variabilité génétique très faible, même s'il semblerait que le lin à graines (voir ci-dessous) a une plus grande variabilité génétique (Doré et Varoquaux, 2006).

Au sein de cette même espèce, on distingue deux groupes de variétés : les variétés à fibres et les variétés à graines. Il est à noter que la notion agricole du mot « variété » n'est apparue qu'au début du 19ème siècle, le terme utilisé en agriculture auparavant était « catégorie **»,**mot qui servait à désigner un niveau taxonomique inférieur à l'espèce (Bonneuil et Hochereau, 2008).

Bien que ces variétés présentent des différences morphologiques : forme des capsules, taille, ramifications, il s'agit bel et bien de la même espèce ; c'est ce qui est appelé des « cultigroupes ». En effet, le lin à fibres est cultivé avec une haute densité de semis qui se situe entre 1800 à 2200 graines/ m^2, tandis que le lin à graines est semé avec une densité comprise entre 350 et 400 graines/m^2. Dans le premier cas, les plantes qui en résultent sont très hautes,

possèdent une seule tige, alors que dans le second cas, les plantes de lin sont plus courtes avec une tige qui porte plusieurs ramifications (Figure 29 A).

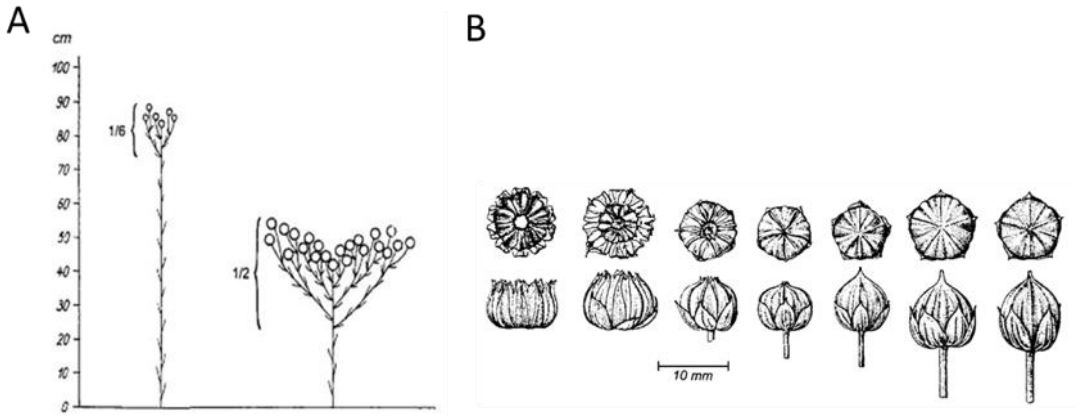


Figure 29 : Représentation schématique des différences existant entre les variétés à fibres et les variétés à graines du lin. (A) différence de taille et de ramification et (B) les différentes formes de capsules observées. D'après (Diederichsen and Richards, 2003).

En ce qui concerne la forme des capsules, la différence est observée lors de la maturation des graines, où les capsules des variétés à fibre sont légèrement ouvertes, et n'atteignent quasiment jamais le nombre maximum de graines -qui est de dix-, tandis que les capsules des variétés à graines restent bien fermées, avec souvent dix graines à l'intérieur (Figure 29 B).

Les variétés à fibres poussent généralement sur une terre argileuse dans un climat relativement tempéré et humide. Ces variétés sont plutôt des cultures de printemps, cependant des variétés d'hiver existent aussi et elles sont de plus en plus nombreuses. En effet, la culture en hiver permet d'éviter les périodes de sécheresse pendant les étapes d'installation et de croissance de la plante (Figure 30).

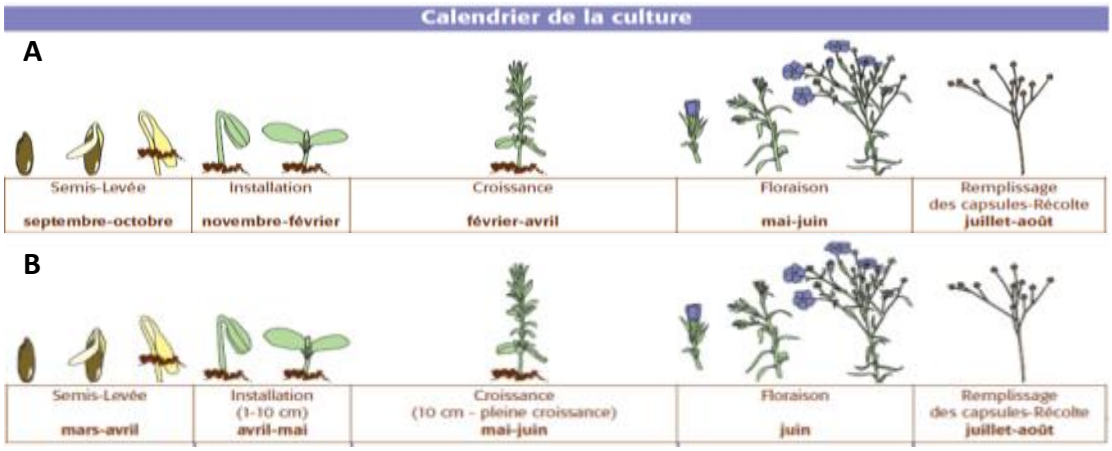


Figure 30 : Calendrier de culture des variétés d'hiver (A) et de printemps (B). Adapté de http://www.cetiom.fr/.

La répartition géographique de la culture du lin à fibres est restreinte à quelques pays, dont les principaux producteurs sont les pays de l'Europe, à leur tête la France. Les variétés à graines cultivées pour leurs graines oléagineuses nécessitent moins d'humidité et elles occupent davantage de surfaces de culture ; le Canada en est le premier producteur mondial devant la Russie et la Chine (FAOSTAT, 2013).

Chaque année, de nouvelles variétés sont inscrites dans le catalogue européen des espèces et des grandes cultures. À ce jour, 284 variétés de lin ont été recensées dans ce catalogue, mais uniquement 189 sont encore cultivées et commercialisées suivant les normes européennes (http://www.gnis.fr/index/action/page/id/256/title/Catalogues_europeens). Parmi ces 189 variétés, 69 sont retrouvées dans le catalogue officiel français GEVES (Groupe d' Etude et de contrôle des VariétésEt des Semences)(http://cat.geves.info/Page/ListeNationale). Ces catalogues, français ou européens évoluent tous les ans en incluant de nouvelles variétés et/ou en retirant d'autres du marché en fonction des différents critères de sélection. En effet, la quantité des fibres produites n'est pas le seul critère et la résistance aux maladies, la précocité des variétés et la qualité des fibres produites sont également prises en compte.

En France, la création des variétés et la production de semences du lin sont assurées principalement par les coopératives de teillage, ce qui est dû au fait que les graines de lin sont des coproduits, qui sont récupérés lors de l'extraction des fibres, ce qui est valable au moins pour les variétés à fibres. La plus ancienne variété de lin à fibres isolée a été « Frison à fleur blanche» en 1816 par les russes. Ensuite les programmes de création variétale ont été mis en place dans les années 1920 par les irlandais et les néerlandais pour les variétés de lin à fibre. Un

an après, ces derniers ont réussi à créer une variété qu'ils ont baptisée « Concurrent ». Celle-ci a été commercialisée jusqu'en 1952. Durant cette même année en France, c'est la variété néerlandaise Weira qui été cultivée grâce à son bon rendement. D'autres variétés néerlandaises ont vu le jour, telle que Hera (1970) et Natasja (1978) avant la rentrée dans la course de la sélection variétale des français, qui produirent la variété « Ariane » en 1978. Outre sa résistance à la fusariose, cette variété se caractérisait par son rendement 20% supérieur aux autres variétés disponiblwes à l'époque. Grâce à ses qualités, Ariane a été cultivée dans toute l'Europe de l'ouest pendant près de 20 ans.

Les variétés « Belinka » (1982) et « Hermès » (1992) ont été sélectionnées pour la qualité de fibres et pour le rendement en fibres et/ou en graines. La variété « Diane » (1995) est caractérisée par sa richesse en fibres et elle est également la première variété résistante à la maladie de la Brûlure. La première variété de lin fibre hiver (Adelie), a été sélectionnée en 2000, suivi par la variété « Drakkar » (2003). Ce processus continue et la sélection variétale vise actuellement la production de variétés de plus en plus résistantes aux différentes maladies, avec des rendements et des qualités optimisés.

En ce qui concerne le lin à graines, la sélection et la création variétale se sont mises en place au moins 30 ans après celle du lin à fibres. En effet, la première variété nommée « Ocean » avec une production satisfaisante a été obtenue en 1976. La variété « Oliver » a été la première variété oléagineuse de type hiver (Doré et Varoquaux, 2006).

Objectifs de la thèse

Mes travaux de thèse visent à approfondir nos connaissances sur certains facteurs pouvant impacter/moduler la structure de la paroi des fibres cellulosiques de lin. Une meilleure compréhension du lien potentiel entre la structure de la paroi et les performances des fibres devrait permettre de mieux maîtriser la variabilité de la qualité associée à cette ressource naturelle d'intérêt économique. Deux sources potentielles de variabilité ont été explorées – la première (variabilité génotypique) était étudiée dans le cadre du projet européen KBBE « FIBRAGEN », la deuxième (stress hydrique) dans le contexte du projet ANR « NoStressWall ».

Variabilité génotypique (effet variété)

Ce premier objectif s'est réalisé dans le cadre du projet KBBE FIBRAGEN (Flax for Improved Biomaterials through Applied Genomics) qui vise à identifier et/ou sélectionner les variétés de lin les mieux adaptées pour leurs utilisations dans la fabrication de biomatériaux. Pour cela, l'identification des marqueurs pour des caractères liés aux propriétés des fibres de lin représente une étape essentielle.

Dans un premier temps, des partenaires du projet FIBRAGEN ont évalué plusieurs propriétés moléculaires et physiques, y compris la composition chimique de différents germoplasmes de lin. Malgré une variabilité génétique relativement faible (Doré and Varoquaux, 2006), des études transcriptomiques précédentes réalisées au sein de notre laboratoire avaient mis en évidence des expressions différentielles entre les variétés Drakkar et Belinka de plusieurs gènes en rapport avec la biosynthèse et le développement de la paroi cellulaire (Fenart et *al.*, 2010). Ainsi un premier objectif de ma thèse a été de réaliser une étude comparative de transcriptomique entre plusieurs variétés de lin pour lesquelles le consortium FIBRAGEN possèdait des données mécaniques/chimiques. En parallèle à cette étude transcriptomique, j'ai également initié une approche de protéomique. Dans un premier temps nous avons identifié des protéines dans 4 organes/tissus végétatifs afin d'établir un « protéome atlas » du lin. Ensuite nous avons entrepris une étude de protéomique comparative entre une variété de lin à fibres et une variété à graines.

L'impact du stress hydrique sur la paroi cellulaire du lin

Comme indiqué ci-dessus, le stress hydrique impact la paroi cellulaire des végétaux, mais son effet est tissu- et espèce-dépendent. Chez le lin, l'impact de ce stress est très peu connu en dehors d'une étude sur l'état physiologique de la plante (Guo et *al.*, 2012) et une étude transcriptomique récente apparue pendant mes travaux de thèse (Dash et *al.*, 2014). D'un point de vue de l'impact du stress hydrique sur le développement des fibres il n'y avait qu'une seule publication (Chemikosova et *al.*, 2006). Cependant, les liniculteurs décrivent le syndrome des « fibres plates » qui survient lors des saisons sèches. Il semblerait que le stress hydrique altère la biosynthèse de la paroi secondaire des fibres produisant des fibres « vides », facilement aplaties, et donc de très mauvaise qualité.

Le deuxième objectif majeur de mes travaux de thèse était donc de réaliser une étude transcriptomique globale pour mieux comprendre l'impact d'un stress hydrique sur le lin avec une attention particulière sur le métabolisme pariétal. Ces travaux étaient réalisés dans le cadre du projet ANR NoStressWall (pour NOvel information on the effect of drought STRESS on the cell WALL). Ce projet a été pensé dans le but global d'élargir les connaissances sur l'impact du stress hydrique sur la physiologie de la plante en général, et sur la qualité et la dynamique de la paroi cellulaire en particulier, chez deux espèces d'un grand intérêt économique, le lin et le Brachypodium (*Brachypodium distachyon*), espèce modèle utilisée pour étudier la biologie des graminées. Pour atteindre ce but, le projet a été organisé sous forme de trois volets principaux :

1- Générer et intégrer les données des analyses transcriptomiques, protéomiques et métabolomiques, ainsi que les analyses de structure des parois cellulaires et les modifications induites par le stress hydrique.

2- Utiliser la génétique inverse pour identifier des mutants spécifiques en lien avec la paroi cellulaire, dans des collections de mutants chimiques du lin et de Brachypodium.

3- Caractériser de manière fonctionnelle ces mutants spécifiques.

Mon rôle dans le projet NSW, s'est inscrit dans le cadre du premier objectif de ce projet et j'étais responsable de la production du matériel végétal, des mesures physiologiques, des analyses histologiques et des analyses transcriptomiques chez le lin.

Chapitre 2 : Article 1

Chapitre 2 : Article 1

Avant-propos

Les fibres de lin possèdent une architecture pariétale « atypique », principalement à cause de leur faible teneur en lignine. Pour mieux comprendre les mécanismes associés à la formation et au développement de ce type de paroi cellulaire, notre laboratoire s'est investi dans le développement de plusieurs outils d'analyse génomique et fonctionnelle.

Un premier travail de fabrication d'un outil d'analyse transcriptomique chez le lin, a été réalisé en 2010 dans le cadre du projet ANR GENOLIN (Fenart et *al.*, 2010) par la fabrication d'une puce à ADN spécifique du lin. En effet, à partir de 9 échantillons d'ARN provenant de différents tissus et récoltés à des différents stades de développement, appartenant à deux variétés distinctes, une collection de 1 066 481 ESTs a été séquencée. Ces séquences ont été assemblées en 59 000 unigenes dont 48 000 ont servi dans le design des sondes. Chaque unigene a été représenté par 8 sondes non-chevauchantes avec une taille comprise entre 25 et 30 nucléotides sur des puces à ADN de type Nimbelgen 385K uniplex. Ces puces à ADN (appelées puces à 25-mers) ont été utilisées avec succès dans plusieurs études chez le lin (Fénart et *al.*, 2010 ; Huis *et al.*, 2012; Day *et al.*, submitted).

Pour améliorer cet outil d'analyse transcriptomique, nous avons testé les performances des hybridations sur des puces ADN construites à partir des sondes longues de 60 mers. Une comparaison entre les résultats fournis par les microarrays 25-mers et 60-mers a été réalisée, en utilisant des conditions identiques de préparation des cibles, marquage, hybridation, analyse d'images et traitement des données. La comparaison a porté sur les paramètres d'hybridation, la qualité, la précision, ainsi que sur les profils d'expression des gènes. Les deux types de microarrays ont donné des résultats reproductibles, précis et cohérents pour l'interprétation des profils d'expression des gènes, néanmoins les puces 60-mers ont présenté une plus grande efficacité d'hybridation et une plus importante sensibilité de détection des unigènes et ont été recommandées pour les approches à venir.

Ce travail auquel j'ai pu apporter ma contribution (validation des résultats par qRT-PCR) a fait l'objet d'une publication (Fenart *et al.*, 2012) et constitue le premier chapitre de ma thèse.

Fenart et al. BMC Research Notes 2013, 6:43
http://www.biomedcentral.com/1756-0500/6/43

BMC Research Notes

RESEARCH ARTICLE Open Access

Intra-platform comparison of 25-mer and 60-mer oligonucleotide Nimblegen DNA microarrays

Stephane Fenart[1,2], Malika Chabi[1], Sophie Gallina[3], Rudy Huis[1], Godfrey Neutelings[1], Nathalie Riviere[4], Brigitte Thomasset[5], Simon Hawkins[1] and Anca Lucau-Danila[1*]

Abstract

Background: We performed a Nimblegen intra-platform microarray comparison by assessing two categories of flax target probes (short 25-mers oligonucleotides and long 60-mers oligonucleotides) in identical conditions of target production, design, labelling, hybridization, image analyses, and data filtering. We compared technical parameters of array hybridizations, precision and accuracy as well as specific gene expression profiles.

Results: Comparison of the hybridization quality, precision and accuracy of expression measurements, as well as an interpretation of differential gene expression in flax tissues were performed. Both array types yielded reproducible, accurate and comparable data that are coherent for expression measurements and identification of differentially expressed genes. 60-mers arrays gave higher hybridization efficiencies and therefore were more sensitive allowing the detection of a higher number of unigenes involved in the same biological process and/or belonging to the same multigene family.

Conclusion: The two flax arrays provide a good resolution of expressed functions; however the 60-mers arrays are more sensitive and provide a more in-depth coverage of candidate genes potentially involved in different biological processes.

Keywords: Nimblegen, DNA arrays, Gene expression

Background

Technologies for performing genome-wide expression analyses have rapidly multiplied in recent years and different cross-platform studies have focused on target type, target production and design, labelling or hybridization protocols [1-4] as well as mathematical approaches [5-8]. Despite the tremendous progress in Next Generation Sequencing (NGS) technology and the increasing use of RNAseq approaches, different microarray platforms continue to generate large amounts of high quality expression data for a wide range of animal and plant species and are extensively applied in medical decision-making research. In general, arrays can contain oligonucleotide probes of 25, 30, 40, 50, 60, 65, 70–80 bases in length [9]. For example, *in situ* synthesized arrays for human, mouse, yeast, rat, Arabidopsis, Drosophila, *C. elegans*, zebrafish and other species, can use 25-mers probes (Affimetrix platform), 50-mers probes (Illumina platform), 60-mers probes (Agilent platform), and 50-75-mers probes (Nimblegen platform) [10].

Since 1999 Roche NimbleGen provides high-density arrays for advanced gene expression analysis, synthesized by digital light processing and rapid, high-yield photochemistry using Maskless Array Synthesis (MAS) technology. These arrays present the advantage of a custom design allowing specification of the regions of interest or the targeted probes for a tailored array solution in any organism (http://www.nimblegen.com). Nimblegen 25-mers arrays were successfully used in gene expression analyses in bacteria [11], yeast [12], and human [13]. Nimblegen 36-mers arrays were used in rice [14], and 50- to 75-mers were also used in bacteria [15,16], zebrafish [17], *Mus musculus* [18], human [19], alga [20], poplar [21], Arabidopsis [22], rice [23] and many other species.

We have recently developed a flax high-density oligo-microarray platform using Nimblegen technology [24]. Flax (*Linum usitatissimum* L.) is one of mankind's oldest cultivated plants and is grown for both its cellulose-rich fibers and for its seeds rich in alpha linolenic acid (ALA, C18:3)

* Correspondence: Anca.Lucau@univ-lille1.fr
[1]Université Lille Nord de France, Lille 1, UMR INRA 1281, SADV, F- 59650 Villeneuve d'Ascq cedex, France
Full list of author information is available at the end of the article

[25]. Our current flax platform is based upon a uniplex 385K system consisting of 8 short (25-mers) oligonucleotides per unigene and a total of 48,021 unigenes per slide. This platform represents the first high-density flax microarray system and is currently providing extremely useful biological information [24,26]. Nevertheless, we wanted to know whether adifferent design based upon long oligonucleotides (60-mers) would improve the performances of our gene expression analyses and consequently increase the yield of meaningful biological information. In order to do this we compared two categories of flax target probes: short (25-mers) oligonucleotides and long (60-mers) oligonucleotides in identical conditions of target production, design, labelling, hybridization, image analyses, and data filtering. This comparison was realized with two different flax samples and each RNA sample was used for the two categories of arrays. Experiments were realized in order to discriminate specific gene expression profiles of two different flax tissues, and results were cross-validated using an independent method (qRT-PCR). In this paper technical parameters of array hybridizations are compared and their relevance for the generation of biologically useful information are discussed.

Results and discussion

The Nimblegen array system is based upon the hybridization of a single labelled sample (derived from RNA), followed by one-channel detection. The intensity of the hybridization signal is then used to determine target concentration. We used two contrasted samples, one from flax inner stem tissue and the other from the outer stem tissue (Additional file 1). These two tissues are easily separated without cross-contamination as previously demonstrated [27].

Three independent hybridizations were performed for each sample using the two array types (25-mers and 60-mers). After verifying the hybridization quality for all experiments the results obtained using the 2 arrays were compared by evaluating the precision and accuracy of expression measurements; a sub-list of 9 genes was used for the comparison of microarray data and qRT-PCR data.

Hybridization quality

Our results (GSE37980) showed that all probes present on the two types of array were capable of hybridizing successfully (signal>background). The sensitivity of both array types was demonstrated by the wide signal dynamic range obtained (log2 values of 6 to 15 for 25-mers arrays and 4 to 15 for 60-mers arrays). Hybridization quality was verified using experimental metrics reports (NimbleScan v2.5) as recommended by Roche/Nimblegen. This program generates summary statistics (interquartile density, signal range, uniformity mean, uniformity CV (coefficient of variation), number of empty features on the array, mean empty, the number of random control features present on the array, mean random). All metrics were within the recommended value range indicating that hybridization quality was satisfactory for all experiments and any potential artifacts during hybridization were registered for both array types and for all samples. Raw expression data on all flax hybridization experiments were normalized through RMA (Robust Multi-array Average) algorithms included in the NimbleScan software and 46,589 common targets on both array types were taken into account for further comparisons.

Comparison of the precision of expression measurements

Precision (also called reproducibility or repeatability) represents the degree to which repeated measurements of the same sample hybridization will show the same or similar results [28]. In order to compare the precision of the measurements derived from each array type we used the following criteria: 1) the distribution of inter-slide variation measures; 2) inter-slide correlation of expression profiles.

Comparison of the hybridization signal intensities for all experiments in the two array types (Figure 1A), and of expression measurements between each pair-wise combination of inner *vs.* outer tissues (Figures 1B and 1C), show that the data from both array types is highly reproducible. Nevertheless, the 60-mers arrays presented globally higher signal intensities and lower variation measures compared to 25-mers array type as shown by the coefficient of variation (CV) and standard deviation (SD) values for both array types (Figure 1C).

Variations in signal intensities of probes corresponding to different regions of the same mRNA target have previously been observed [29,30], and highly sequence dependent [31-33]. The hybridization efficiency between a probe and its targets is determined by the balance between the binding strength of the probe-target duplex and the formation of probe-probe dimers and secondary structures in either probes or targets [34,35]. The duplex melting temperature is generally considered as one of the most popular measures in the evaluation of microarray probes. It gives the temperature at which half of all probes form a duplex with their target while the other half are unbound, assuming a simple two state transition [34]. General thermodynamic models of probe-target hybridization have also recently been used to compare 25- and 45 to 75-mers tiling Nimblegen human arrays in order to calculate the thermodynamic parameters and model choice [36]. Differences in the probe sequence seems to explain the specific variations of microarray signal intensities as the melting temperature is different for each probe set. In our experiments the 25-mers arrays were hybridized at 38°C, and the 60-mers arrays at 42°C, conforming to Nimblegen recommendations.

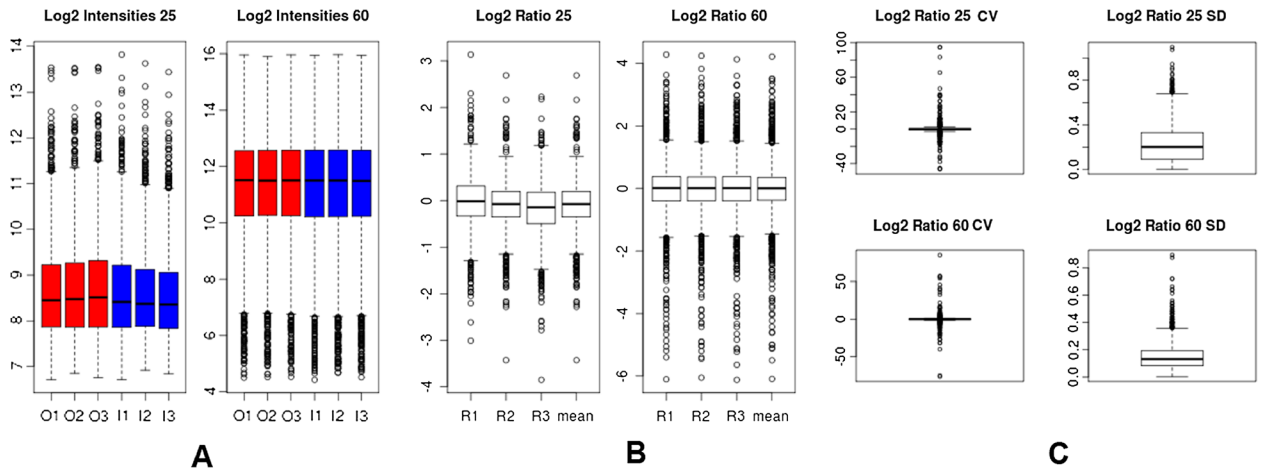


Figure 1 Boxplots of the distributions of inter-slide hybridization signal intensities and log2 ratios. A: hybridization signal intensities for all experiments in the two array types (I-inner tissues; O-outer tissues); **B**: expression measurements between each pair-wise combination of inner *vs.* outer tissues (R1, R2, R3 - replicates 1, 2, 3); **C**: distribution of the average inter-slides variations measured as the coefficient of variation (CV) or standard deviation (SD) of log2 ratios for 25-mers (25) and 60-mers flax Nimblegen arrays (60). Data from both array types are highly reproducible, nevertheless, the 60-mers arrays presented globally higher signal intensities and lower variation measures compared to 25-mers array type.

Inter-array comparison of expression profiles (Figure 2) showed that a strong correlation exists between the two arrays. Only a few exceptions were detected for up-regulated profiles (log2 ratio >1) on 60-mers arrays and for down-regulated profiles (log2 ratio <−1) on 25-mers arrays.

Comparison of the accuracy of expression measurements

Accuracy is defined as the degree of conformity of the measured quantity to its actual (true) value [28,37]. To evaluate this parameter we used: 1) the number of targets showing differences in expression values between each pair-wise combination of replica slides, and 2) the concordance between relative expression values obtained on arrays with those obtained by qRT-PCR for a subset of 9 genes. The number of targets showing significant differences in expression values between each pair-wise combination of inner *vs.* outer tissues (Figure 3) are given for three different thresholds: -1< log2 ratio >1 (Figures 3A and 3B), -2< log 2ratio >2 (Figures 3C and 3D), and −3< log2 ratio >3 (Figures 3E and 3F). Significantly expressed targets were detected by both arrays at all threshold values used.

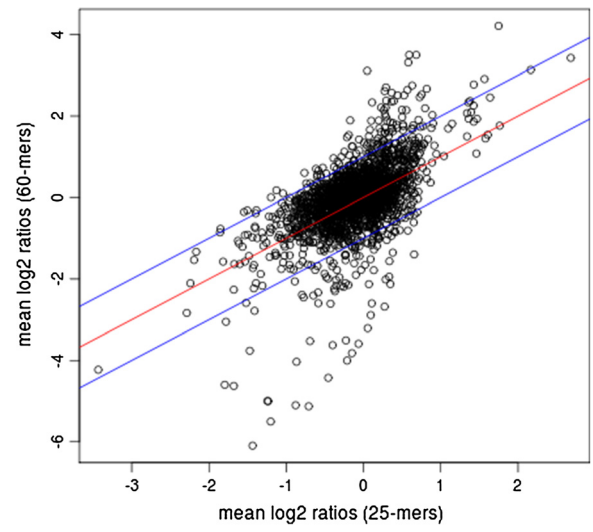
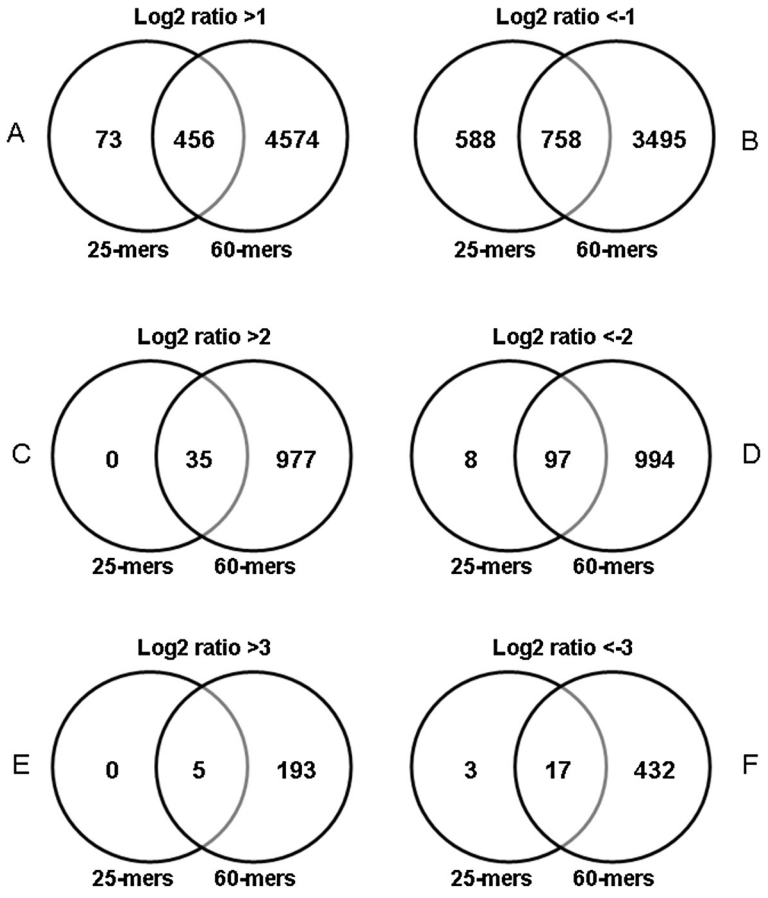


Figure 2 The inter-array type's correlation of expression profiles. Scatter plot was realized using means of log2 ratio of probe intensities in all replicates and shows a strong correlation between the two array types.

The 60-mers arrays showed a much greater sensibility in significant target expression and the number of up- or down-regulated targets was between 4x and 39x more important than with 25-mers arrays depending upon the threshold value used (Additional file 2). This sensibility of detection could be related to the higher intensities of signals in 60-mers arrays (Figure 1A) and the hybridization efficiency that is sequence-dependent (see Comparison of the precision of expression measurements).

The 25-mers arrays seemed to produce accurate measurements as a high number of identified targets were confirmed by 60-mers arrays (between 56 and 100% of significant expression values obtained in 25-mers arrays were detected in 60-mers arrays at the same threshold). Targets that were specifically detected using the 25- or 60-mers arrays generally presented relatively similar log2 ratio values even though they did not necessarily occur within the same threshold range. Only three targets (out of the 46,589 unigenes targeted) showed discordant expression values being significantly up-/down-regulated in one array type as compared to the other (Figure 3G).

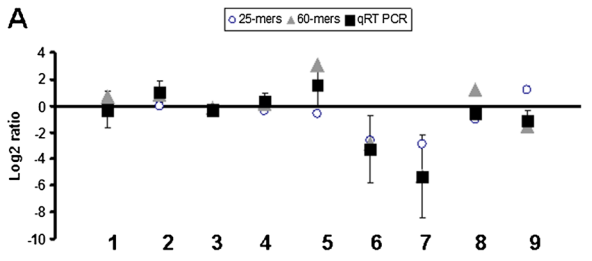


G

	25-mers	60-mers	qRT PCR
GENOLIN_C50701 (Similar to nucleolar complex 2)	-1.06	1.21	-0.58
GENOLIN_C29324 (Cupin)	1.20	-1.54	-1.11
GENOLIN_C53518 (lcl genolin_c53518 105nt)	1.10	-1.06	ND

Figure 3 Venn diagrams representing the number of targets showing significant differences in expression values between each pair-wise combination of inner *vs.* outer tissues represented for three different thresholds: -1< log2 ratio >1 (A and B), -2< log2 ratio >2 (C and D) and −3< log2 ratio >3 (E and F). G: unigenes showing significantly opposed expression values on the two arrays.

In an attempt to understand this discordance we measured the expression levels of 2 of these genes (C29324: up-regulated on the 25-mers array and down-regulated on the 60-mers array, and C50701: down-regulated on the 25-mers array and up-regulated on the 60-mers array) by qRT-PCR. Our results (Figures 3G and 4) show that qRT-PCR measurements indicate that C50701 is under-expressed in stem inner tissues in agreement with the 25-mers value but in disagreement with the 60-mers value. In contrast the qRT-PCR results indicated that C29324 was up-regulated in inner stem tissues in agreement with the 60-mers results, but not the 25-mers results. One possible explanation for the observed differences could be the existence of alternative splicing variants. Comparison of splicing predictions between genomic (www.phytozome.org) and EST (http://urgi.versailles.inra.fr/Species/Flax/Download-sequences) databases suggest that splice variants could exist for C50701. Nevertheless alignment of 25-mer, 60-mer probes and qRT-PCR primers (Additional file 3) showed that none would be capable of distinguishing the different potential splice variants. Further investigation using microarrays specifically designed as tiling or splice junction arrays could provide further information. Comparison of the genomic sequence, probes and primers for the C29324 unigene (Additional file 3) also provided a possible explanation for the observed discordance between 25-mers and 60-mers values. The genomic sequence (*Lus*10011816) and the C29324 unigene show good alignment in the central region of the unigene but are not aligned at both extremities suggesting that the C29324 contig is not correctly assembled. Both 60-mers probes and qRT-PCR primers target the central region (correct) of the



B

	Gene ID	group	25-mers	60-mers	qRT PCR
1	C24118	A	0,30	0,69	-0,29
2	C3323	A	-0,01	0,81	1,07
3	C21991	A	-0,24	-0,16	-0,34
4	C2533	A	-0,39	0,17	0,31
5	C57711	B	-0,59	3,10	1,56
6	C602	C	-2,58	-2,83	-3,25
7	C822	C	-2,91	-5,25	-5,33
8	C50701	D	-1,06	1,21	-0,58
9	C29324	D	1,20	-1,54	-1,11

Figure 4 Correlations of 25-mers and 60-mers array data with sample-matched qRT-PCR data. Standard deviation of qRT-PCR data were represented as bars. Tested genes were: showing **A**) no significant (1<log2 ratio>−1) expression values on both arrays (C24118, C3323, C21991, C2533), **B**) significant expression value (log2 ratio>1 or log2 ratio<−1) on one array, but not the other (C57711), C) significant expression values on both arrays (C602, C822) and D) significant but opposed expression values (C50701, C29324). For 5 out of the 9 tested genes, qRT-PCR determined expression values were not significantly different from those determined by both flax arrays. For three other genes, qRT-PCR expression values were significantly different from 25-mers data, but not from 60-mers data. The qRT-PCR expression value of one gene was significantly different from 60-mers data but not from 25-mers data.

gene/transcript whereas the 25-mers target the extremities that are problematic. In conclusion, only 3 out of the 46,589 unigenes (0.0064%) targeted by the 25-mers-arrays and the 60-mers arrays showed discordant significant expressions thereby confirming the overall conformity of probe design for the 2 array platforms. Verification by qRT-PCR and analyses of sequence data for 2 unigenes showing discordant expression values indicated possible explanations for the observed contradictions.

We have previously evaluated the accuracy of the 25-mers platform by qRT-PCR cross-validation using 9 genes [24]. We therefore adopted the same approach to validate the 60-mer platform using a subset of 9 genes showing i) no significant (1<log2 ratio>−1) expression values on both arrays (C24118, C3323, C21991, C2533), ii) significant expression value (log2 ratio>1 or log2 ratio<−1) on one array, but not the other (C57711), iii) significant expression values on both arrays (C602, C822) and iv) significant but opposed expression values (C50701, C29324) (Figure 4). For 5 out of the 9 tested genes (C24118, C3323, C21991, C602, and C2533), qRT-PCR determined expression values were not significantly different from those determined by both flax arrays. For three other genes (C822, C57711, C29324), qRT-PCR expression values were significantly different from 25-mers data, but not from 60-mers data suggesting that the 60-mers array performs better than the 25-mers. The qRT-PCR expression value of only one gene (C50701) was significantly different from 60-mers data but not from 25-mers data. Although examination of genomic and transcript data suggested that different splice variants might exist for this gene as indicated above, neither the 25-mers, nor the 60-mers probes would distinguish the different forms and it is therefore difficult to explain why the 25-mers apparently give a more accurate measure of expression levels for this unigene. The design and use of different qRT-PCR primers and further sequence analyses would enable to clarify this point.

The correlation coefficient was calculated separately between the 60-mers results and the qRT-PCR results for 6 selected genes. We deliberately excluded the unigenes C50701 and C29324 that gave discordant results between the platforms and/or the qRT-PCR data, probably resulting from assembly problems as indicated above. We also decided to exclude the unigene C57711 because of the discordant 25-mers and 60-mers values, but also because we were unable to identify the corresponding genomic sequence. The obtained value ($r = 0.9832$) indicated a highly statistically significant correlation (Figure 5). A similar calculation for the 25-mers platform (Figure 5) also gave a highly significant (but lower) r value (0.9414). Taken together these data indicate a good accuracy for both array types and are in agreement with other similar studies demonstrating that experimental errors were not a significant source of unwanted variability in expression profiling obtained by Affymetrix U74Av2 arrays transcriptome experiments [1] or custom made microarrays [3].

Differential gene expression in flax tissues

Our results showed that the 60-mers array detected a higher number of unigenes differentially expressed between the two flax samples and was therefore more sensitive than the 25-mers array. However, we wondered whether the increased sensitivity also represented an augmentation in the biologically-relevant information. As a first step to answering this question we functionally classified genes showing significant differential expression on the two arrays using GO (Gene Ontology in biological process category) annotations based on blast results and GOA and TAIR gene cross-referenced files [24,26]. Functional categories of up- and down-regulated genes in inner *vs.* outer tissues at −1<log2 ratio<1 in the two array types are represented in Figure 6 and Additional file 2. Even if the total number of significantly expressed genes is very different in the two array types

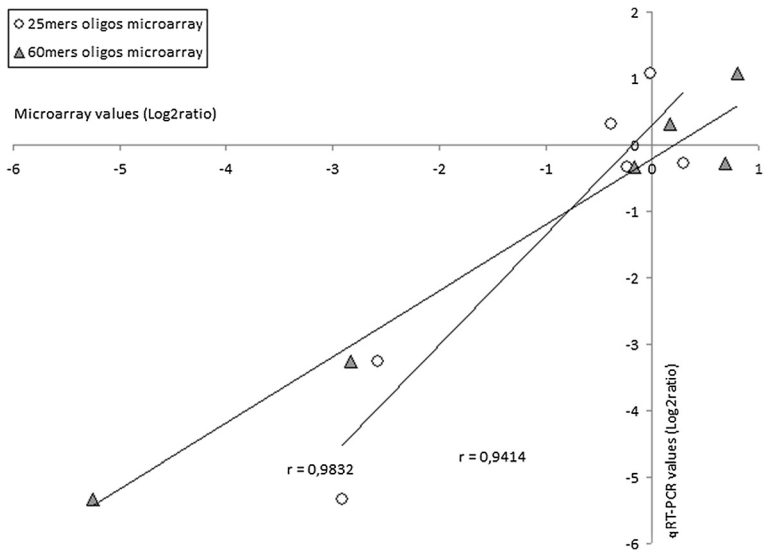


Figure 5 Correlations between qRT-PCR and microarray results. Statistically significant correlations were obtained for both 60-mers arrays (r = 0.9832), and 25-mers arrays (r = 0.9414).

(number up-regulated genes: 529 for 25-mers, 5,030 for 60-mers; number down-regulated genes: 1,346 for 25-mers; 4,253 for 60-mers), the percentages of annotated genes involved in different functional groups are very similar for the 2 arrays. For example, 7.96%, (9 genes) and 7.84%, (43 genes) of all genes significantly more expressed in stem inner tissues were assigned to the class 'secondary metabolites' in the 25-mers and 60-mers data sets, respectively. Similarly, 18.58%, (102 genes) and 18.8% (277 genes) were assigned to the class 'response to stress', and transport 15.93%, (18 genes) and 13.32%, 87 genes) were assigned to the class 'transport' in the 25-mers and 60-mers data sets, respectively. The high similarity between functional class percentage values was also observed for genes showing a significant higher expression in stem outer tissues. For example, 12.56% (86 genes) and 11.41% (170 genes) were assigned to the class 'photosynthesis', and 19.05% (132 genes) and 18.35% (275 genes) were assigned to the class 'response to stress' in the 25-mers and 60-mers data sets, respectively. Generally, these observations are in close agreement with the known physiological roles of these two different tissues [26,38,39] and confirm the biological consistency of data reported by both array types. Taken together, these observations suggest that both arrays are able to provide a biologically-coherent global view of the flax stem transcriptome.

In order to better assess whether the increase in the number of significantly expressed genes detected by the 60-mers arrays as compared to the 25-mers arrays represented biologically relevant information we decided to focus on genes encoding enzymes responsible for the biosynthesis of lignin monomers (monolignols) and/or their oxidation (laccases). Our results (Figure 7A) show that the 25-mers array detected 3 significantly expressed unigenes corresponding to 3 multigenes families (*Phenylalanine Ammonia Lyase*: *PAL*, *4-Coumarate Ligase*: *4CL*, *Cinnamyl Alcohol Dehydrogenase*: *CAD*) encoding enzymes involved in monolignol biosynthesis. When the 60-mers array was used additional significantly expressed unigenes corresponding to each of these 3 multigene families were detected (5 *PAL* unigenes, 2 *4CL* unigenes and 2 *CAD* unigenes). Nevertheless, the significantly expressed unigene detected by the 25-mers array also corresponded to the most significantly expressed unigene detected by the 60-mers array. In addition, the 60-mers array, but not the 25-mers array, also detected significantly expressed unigenes corresponding to 3 further multigene families encoding enzymes involved in monolignol biosynthesis (*Cinnamate 4-Hydroxylase*: *C4H*, *Caffeic Acid O-methyltransferase: COMT*, *Cinnamoyl CoenzymeA Reducatse*: *CCR*). Although the *C4H* unigene expression level determined by the 25-mer array was just below the cut-off value (0.98), the *COMT* unigene expression level was considerably inferior (0.37). Similarly all *CCR* unigene expression levels determined by the 25-mers array were well below the threshold value. Similar observations could be made for those unigenes encoding enzymes (laccases) potentially involved in the oxidative polymerisation step of the lignification process. Interestingly however, one laccase unigene (C37539) showed a significant expression with the 25-mers array but not the 60-mers array.

In order to understand the possible reason for the higher sensitivity of the 60-mers arrays, we focused on *PAL* unigenes (Figure 7B) and examined signal

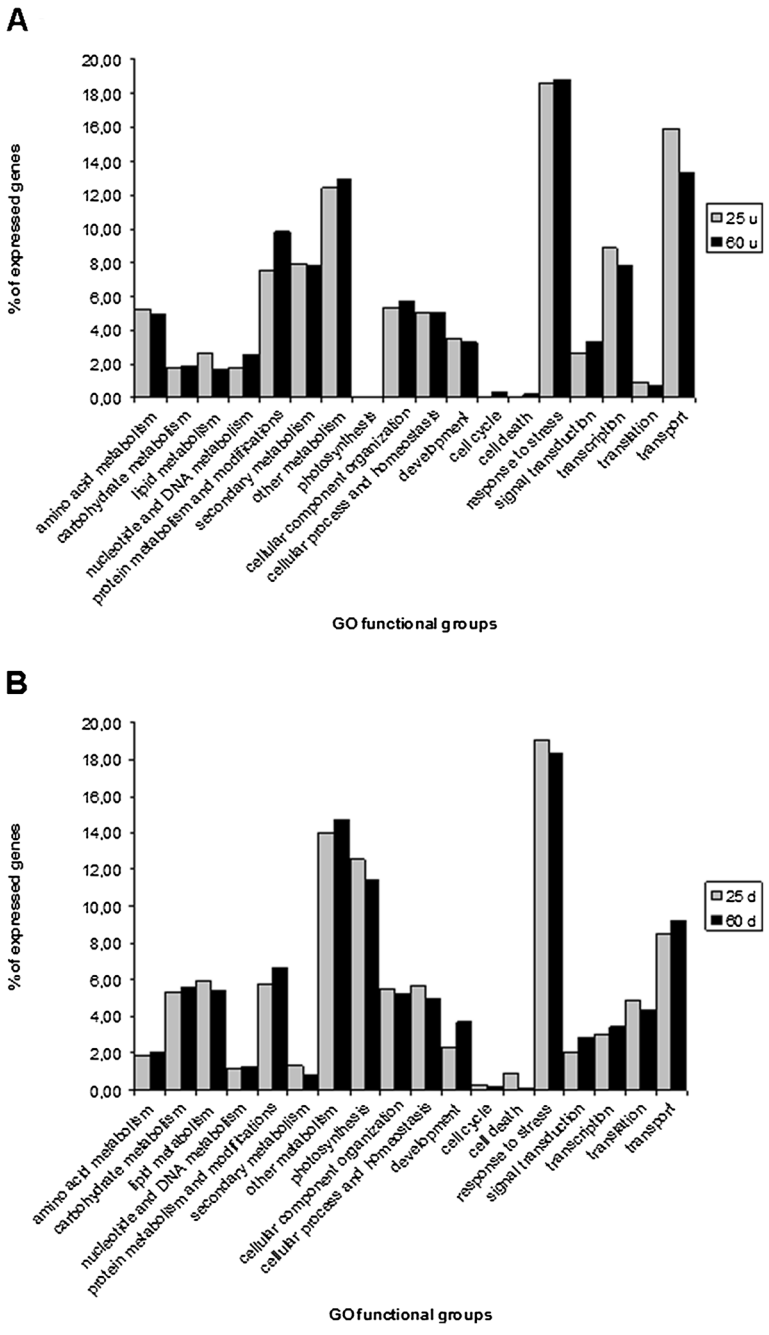


Figure 6 GO functional classification of up-regulated genes (log2 ratio>1) (A) and down-regulated genes (log2 ratio<−1) (B) in inner *vs.* outer tissues identified by 25- and 60-mers flax Nimblegen arrays.

intensities, unigene lengths, probe Tm, as well as the probe position and coverage of the EST (Additional file 4). Signal intensities were consistently higher on 60-mers arrays, presumably since the probe Tm average was higher, resulting in lower background. No relation was found between unigene length and array sensitivity. Both arrays covered similar unigene region lengths, generally the 60-mers probes cover 240 bp (4 duplicate probes per unigene) and the 25-mers probes cover 200 pb (8 25-bp probes per unigene) (Additional file 4).

All these observations support the hypothesis that hybridization efficiency depends on probe thermodynamic parameters as previously suggested [33]. Similar results were found with the Agilent platform [40] when 25- and 60-mers arrays were compared. Agilent 60-mers arrays tended to have higher sensitivity, with an average lower

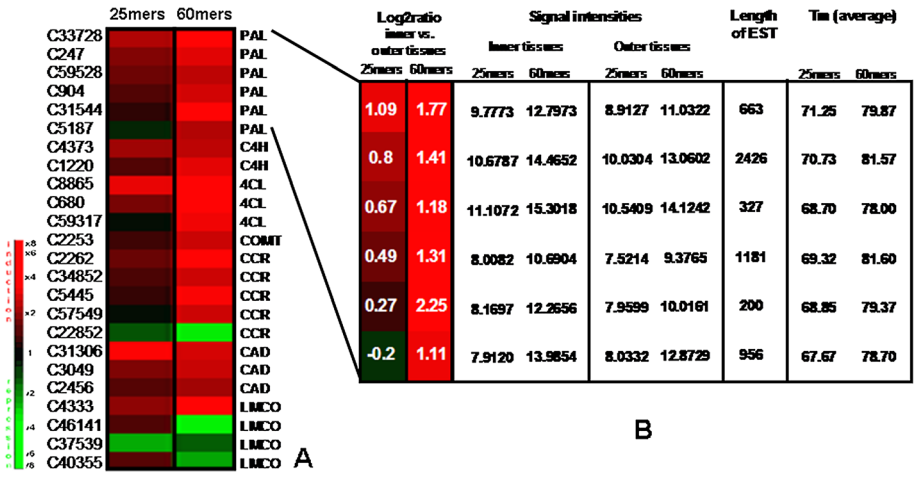


Figure 7 A: Cluster representing expression profiles of unigenes involved in phenylpropanoid metabolism. First and second columns represent log2 ratio of inner *vs.* outer tissues in 25- and 60-mers arrays respectively. **B**: Cluster representing expression profiles of *PAL* (*Phenylalanine Ammonia Lyase*) unigenes associated with signal intensities, unigene length and average melting temperartures (Tm) for probes.

detection limit as compared to 25-mers. In contrast, reproducibility of log2 ratio values, system noise and accuracies of log2 ratio determination were comparable between these two microarray types. Similarly, the overall biological information obtained with these 2 arrays was similar in agreement with our observations in flax stem tissues.

Conclusions

Our study compared two different flax Nimblegen high density microarray platforms based on a short- and long-oligonucleotide design. Our results showed that both array types yielded reproducible, accurate and comparable data that are coherent for expression measurements and identification of differentially expressed genes. Nevertheless, we found that the 60-mers arrays gave higher hybridization efficiencies and therefore were more sensitive allowing the detection of a higher number of unigenes involved in the same biological process and/or belonging to the same multigene family. The two flax arrays provide a good resolution of expressed functions; however the 60-mers arrays are more sensitive and provide a more in-depth coverage of candidate genes potentially involved in different biological processes.

Methods

Plant material

Linum usitatissimum L. (cv. Barbara) plants were grown in a growth chamber (light/night cycles 16h (22°C)/8h (19°C), 50% humidity and light intensity of 400 μE s-1 m-2) and harvested after nine weeks of grown. The outer fiber-bearing tissues were peeled off and inner tissues (xylem) from a 15 cm long stem section were cut into short fragments before both tissues were frozen in liquid nitrogen as previously described [24,26].

RNA extraction

Total RNA was isolated from pooled flax inner- and outer-stems using the NucleoSpin® RNA Plant kit (Macherey-Nagel) following manufacturer's guidelines. To obtain sufficient amount of RNA for microarray analysis (10 μg), a minimum of three extractions with up to 150 mg of fresh tissue were necessary for each sample. To eliminate DNA contamination, on column treatment was done using the RNAse-free DNAse included in the kit. RNA integrity and concentration were evaluated with RNA StdSens Chips using the ExperionTM automated eletrophoresis system (Bio-Rad). For each sample, the three RNA extracts were pooled and final concentrations were adjusted to 1 μg/μL.

Microarray design and oligo synthesis

Two types (25-mers, 60-mers) of high-density flax microarrays based on the Nimblegen 385K design format (Nimblegen Systems, Inc., Madison, WI, USA) each containing a total of 384,168 oligonucleotides were designed. The 25-mers array utilized 8 distinct, 25 bp-long oligos for each of the 48,021 contigs and the 60-mers array utilized 4 duplicate, 60 bp-long oligonucleotides for 46,589 contigs. Microarray contigs were selected from a collection of 59,000 contigs obtained by assembling the GS FLX sequences [24]. The 46,589 contigs targeted by the 60-mers array were also targeted by the 25-mers array allowing direct comparison between the two designs.

cDNA synthesis, labeling and hybridization

Double-stranded cDNA (ds-cDNA) was synthesized from 10 μg of total RNA using an Invitrogen SuperScript ds-cDNA synthesis kit in the presence of 250 ng random hexamer primers. ds-cDNA was cleaned and labeled in accordance with the Nimblegen Gene Expression Analysis protocol (Nimblegen Systems, Inc., Madison, WI, USA). Briefly, ds-cDNA was incubated with 4 μg RNase A (Promega) at 37°C for 10 min and cleaned using phenol:chloroform:isoamyl alcohol, followed by ice-cold absolute ethanol precipitation. For Cy3 labeling of cDNA, the Nimblegen One-Color DNA labeling kit was used according to the manufacturer's guideline detailed in the Gene Expression Analysis protocol (Nimblegen Systems, Inc., Madison, WI, USA). One μg ds-cDNA was incubated for 10 min at 98°C with 2 OD of Cy3-9mer primer. Then, 100 pmol of deoxynucleoside triphosphates and 100U of the Klenow fragment (New England Biolabs, Ipswich, MA, USA) were added and the mix incubated at 37°C for 2h30. The reaction was stopped by adding 0.1 volume of 0.5 M EDTA, and the labeled ds-cDNA was purified by isopropanol/ethanol precipitation. Microarrays were hybridized at 38°C (25-mers arrays) and at 42°C (60-mers arrays) during 16 to 18h with 6μg of Cy3 labelled ds-cDNA in Nimblegen hybridization buffer/hybridization component A in a hybridization chamber (Hybridization System - Nimblegen Systems, Inc., Madison, WI, USA). Following hybridization, washing was performed using the Nimblegen Wash Buffer kit (Nimblegen Systems, Inc., Madison, WI, USA).

Data analysis

Slides were scanned at 5 μm/pixel resolution using an Axon GenePix 4000B scanner (Molecular Devices Corporation, Sunnyvale, CA, USA) piloted by GenePix Pro 6.0 software (Axon). Scanned images (TIFF format) were then imported into NimbleScan software (Nimblegen Systems, Inc., Madison, WI, USA) for grid alignment and expression data analyses. Expression data were normalized through quantile normalization [41] and the Robust Multichip Average (RMA) algorithm [42] included in the NimbleScan software. Identification of genes displaying a change in expression over repetitions was accomplished with a script utilizing library functions in R with a false discovery rate (FDR) of less than 5%. The SAM [43] was used to identify differentially expressed genes over different conditions. Analysis was completed with the Tree-view clustering program [44]. Functional annotation of differentially-expressed genes was based on Gene Ontology (http://www.geneontology.org/). All the microarray data have been submitted to the Gene Expression Omnibus (GEO) database [45] with the accession number is GSE37980.

Quantitative reverse transcriptase-PCR (qRT-PCR) analysis

For qRT-PCR analyses, 1 μg of total RNA was reverse-transcribed to single stranded cDNA using the IScript cDNA synthesis kit (Bio-Rad) according to the manufacturer's instructions. The qRT-PCRs were carried out in 96-wells plates with a MyIQ real time PCR detection system (Bio-Rad) using iQSYBR Green PCR Kit (Bio-Rad) in a reaction volume of 20 μL (5 μL diluted cDNAs, 10 μL of 2× SYBR Green mix and primer pairs at 0.4 μM). Aliquots from the same cDNA solutions were used with all primer sets in each experiment. All PCR reactions were performed under the following conditions: 95°C for 15 min, 40 cycles of 10 s at 95°C and 30 s at 60°C. For each primer pair, a melting curve was generated in order to confirm the specificity of the amplification. The primer sequences used for all target genes are presented in Additional file 5.

Each experiment was repeated on three biological replicates, each one represented by three technical repetitions. PCR reactions on samples lacking the cDNA template or the reverse transcriptase during the cDNA synthesis were also performed as negative controls for each primer pair. The efficiency (E) value of each reaction was between 0.85 and 1.17 with R2 values higher than 0.99.

Data were analysed using Bio-Rad iQ5 software. For each primer pair, a melting curve was generated in order to confirm the specificity of the amplification. The PCR efficiencies (E) for each reaction were between 0.85 and 1.17 with R^2 values higher than 0.99. The expression of each gene was normalized by using 2 reference genes, *ETIF1* and *ETIF4F*, shown to be expressed in a stable manner in flax stem tissues [46].

Additional files

Additional file 1: Representative image of flax stem tissues.

Additional file 2: Average of differentially expressed genes in inner *vs.* outer tissues in 25 and 60-mers arrays.

Additional file 3: Alignment of microarray probes, qRT-PCR primers and unigene sequences for C50701 and C29324.

Additional file 4: Coverage of phenylalanine ammonia lyase ESTs by 25-mers and 60-mers probes.

Additional file 5: Primer design for qRT-PCR.

Abbreviations

NGS: Next generation sequencing; RMA: Robust multi-array average; CV: Coefficient of variation; SD: Standard deviation; qRT-PCR: quantitative Real Time Polymerase Chain Reaction; GO: Gene ontology; GOA: Gene ontology annotation; TAIR: The arabidopsis information resource; PAL: Phenylalanine ammonia lyase; 4CL: 4-Coumarate ligase; CAD: Cinnamyl alcohol dehydrogenase; C4H: Cinnamate 4-Hydroxylase; COMT: Caffeic acid O-methyltransferase; CCR: Cinnamoyl coenzymeA reducatse; EST: Expressed sequence tag; GEO: Gene expression omnibus.

Competing interests

The authors declare that they have no competing interests. No funding either now or in the future, no stocks or shares was or will be perceived, no interest or financing is known from an organisation for this manuscript. No

patents relating to the content of the manuscript are known, and no non-financial competing interests are to declare in relation to this manuscript.

Authors' contributions

SF and RH realized the plant cultures, RNA extractions, and transcriptomic analyses. MC and GN performed the qRT-PCR analyses, SG executed the biocomputing analyses of the results, NR and BT contributed to the microarrays design and construction, SH participated in the writing and global interpretations of paper, AL organized the various strategies of analyses and interpretation and drafted the paper. All authors read and approved the final manuscript.

Authors' information

SF is a researcher in post doctorat position; MC is a PhD student; SG is bioinformatics ingeneer; RH is researcher; GN is assistant professor in charge of qRT-PCR platform; NR is the head of Biogemma's transcriptomics group; BT is professor; SH is professor and the Plant FibreTeam leader; AL is assistant professor in charge of transcriptomic platform.

Acknowledgements

Authors gratefully acknowledge financial support of the French Nord-Pas de Calais Region (Plant Teq 4 project), RH gratefully acknowledges the support of the French government (PhD grant) and MC gratefully acknowledges the support of the ANR (Agence National de Recherche) (PhD grant KBBE FIBRAGEN project). This study was also supported by the ANR project GENOLIN.

Author details

[1]Université Lille Nord de France, Lille 1, UMR INRA 1281, SADV, F- 59650 Villeneuve d'Ascq cedex, France. [2]Roche Diagnostics France, 2, Avenue du Vercors, 38242 Meylan, France. [3]Université Lille Nord de France, Lille 1, GEPV, Laboratoire de Génétique & Évolution des Populations végétales, CNRS UMR 8198, F- 59650 Villeneuve d'Ascq cedex, France. [4]BIOGEMMA, Z.I, du Brezet, 8 rue des Frères Lumières, 63028 Clermont-Ferrand cedex 2, France. [5]UMR CNRS 6022, GEC, Université de Technologie de Compiègne, BP 20529, 60205 Compiègne cedex, France.

Received: 19 June 2012 Accepted: 30 January 2013
Published: 4 February 2013

References

1. Bakay M, Chen YW, Borup R, Zhao P, Nagaraju K, Hoffman EP: **Sources of variability and effect of experimental approach on expression profiling data interpretation.** *BMC Bioinformatics* 2002, **3:**4.
2. Ramdas L, Coombes KR, Baggerly K, Abruzzo L, Highsmith WE, Krogmann T, Hamilton SR, Zhang W: **Sources of nonlinearity in cDNA microarray expression measurements.** *Genome Biol* 2001, **2**(11):1–7.
3. Dobbin KK, Kawasaki ES, Petersen DW, Simon RM: **Characterizing dye bias in microarray experiments.** *Bioinformatics* 2005, **17:**17.
4. Woo Y, Affourtit J, Daigle S, Viale A, Johnson K, Naggert J, Churchill G: **A comparison of cDNA, oligonucleotide, and Affymetrix GeneChip gene expression microarray platforms.** *J Biomol Tech* 2004, **5:**276–284.
5. Culhane AC, Perriere G, Higgins DG: **Cross-platform comparison and visualisation of gene expression data using co-inertia analysis.** *BMC Bioinformatics* 2003, **4:**59.
6. Bloom G, Yang IV, Boulware D, Kwong KY, Coppola D, Eschrich S, Quackenbush J, Yeatman TJ: **Multi-platform, multi-site, microarray- based human tumor classification.** *Am J Pathol* 2004, **164:**9–16.
7. Wright G, Tan B, Rosenwald A, Hurt EH, Wiestner A, Staudt LM: **A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma.** *Proc Natl Acad Sci USA* 2003, **100:**9991–9996.
8. Stevens JR, Doerge RW: **Combining Affymetrix microarray results.** *BMC Bioinformatics* 2005, **6:**57.
9. Barrett JC, Kawasaki ES: **Microarrays: the use of oligonucleotides and cDNA for the analysis of gene expression.** *Drug Discov Today* 2003, **8:**134–141.
10. *Nimblegen platform*. http://www.dnamicroarray.com/services_features.html.
11. Ulijasz AT, Andes DR, Glasner JD, Weisblum B: **Regulation of iron transport in Streptococcus pneumoniae by RitR, an orphan response regulator.** *J Bacteriol* 2004, **186:**8123–8136.
12. Morschhäuser J, Barker KS, Liu TT, BlaB-Warmuth J, Homayouni R, Rogers PD: **The transcription factor Mrr1p controls expression of the MDR1 efflux pump and mediates multidrug resistance in Candida albicans.** *PLoS Pathog* 2007, **3**(11):e164.
13. Chiu S, Kim K, Haus KA, Espinal GM, Millon LV, Warden CH: **Identification of positional candidate genes for body weight and adiposity in subcongenic mice.** *Physiol Genomics* 2007, **31**(1):75–85.
14. Stolc V, Li L, Wang X, Li X, Su N, Tongprasit W, Han B, Xue Y, Li J, Snyder M, Gerstein M, Wang J, Deng XW: **A pilot study of transcription unit analysis in rice using oligonucleotide tiling-path microarray.** *Plant Mol Biol* 2005, **59**(1):137–149.
15. Abomoelak B, Hoye EA, Chi J, Marcus SA, Laval F, Bannantine JP, Ward SK, Daffé M, Liu HD, Talaat AM: **MosR, a novel transcriptional regulator of hypoxia and virulence in Mycobacterium tuberculosis.** *J Bacteriol* 2009, **191**(19):5941–5952.
16. Subramoni S, Venturi V: **PpoR is a conserved unpaired LuxR solo of Pseudomonas putida which binds N-acyl homoserine lactones.** *BMC Microbiol* 2009, **9:**125.
17. Bai X, Kim J, Yang Z, Jurynec MJ, Akie TE, Lee J, LeBlanc J, Sessa A, Jiang H, DiBiase A, Zhou Y, Grunwald DJ, Lin S, Cantor AB, Orkin SH, Zon LI: **TIF1gamma controls erythroid cell fate by regulating transcription elongation.** *Cell* 2010, **142**(1):133–143.
18. Badders NM, Goel S, Clark RJ, Klos KS, Kim S, Bafico A, Lindvall C, Williams BO, Alexander CM: **The Wnt receptor, Lrp5, is expressed by mouse mammary stem cells and is required to maintain the basal lineage.** *PLoS One* 2009, **4**(8):e6594.
19. Borghese B, Mondon F, Noël J, Fayt I, Mignot T, Vaiman D, Chapron C: **Gene expression profile for ectopic versus eutopic endometrium provides new insights into endometriosis oncogenic potential.** *Mol Endocrinol* 2008, **22**(11):2557–2562.
20. Dittami S, Scornet D, Petit J, Ségurens B, da Silva C, Corre E, Dondrup M, Glatting K, König R, Sterck L, Rouzé P, van de Peer Y, Cock J, Boyen C, Tonon T: **Global expression analysis of the brown alga Ectocarpus siliculosus (Phaeophyceae) reveals large-scale reprogramming of the transcriptome in response to abiotic stress.** *Genome Biol* 2009, **10**(6):R66.
21. Ramírez-Carvajal GA, Morse AM, Dervinis C, Davis JM: **The cytokinin type-B response regulator PtRR13 is a negative regulator of adventitious root development in Populus.** *Plant Physiol* 2009, **150:**759–771.
22. Adamczyk BJ, Fernandez DE: **MIKC* MADS domain heterodimers are required for pollen maturation and tube growth in Arabidopsis.** *Plant Physiol* 2009, **149**(4):1713–1723.
23. Kim EH, Kim YS, Park S, Koo YJ, Choi YD, Chung Y, Lee I, Kim J: **Methyl jasmonate reduces grain yield by mediating stress signals to alter spikelet development in rice.** *Plant Physiol* 2009, **149**(4):1751–1760.
24. Fenart S, Ndong YPA, Duarte J, Rivière N, van Wuytswinkel O, Lucau A, Cariou E, Neutelings G, Gutierrez L, Chabbert B, Guillot X, Tavernier R, Hawkins S, Thomasset B: **Development and validation of a flax (Linum usitatissimum L.) gene expression oligo microarray.** *BMC Genomics* 2010, **11:**592.
25. Deyholos MK: **Bast fiber of flax (Linum usitatissimum L.): Biological foundations of its ancient and modern uses.** *Israel J Plant Sci* 2007, **54:**273–280.
26. Huis R, Morreel K, Fliniaux O, Lucau-Danila A, Fénart S, Grec S, Neutelings G, Chabbert B, Mesnard F, Boerjan W, Hawkins S: **Natural hypolignification is associated with extensive oligolignol accumulation in flax stems.** *Plant Physiol* 2012, **158**(4):1893–1915.
27. 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.
28. Draghici S, Khatri P, Eklund AC, Szallasi Z: **Reliability and reproducibility issues in DNA microarray measurements.** *Trends Genet* 2006, **22**(2):101–109.
29. Bruun GM, Wernersson R, Juncker AS, Willenbrock H, Nielsen HB: **Improving comparability between microarray probe signals by thermodynamic intensity correction.** *Nucleic Acids Res* 2007, **35**(7):e48.
30. Held GA, Grinstein G, Tu Y: **Relationship between gene expression and observed intensities in DNA microarrays-a modeling study.** *Nucleic Acids Res* 2006, **34**(9):e70.
31. Zhang L, Miles MF, Aldape KD: **A model of molecular interactions on short oligonucleotide microarrays.** *Nat Biotechnol* 2003, **21**(7):818–821.
32. Wu Z, Irizarry R, Gentleman R, Murillo FM, Spencer F: **A model based background adjustement for oligonucleotide expression arrays.** *JASA* 2004, **99:**909–917.

33. Royce TE, Rozowsky JS, Gerstein MB: **Toward a universal microarray: prediction of gene expression through nearest-neighbor probe sequence identification.** *Nucleic Acids Res* 2007, **35**(15):e99.
34. Matveeva OV, Shabalina SA, Nemtsov VA, Tsodikov AD, Gesteland RF, Atkins JF: **Thermodynamic calculations and statistical correlations for oligo-probes design.** *Nucleic Acids Res* 2003, **31**(14):4211–4217.
35. Khomyakova E, Livshits MA, Steinhauser MC, Dauphinot L, Cohen-Kaminsky S, Rossier J, Soussaline F, Potier MC: **On-chip hybridization kinetics for optimization of gene expression experiments.** *Biotechniques* 2008, **44:**109–117.
36. Mueckstein U, Leparc GG, Posekany A, Hofacker I, Kreil DP: **Hybridization thermodynamics of nimblegen microarrays.** *BMC Bioinformatics* 2010, **11:**35.
37. Moreau Y, Aerts S, De Moor B, De Strooper B, Dabrowski M: **Comparison and meta-analysis of microarray data: from the bench to the computer desk.** *Trends Genet* 2003, **19**(10):570–577.
38. Roach MJ, Deyholos MK: **Microarray analysis of flax (Linum usitatissimum L.) stems identifies transcripts enriched in fibre-bearing phloem tissues.** *Mol Genet Genomics* 2007, **278:**149–165.
39. Roach MJ, Mokshina NY, Badhan A, Snegireva AV, Hobson N, Deyholos MK, Gorshkova TA: **Development of cellulosic secondary walls in flax fibers requires β-galactosidase.** *Plant Physiol* 2011, **156**(3):1351–1363.
40. Fulmer-Smentek SB: **Performance comparison of Agilent's 60-mer and 25-mer *in situ* synthesized oligonucleotide microarrays.** *Agilent Technologies* 2003, 5977–5988. EN. http://www.chem.agilent.com/temp/radEB25E/00042209.pdf.
41. Bolstad BM, Irizarry RA, Astrand M, Speed TP: **A comparison of normalization methods for high density oligonucleotide array data based on variance and bias.** *Bioinformatics* 2003, **19**(2):185–193.
42. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP: **Exploration, normalization, and summaries of high density oligonucleotide array probe level data.** *Biostatistics* 2003, **4**(2):249–264.
43. Tusher VG, Tibshirani R, Chu G: **Significance analysis of microarrays applied to the ionizing radiation response.** *Proc Natl Acad Sci USA* 2001, **98**(9):5116–5121.
44. Eisen MB, Spellman PT, Brown PO, Botstein D: **Cluster analysis and display of genome-wide expression patterns.** *Proc Natl Acad Sci USA* 1998, **95:**14863–14868.
45. *Gene Expression Omnibus database*; http://www.ncbi.nlm.nih.gov/geo.
46. Huis R, Hawkins S, Neutelings G: **Selection of reference genes for quantitative gene expression normalization in flax (Linum usitatissimum L.).** *BMC Plant Biol* 2010, **10:**71.

doi:10.1186/1756-0500-6-43
Cite this article as: Fenart *et al.*: Intra-platform comparison of 25-mer and 60-mer oligonucleotide Nimblegen DNA microarrays. *BMC Research Notes* 2013 **6**:43.

Chapitre 3 : Mise en évidence de la variabilité génotypique associée aux différences de métabolisme pariétal chez le lin

Chapitre 3 : Mise en évidence de la variabilité génotypique associée aux différences de métabolisme pariétal chez le lin

Avant-propos

La création variétale du lin est basée sur la sélection de variétés en fonction des quantités/qualités des produits (fibres, graines, huile) obtenus, mais également autres critères tels que la résistance aux maladies, la tolérance au froid, (variétés d'hiver) de résistance à la verse etc. Dans le cas du lin fibre, la sélection est associée à une évaluation de la quantité et de la « qualité » des fibres (performances mécaniques, physiques, chimiques…). Cette dernière est associée à la composition et à la structure de la paroi cellulaire mise en place pendant la croissance de la plante et dépendant du métabolisme pariétal de la plante. Cependant, les liens entre la « qualité » des fibres, la variété dont sont issues les fibres, et le métabolisme pariétal de la variété en question sont extrêmement mal renseignés.

C'est dans ce contexte que leprojet KBBE FIBRAGEN (Flax for Improved Biomaterials through Applied Genomics, 2011 - 2015) a été conçu. L'objet majeur de ce projet était d'obtenir des résultats permettant de mettre en place les connaissances nécessaires pour sélectionner et/ou créer les variétés les mieux adaptées (en termes de qualité de fibre) pour produire les fibres utilisées pour renforcer les agro-matériaux composites. Le consortium du projet FIBRAGEN était composé 8 partenaires venant de 4 pays : France : Limagrain Europe, Terre de lin, Linea Semences de Lin, Université science et technologie Lille 1, UMR 614 INRA-URCA, Fractionnement des Agro Ressources et Environnement. Espagne : AIMPLAS. Allemagne : université des sciences appliquées Bremen. Canada : université d'Alberta.

Le projet était composé de 4 WPs scientifiques (WP1- Structure et composition des fibres ; WP2-Interface et performances des fibres ; WP3- SNP et QTL cartographie ; WP4- Biologie intégrative des fibres) et une partie de mon travail de thèse se situait dans le WP4 (Biologie intégrative des fibres) avec des analyses histologiques, transcriptomiques et protéomiques. Dans un premier temps, nous avons réalisé une étude protéomique détaillée sur 4 organes/tissus végétatifs du lin afin de créer un « protéome atlas » chez le lin avec un focus sur les protéines en

rapport avec le métabolisme pariétal. Ce travail est actuellement soumis au journal Molecular and cellular proteomics et le manuscript est présenté ci-dessous. Dans un deuxième temps, nous avons réalisé une comparaison entre les transcriptomes des tissus externes (riches en fibres) obtenus à partir de 7 variétés de lin cultivées en champs. Afin de tenir compte de l'effet de l'environnement sur les cultures, l'expérience a été réalisée sur deux années successives. Pour complémenter ces résultats, nous avons également comparé les protéomes de 2 variétés contrastées (Diane, variété à fibres printemps et Oliver, variété à huile hiver). Ce travail est présenté ci-après, et un papier intégrant ces résultats à ceux obtenus par d'autres partenaires du projet FIBRAGEN est actuellement en préparation.

Partie 1: Flax proteome atlas

This manuscript will be submitted to the Molecular and Cellular Proteomics. For space reasons, only the supplementary data 6 are presented in the printed version of this paper.

Manuscript Title: Organ-specific proteomics and targeted cell wall analyses provide novel information on hemicellulose metabolism in flax

Manuscript No: MCP/2016/063727

Manuscript Type: Research Article

Date Submitted by the Author: 20 Sep 2016

Complete List of Authors: Malika Chabi, Estelle Goulas, Celine Leclercq, Isabelle de Waele, Christophe Rihouey, Ugo Cenci, Arnaud Day, Anne-Sophie Blervacq, Godfrey Neutelings, Ludovic Duponchel, Patrice Lerouge, Jean-Francois Hausman, Jenny Renaut, and Simon Hawkins

Keywords: Cell biology*; Macromolecular Assemblages*; Networks*; Plant Biology* ; Protein Targeting*; Plant cell wall; fibers

Organ-specific proteomics and targeted cell wall analyses provide novel information on hemicellulose metabolism in flax.

Malika Chabi, Estelle Goulas, Celine C. Leclercq, Isabelle de Waele, Christophe Rihouey, Ugo Cenci, Arnaud Day, Anne-Sophie Blervacq, Godfrey Neutelings, Ludovic Duponchel, Patrice Lerouge, Jean-François Hausman, Jenny Renault, Simon Hawkins*

Univ. Lille, CNRS, UMR 8576 – UGSF - Unité de Glycobiologie Structurale et Fonctionnelle, F 59000 Lille, France (MC, EG, AD, A-SB, GN, SH) ; Department Environmental Research and Innovation (ERIN), Luxembourg Institute of Science and Technology (LIST), L-4422 Belvaux, Luxembourg (CL, J-FH, JR) ; Laboratoire Glyco-MEV EA 4358, Institut de Recherche et d'Innovation Biomédicale, Normandie Université, Mont-Saint-Aignan, France (PL) ; Laboratoire Polymère Biopolymère Surface, UMR6270 CNRS, Institut de Recherche et d'Innovation Biomédicale, Normandie Université, Mont-Saint-Aignan, France (CR) ; University Nord de France, Lille1, LASIR (UMR CNRS A8516), 59655 Villeneuve d'Ascq Cedex, France (LD, IdW) ; Department of Biochemistry and Molecular Biology and Centre for Comparative Genomics and Evolutionary Bioinformatics Dalhousie University, Halifax, Canada (UC).

*Corresponding author: Simon Hawkins; e-mail: simon.hawkins@univ-lille1.fr; tel: +33 (0) 3 43 40 30; Fax: Fax : +33 (0) 3 20 43 65 55

Running Tittle: Flax proteome atlas and hemicelluloses

Corresponding author: Simon Hawkins

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

Tel: + 33 (0)3 20 43 40 30

E-mail: simon.hawkins@univ-lille1.fr

Abbreviations:

AA: Auxillary activities

AGPs: Arabinogalactan proteins

AIR: Alcohol insoluble residue

ALA: Alpha linolenic acid

CAZy: Carbohydrate Active Enzyme

CBM: Carbohydrate binding modules

CE: Carbohydrate esterases

CSC: Cellulose Synthase complex

CWP: Cell wall proteins

FT-IR: Fourier Transform Infrared

GH: Glycosyl hydrolases

G-layer/fibers: Gelatinous layer/fibers

GT: Glycosyl transferases

ID: Proteins with interacting domains

LM: Lipid metabolism

M: Miscellaneous

M&S: Murashige and Skoog

NCP: Non-cellulosic polysaccharides

OR: Oxidoreductases

P: Proteases

PAC: Proteins acting on carbohydrates

PIC: Protease inhibitor cocktail

PL: Polysaccharide lyases

PVPP: Polyvinylpolypyrrolidone

S: Signalling

SDG: secoisolariciresinol di-glucoside

SP: Structural proteins

UF: Proteins with unknown functions

UGT: UDP-glycosyltransferase

XEH: Xyloglucan endo-hydrolase

XET: Xyloglucan endotransglycosylase

XTH: Xyloglucan endo-transglycosylase/hydrolase

Summary

Proteomic analyses of four different flax organs/tissues (inner-stem, outer-stem, leaves and roots) enriched in proteins from 3 different sub-compartments (soluble-, membrane-, and cell wall-proteins) resulted in the identification of 2355 total proteins corresponding to 1242 non-redundant proteins. Combination of these results with publically available data on flax seed proteins and whole stems enabled the construction of a flax proteome atlas containing 2996 non-redundant total proteins (7.5 % predicted proteins). Functional classification revealed sample-specific differences related to the different biological roles of the flax organs/tissues. Classification of proteins according to MapMan, CAZy, WallProtDB, and expert curation identified a total of 465 non-redundant proteins localized in the cell wall and/or associated with cell wall biosynthesis, remodeling and other cell wall related processes. Protein interaction analyses by STRING underlined differences between different flax organs/tissues in agreement with contrasted cell wall structure. Enzymatic fingerprinting, immunolocalisation and FT-IR microspectroscopy indicated that flax fiber primary/S1 cell walls contained xyloglucans with typical substituted side chains as well as glucuronoxylans in much lower quantities. Phylogenetic analyses revealed an important paralogy in the class IIIA xyloglucan endo-transglycosylase/hydrolase (XTH) family associated with xyloglucan endo-hydrolase activity. These results suggest a likely central role of xyloglucans and endotransglucosylase/hydrolase activity in flax fiber formation and cell wall remodeling processes.

Introduction

Flax (*Linum usitatissimum* L.) is an economically important species cultivated for both its cellulose-rich bast fibers and for its seeds used as a source of animal feed and oils containing unsaturated fatty acids such as Alpha linolenic acid (ALA) (1). In addition flax seeds also contain high amounts of the lignan secoisolariciresinol di-glucoside (SDG) that shows a number of biological activities of interest for human health (2). As in other plant species improvement in the production and quality of the wide range of products obtained from flax, as well as a better tolerance to abiotic stress and pathogens requires a more thorough understanding of many aspects of flax biology including cell wall metabolism. This is currently being achieved through the use of different genomics, transcriptomics and functional approaches
(3)(4)(5)(6)(7)(8)(9)(10).

In contrast, the currently available proteomics data available for this species is limited to certain organs and/or cell types and a more comprehensive overview of the flax proteome is therefore lacking(11)(12)(13). Most transcriptomics data is based upon the measurement of statistically significant differences in transcript accumulation between different biological samples. However such differences do not necessarily represent absolute proof that the corresponding protein is more abundant in one sample as compared to another sample since other factors (translation efficiency, protein turnover, post-translational modifications) affect protein quantity and activity (14). Similarly, the absence of significant differential transcript accumulation between different samples does not mean that the protein is absent. Proteomics is therefore an important complementary tool that has been successfully used in plant biology to improve our understanding of different processes (15)(16). We have previously used this approach to identify a total of 1135 non-redundant proteins in flax stems of which 152 were considered to be cell wall proteins based on the use of TargetP, Predator and WoLF PSORT

algorithms (13). Knowledge about the proteins localized in this compartment is important since, in addition to polysaccharidic (cellulose, hemicellulose, pectin) and phenolic (lignin) polymers cell walls can also contain between 5 and 10% proteins that are involved in both cell wall structuring and other biological processes including defense and signaling.

In order to provide more comprehensive proteomics data on flax in general, and more specifically on cell wall metabolism in this species we decided to identify proteins present in 3 different vegetative flax organs (stems, leaves and roots). Given the highly contrasted cell wall structures present in stem inner- and outer-tissues we analyzed these 2 tissues separately (17)(18). This information was completed by publically available data on flax seed and whole stem proteomics (12)(13) to produce a flax proteome atlas that we believe will be of use to the scientific community working on this species.

Since our group is mainly interested in learning more about cell wall biology in flax we then decided to undertake a detailed analyses of proteins associated with cell wall metabolism in the different organs/tissues. Despite the fact that cell wall polymers are localized in the cell wall, a number of proteins associated with their biosynthesis and remodeling are not localized in this compartment. For example the phenylpropanoid enzymes involved in lignin monomer (monolignol) biosynthesis are localized in the cytosol and the cellulose synthase enzymes are associated with the cellulose synthase complex (CSC) in the plasma membrane (19)(20). Although the majority of lignin biosynthesis enzymes can be extracted in the soluble protein fraction, the obtenion of membrane-associated proteins is more complicated and we therefore also used a protocol previously utilized to obtain membrane proteins in poplar (21).

In addition to analyzing proteins associated with cell wall metabolism in flax we also decided to complete some gaps in our knowledge of flax fiber cell wall structure. The fiber cell wall has been extensively investigated by different authors and is characterized by an extremely thick secondary layer that is sometimes referred to as a gelatinous layer (G-layer) because of the similarity with the corresponding layer in tension wood G-fibers (22)(23)(17)(24). Analyses of the flax fiber secondary cell wall shows that it is made up of approximately 70 % cellulose, 5-15 % non-cellulosic polysaccharides (NCPs) consisting of beta-1,4-galactans and arabinogalactans, as well as extremely low amounts of lignin (25)(26)(18)(27). Although flax fibers also contain hemicelluloses these polymers have received less attention despite their likely central importance in both load-bearing and cell expansion processes (28)(29)(30)(31). We therefore decided to characterize fiber cell wall hemicellulose content by a combination of enzymatic fingerprinting, immunolocalisation and FT-IR microspectroscopy.

Altogether our results represent an important contribution to flax proteomics in general and more specifically to our understanding of cell wall biology in this species. This knowledge will provide a clearer vision of the dynamic mechanisms involved in bast fiber construction and of the link between cell wall structure and mechanical properties in different plant fiber species. In addition a better understanding of the molecular mechanisms that enable plant cells to build cellulose-rich walls will provide fundamental knowledge necessary to engineer plant species for the production of more efficient lignocellulosic biomass for biofuels.

Experimental Procedures

Plant growth

Seeds of the flax *(Linum usitatissimum* L.) cultivar Diane were germinated on moistened paper for 24 h at 25 °C in the dark, and then transferred to hydroponic culture and grown for a

total of 70 d on increasing concentrations of M&S nutrient solution in the absence of sucrose and agar according to the following regime (32). Days 0 – 30 (1/20 M&S renewed every 2 d, photoperiod 12/11 h day/night, temperature 14/12 °C day/night); days 30 – 60 (1/15 M&S renewed every 1 d, photoperiod 13/11 h day/night, temperature 16/14 °C day/night); days 60 – 70 (1/10 M&S renewed every 1 d, photoperiod 14/10 h day/night, temperature 18/16 °C day/night). Light intensity was 210.18 $\mu mol.s^{-1}.m^{-2}$, and hygrometry was 60%. Leaves, roots and stems were harvested separately after 70 d and rapidly frozen in liquid nitrogen, and stored at −80°C. For the stems, outer tissues and inner tissues were separated as described previously (18).

Protein extraction and fraction enrichment

Leaves, roots, outer stems and inner stems from 20 plants per biological replicate were ground independently in liquid nitrogen to a fine homogenized powder using a grinder and then pooled. Two g powder was used for the extraction of soluble and cell wall proteins and 7 g were used to obtain a fraction enriched on membrane proteins. Altogether 24 samples were obtained: 4 organs/tissues x 3 subfractions x 2 biological replicates.

Soluble (S) and cell wall (C) protein enrichment:

Soluble proteins were extracted according as previously described (33). Briefly, 2 g ground sample were further ground for 5min in 10 ml Tris-HCl buffer (50 mM, 0.06% PIC (Protease inhibitor cocktail), pH 7.5; Sigma-Aldrich) and then centrifuged (10 min at 4 °C, 16000g). The pellet was recovered and stored at -80 °C for cell wall protein (CWP) extraction (see below). The supernatant was recovered and incubated (15 min, room temperature, gentle agitation) with protamine sulfate (1mg/1ml) before being centrifuged (10 min, 18000g). The resulting supernatant was recovered and soluble proteins (S) were precipitated by incubation

with 72 % TCA (trichloroacetic acid, 1 h, -20 °C). The pellet was washed once with cold acetone and dried at room temperature (5 mins) before being stored at – 20 °C.

CWPs were extracted from the CWP pellet as previously described (34). Briefly, the pellet was washed with 500 ml sodium acetate buffer (5mM, pH 4.6) and filtered through a nylon membrane (40×40 microns, Millipore Corporation, Bedford, MA, USA). The pellet was washed again with sodium acetate buffer (5mM, pH 4.6) and then incubated with 20 ml 1.5M NaCl (30 min at 4°C with gentle agitation). The solid cell wall residue was then further extracted by i) incubation with 20mL $CaCl_2$ buffer (5mM sodium acetate, 200 mM $CaCl_2$, 30min at 4°C) and ii) incubation with 20mL LiCl buffer (5mM sodium acetate, 2 M LiCl, 30min at 4°C). The liquid fractions from all CWP extractions were combined and CWPs precipitated by adding 10% TCA.

Membrane (M) protein enrichment

Membrane (M) proteins were extracted from 7 g ground sample according to Song et al. (2011). Briefly, a small amount of 1 % PVPP (polyvinylpolypyrrolidone) was added to ground material before being incubated in 50 ml extraction buffer (0.5 M Tris–HCl, pH 8.5, 0.7 M sucrose, 0.1 M KCl, 50 mM EDTA, 1 mM PMSF (phenylmethylsulfonyl fluoride), 2% (v/v) β-mercaptoethanol, 0.1% PIC), 5 min at 4°C). Next, the homogenate was centrifuged (10 min, 12,000g, 4°C) and filtered. The liquid fraction was diluted by an equal volume of ice-cold water and centrifuged (150,000g, 40 min). The resulting pellet was then washed three times with ice-cold water and dissolved in SDS buffer (0.5 M Tris–HCl pH 8.5, 2% (v/v) β-mercaptoethanol, 30% (v/v) glycerol, 4% SDS, 1 mM PMSF, 0.1% PIC) and heated for 5 min at 80°C, before being centrifuged (12,000g 30 min at room temperature). The supernatant was extracted three times with an equal volume of water-saturated phenol. Proteins were precipitated at -20 °C overnight from the phenol

phase by adding 5 volumes of cold methanol containing 0.1 M ammonium acetate. After precipitation the proteins were pelleted by centrifugation (12,000g, 10 min at 4 °C) and washed three times with 90% cold methanol, and once with 90% cold acetone before being vacuum dried (5 min).

Samples were subsequently prepared for analyses by suspension in rehydration buffer (7M Urea, 2M Thiourea, 100 mM DTT, 2 % CHAPS). Protein content was determined using a reducing agent compatible (BioRad) and a detergent compatible (Bradford) assay.

Twenty-five µg total proteins per sample were partially separated by electrophoresis on a precast ready gel according to the manufacturer's instructions (Criterion ™ XT precast 1D gel, 4-12% Bis-Tris, 1 mm x 12 wells, Bio-Rad). The gel was stained with Instant Blue (Gentaur BVBA, Kampenhout, Belgium) and 5 bands were excised from each sample lane, cut into 1-2 mm cubes and transferred into a microplate. Proteins were reduced with 10mM DTT (in 100mM ammonium bicarbonate) for 30min at 56°C, then alkylated with 55mM iodoacetamide (in 100mM ammonium bicarbonate) for 20min at room temperature. Finally, gel pieces were de-stained and then digested overnight by modified trypsin enzyme (sequencing mass grade, Promega). Peptides were recovered and separated on a NanoLC ™-2D System (Eksigent, Sciex, Belgium) coupled to a TripleTOF® 5600+ mass spectrometer (Sciex, Belgium). Peptide desalting and enrichment were achieved using a pre-column (C18 PepMap ™, 5µm, 5mm * 300µm i.d., Thermo scientific, Bremen, Germany). Peptides were separated and eluted on a C18 reverse phase column (PepMap ™ 100, 3µm, 100Å, 75µm i.d. x 15 cm, Thermo scientific) using a linear binary gradient (solvent A: 0.1% FA (formic acid); solvent B: 80 % ACN 0.1 % FA; 5 min 5 % B, 40 min 5% to 55% B, flow rate of 300nl/min). The peptides were injected into the TripleTOF® 5600+ with a NanoSpray III source using a 10µm i.d. emitter (New Objective, Woburn, MA). The source parameters used vary depending

on optimized conditions on each day, the values were for gas1= [1-6], gas2=0, curtain gas=[20-30], the ion spray voltage approximately 2.2keV. For each sample 2 biological and 2 technical replicates were randomly analyzed.

MS analysis was performed in information-dependent acquisition mode. MS spectra were acquired using 250 ms accumulation time per spectrum with a mass range of 300-1250 Da. The top 20 precursor ions were selected in each MS scan for subsequent MS/MS scans with high sensitivity during 250 ms of accumulation time (range from m/z 100-1250 Da) and the voltage was automatically adjusted with the system of rolling collision energy. The dynamic exclusion was set at 10 s. Systems were controlled by Analyst software (version TF1.7, Sciex). An automatic mass recalibration was performed using a beta-galactosidase digest during the sequence of samples.

CID spectra were processed by Mascot (version 2.4.2) using Protein Pilot (version 4.5, Sciex) by searching against the *Linum usitatissimum* database (v1.0, http://www.phytozome.net released on 10th December 2014) and NCBI *Viridiplantae* database (released on 30th October 2014). The searches were performed with the following parameters: enzyme: trypsin, 2 missed cleavages, mass accuracy precursor: 20 ppm, mass accuracy fragments: 0.3 Da, fixed modifications: Carbamidomethyl (C), dynamic modifications: Dioxidation (W), Oxidation (HW), Trp-> Kynurenin (W), Oxidation (M). Three supplemental filters were applied to Mascot results: i) a peptide confidence ($p > 0.05$), ii) an individual ion score (calculated by Mascot) of the considered research and iii) a minimum of two significant peptides per proteins. Only proteins fulfilling these criteria were retained.

Protein identification and functional classification

Flax proteins were identified directly via the inclusion in the Mascot database of peptides sequences based on the sequenced *Linum usitatisimum* genome in Phytozome (http://www.phytozome.net). Only proteins present in both biological repetitions independently run by MS analysis were retained for further study. The predicted protein annotation was checked and completed with InterPro v.48 (http://www.ebi.ac.uk/interpro/) and CAZy databases (http://www.cazy.org/). Functional classification was performed using MapMan (http://mapman.gabipd.org/web/guest/mapmanstore;) (35). Subcellular locations were predicted using Predotar v. 1.03 (https://urgi.versailles.inra.fr/Tools/Predotar), TargetP1.1 (www.cbs.dtu.dk/services/TargetP/), and SignalP 4.1. (www.cbs.dtu.dk/services/SignalP/) (36). Annotation of enzymes that assemble, modify and breakdown oligo- and polysaccharides was done according to the CAZy database (37) and the scientific literature. Peroxidases were named according to the PeroxiBase (38). The predicted cell wall location of proteins was verified using WallProt database (www.polebio.lrsv.ups-tlse.fr/WallProtDB/) and proteins were functionally classified according to classes established in this database.

Phylogenetic analysis

Homologues of Linum usitassinum CWPs were selected by BLASTp, (e-value cutoff of 1e-20, percentage coverage ≥ 50% minimum, maximum of 1000 sequences) against a database of selected proteomes from phytozome v.10 (*L. usitassinum, P. trichocarpa, M. truncatula, R. communis, G. max, P. persica, A. thaliana, B. distachyon, P. patens, C. reinhardtii, V. carteri f. nagariensis, M. pusilla CCMP1545, O. lucimarinus, M. esculenta, A. coerulea, C. sinensis, E. grandis, G. raimondii, S. purpurea,V. vinifera, P virgatum*). The selected set of homologous sequences was aligned using multiple sequence comparison by log expectation (MUSCLE) (Edgar, 2004). Block selection was performed with BMGE (Criscuolo and

Gribaldo 2010). Reduced alignments were analyzed under maximum likelihood with the LG4X (Le et al., 2012) and 1000 bootstrap repetitions with IQ-TREE software (Nguyen et al, 2015).

Interaction networks of cell wall-related proteins

Statistical analyses for identification of flax cell wall-related protein networks were performed by using STRING v10 (http://string-db.org/) (39). The flax total CWP list was blasted against the *Arabidopsis thaliana* STRING database containing relationships supported by associations derived from four sources: genomic context, high-throughput experiments, co-expression and previous knowledge. Interactions were determined with 0.400 minimum required interaction score. If there was more than one isoform per gene, the longest isoform was selected unless more information suggesting that other isoform is regarded as canonical.

Preparation of flax hemicellulose fraction

Alcohol-insoluble cell wall residue (AIR) was prepared by grinding 5 g inner- and outer-stem tissues in liquid N_2 and extracting in 20 ml 70 % EtOH (V/V) for 15 min at 70 °C prior to centrifugation (10 min, 5,000g) and pellet recuperation. This step was repeated 3 times and the pellet (= AIR) recovered and lyophilized (Fry, 1988). A hemicellulose fraction was obtained by sequential extraction. Pectins were first removed from AIR by heating (100 °C, 2h) in ammonium oxalate (0.5 % P/V) followed by centrifugation, elimination of the supernatant and washing of the pellet (x2) in 20 ml 4M NaOH. This step was repeated twice and the pellet was then incubated overnight in 4M NaOH at 4 °C to extract hemicelluloses. The basic extract was then neutralized by 6N HCl, centrifuged (10 min, 5,000g) and the supernatant recovered and dialyzed twice before being lyophilized.

Generation of oligo-hemicelluloses

The lyophilized fraction containing flax hemicelluloses was dissolved in 2.5 ml ammonium acetate (10 mM, pH 5.5) and 0.5 ml digested by either i) 2.5 units of endo-β-glucanase (Megayme) for 18h at 37°C to generate oligo-xylo-glucans or ii) 1 unit of endo-xlanase (Megazyme) for 18h at 28 °C to generate oligo-xylanes. The digestions were stopped by adding 3 volumes 95% EtOH and the liquid phase (containing the oligo-xyloglucans/xylans) lyophilized and then dissolved in ultra-pure water.

MALDI-TOF-MS

One µL of oligo-xyloglucans/xylans was co-crystalized with 1 µL 2,5 dihydrobenzoic acid prepared by dissolving 5mg DHB in 500µL acetonitrile/0.1% TFA (70:30, V/V) on a stainless steel support. Mass spectra were recorded on a Micromass TOF E MS (Manchester, UK) in positive mode. Spectra were calibrated using the Sequazyme Peptide Mass Standards Kit (Applied Biosystems).

Light Microscopy

Flax stem samples were excised at 10 cm above the insertion point of cotyledon and fixed in a formaldehyde: acetic acid: ethanol solution (FAE, 3.5:6.5:90, *v/v*.), progressively dehydrated in an alcohol series and infiltrated with Technovit 7100 resin (Kulzer). Sections (3 µm) were cut using a Leica RM 2065 microtome and stained with a 0.5% (w/v) aqueous solution of Toluidine Blue O (TBO) (Sigma) for examining tissue organization. Sections were examined with a Leica DM2000 microscope coupled to a Leica DFC320 camera. Images were analyzed with the Leica Application Suite program.

Immunolocalization

Resin was removed by acetone, and sections were rehydrated in a 5 min graded ethanol series followed by immersion in PBS. For immunolocalization, non-specific sites were saturated in 2 x 15 min PBST (2% BSA (w/v) and 0.05% tween 20 (v/v) in PBS) followed by immunolocalization of xylan epitopes with rat monoclonal LM10-11, xyloglucan epitopes with LM-15-24 and mannan/heteromannan with LM21 primary antibodies (Plant Probes, Leeds, U.K.). Experiments were performed on 5-7 µm semi-thin sections, using diluted 1/10 antibodies in a PBST buffer (137 mM NaCl, 2.7 mM KCl, 16.3 mM Na_2HPO_4, 1.5 mM KH_2PO_4, pH 6.8, 2 % (w/v) BSA, 0.01 % (v/v)Triton X100). After rinsing in three 10-min changes of PBS at room temperature, sections were incubated with 1:100 diluted (PBS) goat anti-rat antibodies conjugated with FITC (Fluorescein IsoThioCyanate, Jackson Immunoresearch Labs., West GROVE-PA, US), for 1 h 30 at 37˚C and finally rinsed three times in PBS. Sections were mounted in Fluoromount^{TM} (EMS, Fort Washington-PA, USA) and conserved at 4˚C. Observations were done on a Leica microscope under ultraviolet irradiation, at λexc =480 nm: λem = 505 nm for FITC λexc = 365 nm: λem = 420 nm for cellulose autofluorescence, or in bright field with phase contrast. In control sections, primary antibodies were replaced by preimmune rat serum or water. No signal was obtained on control sections.

FT-IR microspectroscopy

Flax stem samples were dehydrated in a graded ethanol series (2 x 30 min, RT), followed by Histoclear (VWR, 2 x 1h), followed by Paraplast+ (Sigma) embedding. Ten micrometers thick transversal sections were obtained using a rotary microtome (Leica RM2065) and placed on CaF_2 windows (Crystran, U.K.) before being deparaffinized (Histoclear, 2 x 10 min, RT) rinsed in absolute ethanol (2x10 min, RT), air dried and stored under dust-free conditions.

FTIR spectra were acquired in transmission mode with a Bruker FTIR microscope Hyperion 3000 coupled to a Bruker Vertex 70 spectrometer equipped with a liquid nitrogen cooled MCT detector. Two hundred and fifty-six scans were recorded at 4 cm^{-1} spectral resolution in the 650cm^{-1} - 4000 cm^{-1} spectral range using a 36x Cassegrainian objective (NA=0.5). The area of interest (10 μm x 50 μm) was selected by adjusting the knife-edge aperture so that only the area of interest is exposed for analysis. Only absorption band maxima in the 1800-900 cm^{-1} spectral range were used to identify cell wall polymers. Five spectra were averaged from two CaF_2 windows and two different stem sections were used. The second derivative of the spectra was computed using the well-known Savitsky-Golay algorithm in order to extract more spectral features (40).

Results

Protein identification and functional classification in different flax organs/tissues

We used 3 different extraction protocols to obtain fractions enriched in soluble proteins (S), cell wall proteins (C) and membrane proteins (M) from 4 different flax organs/tissues (leaves, roots, outer- and inner-stems) containing contrasted tissues and cell types (Fig. 1). Trypsin-digested peptides from the 12 different samples were then individually analyzed by liquid chromatography tandem mass spectrometry (nanoLC-MS/MS) and proteins identified using the sequenced flax genome as reference (3). Altogether a total of 2355 proteins were identified in the 12 different samples corresponding to 1242 non-redundant proteins (Supp. Data 1 and 2). Of these 1242 proteins, 219 were uniquely found in the S fraction, 33 in the C fraction and 740 in the M fraction (Fig. 2a). Two hundred and twelve proteins were common to 2 fractions and 38 proteins were common to all 3 fractions (Fig. 2a). The use of specific

extraction protocols in addition to the soluble fraction step enabled us to recover 62 (cell wall fraction) and 769 (membrane fraction) additional proteins, of which 29 proteins were common to the C and M fractions. One hundred and sixteen, 100, 189 and 186 proteins (all fractions) were uniquely identified in leaves, roots, outer-stems and inner-stems respectively (Fig. 2b). In all organs/tissues, 322, 196 and 133 proteins were common to 2, 3 or all 4 samples (Fig. 2b). Inclusion of previously published proteomics data on flax seeds and whole stems allowed us to recover 1754 additional proteins bringing the total of non-redundant proteins in the flax proteome atlas to 2996 (Supp. Data 3) (12)(13).

To obtain a global view of the biological processes in which the identified flax proteins were involved we classified them into different MapMan Bins (Fig. 3a.) (http://mapman.gabipd.org/web/guest/mapmanstore;) (35). The class "proteins" was particularly highly represented in all samples with more than 21% total proteins, followed by the class unassigned. Specific organs were particularly rich in certain categories (e.g. the class "photosynthesis" in leaves; the class "lipid metabolism" in seeds, or the class "miscellaneous" in roots) reflecting their biological specificity and/or our lack of functional knowledge. 55 proteins were present in all 6 organs and represent potential "housekeeping proteins" as previously observed for the "mitochondrial ATP synthase D chain" and "H^+-ATPase" proteins (41). Functional classification of these housekeeping proteins indicated that 17 out of the 35 MapMan bins were represented (Fig. 3b).

Proteins associated with cell wall metabolism

Since we are primarily interested in understanding more about flax cell wall formation we then concentrated on proteins potentially involved in this process. Plant cell walls mainly consist of polysaccharides and we firstly analyzed the proteome atlas to identify different

CAZy (Carbohydrate Active Enzymes) families(37).. Overall we identified 207 CAZy proteins distributed among the 7 CAZy classes: glycoside hydrolases (GH), glycosyl transferases (GT), polysaccharide lyases (PL), carbohydrate esterases (CE), auxiliary activites (AA), carbohydrate binding modules (CBM) and expansin (Expn) (Fig. 4, Supp. Data 4). Analyses of CAZy proteins from combined organ/tissue data indicated that GH was the most represented class (39 %), followed by AA (30 %), GT (16 %), CE (9 %), CBM (4 %) and 1 % (PL, Expn). When organs/tissues were analyzed individually, GHs and AAs were major classes represented in all samples and ranging from 34 % total CAZy proteins (whole stem) to 48 % (roots) for GHs and from 28 % (seeds) to 45 % (leaves) for AAs. Like GHs and AAs, GTs represented another major group in inner stems (31 %) and seeds (21 %), but showed much greater variation in other organs/tissues (4 % in roots, 9 % in outer stems, 10 % in whole stems) and were absent from leaves. Overall protein profiles indicated a similar distribution for outer stems, whole stems and seeds. Roots and leaves also showed similar profiles, mainly due to the absence/low abundance of GTs. Inner stems showed the most simple profile containing only GHs, GTs and AAs in roughly equal proportions.

Examination of individual families within each CAZy class (Fig. 5, Table 1) allowed us to more clearly visualize those families and enzymatic activities related to cell wall metabolism. Overall, of the 207 identified flax CAZy proteins, 178 were potentially associated with cell wall metabolism. For GHs we identified proteins belonging to 22 different families of which 14 have previously been associated with cell wall metabolism (42). Only 2 cell wall GH families (GH2: beta-mannosidase/glucosidase, glucanase and GH36: alpha-galactosidase) were not represented. Highly represented cell wall GH families included GH1, GH3, GH19, and GH35. Proteins belonging to families 1 and 3 were identified in all samples (Fig. 6) while the number of proteins belonging to other families showed important differences between the

different organs/tissues. For GTs we identified proteins belonging to 10 different families of which 7 have previously been associated with cell wall metabolism. Cell wall GT families that were the most represented included GT2 (cellulose), GT4 (cellulose – sucrose synthase), and GT75 (GAX) (43). Other plant GT families involved in hemicellulose (GT47), callose (GT48) and pectin (GT8) biosynthesis were also present. Another highly represented family includes the GT1 family that includes UDP-glycostransferases (UGTs) glycosylating secondary metabolites and hormones and that may be involved in regulating lignin monomer availability (44)(45). As for GHs, GT families were more represented in certain organs/tissues than in others. In the class AA, 6 families were present with the AA2 family containing peroxidases and the AA1 family containing laccase enzymes involved in lignin biosynthesis and cell wall reticulation being the most abundant. AA2 Family proteins were present in all organ/tissue samples analyzed and were most represented in whole stems, roots, seeds and outer stem tissues. AA1 family proteins were absent in leaves and roots. We also identified proteins involved in pectin acetylation (CE13) (46), methylesterification (CE8) and hydrolysis (PL4). Once again family representation varied according to organ/tissue type. Finally proteins belonging to 5 CBM families and 1 expansin family were also present.

Plant cell walls also contain other non CAZy proteins that are involved in the biosynthesis, deposition and remodeling of cell wall structural components, as well as in other important cell wall-related biological roles such as defense and signaling. Many of these proteins are physically located in the cell wall and/or apoplasm and are therefore secreted. We identified such proteins in our data using several criteria (presence of peptide signal, secretion predicted by TargetP, SignalP and Predotar software, no more than 1 transmembrane domain). Altogether 355 proteins (including 103 CAZy proteins) representing 11.18 % total proteins satisfied these criteria and were considered as cell wall proteins (CWPs). The proteins were

then grouped into the 9 classes of the WallProtDB (http://www.polebio.lrsv.ups-tlse.fr/WallProtDB/) (47) (Fig. 7, Supp. Data 5). Global analyses (Fig. 7a) of pooled data from all organs/tissues showed that the largest class corresponded to 'proteins with unknown functions' (UF, 25 %), followed by 'proteins acting on cell wall polysaccharides' (PAC, 19 %), 'miscellaneous' (M, 13 %); 'signaling' (S, 13 %), 'oxidoreductases' (OR, 12 %), 'proteases' (P, 12 %), 'proteins with interacting domains' (ID, 3 %); 'lipid metabolism' (LM, 3 %) and 'structural proteins' (SP, < 1 %). When analyzed individually (Fig. 7b), most organs/tissues showed a similar CWP profile to the global pattern with for example the UF fraction constituting the class the most represented in all samples (17 – 34 %) except for roots where the most represented class was PAC (27 %), followed by UF (17 %). Structural proteins (SP) were only found in 3 samples (OS, WS and S).

In addition to CAZy proteins and secreted/cell wall proteins we also analyzed our data set to identify proteins involved in lignin metabolism. Our results (Fig. 8, Supp data 6) indicate that proteins belonging to 9 out of the 11 canonical monolignol biosynthesis enzyme families were present in different flax samples. Different individual proteins of the same family could be present in the same or different organs/tissues except for leaf samples were no monolignol biosynthesis enzymes were detected. Lignin formation occurs by a 2-step oxidative polymerization process in which monolignols are first oxidized to highly reactive radicals by laccases and/or peroxidases before undergoing spontaneous polymerization (19). Flax peroxidases potentially involved in this process are indicated in Supplementary Data 6.

Proteins identified in the MapMan cell wall bin, but not included in the groups discussed above are listed in Supplementary Data 7. This group contains enzymes involved in the synthesis/inter-conversion of UDP-activated cell wall polymer sugars, as well as cell wall

degradation and a member of the COBRA family of proteins related to cellulose production(48).

Altogether we identified 465 non-redundant proteins potentially involved in cell wall metabolism in flax. We named this set of proteins 'Total cell wall proteins' (Total CWPs = cell wall CAZy proteins + CWPs + Lignin enzymes + remaining MapMan CWPs) (Fig.9, Supp. Data 8.). Examination of protein repartition by organ showed that 282 proteins were present in whole stem samples, of which 176 were only present in this organ, followed by seeds (194 total proteins, 138 specific), roots (99 total proteins, 28 specific) and lastly leaves (58 proteins, 6 specific). Of the 282 total proteins identified in the stem, 131 total proteins (32 specific) were identified in outer-stem tissues and 119 total proteins (32 specific) were identified in inner stem tissues.

Interacting cell wall protein networks in flax

To examine flax cell wall-related protein networks we used the STRING v10 (http://string-db.org/) database that identifies relationships supported by associations derived from four sources: genomic context, high-throughput experiments, co-expression and previous knowledge(39). Interrogation of the STRING database with the 465 flax Total CWPs using *Arabidopsis thaliana* as reference organism led to the identification of 311 corresponding proteins (Supp. Data 9).

When protein-protein networks were examined individually for the 6 flax organs/tissues, 8 major clusters could be identified (Fig. 10, Table 2). Although the majority of proteins present with a given cluster are associated with the process/polymer indicated one/more proteins may be associated with another process (e.g. the IRX6 protein present in the inner stem cellulose

cluster corresponds to the CCR protein involved in lignin monomer biosynthesis). Generally the identified clusters could be related to the biological role(s) of the different organs/tissues (see discussion). For each organ/tissue, all clusters were inter-linked, except for the 'protease' cluster in inner stems and roots that formed a separate group. Only proteins interacting with at least one other protein are represented in the STRING figures and examination of the relative proportions of 'connected' (present in the diagrams) and 'non-connected' (not present) showed that values varied between 29.7 % (all organs/tissues) to 50.6 % (roots) underlying our current lack of knowledge about protein-protein networks in plants (Table 3). Of the 'non-connected' proteins, between 22.8 % (inner stem) and 38.5 % (roots) could be directly related to cell wall polymer metabolism (Table 3).

XTH paralogy and characterization of flax stem hemicelluloses

In order to obtain a better overview of the flax Total CWPs we generated maximum likelihood phylogenetic trees (data not shown). Inspection of the different trees allowed us to identify 2 protein families showing an important paralogy in comparison with the other species used for the analysis. The first corresponded to the GH19 family containing chitinase-like (CTL) as previously reported (3)(6)(49). The second paralogy concerned the xyloglucan endo-transglycosylase/hydrolase family GH16 IIIA (37) that is associated with hemicellulose metabolism and cell wall expansion in plants (50) (Fig. 11, Table 4, Supp. Data 10). Our results indicated the presence of 10 flax proteins in the class IIIA compared to 3 proteins in arabidopsis and *Populus thiocarpa*. The flax XTH sequences group together and show bootstrap values suggesting a strong paralogy that might be related to cell wall specificities in this species. Overall, the flax genome contained a greater number of *XTH* genes (67) compared to poplar (44) and arabidopsis (33) (3)(51).

The existence of an important XTH family paralogy, as well as the fact that we had identified proteins belonging to 10 CAZy families (8 GHs, 2GTs, Figs 5,6, Table 1) potentially associated with cell wall hemicellulose metabolism prompted us to take a closer look at this polymer. Although detailed information on the structure of cellulose, pectin and lignin in flax is available less information is currently available on the detailed structure of flax hemicelluloses despite the probable importance of these polymers in structuring the cell wall. We therefore analyzed the structure of flax hemicelluloses in inner- and outer-stem tissue samples of 2-month-old flax plants by enzymatic fingerprinting. Analyses using MALDI-TOF-MS (Fig. 12) of a hemicellulose-rich fraction from inner stem tissues digested with endo(1-4)-β-D-xylanase (Fig. 12A) revealed 2 mains ions with *m/z* values of 759 and 781 corresponding to the $(M+Na)^+$ and $(M+2Na-H)^+$ adducts, respectively, of an oligosaccharide with a 4-*O*-MeGlcA residue linked to 4 pentose residues. Considering the monosaccharide composition of the fraction (Supp Data 11), this fragment is most likely derived from the enzymatic degradation of a glucuronoxylan. Analyses of the same sample digested by endo-glucanase (Fig. 12B) allowed the identification of 4 ions at *m/z* values corresponding to different xyloglucan oligosaccharides (1085: XXXG; 1247: XLXG/XXLG, 1393: XXFG, 1555: XLFG). These fragments are derived by enzymatic cleavage of non-substituted glucose residues of a fucosylated XXXG-type xyloglucan. Ions corresponding to the previously identified glucuronoxylan oligosaccharide (759, 781) as well as a new ion at *m/z* = 891 most likely indicative of a 4-*O*-MeGlcA residue linked to 5 pentose (xylose) residues were also present. The presence of xylan-derived oligosaccharides in the endo-glucanase-digested fraction would seem to suggest that the enzyme used is contaminated by xylanase activity as previously observed (52). MALDI-TOF-MS analyses of the hemicellulose-rich fraction from outer stem tissues digested with endo(1-4)-β-D-xylanase (Fig. 12C) revealed the presence of a major peak with m/z = 759 corresponding to the $(M+Na)^+$ adduct of a 4-*O*-MeGlcA residue

linked to 4 pentose (xylose) residues. Analyses of endo-glucanase-digested samples (Fig. 12D) allowed the identification of the same 4 peaks identified in inner-stem samples, but at much higher levels indicating that outer-tissue hemicelluloses are richer in xyloglucans. Two smaller peaks with *m/z* = 791 and 953 also revealed the presence of XXG and XXGG/GXXG xyloglucan oligosaccharides. No peaks corresponding to xylan oligosaccharides could be detected. Overall these results indicate that the hemicelluloses present in flax stem inner tissues are mainly xylans together with lesser amounts of xyloglucans. In contrast, outer stem tissue hemicelluloses are mainly xyloglucans.

These results were confirmed by immunlocalisation of different hemicellulose epitopes in flax stems (Fig.13). Cell walls in flax xylem tissue were extensively labeled by xylan LM10 and LM11 antibodies (Fig. 13b, c). While no labeling was observed in bast fiber cell walls with LM10 antibodies (Fig. 13b), a relatively intense labeling was present in the middle lamella/primary cell wall of bast fibers with LM11 antibodies that recognize more heavily substituted xylans. The middle lamella/primary cell wall of bast fibers was heavily labeled by the xyloglucan antibodies LM15 and LM24 that recognize the XXXG and XXLG/XLXG oligosaccharide epitopes, respectively (Fig. 13d, e). LM15 antibodies also recognized epidermal cell walls and gave a weak labeling in xylem walls (Fig. 13d). The LM24 antibodies weakly labeled xylem and pith cell walls (Fig. 13e). Finally, the LM21 antibody that recognizes mannans/heteromannans gave a strong signal in bast fiber secondary (but not primary) cell walls, as well as in the first few cell layers of the most recently formed xylem tissue (Fig. 13f).

The presence of hemicelluloses in the flax fiber cell wall was further confirmed by using FT-IR micro-spectroscopy that enabled us to identify significant bands characteristic of xylans

(1313, 1452, cm^{-1}) arabinoxylans (1053 cm^{-1}), xyloglucans (1201, 1313, 1371, 1452 cm^{-1}) and mannans (1053 cm^{-1}) (Fig. 14). The presence of other characteristic bands indicated the presence of cellulose (1429, 1201, 1161, 1103 cm^{-1}), pectin (1275, 983 $cm\text{-}^{1}$) and (lignin (1514 cm^{-1}) (Supp. Data 12).

Discussion

Flax is an economically important species and is grown for both its bast fibers used in the textile- and composite material industries and for its seeds (linseed) rich in alpha linolenic acid (ALA) (1). In this study we combined experimental proteomics data from 4 flax tissues/organs with publically available data to generate a flax proteome atlas consisting of 2996 non-redundant proteins (7.5 % predicted proteins) with information on proteins present in 6 different organs/tissues (whole stems, inner-/outer-stems, leaves, roots and seeds). Functional characterization of proteins using MapMan provided a general view of metabolism in the different samples as well as identifying 'housekeeping proteins' that were present in all samples. Altogether this data will represent a useful resource for the scientific community working on different aspects of flax biology.

The use of different protocols designed to enrich proteins in separate sub-compartments (soluble-, membrane- and cell wall-proteins) enabled us to not only significantly increase the total number of flax proteins obtained but also provides important information about the subcellular location of the identified proteins. Plant development and responses to environmental stress and pathogens are complex processes that depend upon coordinated interactions between keys sets of proteins and metabolites in different compartments and subcellular proteomics is therefore proving to be a powerful tool for improving our understanding of these processes (53).

Flax cell wall proteomics

Proteomics targeting the cell wall started about 10 years ago and there are now more than 40 papers reporting studies of this particular compartment (54). Cell wall proteins physically located within the cell wall are more or less easy to extract depending upon their interactions with other components of the cell wall. Proteins that are only loosely bound in the cell wall are generally extracted in the soluble fraction whereas proteins that are more tightly bound (Van de Waals, hydrogen bonds, ionic, hydrophobic interactions) can be extracted by using successive salt extractions (55). Other proteins can be covalently linked thereby forming insoluble networks that are not be easily extracted from the cell wall (56). We have previously used a sequential salt-based extraction protocol to identify cell wall proteins in whole flax stems (13). In this paper we combined this approach with a protocol (21) aimed at recovering a fraction enriched in membrane proteins. Such an approach is justified since the term 'cell wall protein' can be ambiguous. One interpretation requires that the protein is physically located within the cell wall and is based for example on the selection of proteins having a predicted signal peptide but no intracellular retention signal (47). However, a number of proteins that are not physically located in the cell wall are intimately involved in the biosynthesis of major cell wall polymers and it is therefore of immense biological interest to include information about these proteins in a study of cell wall metabolism. Glycosyltransferases involved in hemicellulose, pectin and cellulose biosynthesis, as well as lignin monomer biosynthesis enzymes are typical examples of proteins not physically located in the cell wall but that obviously make a major contribution to cell wall structure. Although the majority of lignin biosynthesis enzymes are cytosolic and therefore recovered in the soluble fraction, glycosyltransferases are physically associated with the plasma membrane (cellulose synthases) and/or the endoplasmic reticulum/Golgi apparatus and the use of a

protocol aimed at recovering these proteins is particularly relevant. Moreover other proteins known to be involved in cell wall processes such as arabinogalactan proteins (AGPs), wall-associated kinases, or formin1 (AtFH1) are able to form molecular bridges between the plasma membrane and the cell wall (58). Interconnections between cell wall and plasma membrane microstructure have also been proposed to play a crucial role in several physiological processes such as trafficking, signal transduction, cell growth and cell physiological responses to the environment (59). By using transmission electron microscopy it was shown that apoplastic vesicular membrane structures are more frequent beneath curved cell wall regions than linear regions of interdigitated pavement cells and guard cell ends in young leaf epidermis (60). Likewise, functional analysis of the membrane protein TED7 also suggested that it is directly bound to a subunit of the secondary cell wall-related cellulose synthase complex, probably in order to promote secondary cell wall formation in xylem vessel elements (61). Altogether, these observations confirm that the study of proteins present in the membrane sub-compartment is relevant to proteomic studies of cell wall biology. We therefore used a multiple-source based interpretation of the term ‘cell wall protein’ that included i) those proteins physically located in the cell wall and divided into the 9 WallProtDB functional classes previously established (47)ii) CAZy proteins related to cell wall metabolism (but excluding starch-related proteins), iii) lignin monomer (monolignol) biosynthesis enzymes and iv) MapMan cell wall bin proteins not included in the previous 3 classes. Altogether the CWP set contained 465 proteins of which 231 had not been previously identified in previously published data on seeds and whole stems (12)(13) .As for the total proteome data set the use of a specific protocol aimed at obtaining a fraction enriched in membrane proteins allowed us to recover 106 additional proteins validating the interest of this approach.

Overall proteins associated with the biosynthesis and remodeling of the major cell wall polymers (cellulose, hemicelluoses, pectins, lignin) could be found in all organs/tissues although in different proportions. Based on the relative proportion of the different CAZy families, inner- and outer-stem tissues and seeds appeared to be more similar to each other than to either roots or leaves. Although it is likely that a part of the observed differences can be related to the ease or difficulty with which certain proteins can be extracted from the various tissues, the fact that the overall number of proteins identified in the different organs/tissues is similar would suggest that the observed differences also reflect differences in cell wall metabolism. Analyses of protein interactions by STRING confirmed this point and generally the identified clusters reflect the biological role(s) of the different organs/tissues. A cellulose cluster was only identified in inner stem tissues and is probably related to the extensive secondary cell wall formation involved in xylem formation. Similarly the identification of a lignin cluster in inner stem tissues and whole stems can probably also be related to xylem lignin formation and is in agreement with both transcriptomic and biochemical data of flax stem tissues (18)(62)(4). Interestingly, the lignin cluster is also present in the seed protein data set where it is presumably related to the biosynthesis of monolignols necessary for lignin production (2)(63). In contrast the observed lignin cluster in outer stem tissues is more difficult to explain given the absence of highly-lignified cell walls in this tissue. One possible explanation is that the 3 phenylpropanoid biosynthesis enzymes and the 7 peroxidases making up the cluster might be involved in the formation of non-lignin phenylpropanoids responsible for UV protection and localized in epidermal cells (64). The observation that the pectin/hemicellulose cluster was only identified in outer stem tissues in also logic given the presence of bast fibers with thick secondary cell walls that undergo extensive non-cellulosic polysaccharide (NCP) remodeling and restructuring during their maturation (17)(5)(24). STRING analyses also underlined the interaction between proteins

directly involved in cell wall polymer biosynthesis and remodeling and other proteins involved in chaperone activity/protein folding. Interestingly, mutations in the genes coding some of these proteins (SHD, RSW3) are known to be associated with modified cell wall formation and/or defects in cell expansion (65)(66).

Flax stem hemicelluloses and XTH paralogy

Flax bast fibers are characterized by an extremely thick cell wall that is made up of a thin primary cell wall and secondary S1 layer, together with a much thicker cellulose-rich S2 layer that is also referred to as a gelatinous layer (G-layer) as it presents many similarities with the same layer in tension wood G fibers (17)(18)(24). Although it is the cellulose microfibrils that provide the main load-bearing structure in the flax fiber cell wall and are responsible for it's remarkable mechanical properties, it is likely that other non-cellulosic polysaccharides (NCPs) also contribute to the overall performance of the flax fiber. In the S2 (G-layer) of tension wood gelatinous fibers, the main hemicellulose fraction consists of xyloglucans whereas in flax fibers, the corresponding hemicellulose is a beta-1,4-galactan that is believed to be progressively degraded by a GH35 family beta galactosidase leading to a closer association between cellulose microfibrils in the maturing S2 (G-layer) (5)(24). The overall hemicellulose content of flax fibers varies from 4.7 % (unretted fibers) to 8.9 % (retted fibers) and contains both xylans and xyloglucans (28)(29). Our chemical, immunological and spectroscopic results confirmed the presence of xyloglucans, as well as of some substituted xylans in flax outer tissues/fibers. In contrast to the beta-1,4-galactan that is mainly localized in the S2/G-layer of flax fibers the xyloglucan (and xylan) were localized to the primary cell wall and S1 layers. Labelling by LM21 antibody also indicated the presence of beta-1,4-mannans (glucomannans/galactomannans) mainly in the older parts of the S2/G-layer, as well as more weakly in the S1 layer. Fingerprinting analyses of hemicelluloses from outer–stem

tissues indicated that the presence of a fucosylated XXXG-type xyloglucan. This contrasts with similar analyses of hemicellulose fractions from linseeds where XLLG was shown to be the most represented oligo-xyloglucan and underlines organ-/tissue-specificities (67). Overall the observed population of flax oligo-xyloglucans is similar to that of Arabidopsis suggesting that the xyloglucans in the flax fiber primary- and S1 secondary-cell wall are similar to those in other Eudicots despite the particular structure of the fiber S2/G-layers (68). The flax xylan oligosaccharide fragment composition released by an endo-xylanase was similar in seeds, and inner- and outer-stem tissues are very similar with a xylose tetrasaccharide fragment carrying one 4-*O*-MeGlcA residue being the most abundant in all 3 tissues. This fragment is derived from the endoxylanase hydrolysis of a glucuronoxylan. Once again this pattern of substitution is similar to that observed in Arabidopsis (69). Overall these results confirm the presence of fucosylated xyloglucans, and to a more limited extent glucuronoxylans, in flax fiber cell walls.

Hemicelluloses are synthesized by different glycosyltransferases (GTs) and undergo modifications during cell wall maturation by glycosylhydrolases (GHs) (70)(71). We did not detect any proteins belonging to GT families involved in xyloglucan biosynthesis in outer-stem tissues. This could be related to the fact that our samples were obtained from 2-month-old adult plants containing mature fibers in which primary cell wall biosynthesis processes are most likely finished. In contrast we did identify proteins belonging to several GH families associated with xyloglucan- (GH1, 9, 16, 31) and xylan- (GH3, 51) remodeling. Proteins belonging to GH16 and GH51 were only found in outer-stem tissues whereas the other proteins were found in both inner- and outer-stem tissues. Closer examination of the GH16 family protein indicated that it was a xyloglucan endo-transglycosylase/hydrolase (XTH). This enzyme cuts (hydrolase activity) xyloglucan hemicellulose polymers linking cellulose

microfibrils and is believed to favor cell expansion during cell growth although recent work has cast doubt upon this generally accepted idea (31)(72)(73)(74). Subsequently, the XTH relinks (endotransglycosylation activity) the cut end of the xyloglucan to another xyloglucan polymer thereby relinking cellulose microfibrils and contributing to stopping cell expansion (72)(50).

As well as being characterized by thick, cellulose-rich cell walls, flax fibers are also extremely long cells and can reach lengths of up to 70 mm. Their development therefore requires careful coordination between cell loosening mechanisms allowing rapid cell expansion and cell wall polymer biosynthesis necessary for secondary cell wall formation (75)(24). Both XTH activity and gene expression have previously been correlated with G-fiber formation in tension wood in poplar and, given the presence of xyloglucans in the flax fiber cell wall, it is perhaps not surprising to an XTH in fiber-containing tissues in flax (76)(50)(77)(24). In addition *XTH* gene expression has also been correlated with fiber-containing outer-stem tissues in flax in a number of previous studies (78)(79)(62)(4). Interestingly, the flax XTH gene family contains 61 predicted genes and is much larger than in many other species such as arabidopsis (33), poplar (41) and tomato (25) (3)(51). Based on enzymatic and phylogenetic studies plant XTHs are divided into 2 main groups, the class I/II and the class IIIA (72). Class I/II XTHs show xyloglucan endotransglycosylase activity (XET; EC 2.4.1.207) and are classically believed to be involved in cutting/rejoining xyloglucans during cell extension processes (80)(51). In contrast class IIIA XTHs show xyloglucan endo-hydrolase activity (XEH; EC 3.2.1.115) and are involved in cutting/degrading xyloglucans, but show no/little transglycosylation activity. The class IIIA includes the archetypal nasturtium (*Tropaeolum majus*) xyloglucanase1 TmNXG1 that hydrolyzes seed storage xyloglucan during germination (81). The flax XTH protein identified

in our studies belongs to the class IIIA and is therefore likely to show xyloglucan endo-hydrolase activity rather than a transglycosylation activity. Interestingly, phylogenetic analyses of the class IIIA family in flax revealed an important paralogy with 10 flax class IIIA XTHs as compared to 3 in both arabidopsis and poplar. Although flax has undergone 2 whole genome duplications during its evolution, it is unlikely that the observed XTH parlogy is related to these whole genome events since we did not observe comparable duplications in the other cell wall families apart from that previously noted for the GH19 family containing chitinase-like (CTL) genes (3)(6)(49). The flax genes are orthologs of AtXTH31 and AtXTH32 that are expressed in regions of plant bodies undergoing expansion and for which (ATXTH31 only) endo-hydrolase activity has been experimentally demonstrated (80).

XTH activity in aspen tension wood fibers is proposed to be necessary for both cell elongation and for maintaining/repairing xyloglucan links between the S layers and G-layer thereby contributing to fiber cell wall structural integrity (24)(82). Our results identifying xylans, as well as ferulic acid residues in the flax fiber cell wall could also suggest that xylan hemicelluloses contribute to cell wall integrity in these cells. In addition it has also been suggested that XTHs contribute to trimming xyloglucans attached to newly formed cellulose microfibrils facilitating the assembly of cellulose macrofibrils (83). However, unlike tension wood fibers, the major NCP of the S2/G-layer in flax is not xyloglucan, but a beta-1,4 glucan (24). Although such an observation could suggest that the flax XTHs are more likely involved in the cell elongation process it is important to keep in mind recent studies casting doubt upon the role of xyloglucans in cell expansion (73)(74). Alternatively, it is possible that the flax XTH is part of a xyloglucan recycling pathway as has been suggested in arabidopsis (84)(80). Overall these results highlight the likely important role of xyloglucans and XTHs in flax fiber cell wall metabolism. The availability of flax mutant populations and the recent development

of a flax VIGS system open the door to future functional studies that should allow us to get a better understanding of these processes (7)(9).

References

1. Singh, K. K., Mridula, D., Rehal, J., and Barnwal, P. (2011) Flaxseed: A Potential Source of Food, Feed and Fiber. *Crit. Rev. Food Sci. Nutr.* 51, 210–222
2. Touré, A., and Xueming, X. (2010) Flaxseed lignans: Source, biosynthesis, metabolism, antioxidant activity, Bio-active components, and health benefits. *Compr. Rev. Food Sci. Food Saf.* 9, 261–269
3. Wang, Z., Hobson, N., Galindo, L., Zhu, S., Shi, D., McDill, J., Yang, L., Hawkins, S., Neutelings, G., Datla, R., Lambert, G., Galbraith, D. W., Grassa, C. J., Geraldes, A., Cronk, Q. C., Cullis, C., Dash, P. K., Kumar, P. A., Cloutier, S., Sharpe, A. G., Wong, G. K.-S., Wang, J., and Deyholos, M. K. (2012) The genome of flax (*Linum usitatissimum*) assembled *de novo* from short shotgun sequence reads. *Plant J.* 72, 461–473
4. Huis, R., Morreel, K., Fliniaux, O., Lucau-Danila, A., Fenart, S., Grec, S., Neutelings, G., Chabbert, B., Mesnard, F., Boerjan, W., and Hawkins, S. (2012) Natural Hypolignification Is Associated with Extensive Oligolignol Accumulation in Flax Stems. *PLANT Physiol.* 158, 1893–1915
5. Roach, M. J., Mokshina, N. Y., Badhan, A., Snegireva, A. V., Hobson, N., Deyholos, M. K., and Gorshkova, T. A. (2011) Development of Cellulosic Secondary Walls in Flax Fibers Requires -Galactosidase. *Plant Physiol.* 156, 1351–1363
6. Mokshina, N., Gorshkova, T., and Deyholos, M. K. (2014) Chitinase-like (CTL) and cellulose synthase (CESA) gene expression in gelatinous-type cellulosic walls of flax (Linum usitatissimum L.) bast fibers. *PLoS One* 9, e97949

7. Chantreau, M., Grec, S., Gutierrez, L., Dalmais, M., Pineau, C., Demailly, H., Paysant-Leroux, C., Tavernier, R., Trouvé, J.-P., Chatterjee, M., Guillot, X., Brunaud, V., Chabbert, B., Van Wuytswinkel, O., Bendahmane, A., Thomasset, B., and Hawkins, S. (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,
8. Chantreau, M., Portelette, A., Dauwe, R., Kiyoto, S., Crônier, D., Morreel, K., Arribat, S., Neutelings, G., Chabi, M., Boerjan, W., Yoshinaga, A., Mesnard, F., Grec, S., Chabbert, B., and Hawkins, S. (2014) 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 Online* 26, 4462–4482
9. Chantreau, M., Chabbert, B., Billiard, S., Hawkins, S., and Neutelings, G. (2015) Functional analyses of cellulose synthase genes in flax (*Linum usitatissimum*) by virus-induced gene silencing. *Plant Biotechnol. J.* 13, 1312–1324
10. Ibragimova, N. N., Mokshina, N. E., and Gorshkova, T. A. (2012) Cell Wall Proteins of Flax Phloem Fibers. *Russ. J. Bioorganic Chem.* 38, 117–125
11. Hotte, N. S., and Deyholos, M. K. BMC Plant Biology A flax fibre proteome: identification of proteins enriched in bast fibres.
12. Barvkar, V. T., Pardeshi, V. C., Kale, S. M., Kadoo, N. Y., Giri, A. P., and Gupta, V. S. (2012) Proteome profiling of flax (Linum usitatissimum) seed: Characterization of functional metabolic pathways operating during seed development. *J. Proteome Res.* 11, 6264–6276
13. Day, A., Fénart, S., Neutelings, G., Hawkins, S., Rolando, C., and Tokarski, C. (2013) Identification of cell wall proteins in the flax (Linum usitatissimum) stem. *Proteomics* 13, 812–825

14. Haider, S., and Pal, R. (2013) Integrated analysis of transcriptomic and proteomic data. *Curr. Genomics* 14, 91–110

15. Rossignol, M., Peltier, J. B., Mock, H. P., Matros, A., Maldonado, A. M., and Jorrín, J. V. (2006) Plant proteome analysis: A 2004-2006 update. *Proteomics* 6, 5529–5548

16. Cui, S., Hu, J., Guo, S., Wang, J., Cheng, Y., Dang, X., Wu, L., and He, Y. (2012) Proteome analysis of Physcomitrella patens exposed to progressive dehydration and rehydration. *J. Exp. Bot.* 63, 711–726

17. 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

18. 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

19. Boerjan, W., Ralph, J., and Baucher, M. (2003) Lignin biosynthesis. *Annu. Rev. Plant Biol.* 54, 519–46

20. McFarlane, H. E., Döring, A., and Persson, S. (2014) The cell biology of cellulose synthesis. *Annu. Rev. Plant Biol.* 65, 69–94

21. Song, D., Xi, W., Shen, J., Bi, T., and Li, L. (2011) Characterization of the plasma membrane proteins and receptor-like kinases associated with secondary vascular differentiation in poplar. *Plant Mol. Biol.* 76, 97–115

22. van Hazendonk, J. M., Reinerik, E. J. M., de Waard, P., and van Dam, J. E. G. (1996) Structural analysis of acetylated hemicellulose polysaccharides from fibre flax (Linum usitatissimum L.). *Carbohydr. Res.* 291, 141–154

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

24. Mellerowicz, E. J., and Gorshkova, T. A. (2012) Tensional stress generation in gelatinous fibres: A review and possible mechanism based on cell-wall structure and composition. *J. Exp. Bot.* 63, 551–565

25. Davis, E. A., Derouet, C., Herve Du Penhoat, C., and Morvan, C. (1990) Isolation and an N.M.R. study of pectins from flax (linum usitatissimum L.). *Carbohydr. Res.* 197, 205–215

26. Girault, R., Bert, F., Rihouey, C., Jauneau, A., Morvan, C., and Jarvis, M. (1997) Galactans and cellulose in flax fibres: Putative contributions to the tensile strength. *Int. J. Biol. Macromol.* 21, 179–188

27. Gorshkova, T., and Morvan, C. (2006) Secondary cell-wall assembly in flax phloem fibres: Role of galactans. *Planta* 223, 149–158

28. Mcdougall, G. J. (1993) Isolation and partial characterisation of the non-cellulosic polysaccharides of flax fibre. *Carbohydr. Res.* 241, 227–236

29. Sharma, H. S. S., Faughey, G., and Lyons, G. (1999) Comparison of Physical , Chemical , and Thermal Characteristics of Water- , Dew- , and. 139–143

30. Carpita, N. C., and McCann, M. C. (2000) in *Biochemistry and Molecular Biology of Plants*, pp 52–109.

31. Cosgrove, D. J. (2005) Growth of the plant cell wall. *Nat. Rev. Mol. Cell Biol.* 6, 850–61

32. Murashige, T., and Skoog, F. (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant.* 15, 473–497

33. Goulas, E., Dily, F. Le, Teissedre, L., Corbel, G., Robin, C., and Ourry, A. (2001) Vegetative Storage Proteins in White Clover (Trifolium repens L.): Quantitative and Qualitative Features. *Ann. Bot.* 88, 789–795

34. Feiz, L., Irshad, M., Pont-Lezica, R. F., Canut, H., and Jamet, E. (2006) Evaluation of

cell wall preparations for proteomics: a new procedure for purifying cell walls from Arabidopsis hypocotyls. *Plant Methods* 2, 10

35. Thimm, O., Bläsing, O., Gibon, Y., Nagel, A., Meyer, S., Krüger, P., Selbig, J., Müller, L. A., Rhee, S. Y., and Stitt, M. (2004) MAPMAN: A user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J.* 37, 914–939
36. Emanuelsson, O., Brunak, S., von Heijne, G., and Nielsen, H. (2007) Locating proteins in the cell using TargetP, SignalP and related tools. *Nat. Protoc.* 2, 953–971
37. Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M., and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42,
38. Oliva, M., Theiler, G., Zamocky, M., Koua, D., Margis-Pinheiro, M., Passardi, F., and Dunand, C. (2009) PeroxiBase: A powerful tool to collect and analyse peroxidase sequences from Viridiplantae. *J. Exp. Bot.* 60, 453–459
39. Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., Roth, A., Santos, A., Tsafou, K. P., Kuhn, M., Bork, P., Jensen, L. J., and Von Mering, C. (2015) STRING v10: Protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 43, D447–D452
40. Savitzky, A., and Golay, M. J. E. (1964) Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* 36, 1627–1639
41. Harper, J. F., Manney, L., DeWitt, N. D., Yoo, M. H., and Sussman, M. R. (1990) The Arabidopsis thaliana plasma membrane H(+)-ATPase multigene family. Genomic sequence and expression of a third isoform. *J. Biol. Chem.* 265, 13601–8
42. Franková, L., and Fry, S. C. (2013) Biochemistry and physiological roles of

enzymes that “cut and paste” plant cell-wall polysaccharides. *J. Exp. Bot.* 64, 3519–3550

43. Zeng, W., Jiang, N., Nadella, R., Killen, T. L., Nadella, V., and Faik, A. (2010) A glucurono(arabino)xylan synthase complex from wheat contains members of the GT43, GT47, and GT75 families and functions cooperatively. *Plant Physiol.* 154, 78–97
44. 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
45. 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, 1030–1042
46. Manabe, Y., Verhertbruggen, Y., Gille, S., Harholt, J., Chong, S.-L., Pawar, P. M., Mellerowicz, E. J., Tenkanen, M., Cheng, K., Pauly, M., and Scheller, H. V. (2013) Reduced Wall Acetylation proteins play vital and distinct roles in cell wall O-acetylation in Arabidopsis. *Plant Physiol.* 163, 1107–17
47. San Clemente, H., and Jamet, E. (2015) WallProtDB, a database resource for plant cell wall proteomics. *Plant Methods* 11, 2
48. Roudier, F., Fernandez, A. G., Fujita, M., Himmelspach, R., Borner, G. H. H., Schindelman, G., Song, S., Baskin, T. I., Dupree, P., Wasteneys, G. O., and Benfey, P. N. (2005) COBRA, an Arabidopsis extracellular glycosyl-phosphatidyl inositol-anchored protein, specifically controls highly anisotropic expansion through its involvement in cellulose microfibril orientation. *Plant Cell* 17, 1749–1763
49. Sveinsson, S., McDill, J., Wong, G. K. S., Li, J., Li, X., Deyholos, M. K., and Cronk, Q. C.

B. (2014) Phylogenetic pinpointing of a paleopolyploidy event within the flax genus (Linum) using transcriptomics. *Ann. Bot.* 113, 753–761

50. Nishikubo, N., Takahashi, J., Roos, A. A., Derba-Maceluch, M., Piens, K., Brumer, H., Teeri, T. T., Stålbrand, H., and Mellerowicz, E. J. (2011) Xyloglucan endo-transglycosylase-mediated xyloglucan rearrangements in developing wood of hybrid aspen. *Plant Physiol.* 155, 399–413
51. Shi, Y. Z., Zhu, X. F., Miller, J. G., Gregson, T., Zheng, S. J., and Fry, S. C. (2015) Distinct catalytic capacities of two aluminium-repressed Arabidopsis thaliana xyloglucan endotransglucosylase/hydrolases, XTH15 and XTH31, heterologously produced in Pichia. *Phytochemistry* 112, 160–169
52. Thomson, J. A. (1993) Molecular biology of xylan degradation. *FEMS Microbiol. Lett.* 104, 65–82
53. Wang, X., and Komatsu, S. (2015) Plant subcellular proteomics: Application for exploring optimal cell function in soybean. *J. Proteomics*,
54. Albenne, C., Canut, H., Hoffmann, L., and Jamet, E. (2014) Plant Cell Wall Proteins: A Large Body of Data, but What about Runaways? *Proteomes* 2, 224–242
55. Jamet, E., Albenne, C., Boudart, G., Irshad, M., Canut, H., and Pont-Lezica, R. (2008) Recent advances in plant cell wall proteomics. *Proteomics* 8, 893–908
56. Albenne, C., Canut, H., and Jamet, E. (2013) Plant cell wall proteomics: the leadership of Arabidopsis thaliana. *Front. Plant Sci.* 4, 111
57. Day, A., Fénart, S., Neutelings, G., Hawkins, S., Rolando, C., and Tokarski, C. (2013) Identification of cell wall proteins in the flax (*Linum usitatissimum*) stem. *Proteomics* 13, 812–825
58. Martinière, A., Gayral, P., Hawes, C., and Runions, J. (2011) Building bridges: Formin1 of Arabidopsis forms a connection between the cell wall and the actin

cytoskeleton. *Plant J.* 66, 354–365

59. Martinière, A., and Runions, J. (2013) Protein diffusion in plant cell plasma membranes: the cell-wall corral. *Front. Plant Sci.* 4,

60. Akita, K., Kobayashi, M., Sato, M., Kutsuna, N., Ueda, T., Toyooka, K., Nagata, N., Hasezawa, S., and Higaki, T. (2016) Cell wall accumulation of fluorescent proteins derived from a trans-Golgi cisternal membrane marker and paramural bodies in interdigitated Arabidopsis leaf epidermal cells. *Protoplasma*, 1–11

61. Rejab, N. A., Nakano, Y., Yoneda, A., Ohtani, M., and Demura, T. (2015) Possible contribution of TED6 and TED7, secondary cell wall-related membrane proteins, to evolution of tracheary element in angiosperm lineage. *Plant Biotechnol.* 32, 343–347

62. Fenart, S., Ndong, Y.-P. A., Duarte, J., Rivière, N., Wilmer, J., van Wuytswinkel, O., Lucau, A., Cariou, E., Neutelings, G., Gutierrez, L., Chabbert, B., Guillot, X., Tavernier, R., Hawkins, S., and Thomasset, B. (2010) Development and validation of a flax (Linum usitatissimum L.) gene expression oligo microarray. *BMC Genomics* 11, 592

63. Venglat, P., Xiang, D., Qiu, S., Stone, S. L., Tibiche, C., Cram, D., Alting-Mees, M., Nowak, J., Cloutier, S., Deyholos, M., Bekkaoui, F., Sharpe, A., Wang, E., Rowland, G., Selvaraj, G., and Datla, R. (2011) Gene expression analysis of flax seed development. *BMC Plant Biol.* 11, 74

64. Dixon, R. A., and Paiva, N. L. (1995) Stress-Induced Phenylpropanoid Metabolism. *Plant Cell* 7, 1085–1097

65. Burn, J. E., Hurley, U. A., Birch, R. J., Arioli, T., Cork, A., and Williamson, R. E. (2002) The cellulose-deficient Arabidopsis mutant rsw3 is defective in a gene encoding a putative glucosidase II, an enzyme processing N-glycans during ER quality control.

Plant J. 32, 949–960

66. Ishiguro, S., Watanabe, Y., Ito, N., Nonaka, H., Takeda, N., Sakai, T., Kanaya, H., Okada, K., Aeschbacher, R., Schiefelbein, J., Benfey, P., Akama, K., Shiraishi, H., Ohta, S., Nakamura, K., Okada, K., Shimura, Y., Argon, Y., Simen, B., Bechtold, N., Ellis, J., Pelletier, G., Bisseling, T., Brand, U., Fletcher, J., Hobe, M., Meyerowitz, E., Simon, R., Brand, U., Hobe, M., Simon, R., Buchner, J., Clark, S., Clark, S., Running, M., Meyerowitz, E., Clark, S., Running, M., Meyerowitz, E., Clark, S., Williams, R., Meyerowitz, E., Cock, J., McCormick, S., Cutforth, T., Rubin, G., Doerner, P., Fletcher, J., Brand, U., Running, M., Simon, R., Meyerowitz, E., Fukaki, H., Wysocka-Diller, J., Kato, T., Fujisawa, H., Benfey, P., Tasaka, M., Gething, M., Sambrook, J., Ishiguro, S., Kawai-Oda, A., Ueda, J., Nishida, I., Okada, K., Jeong, S., Trotochaud, A., Clark, S., Kaufman, R., Kaya, H., Shibahara, K., Taoka, K., Iwabuchi, M., Stillman, B., Araki, T., Kayes, J., Clark, S., Koch, G., Koornneef, M., Eden, J. van, Hanhart, C., Stam, P., Braaksma, F., Feensra, W., Krishna, P., Gloor, G., Laufs, P., Grandjean, O., Jonak, C., Kiêu, K., Traas, J., Laux, T., Mayer, K., Berger, J., Jürgens, G., Lee, A., Lee, A., Liu, Y., Mitsukawa, N., Oosumi, T., Whittier, R., Long, J., Barton, M., Macer, D., Koch, G., Mayer, K., Schoof, H., Haecker, A., Lenhard, M., Jürgens, G., Laux, T., Mayer, M., Bukau, B., Melnick, J., Dul, J., Argon, Y., Nicchitta, C., Normington, K., Kohno, K., Kozutsumi, Y., Gething, M., Sambrook, J., Ohta, S., Mita, S., Hattori, T., Nakamura, K., Okada, K., Shimura, Y., Pratt, W., Rose, M., Misra, L., Vogel, J., Running, M., Hake, S., Rutherford, S., Lindquist, S., Schoof, H., Lenhard, M., Haecker, A., Mayer, K., Jürgens, G., Laux, T., Schröder, G., Beck, M., Eichel, J., Vetter, H., Schröder, J., Tachikawa, H., Miura, T., Katakura, Y., Mizunaga, T., Takahashi, T., Naito, S., Komeda, Y., Torii, K., Trotochaud, A., Hao, T., Wu, G., Yang, Z., Clark, S., Trotochaud, A., Jeong, S., Clark, S., Berg, C. van den, Weisbeek, P., Scheres, B., Waites, R., Simon, R., Walther-Larsen,

H., Brandt, J., Collinge, D., Thordal-Christensen, H., Yabe, N., Takahashi, T., Komeda, Y., Yoshida, H., Haze, K., Yanagi, H., Yura, T., and Mori, K. (2002) SHEPHERD is the Arabidopsis GRP94 responsible for the formation of functional CLAVATA proteins. *EMBO J.* 21, 898–908

67. Ray, S., Paynel, F., Morvan, C., Lerouge, P., Driouich, A., and Ray, B. (2013) Characterization of mucilage polysaccharides, arabinogalactanproteins and cell-wall hemicellulosic polysaccharides isolated from flax seed meal: A wealth of structural moieties. *Carbohydr. Polym.* 93, 651–660
68. Obel, N., Erben, V., Schwarz, T., K??hnel, S., Fodor, A., and Pauly, M. (2009) Microanalysis of plant cell wall polysaccharides. *Mol. Plant* 2, 922–932
69. Rennie, E. A., and Scheller, H. V. (2014) Xylan biosynthesis. *Curr. Opin. Biotechnol.* 26, 100–107
70. Scheller, H. V., and Ulvskov, P. (2010) Hemicelluloses. *Annu. Rev. Plant Biol.* 61, 263–289
71. Pauly, M., Gille, S., Liu, L., Mansoori, N., de Souza, A., Schultink, A., and Xiong, G. (2013) Hemicellulose biosynthesis. *Planta* 238, 627–642
72. Eklöf, J. M., and Brumer, H. (2010) The XTH gene family: an update on enzyme structure, function, and phylogeny in xyloglucan remodeling. *Plant Physiol.* 153, 456–466
73. Park, Y. B., and Cosgrove, D. J. (2012) A revised architecture of primary cell walls based on biomechanical changes induced by substrate-specific endoglucanases. *Plant Physiol.* 158, 1933–1943
74. Park, Y. B., and Cosgrove, D. J. (2015) Xyloglucan and its interactions with other components of the growing cell wall. *Plant Cell Physiol.* 56, 180–194
75. Snegireva, a. V., Ageeva, M. V., Amenitskii, S. I., Chernova, T. E., Ebskamp, M., and

Gorshkova, T. a. (2010) Intrusive growth of sclerenchyma fibers. *Russ. J. Plant Physiol.* 57, 342–355

76. Nishikubo, N., Awano, T., Banasiak, A., Bourquin, V., Ibatullin, F., Funada, R., Brumer, H., Teeri, T. T., Hayashi, T., Sundberg, B., and Mellerowicz, E. J. (2007) Xyloglucan endo-transglycosylase (XET) functions in gelatinous layers of tension wood fibers in poplar - A glimpse into the mechanism of the balancing act of trees. *Plant Cell Physiol.* 48, 843–855

77. Baba, K., Park, Y. W., Kaku, T., Kaida, R., Takeuchi, M., Yoshida, M., Hosoo, Y., Ojio, Y., Okuyama, T., Taniguchi, T., Ohmiya, Y., Kondo, T., Shani, Z., Shoseyov, O., Awano, T., Serada, S., Norioka, N., Norioka, S., and Hayashi, T. (2009) Xyloglucan for generating tensile stress to bend tree stem. *Mol. Plant* 2, 893–903

78. Day, A., Addi, M., Kim, W., David, H., Bert, F., Mesnage, P., Rolande, C., Chabbert, B., Neutelings, G., and Hawkins, S. (2005) ESTs from the fibre-bearing stem tissues of flax (Linum usitatissimum L.): Expression analyses of sequences related to cell wall development. *Plant Biol.* 7, 23–32

79. Roach, M. J., and Deyholos, M. K. (2007) Microarray analysis of flax (Linum usitatissimum L.) stems identifies transcripts enriched in fibre-bearing phloem tissues. *Mol. Genet. Genomics* 278, 149–165

80. Kaewthai, N., Gendre, D., Eklof, J. M., Ibatullin, F. M., Ezcurra, I., Bhalerao, R. P., and Brumer, H. (2013) Group III-A XTH Genes of Arabidopsis Encode Predominant Xyloglucan Endohydrolases That Are Dispensable for Normal Growth. *PLANT Physiol.* 161, 440–454

81. Baumann, M. J., Eklöf, J. M., Michel, G., Kallas, A. M., Teeri, T. T., Czjzek, M., and Brumer, H. (2007) Structural evidence for the evolution of xyloglucanase activity from xyloglucan endo-transglycosylases: biological implications for cell wall

metabolism. *Plant Cell* 19, 1947–1963

82. Gorshkova, T., Mokshina, N., Chernova, T., Ibragimova, N., Salnikov, V., Mikshina, P., Tryfona, T., Banasiak, A., Immerzeel, P., Dupree, P., and Mellerowicz, E. J. (2015) Aspen tension wood fibers contain β-(1→4)-galactans and acidic arabinogalactans retained by cellulose cmcrofibrils in gelatinous walls. *Plant Physiol.* 169, 2048–2063

83. Mellerowicz, E. J., and Sundberg, B. (2008) Wood cell walls: biosynthesis, developmental dynamics and their implications for wood properties. *Curr. Opin. Plant Biol.* 11, 293–300

84. Gunl, M., Neumetzler, L., Kraemer, F., de Souza, a., Schultink, a., Pena, M., York, W. S., and Pauly, M. (2011) AXY8 Encodes an -Fucosidase, Underscoring the Importance of Apoplastic Metabolism on the Fine Structure of Arabidopsis Cell Wall Polysaccharides. *Plant Cell* 23, 4025–4040

Footnotes

This work was supported by the French Government (Ph.D. grant MC), the French Region Nord Pas-de-Calais and the Agence National de Recherche (ANR-10-KBBE-0003_01, project FIBRAGEN)

Figure Legends

Figure 1. Flax organs and tissues used for protein extraction. A. Hydroponic grown flax plants, B-E light microscope images of organ/tissue organization illustrating cell types. Bar = 5 cm (A), 50 μm (B, C, D), 100 μm (E). Five-μm semi-thin sections, stained with Toluidine Blue O (0.1% w/v water).

Figure 2. Repartition of total non-redundant proteins identified in 3 cellular compartments (A) and 4 plant organs/tissues (B). Fractions enriched in soluble proteins (S), membrane proteins (M) and cell wall proteins (C). Inner stem (IS), outer stem (OS), leaves (L) and roots (R).

Figure 3. Functional classification of flax proteins. (A) Histogram showing repartition of flax proteins from 6 different organs/tissues in Mapman functional classes and (B) Venn diagram illustrating repartition of housekeeping proteins (present in all 6 samples) in Mapman classes. IS: Inner stem; OS: Outer stem; WS: Whole stem; L: Leaves; R: Roots: S: Seeds.

Figure 4. Percentage distribution of identified flax proteins in different CAZy families. Overall percentages in grouped flax organs (A) and individual flax organs/tissues (B). GH: Glycoside Hydrolases, GT: GlycosylTransferases; PLs: Polysaccharide Lyases; CE: Carbohydrate Esterases; AA: Auxiliary Activities (AA); CBM: Carbohydrate-Binding Modules; EXPN: Expansin. IS: Inner stem; OS: Outer stem; WS: Whole stem; L: Leaves; R: Roots; S: Seeds.

Figure 5. Number of flax CAZy proteins (grouped organs) belonging to different sub-families. GH: Glycoside Hydrolases; GT: GlycosylTransferases; PL: Polysaccharide Lyases; CE: Carbohydrate Esterases; AA: Auxiliary Activities; CBM: Carbohydrate-Binding Modules; EXPN: Expansin. *CAZy families containing proteins involved in cell wall metabolism.

Figure 6. Number of flax CAZy proteins (individual organs) belonging to different subclasses. (A) Glycoside Hydrolases, (B) GlycosylTransferases; (C) Auxiliary Activities,

Carbohydrate-Binding Modules, Carbohydrate Esterases, Expansin, Polysaccharide Lyases. *CAZy families containing proteins involved in cell wall metabolism.

Figure 7. Repartition of predicted flax cell wall proteins classified into the 9 WallProtDB classes (47). (A) percentage proteins in each class (grouped organs), (B) number of proteins present in a particular WallProtDB class (individual organs) and (C) percentage protein in each class (individual organs). PAC: proteins acting on carbohydrates; OR: oxidoreductases; LM: proteins related to lipid metabolism; S: signaling; P: proteases; PID: proteins with interacting domains with proteins or polysaccharides; M: miscellaneous; SP: structural proteins; UF: unknown function.

Figure 8. Monolignol biosynthetic pathway indicating presence of the enzyme in a particular organ/tissue (colored box) or absence (no box). CAD: cinnamyl alcohol dehydrogenase; CCoAOMT: caffeoyl-CoA *O*-methyl transferase; CCR: cinnamoyl-CoA reductase; C3H: *p*-coumaroyl shikimate/quinate-3-hydroxylase; C4H: cinnamate 4-hydroxylase; CSE: Caffeoyl shikimate esterase; 4CL : 4-coumarate-CoA ligase; COMT: caffeic acid/5-hydroxyconiferaldehyde *O*-methyltransferase; F5H: ferulic acid/coniferaldehyde 5-hydroxylase; HCT: *p*-hydroxycinnamoyl-CoA: shikimate/quinate *p*-hydroxycinnamoyltransferase; PAL: phenylalanine ammonia-lyase.

Figure 9. Repartition of the 465 flax Total CWPs identified in 4 different organs (A) and stem tissues (B). Whole stem (WS), Seeds (S), leaves (L), roots (R), Inner stem (IS), Outer stem (OS).

Figure 10. Cell wall protein-protein interaction network clusters for flax inner stems (IS), outer stems (OS), whole stem (WS), leaves (L), roots (R) and seeds (S). Interactions were determined using STRING v10.0 (http://string-db.org/). Confidence view, stronger associations are represented by thicker lines. Clusters of lignin proteins (L), sugar precursor proteins (S), cellulose proteins (C), protein folding proteins (PF), defense proteins (D), seed proteins (SP), protease proteins (P) and pectin/hemicellulose proteins (P/H) are indicated by dotted lines.

Figure 11. Maximum likelihood phylogenetic tree of the xyloglucan endo-transglycosylase/hydrolase family IIIA created using a selection of plant species present in phytozome database v10. Branch support values obtained with 1000 bootstrap repetitions. The tree is rooted with the non-IIIA sequence. The scale bar shows the inferred number of amino acid substitutions per site. Protein sequences found in *Linum usitassinum*, *Populus thiocharpa*, and *Arabidopsis thaliana* are colored in red, blue and purple, respectively.

Figure 12. MALDI-TOF-MS spectra of fragment ions $(M+Na)^+$ corresponding to oligosaccharides released by enzymatic digestion of 2-month-old flax stems. Inner stem tissue sample digested by endoxylanase (A) or endo-beta-glucanase (B), outer stem tissue sample digested by endoxylanase (C) or endo-beta-glucanase (D). Figures refer to *m/z* values of xylan-derived fragments (759: X_4-4OMeGlcA + Na+, 781: X_4-4OMeGlcA + 2Na+ - H+, 891: X_5-4OMeGlcA + Na+) and xyloglucan-derived fragments (1085: XXXG + Na+, 1247: XLXG/XXLG + Na+, 1393: XXFG + Na+, 1595: XLFG + Na+).

Figure 13. Immunolocalization of flax stem hemicellulose epitopes. Pre-immune serum (A); anti-xylan LM10 antibodies (B); anti-xylan arabinoxylan LM11 antibodies (C); anti-

xyloglucan LM15 antibodies (D); anti-xyloglucan LM24 antibodies (E); anti-mannan/heteromannan LM21 antibodies (F). cz: cambial zone; ep: epidermis; f: fiber; x: xylem; r: xylem ray; v: xylem vessel/tracheid. Bar = 100 mm (except for zoom).

Figure 14. Middle infra-red spectra at 1800-900 cm^{-1} of bast fiber cell walls (A) and secondary derivative (B) showing significant peaks characteristic of hemicelluloses (C).

Tables

Table 1. List of CAZy families involved in cell wall polysaccharide, (glycol)protein and lignin metabolism identified in flax proteome. Families marked with an asterisk are potentially involved in hemicellulose metabolism. References: 1) Minic et al., 2008; 2) Frankova et al., 2014; 3) Derba-Maceluch et al., 2015; 4) Knoch et al., 2014; 5) Lim et al., 2001; 6) McFarlaine et al., 2014; 7) Scheller and Ulvskov, 2010; 8) Pauly et al., 2013; 9) Atmojodo et al., 2013; 10) Zeng et al., 2010. S = seed, IS = inner stem, OS = outer stem, WS = whole stem, L = leaf, R = root.

Table 2. Major protein-protein interaction clusters identified in different flax organs/tissues and shown in figure 10. IS: inner stem; OS: outer stem, WS: whole stem; L: leaves, R: roots; S: seeds.

Table 3. Numbers and percentages of connected and non-connected proteins in STRING network analyses. Non-connected cell wall = number/percentage of non-connected proteins related to cell wall polymer metabolism.

Table 4. Amino acid sequences in (bold text) and around the loop involved in the specific hydrolytic activity for the class IIIA XTH proteins of *L. usitassinum*, and *A. thaliana* (selected sequences).

Family	Enzyme Activity	Potential cell wall polymer	Reference	Organ
Glycoside hydrolases				
GH1*	β-glucosidase,	Glucan, cellulose, hemicellulose (xyloglucan)	1,2	All
GH3*	β-xylosidase, α-arabinofuranosidase	Hemicellulose (xylan)	1,2	All
GH5_7	β-mannase	Mannan	1,2	WS, S
GH9*	Glucanase	Glucan, Hemicellulose (xyloglucan)	1,2	IS, OS, WS
GH10_3*	Xylanase, Xylan endotransglycosylase	Hemicellulose (xylan)	1,2,3	S
GH16*	Xyloglucan endotransglycosylase/hydrolase	Hemicellulose (xyloglucan)	1,2	OS, R
GH17	glucanase	MLG, Callose	1,2	WS, L, R, S
GH28	polygalcturonase, pectinase	Pectin (homogalacturonan)	1,2	IS, WS
GH31*	alpha glucosidase, alpha-xylosidase	Hemicellulose (xyloglucan)	1,2	IS, OS, L, R, S
GH35	galactosidase	Galactan	1,2	IS, OS, R, S
GH38	α-mannosidase	glycoproteins	2	IS, OS, WS, R
GH51*	α-arabinofuranosidase	Hemicellulose (xylan)	1,2	OS, WS
GH79_2	glucuronidase	AGP	4	OS
GH95*	α-fucosidase	Hemicellulose (xyloglucan), AGPs	2,4	S
Glycosyltransferases				
GT1	UDP-glycosyltransferase	Lignin (monolignols)	5	IS, S
GT2*	Cellulose synthase, cellulose-synthase-like (mannan synthase, glucomannan synthase, glucan synthase), galactosyltransferase	Cellulose, hemicellulose : xyloglucan, (Gluco)mannan, MLG	6,7	IS
GT4	Sucrose Synthase	cellulose	6	IS, OS, WS, R, S
GT8*	Galacturonosyltransferase, glucuronosyltransferase ; Galacturonosyltransferase,	Pectin (HG), hemicellulose : xylan	7,8,9	IS, OS, WS
GT47*	Arabinosyltransferase, Xylosyltransferase, galactosyl transferase, galacturonosyltransferase	Hemicellulose : Pectin (RGI), xyloglucan, xylan	7,8,9	IS, OS
GT48	Glucan synthase	callose	8	S
GT75*	glucuronosyltransferase	Hemicellulose (xylan)	10	WS, S
GT77	Xylosyltransferase, Arabinosyltransferase,	Pectin (RGII), AGP	4,9	IS
Auxillary activity				
AA1	Phenoloxidase/Laccase	Lignin	11	IS, OS, S

AA2	Phenoloxidase/Peroxidase	Lignin		All
Carbohydrate esterases and Polysaccharide lyases				
CE8	Pectin methylesterase	pectin	2	OS, WS, L, R, S
CE13	Pectin acetylesterase	pectin	2	OS, R
PL4_2	Pectate lyase	pectin		OS
Carbohydrate binding modules and Expansin				
CBM22*	Xylanase	Hemicellulose (xylan)	12	WS
CBM43	Glucanse	MLG, Callose	13	WS, S
EXPN	No enzyme activity	Cellulose-hemicellulose network	14	WS, L, S

Table 1. List of CAZy families involved in cell wall polysaccharide, (glycol)protein and lignin metabolism identified in flax proteome. Families marked with an asterisk are potentially involved in hemicellulose metabolism. References: 1) Minic et al., 2008; 2) Frankova et al., 2014; 3) Derba-Maceluch et al., 2015; 4) Knoch et al., 2014; 5) Lim et al., 2001; 6) McFarlaine et al., 2014; 7) Scheller and Ulvskov, 2010; 8) Pauly et al., 2013; 9) Atmojodo et al., 2013; 10) Zeng et al., 2010. S = seed, IS = inner stem, OS = outer stem, WS = whole stem, L = leaf, R = root.

Table 2. Major protein-protein interaction clusters identified in different flax organs/tissues and shown in figure 10. IS: inner stem; OS: outer stem, WS: whole stem; L: leaves, R: roots; S: seeds.

	Organs/tissues					
Cluster	**IS**	**OS**	**WS**	**L**	**R**	**S**
Lignin (L)	✔	✔	✔			✔
Sugar precursors (S)	✔	✔	✔			✔
Cellulose (C)	✔					
Protein folding (PF)	✔	✔	✔	✔	✔	✔
Defense (D)			✔	✔	✔	
Seed proteins (SP)						✔
Proteases (P)	✔				✔	✔
Pectin / hemicellulose (P/H)		✔				

Table 3. Numbers and percentages of connected and non-connected proteins in STRING network analyses. Non-connected cell wall = number/percentage of non-connected proteins related to cell wall polymer metabolism.

Sample	Total String	Connected string	Non-connected	Non-connected cell wall
IS	98	63	35 (35.7 %)	8 (22.8 %)
OS	108	71	37 (34.3 %)	10 (27 %)
WS	111	68	43 (38.7 %)	15 (34.8 %)
L	50	27	23 (46 %)	8 (34.8 %)
R	77	38	39 (50.6 %)	15 (38.5 %)
S	145	86	59 (40.7 %)	16 (27.1 %)
ALL	310	218	92 (29.7 %)	25 (27.2 %)

Table 4. Amino acid sequences in (bold text) and around the loop involved in the specific hydrolytic activity for the class IIIA XTH proteins of *L. usitassinum*, and *A. thaliana* (selected sequences).

Gene ID	Sequence		
*A._thaliana*_19659993	QTNVFVRGSGD	**- R N V I G**	REMKFTLWFDP
*L._usitassium*_23180042	QTNVYVRGSGD	**- G N I I G**	REMQFHLWFDP
*L._usitassium*_23152616	QTNVYVRGSGD	**- G N I I G**	REMQFHLWFDP
*L._usitassium*_23171810	QTNVYVRGSGD	**- R N V V G**	REVQFHLWFDP
*L._usitassium*_23179194	QTNVYVRGSGD	**- R N I V G**	REVQFHLWFDP
*L._usitassium*_23159458	QTNVYVRGSGD	**G G D I I G**	REARFHLWFDP
*L._usitassium*_23164289	QTNVYVRGSGD	**G G D I I G**	REARFHLWFDP
*A._thaliana*_19639833	QTNVYIRGSGD	**- G K I I G**	REMKFRLWFDP
*L._usitassium*_23159481	QTNVYIRGSGD	**- G K I I G**	REMKFHLWFDP
*L._usitassium*_23164234	QTNVYIRGSGD	**- G K I I G**	REMKFHLWFDP
*L._usitassium*_23146127	QTNVYIRGSGD	**- G R I I G**	REMKFHLWFDP
*L._usitassium*_23169892	QTNVYIRGSGD	**- G R I I G**	REMKFHLWFDP
*L._usitassium*_23154089	QTNLYGNGSTH	**- - - - R G**	REERYSLWFDP
*A._thaliana*_19654973	QTNIYGNGSTH	**- - - - L G**	REERYNLWFDP
*L._usitassium*_23151771	QTNIYGNGSTN	**- - - - L G**	REERYSLWFDP

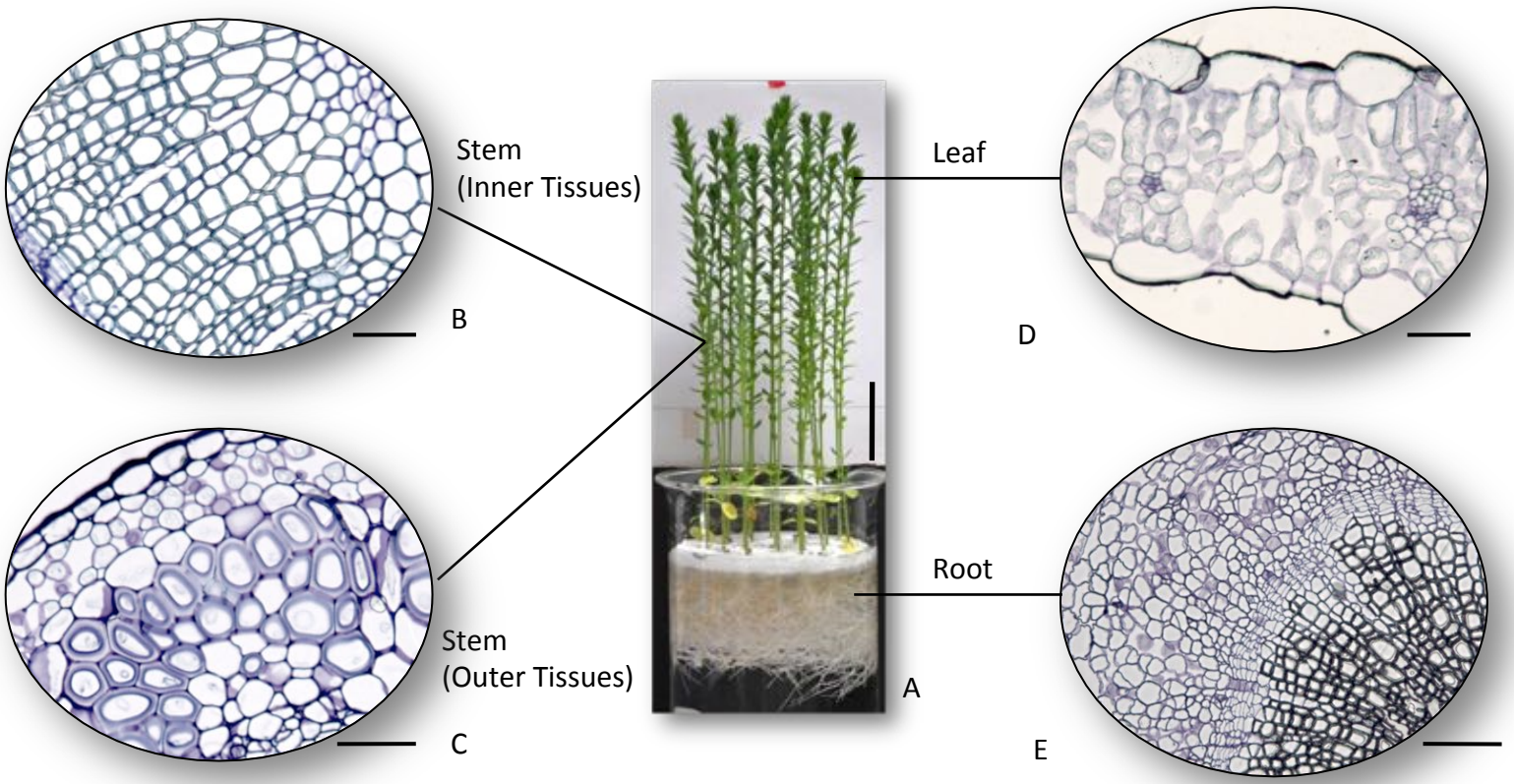


Figure 1. Flax organs and tissues used for protein extraction. A. Hydroponic grown flax plants, B-E light microscope images of organ/tissue organization illustrating cell types. Bar = 5 cm (A), 50 mm (B, C, D), 100 mm (E). Five-mm semi-thin sections, stained with Toluidine Blue O (0.1% w/v water).

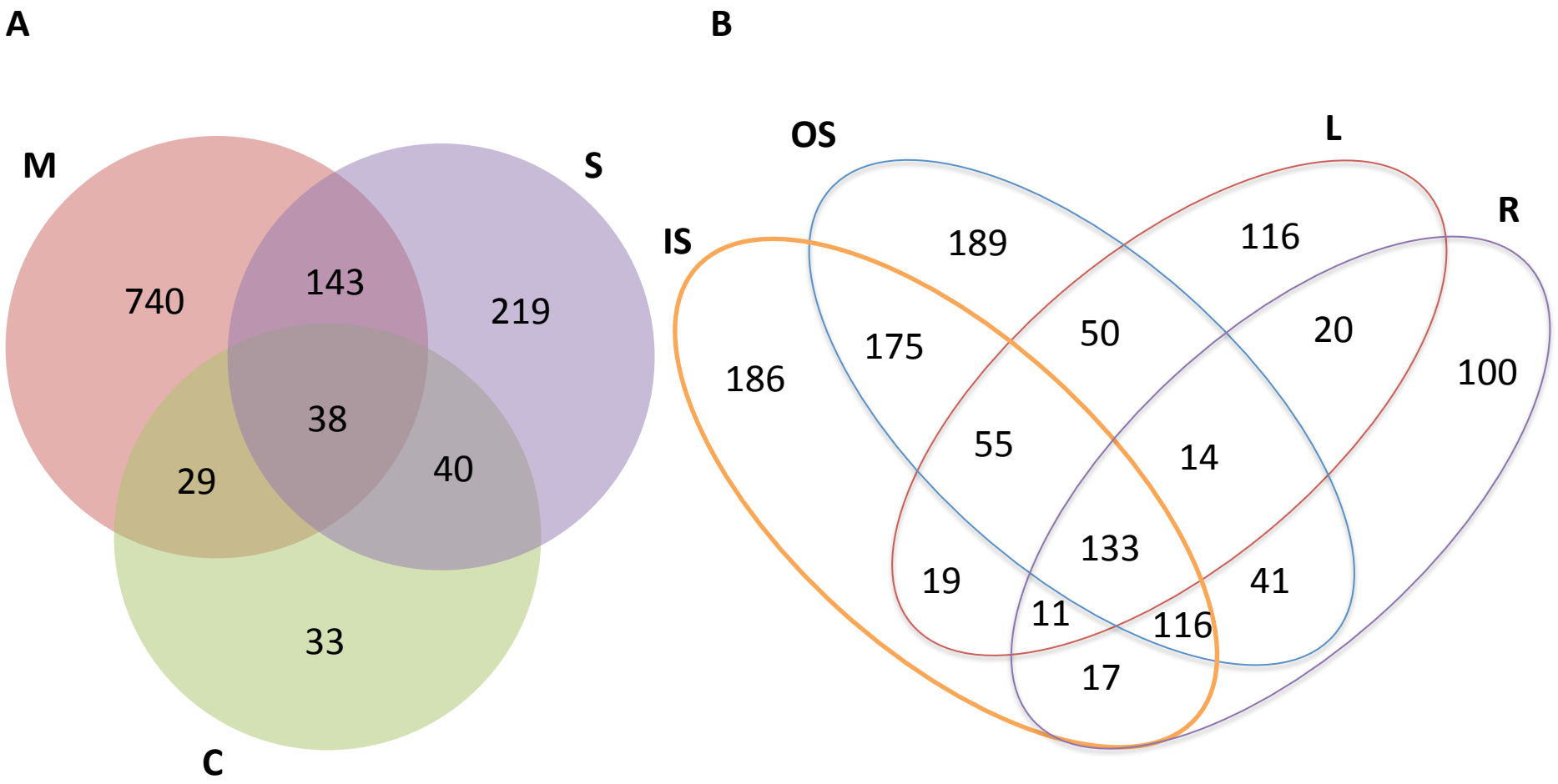


Figure 2. Repartition of total non-redundant proteins identified in 3 cellular compartments (A) and 4 plant organs/ tissues (B). Fractions enriched in soluble proteins (S), membrane proteins (M) and cell wall proteins (C). Inner stem (IS), outer stem (OS), leaves (L) and roots (R).

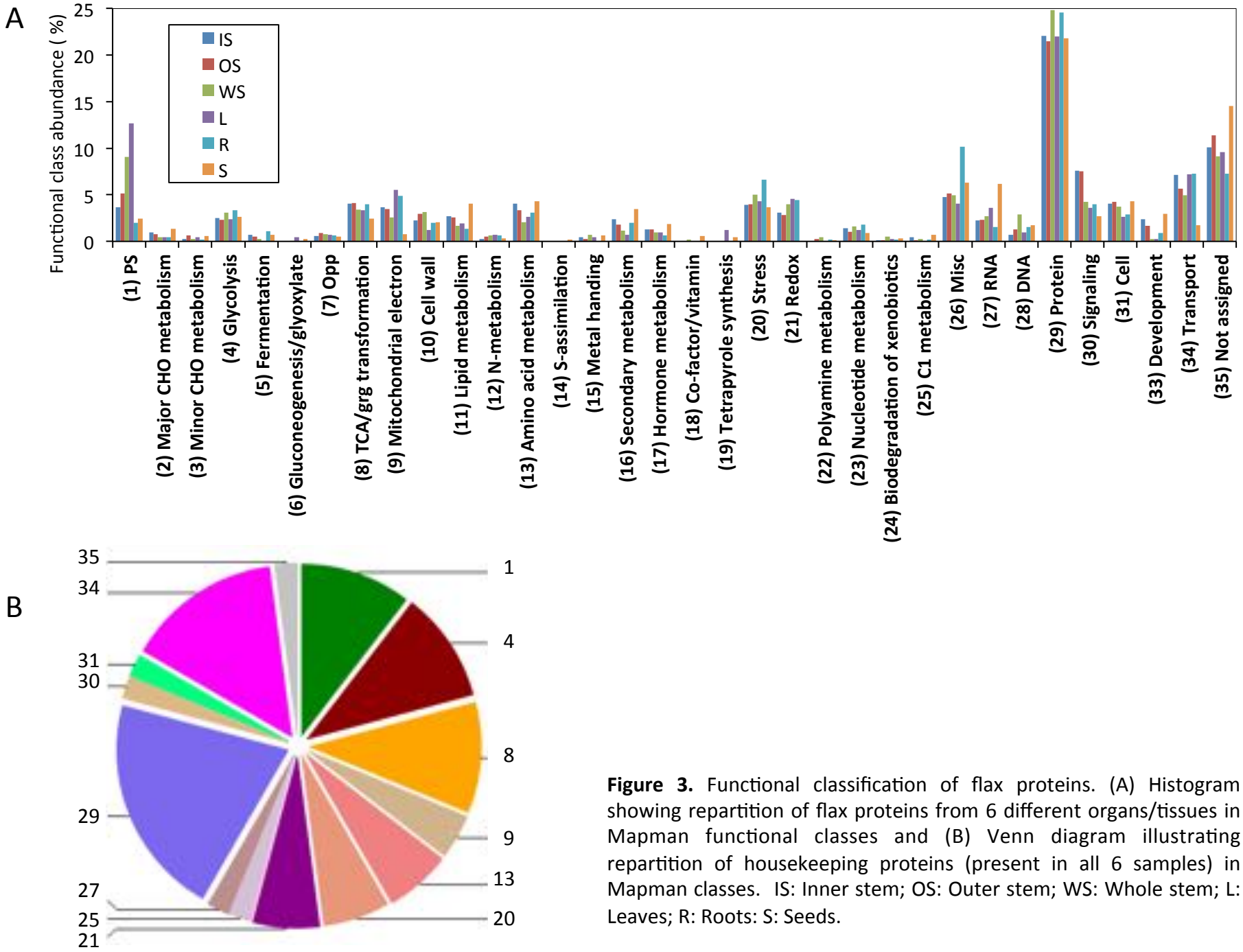
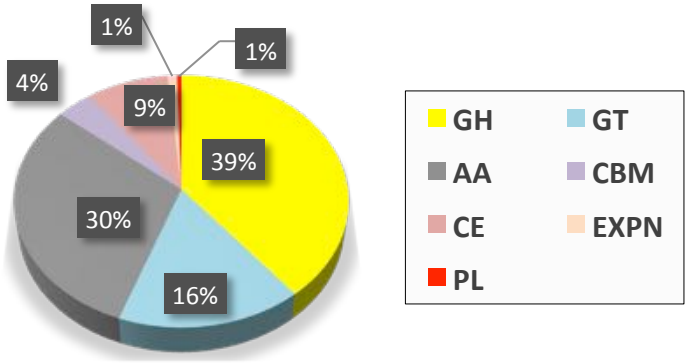


Figure 3. Functional classification of flax proteins. (A) Histogram showing repartition of flax proteins from 6 different organs/tissues in Mapman functional classes and (B) Venn diagram illustrating repartition of housekeeping proteins (present in all 6 samples) in Mapman classes. IS: Inner stem; OS: Outer stem; WS: Whole stem; L: Leaves; R: Roots: S: Seeds.

A



B

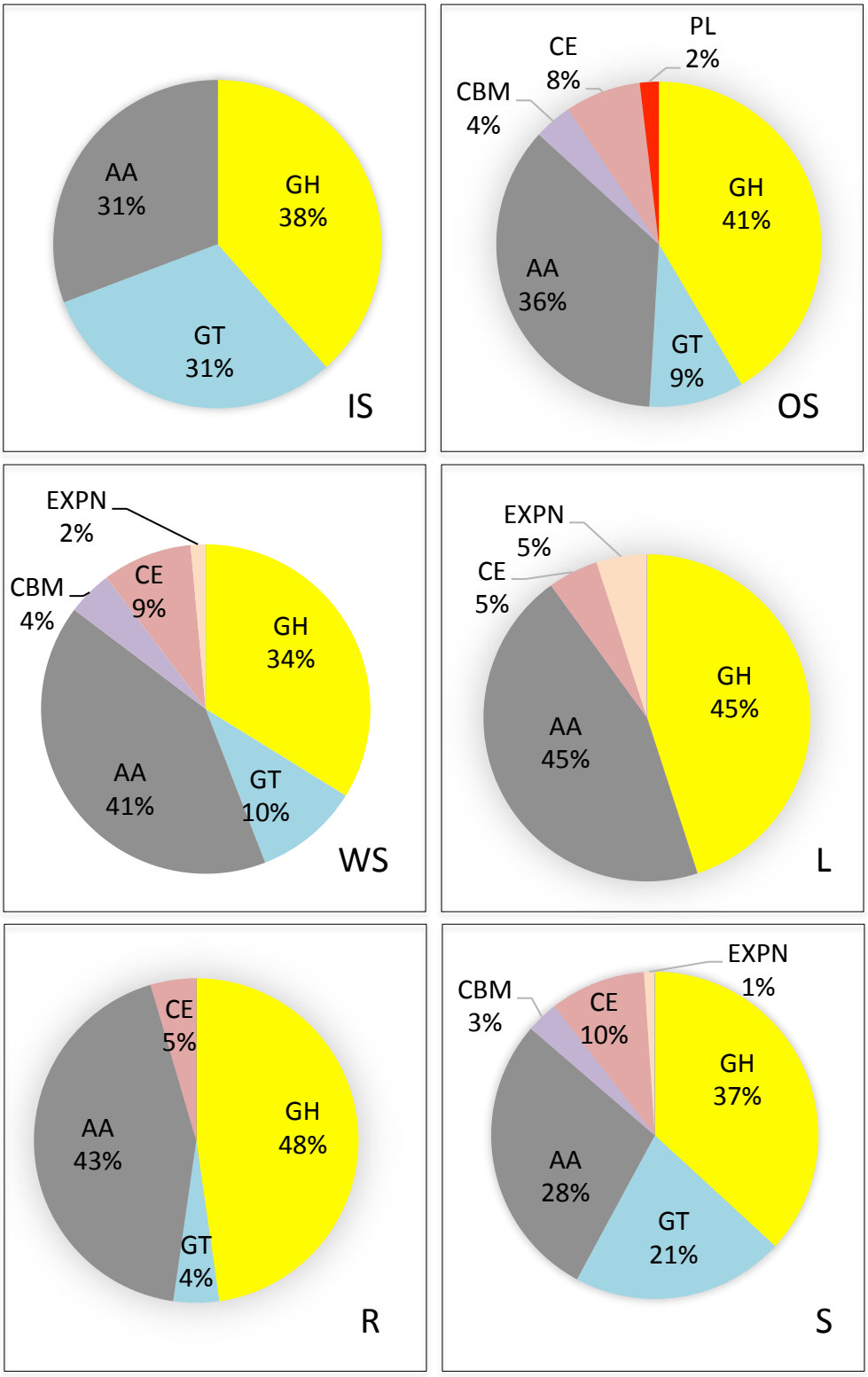


Figure 4. Percentage distribution of identified flax proteins in different CAZy families. Overall percentages in grouped flax organs (A) and individual flax organs/tissues (B). GH: Glycoside Hydrolases, GT: GlycosylTransferases; PLs: Polysaccharide Lyases; CE: Carbohydrate Esterases; AA: Auxiliary Activities (AA); CBM: Carbohydrate-Binding Modules; EXPN: Expansin. IS: Inner stem; OS: Outer stem; WS: Whole stem; L: Leaves; R: Roots; S: Seeds.

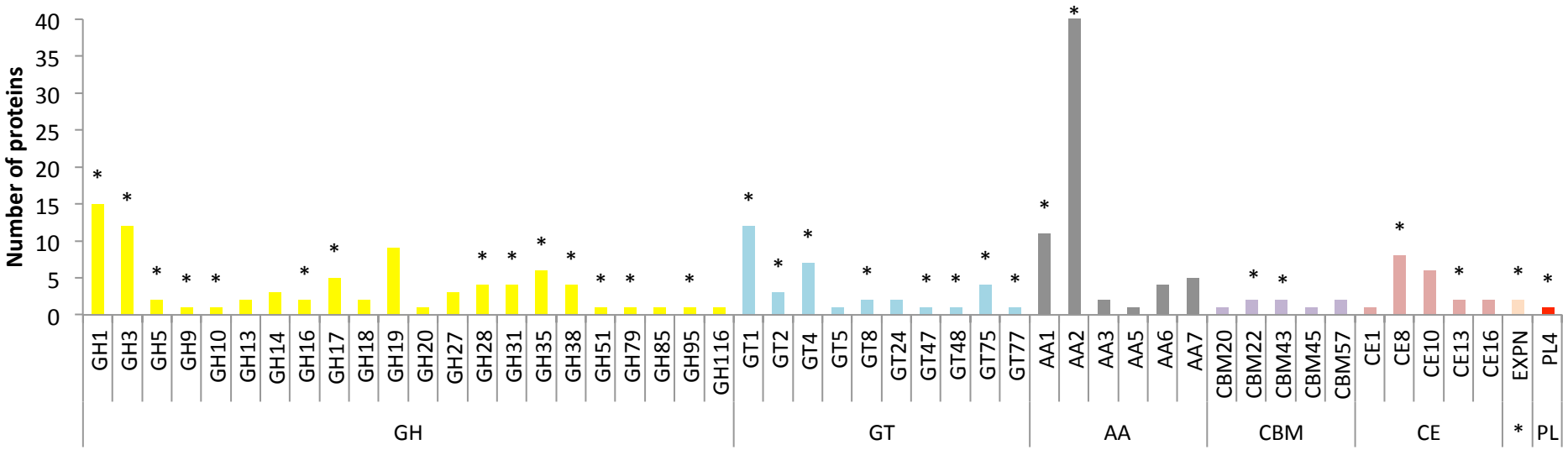


Figure 5. Number of flax CAZy proteins (grouped organs) belonging to different sub-families. GH: Glycoside Hydrolases; GT: GlycosylTransferases; PL: Polysaccharide Lyases; CE: Carbohydrate Esterases; AA: Auxiliary Activities; CBM: Carbohydrate-Binding Modules; EXPN: Expansin. *CAZy families containing proteins involved in cell wall metabolism.

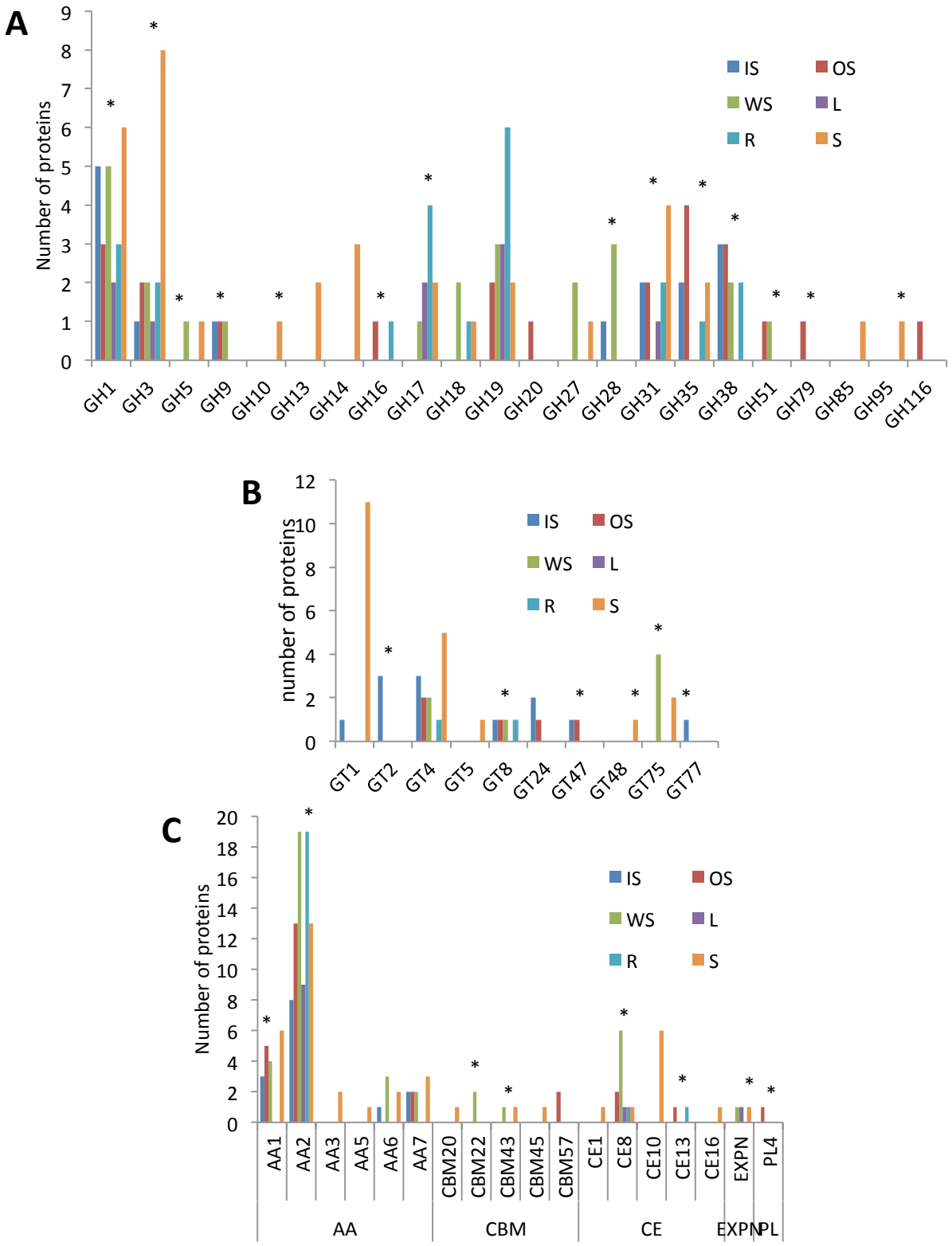


Figure 6. Number of flax CAZy proteins (individual organs) belonging to different subclasses. (A) Glycoside Hydrolases, (B) GlycosylTransferases; (C) Auxiliary Activities, Carbohydrate-Binding Modules, Carbohydrate Esterases, Expansin, Polysaccharide Lyases. *CAZy families containing proteins involved in cell wall metabolism.

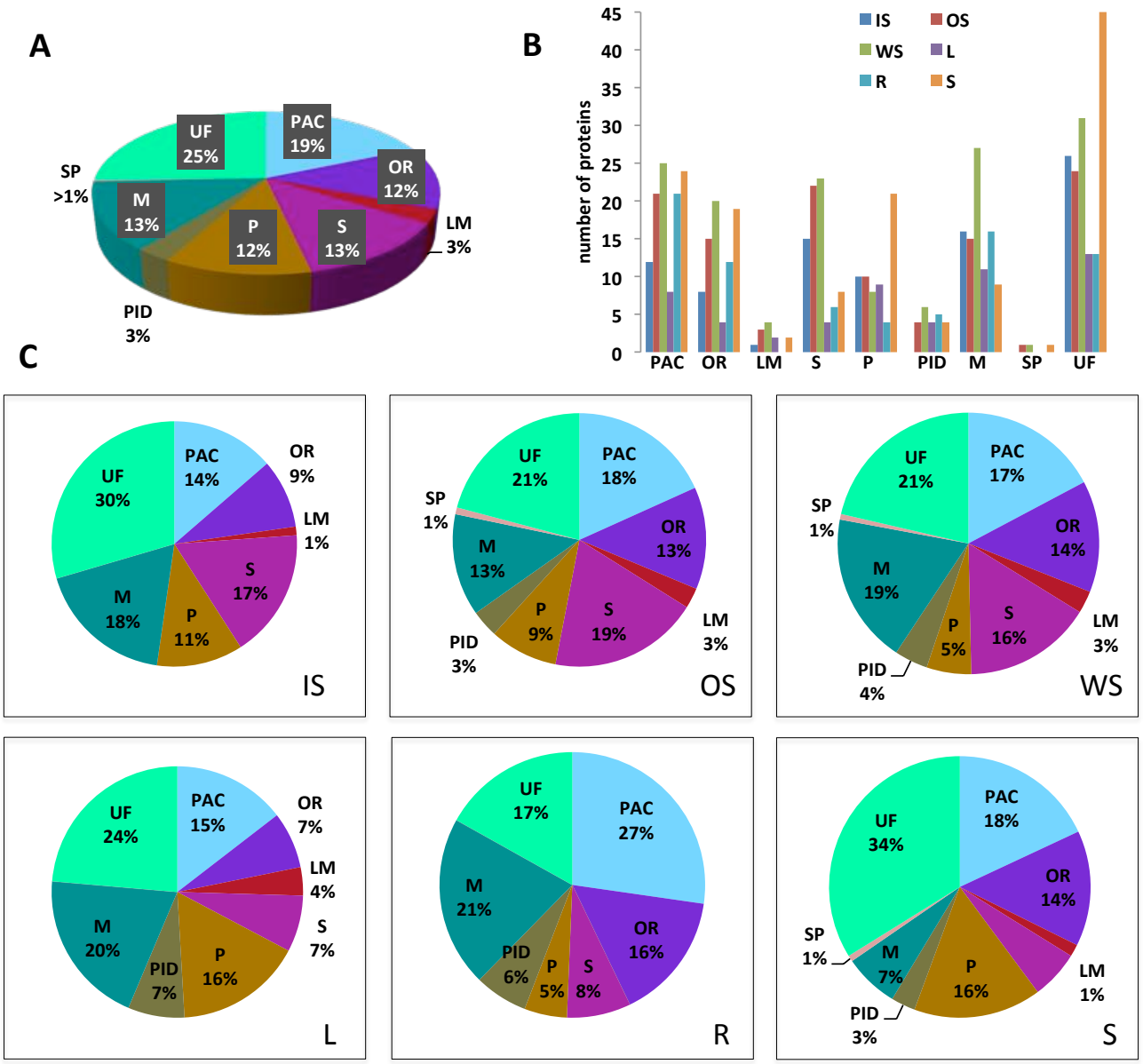

A
UF 25%
PAC 19%
OR 12%
LM 3%
S 13%
P 12%
M 13%
SP >1%
PID 3%
B
number of proteins
IS
OS
WS
L
R
S
PAC
OR
LM
S
P
PID
M
SP
UF
C
UF 30%
PAC 14%
OR 9%
LM 1%
S 17%
M 18%
P 11%
IS
UF 21%
PAC 18%
SP 1%
OR 13%
M 13%
P 9%
S 19%
LM 3%
PID 3%
OS
UF 21%
PAC 17%
SP 1%
OR 14%
M 19%
P 5%
S 16%
LM 3%
PID 4%
WS
UF 24%
PAC 15%
OR 7%
LM 4%
M 20%
S 7%
PID 7%
P 16%
L
UF 17%
PAC 27%
M 21%
OR 16%
PID 6%
P 5%
S 8%
R
UF 34%
PAC 18%
OR 14%
SP 1%
M 7%
LM 1%
P 16%
PID 3%
S

Figure 7. Repartition of predicted flax cell wall proteins classified into the 9 WallProtDB classes (San Clemente and Jamet, 2015). (A) percentage proteins in each class (grouped organs), (B) number of proteins present in a particular WallProtDB class (individual organs) and (C) percentage protein in each class (individual organs). PAC: proteins acting on carbohydrates; OR: oxidoreductases; LM: proteins related to lipid metabolism; S: signaling; P: proteases; PID: proteins with interacting domains with proteins or polysaccharides; M: miscellaneous; SP: structural proteins; UF: unknown function.

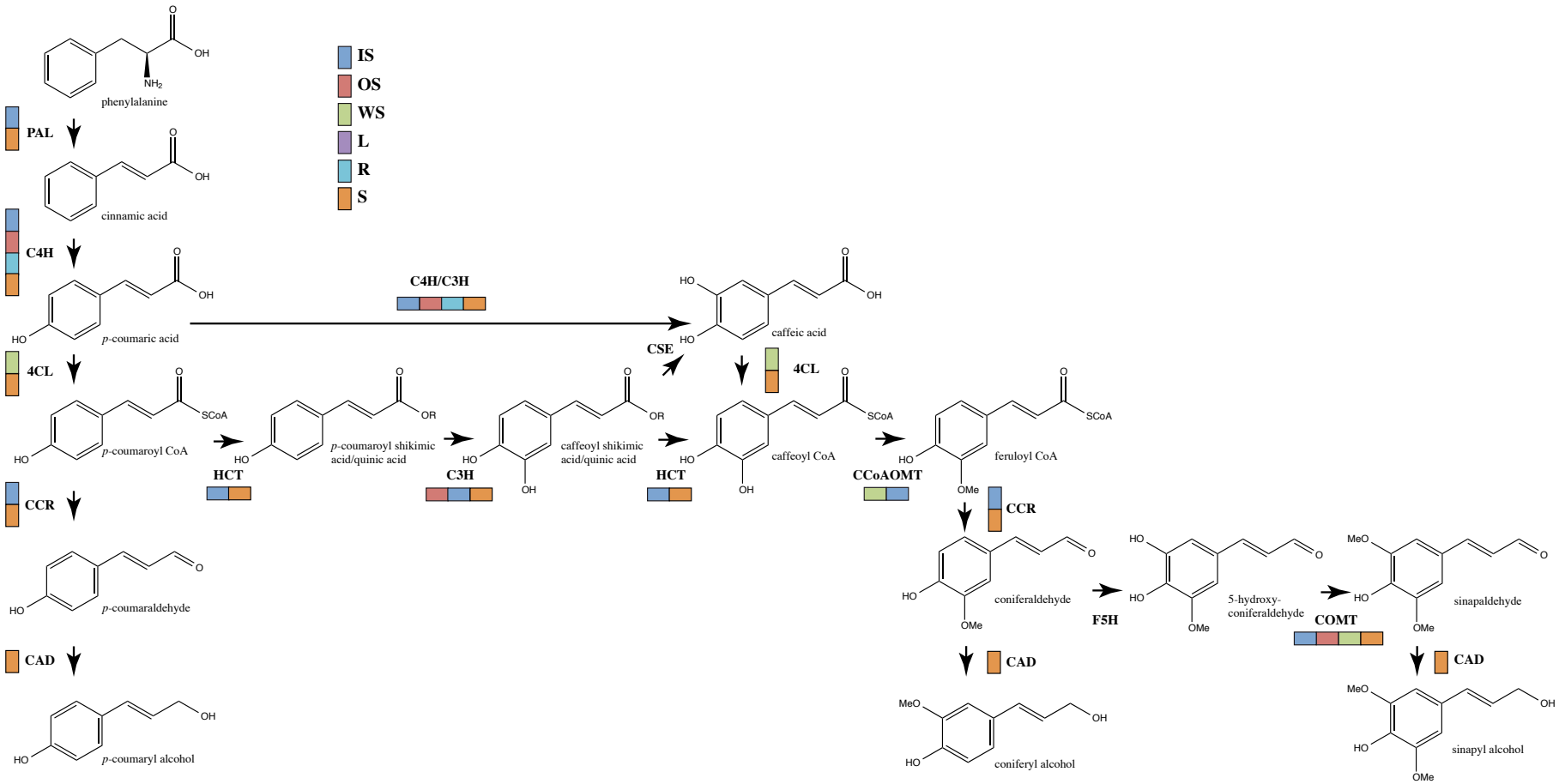


Figure 8. Monolignol biosynthetic pathway indicating presence of the enzyme in a particular organ/tissue (colored box) or absence (no box). CAD: cinnamyl alcohol dehydrogenase; CCoAOMT: caffeoyl-CoA *O*-methyl transferase; CCR: cinnamoyl-CoA reductase; C3H: *p*-coumaroyl shikimate/quinate-3-hydroxylase; C4H: cinnamate 4-hydroxylase; CSE: Caffeoyl shikimate esterase; 4CL : 4-coumarate-CoA ligase; COMT: caffeic acid/5-hydroxyconiferaldehyde *O*-methyltransferase; F5H: ferulic acid/ coniferaldehyde 5-hydroxylase; HCT: *p*-hydroxycinnamoyl-CoA: shikimate/quinate *p*-hydroxycinnamoyltransferase; PAL: phenylalanine ammonia-lyase.

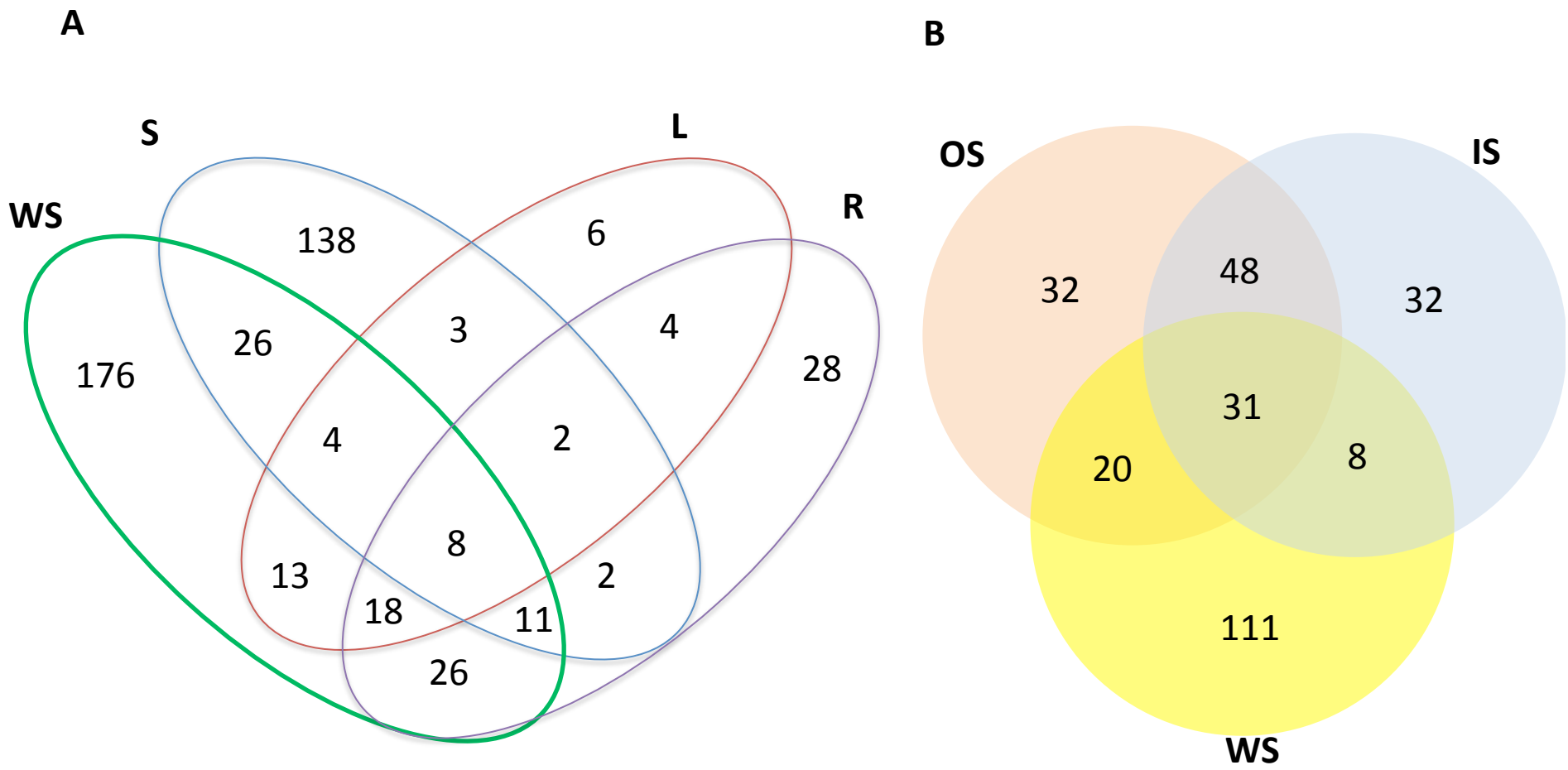


Figure 9. Repartition of the 465 flax Total CWPs identified in 4 different organs (A) and stem tissues (B). Whole stem (WS), Seeds (S), leaves (L), roots (R), Inner stem (IS), Outer stem (OS).

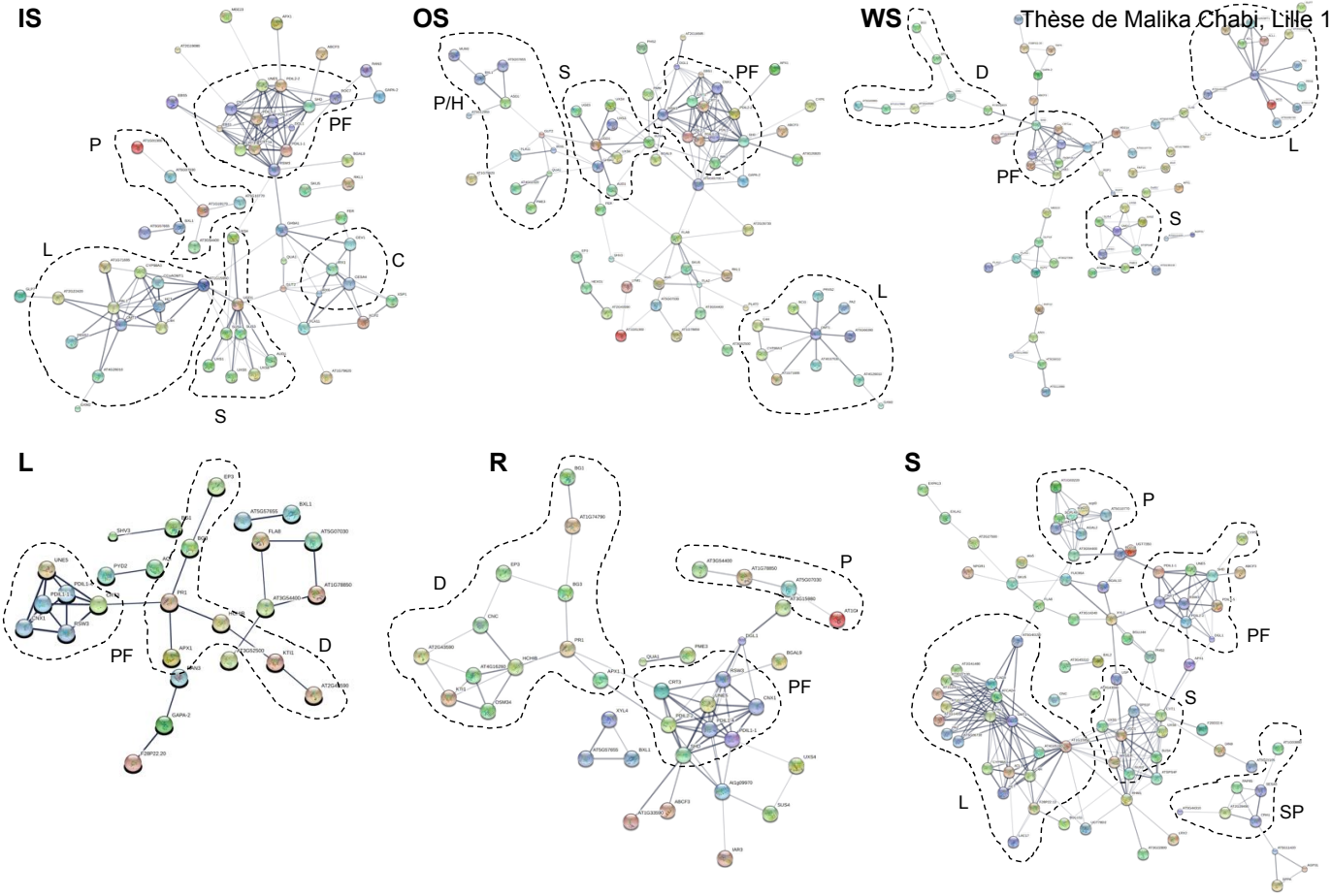

IS
P
PF
L
C
S
OS
P/H
S
PF
L
WS
D
L
PF
S
L
PF
D
R
D
P
PF
S
P
PF
L
SP

Figure 10. Cell wall protein-protein interaction network clusters for flax inner stems (IS), outer stems (OS), whole stem (WS), leaves (L), roots (R) and seeds (S). Interactions were determined using STRING v10.0 (http://string-db.org/). Confidence view, stronger associations are represented by thicker lines. Clusters of lignin proteins (L), sugar precursor proteins (S), cellulose proteins (C), protein folding proteins (PF), defense proteins (D), seed proteins (SP), protease proteins (P) and pectin/hemicellulose proteins (P/H) are indicated by dotted lines.

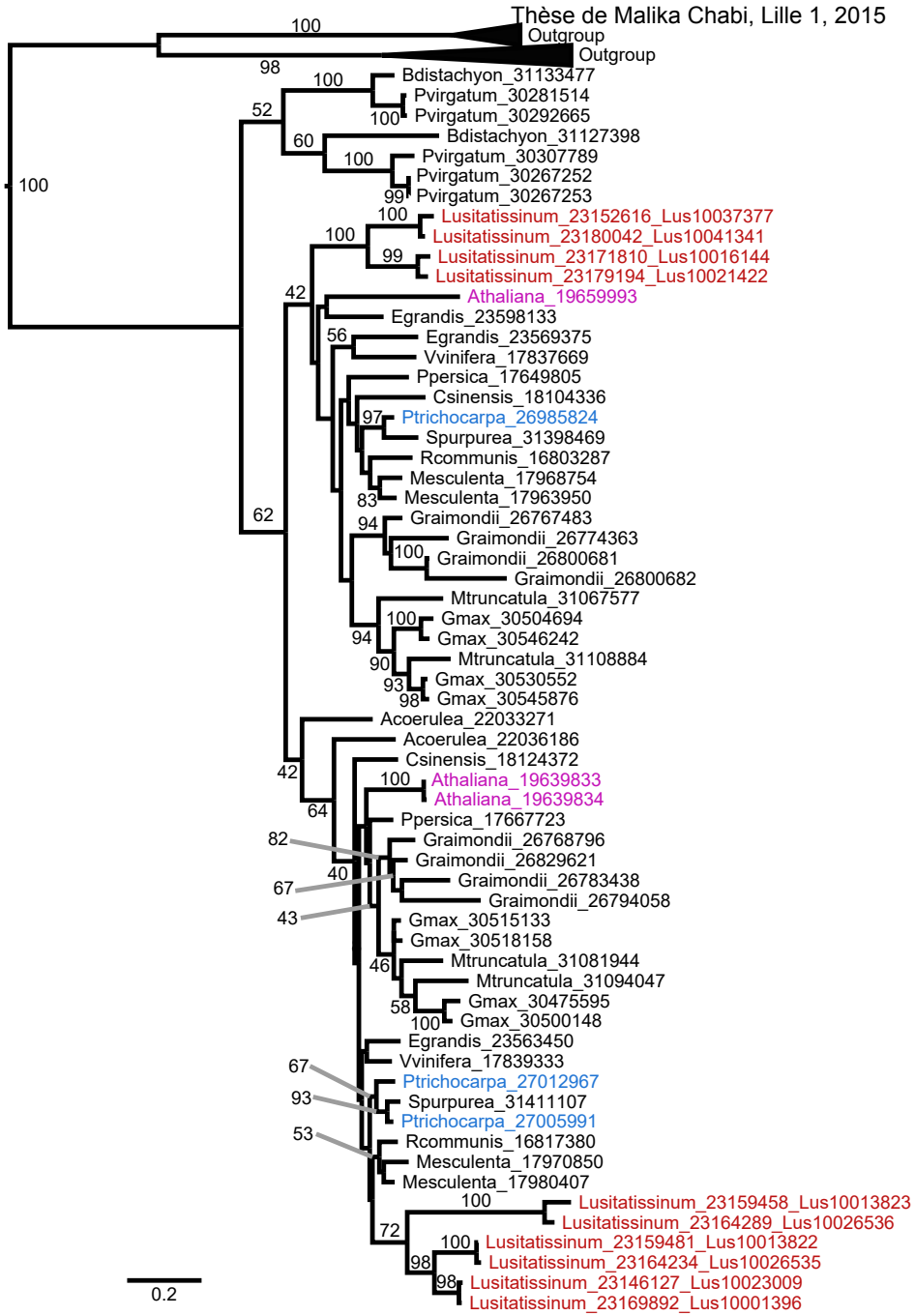

Thèse de Malika Chabi, Lille 1, 2015
Outgroup
Outgroup
Bdistachyon_31133477
Pvirgatum_30281514
Pvirgatum_30292665
Bdistachyon_31127398
Pvirgatum_30307789
Pvirgatum_30267252
Pvirgatum_30267253
Lusitatissinum_23152616_Lus10037377
Lusitatissinum_23180042_Lus10041341
Lusitatissinum_23171810_Lus10016144
Lusitatissinum_23179194_Lus10021422
Athaliana_19659993
Egrandis_23598133
Egrandis_23569375
Vvinifera_17837669
Ppersica_17649805
Csinensis_18104336
Ptrichocarpa_26985824
Spurpurea_31398469
Rcommunis_16803287
Mesculenta_17968754
Mesculenta_17963950
Graimondii_26767483
Graimondii_26774363
Graimondii_26800681
Graimondii_26800682
Mtruncatula_31067577
Gmax_30504694
Gmax_30546242
Mtruncatula_31108884
Gmax_30530552
Gmax_30545876
Acoerulea_22033271
Acoerulea_22036186
Csinensis_18124372
Athaliana_19639833
Athaliana_19639834
Ppersica_17667723
Graimondii_26768796
Graimondii_26829621
Graimondii_26783438
Graimondii_26794058
Gmax_30515133
Gmax_30518158
Mtruncatula_31081944
Mtruncatula_31094047
Gmax_30475595
Gmax_30500148
Egrandis_23563450
Vvinifera_17839333
Ptrichocarpa_27012967
Spurpurea_31411107
Ptrichocarpa_27005991
Rcommunis_16817380
Mesculenta_17970850
Mesculenta_17980407
Lusitatissinum_23159458_Lus10013823
Lusitatissinum_23164289_Lus10026536
Lusitatissinum_23159481_Lus10013822
Lusitatissinum_23164234_Lus10026535
Lusitatissinum_23146127_Lus10023009
Lusitatissinum_23169892_Lus10001396
0.2

Figure 11. Maximum likelihood phylogenetic tree of the xyloglucan *endo*-transglycosylase/hydrolase family IIIA created using all plant species present in phytozome database v10. Branch support values obtained with 1000 bootstrap repetitions. The tree is rooted with the non-IIIA sequence. The scale bar shows the inferred number of amino acid substitutions per site. Protein sequences found in *Linum usitassinum*, *Populus thiocharpa*, and *Arabidopsis thaliana* are colored in red, blue and purple, respectively.

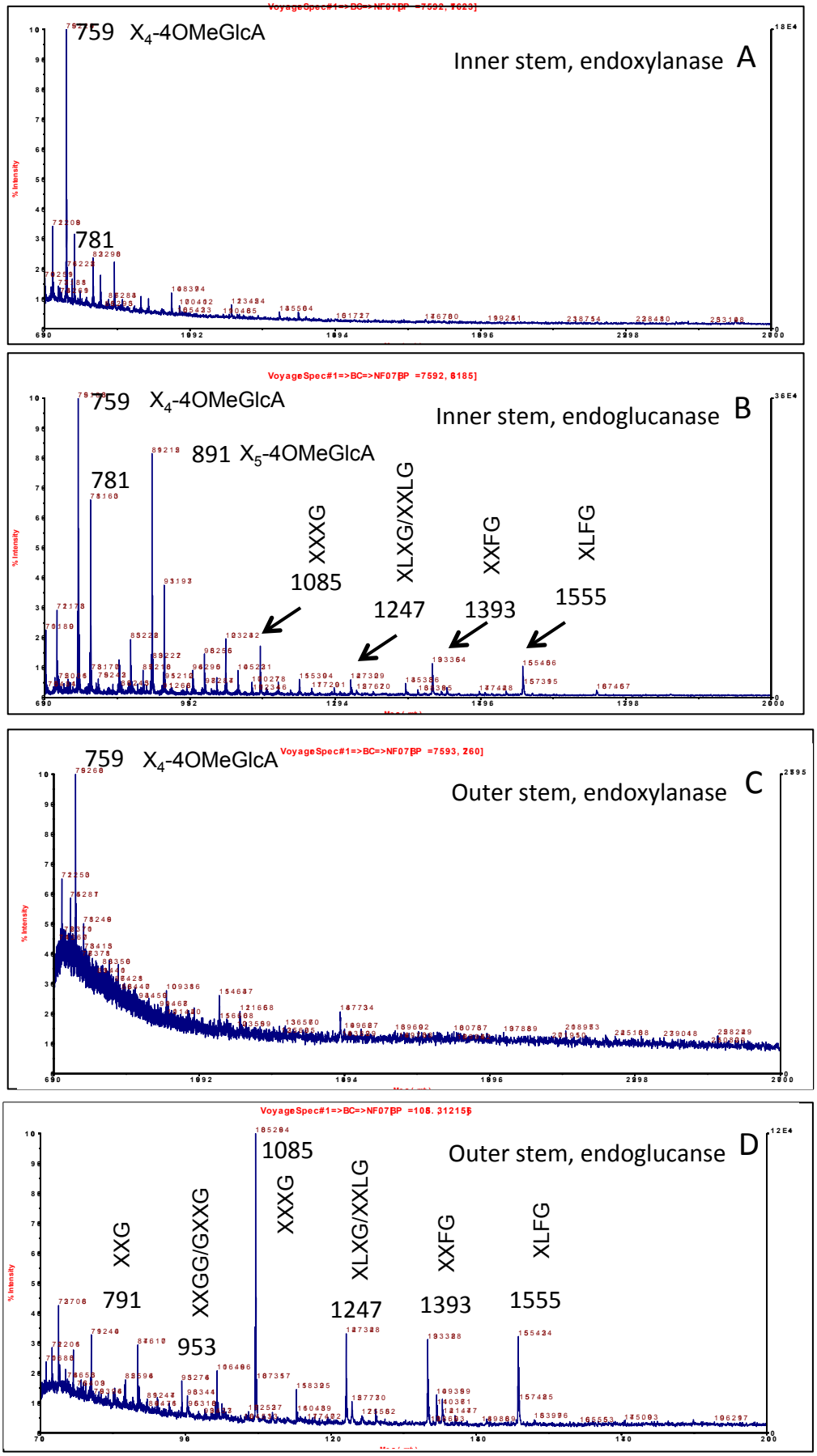


Figure 12. MALDI-TOF-MS spectra of fragment ions $(M+Na)^+$ corresponding to oligosaccharides released by enzymatic digestion of 2-month-old flax stems. Inner stem tissue sample digested by endoxylanase (A) or endo-beta-glucanase (B), outer stem tissue sample digested by endoxylanase (C) or endo-beta-glucanase (D). Figures refer to *m/z* values of xylan-derived fragments (759: X_4-4OMeGlcA + Na+, 781: X_4-4OMeGlcA + 2Na+ - H+, 891: X_5-4OMeGlcA + Na+) and xyloglucan-derived fragments (1085: XXXG + Na+, 1247: XLXG/XXLG + Na+, 1393: XXFG + Na+, 1595: XLFG + Na+).

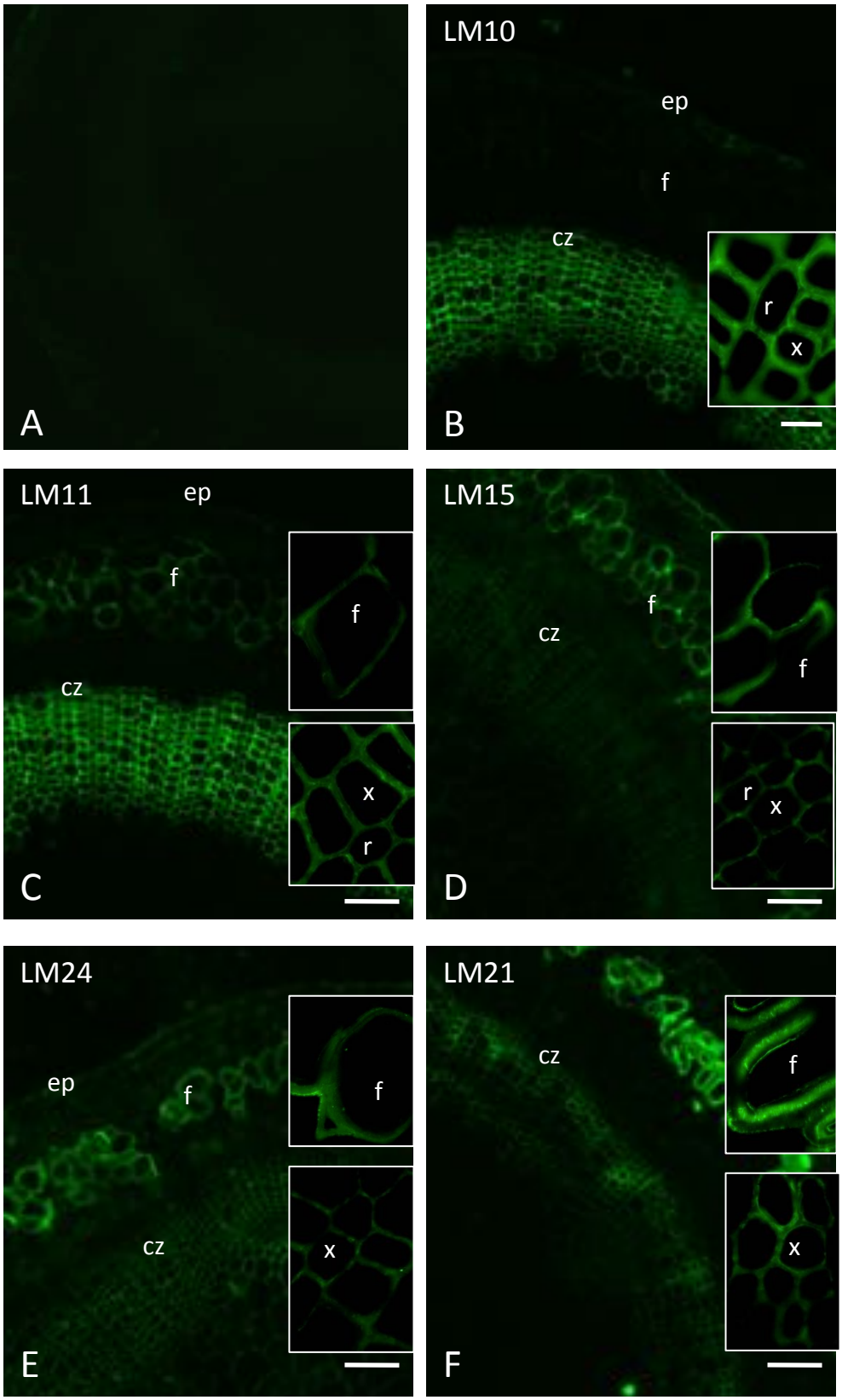
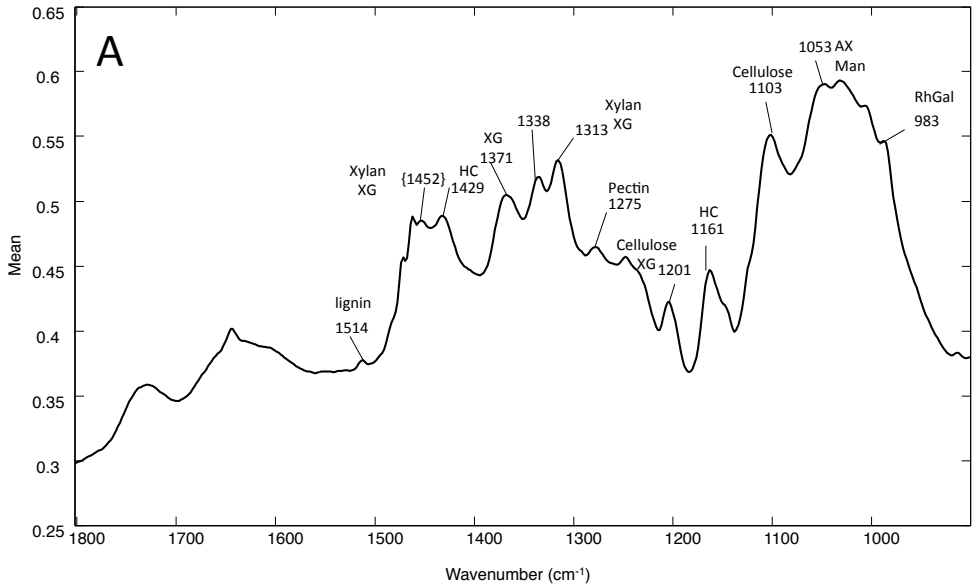
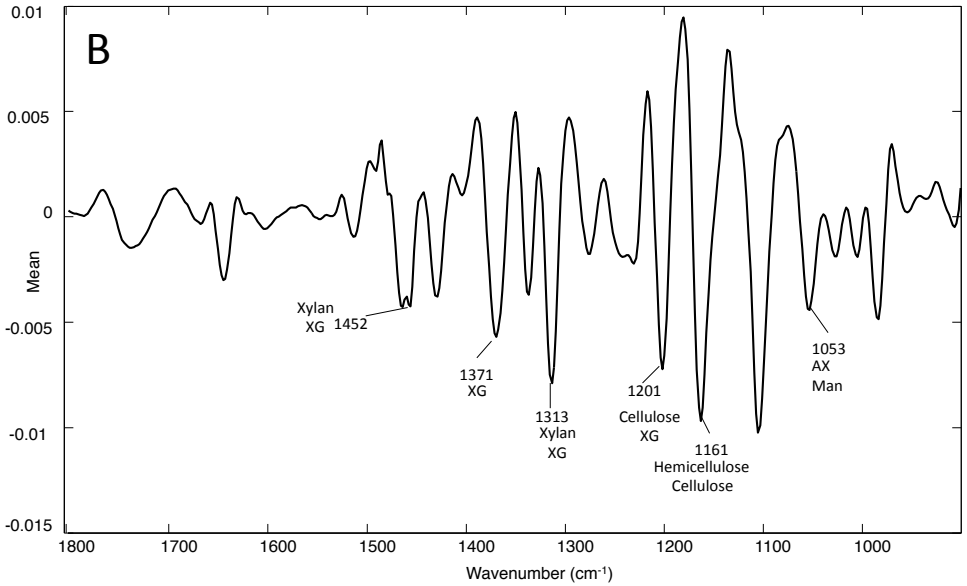


Figure 13. Immunolocalization of flax stem hemicellulose epitopes. Pre-immune serum (A); anti-xylan LM10 antibodies (B); anti-xylan arabinoxylan LM11 antibodies (C); anti-xyloglucan LM15 antibodies (D); anti-xyloglucan LM24 antibodies (E); anti-mannan/heteromannan LM21 antibodies (F). cz: cambial zone; ep: epidermis; f: fiber; x: xylem; r: xylem ray; v: xylem vessel/ tracheid. Bar = 100 μm (except for zoom).





C

Wavenumber (cm-1)	Cell Wall polymers
{1452}	Xylan, Xyloglucan (XG)
1371	Xyloglucan
1313	Xylan, Xyloglucan
1201	Cellulose, Xyloglucan
1161	Hemicellulose, Cellulose
1053	Arabinoxylan (AX), Mannan (Man)

Figure 14: Averaged middle infra-red spectra at 1800-900 cm^{-1} of bast fiber cell walls (A) and secondary derivative (B) showing significant peaks characteristic of hemicelluloses (C). Averaged spectra obtained by combining 19 spectra.

Supplementary Data

For space reasons, only the integrated table of complete flax CWP set composed of 410 proteins, from the supplementary data, is presented in the printed version of this paper.

Protein information				Organs					plantCAZyme	WallProtDB
Gene ID (Lus)	**IDAT**	**Cwdirect link**	**NAME**	**S**	**IS**	**OS**	**L**	**R**		
Lus10010735	0	_	development.unspecified	S	_	_	_	_	_	CW
Lus10002266	unknown	_	not assigned.unknown	_	_	_	L	_	_	CW
Lus10009805	unknown	_	not assigned.unknown	S	_	_	_	_	_	CW
Lus10023525	unknown	_	not assigned.unknown	S	_	_	_	_	_	CW
Lus10040397	unknown	_	not assigned.unknown	S	_	_	_	_	_	CW
Lus10005390	4CL1	lignin	secondary metabolism.lignin biosynthesis.4CL	S	_	_	_	_	_	_
Lus10008677	4CL1	lignin	secondary metabolism.lignin biosynthesis.4CL	S	_	_	_	_	_	_
Lus10015363	AARE	_	protein.degradation	S	_	_	_	_	CE10	_
Lus10023043	ABCF3	_	transport.ABC transporters	S	IS	OS	L	R	_	CW
Lus10018260	AGAL2	Pectin ?	minor CHO metabolism.galactose.alpha-galactosidases	S	_	_	_	_	GH27	CW
Lus10015434	AGP31	_	stress.abiotic.unspecified	S	_	OS	_	_	_	CW
Lus10006505	AIR3	_	protein.degradation.subtilases	S	_	_	_	_	_	CW
Lus10014591	AIR3	_	protein.degradation.subtilases	_	_	_	_	R	_	CW
Lus10032096	AIR3	_	protein.degradation.subtilases	_	_	_	_	R	_	CW
Lus10013537	APX1	_	redox.ascorbate and glutathione.ascorbate	S	IS	OS	L	R	AA2	_
Lus10015970	APX1	_	redox.ascorbate and glutathione.ascorbate	_	IS	OS	L	R	AA2	_
Lus10014128	APX3	_	redox.ascorbate and glutathione.ascorbate	_	IS	OS	L	R	AA2	_
Lus10019781	APX3	_	redox.ascorbate and glutathione.ascorbate	_	IS	OS	L	_	AA2	_
Lus10026565	AQI	_	protein.degradation	_	IS	OS	L	R	_	CW
Lus10006307	ARA12	_	protein.degradation.subtilases	S	_	_	_	_	_	CW
Lus10007765	ARA12	_	protein.degradation.subtilases	_	_	OS	_	R	_	_
Lus10011089	ARA12	_	protein.degradation.subtilases	S	_	OS	_	_	_	_
Lus10038988	ASD1	hemicellulose	cell wall.degradation.mannan-xylose-arabinose-fucose	_	_	OS	_	_	GH51	CW
Lus10003999	AT1G01300	_	protein.degradation.aspartate protease	_	IS	OS	_	R	_	CW
Lus10021936	AT1G03220	_	protein.degradation.aspartate protease	_	_	_	_	R	_	CW

Lus10021938	AT1G03220	_	protein.degradation.aspartate protease	S	_	_	_	_	_	CW
Lus10034036	AT1G03220	_	protein.degradation.aspartate protease	_	_	OS	L	R	_	CW
Lus10036343	AT1G03220	_	protein.degradation.aspartate protease	S	_	OS	_	R	_	CW
Lus10041224	AT1G03220	_	protein.degradation.aspartate protease	S	_	_	_	_	_	CW
Lus10010278	AT1G03230	_	protein.degradation.aspartate protease	_	_	_	L	R	_	CW
Lus10003554	AT1G03890	_	development.storage proteins	S	_	_	_	_	_	CW
Lus10033893	AT1G03890	_	development.storage proteins	S	_	_	_	_	_	CW
Lus10017674	AT1G06260	_	protein.degradation.cysteine protease	S	_	_	_	_	_	CW
Lus10000917	AT1G09730	_	protein.degradation.cysteine protease	S	_	_	_	_	_	CW
Lus10040951	AT1G09870	_	misc.acid and other phosphatases	_	IS	_	_	_	_	CW
Lus10020500	At1g09970	_	signalling.receptor kinases.leucine rich repeat XI	_	_	_	_	R	GH1	_
Lus10002957	AT1G12570	_	not assigned.no ontology	S	_	_	_	_	AA3	CW
Lus10017827	AT1G12570	_	not assigned.no ontology	S	_	_	_	_	AA3	CW
Lus10037079	AT1G13750	_	misc.acid and other phosphatases	_	_	OS	_	_	_	CW
Lus10009898	AT1G14550	lignin	misc.glutathione S transferases	S	_	_	_	_	AA2	CW
Lus10024209	AT1G14550	lignin	misc.glutathione S transferases	S	_	_	_	_	AA2	_
Lus10041651	AT1G15950	lignin	secondary metabolism.lignin biosynthesis.CCR1	S	IS	_	_	_	_	_
Lus10039909	AT1G19170	pectin	cell wall.degradation.pectate lyases	_	IS	_	_	_	GH28	_
Lus10004410	AT1G20030	_	stress.biotic	_	_	_	L	R	_	CW
Lus10004234	AT1G24110	lignin	misc.peroxidases	S	_	_	_	_	AA2	CW
Lus10025384	AT1G29790	_	not assigned.unknown	_	IS	OS	_	_	_	CW
Lus10038445	AT1G30760	_	misc.nitrilases, *nitrile lyases, berberine bridge enzymes	_	IS	OS	_	_	AA7	CW
Lus10002936	AT1G33590	glycoproteins	stress.biotic.PR-proteins	_	_	_	_	R	_	CW
Lus10042502	AT1G44820	_	misc.nitrilases, *nitrile lyases, berberine bridge enzymes	_	_	_	_	R	_	CW
Lus10037797	AT1G47980	_	not assigned.unknown	S	_	_	_	_	_	CW
Lus10005550	AT1G53440	_	signalling.receptor kinases.leucine rich repeat VIII.VIII-2	_	_	OS	_	_	CBM57	_
Lus10031199	AT1G56130	_	signalling.receptor kinases.leucine rich repeat VIII.VIII-2	_	_	OS	_	_	CBM57	_
Lus10015799	AT1G67510	_	signalling.receptor kinases.leucine rich repeat III	S	_	_	_	_	_	CW

Lus10028689	AT1G71695	lignin	misc.peroxidases	_	_	_	L	R	AA2	CW
Lus10028735	AT1G71695	lignin	misc.peroxidases	_	IS	OS	L	R	AA2	CW
Lus10009014	AT1G74020.1	_	secondary metabolism.N misc.alkaloid-like	S	_	_	_	_	_	CW
Lus10036662	AT1G74790	_	cell.organisation	_	_	_	_	R	_	CW
Lus10013572	AT1G75900	_	misc.GDSL-motif lipase	S	_	_	_	_	CE16	_
Lus10018929	AT1G77100	lignin	misc.peroxidases	S	_	_	_	_	AA2	CW
Lus10028631	AT1G77100	lignin	misc.peroxidases	S	_	_	_	_	AA2	CW
Lus10007775	AT1G78060	Hemicellulose (xylan)	cell wall.degradation.mannan-xylose-arabinose-fucose	S	_	_	_	_	GH3	_
Lus10022374	AT1G78060	Hemicellulose (xylan)	cell wall.degradation.mannan-xylose-arabinose-fucose	S	_	_	_	_	GH3	CW
Lus10000579	AT1G78850	_	misc.myrosinases-lectin-jacalin	_	_	OS	L	R	_	CW
Lus10023849	AT1G79620	_	signalling.receptor kinases.leucine rich repeat VIII.VIII-1	_	IS	OS	_	_	_	CW
Lus10013342	AT2G01820	_	signalling.receptor kinases.leucine rich repeat IX	S	_	_	_	_	_	CW
Lus10019882	AT2G03550	_	lipid metabolism.lipid degradation	S	_	_	_	_	CE10	_
Lus10019885	AT2G03550	_	lipid metabolism.lipid degradation	S	_	_	_	_	CE10	_
Lus10039466	AT2G04690	_	not assigned.no ontology	S	_	_	_	_	_	CW
Lus10007577	AT2G16595	_	not assigned.no ontology	_	_	OS	_	_	_	CW
Lus10041713	AT2G18570	_	secondary metabolism.flavonoids.dihydroflavonols	S	_	_	_	_	GT1	_
Lus10024037	AT2G18570	lignin	secondary metabolism.phenylpropanoids.lignin biosynthesis	S	_	_	_	_	GT1	_
Lus10020010	AT2G19080	_	development.unspecified	_	IS	_	_	_	_	CW
Lus10035925	AT2G22420	lignin	misc.peroxidases	_	IS	_	_	_	AA2	_
Lus10004281	AT2G22620	pectin	cell wall.degradation.pectate lyases and polygalacturonases	_	_	OS	_	_	PL4	_
Lus10017756	AT2G26730	_	signalling.receptor kinases.leucine rich repeat III	_	IS	OS	_	R	_	CW
Lus10004869	AT2G27500	callose	misc.beta 1,3 glucan hydrolases	S	_	_	_	_	GH17	CW
Lus10006420	AT2G28490	_	development.storage proteins	S	_	_	_	_	_	CW
Lus10011364	AT2G28490	_	development.storage proteins	S	_	_	_	_	_	CW
Lus10027739	AT2G36780	_	misc.UDP glucosyl and glucoronyl transferases	S	_	_	_	_	GT1	_
Lus10023166	AT2G37050	_	signalling.receptor kinases.leucine rich repeat I	_	IS	OS	_	_	_	CW
Lus10015074	AT2G37130	lignin	misc.peroxidases	S	_	_	_	_	AA2	CW

Lus10012684	AT2G41480	lignin	misc.peroxidases	_	_	_	_	R	AA2	CW
Lus10015555	AT2G41480	lignin	misc.peroxidases	S	_	_	_	_	AA2	CW
Lus10038055	AT2G41480	lignin	misc.peroxidases	S	_	_	_	_	AA2	CW
Lus10003231	AT2G43590	_	stress.biotic	S	_	_	_	R	GH19	CW
Lus10010861	AT2G43590	_	stress.biotic	_	_	_	_	R	GH19	_
Lus10035620	AT2G43590	_	stress.biotic	S	_	OS	L	_	GH19	CW
Lus10035624	AT2G43590	_	stress.biotic	_	_	_	_	R	GH19	CW
Lus10006859	AT2G44220	_	not assigned.unknown	S	_	_	_	_	_	CW
Lus10018829	AT2G46000	_	not assigned.unknown	S	_	_	_	_	_	CW
Lus10031211	AT3G03060	_	protein.degradation.ubiquitin.proteasom	_	IS	OS	_	R	_	CW
Lus10026851	AT3G03330	_	misc.short chain dehydrogenase/reductase (SDR)	_	IS	_	_	_	_	CW
Lus10029501	AT3G08030	_	not assigned.unknown	_	IS	_	_	R	_	CW
Lus10013154	AT3G14067	_	protein.degradation.subtilases	S	_	_	_	_	_	CW
Lus10013187	AT3G14240	_	protein.degradation.subtilases	S	_	_	_	_	_	CW
Lus10002324	AT3G15980	_	cell.vesicle transport	_	_	_	_	R	_	CW
Lus10031824	AT3G20820	_	not assigned.no ontology	_	_	OS	L	_	_	CW
Lus10039362	AT3G22800	glycoproteins	cell wall.cell wall proteins.LRR	S	_	_	_	_	_	_
Lus10043212	AT3G23760	_	not assigned.unknown	S	_	_	_	_	_	CW
Lus10012578	AT3G26720	glycoproteins	misc.gluco-, galacto- and mannosidases.alpha-mannosidase	_	IS	OS	_	R	GH38	CW
Lus10033428	AT3G45310	_	protein.degradation.cysteine protease	S	_	_	_	_	_	CW
Lus10005871	AT3G47000	Hemicellulose (xyloglucan)	cell wall.degradation.cellulases and beta -1,4-glucanases	S	_	_	_	_	GH3	_
Lus10000391	AT3G47040	cellulose	cell wall.degradation.cellulases and beta -1,4-glucanases	S	_	_	_	_	_	_
Lus10025977	AT3G51050	_	not assigned.no ontology	_	IS	_	_	_	_	CW
Lus10016967	AT3G52500	_	protein.degradation	_	_	_	L	_	_	CW
Lus10021293	AT3G52500	_	protein.degradation	_	_	OS	L	_	_	CW
Lus10008654	AT3G54400	_	RNA.regulation of transcription	S	IS	OS	L	R	_	CW
Lus10009813	AT4G01030	_	pentatricopeptide (PPR) repeat-containing protein	S	_	_	_	_	_	CW
Lus10008203	AT4G02320	pectin	cell wall.pectin*esterases.PME	_	_	OS	_	_	CE8	CW

Lus10016135	AT4G05160	lignin	secondary metabolism.lignin biosynthesis.4CL	S	_	_	_	_	_	_
Lus10037780	AT4G16180	_	not assigned.unknown	_	IS	OS	_	R	_	CW
Lus10016883	AT4G16260	Callose, MLG	misc.beta 1,3 glucan hydrolases	_	_	_	_	R	GH17	_
Lus10038437	AT4G20800	_	misc.nitrilases, *nitrile lyases, berberine bridge enzymes	S	_	_	_	_	AA7	CW
Lus10023376	AT4G20820	_	misc.nitrilases, *nitrile lyases, berberine bridge enzymes	S	_	_	_	_	AA7	_
Lus10038436	AT4G20820	_	misc.nitrilases, *nitrile lyases, berberine bridge enzymes	S	_	_	_	_	AA7	CW
Lus10001324	AT4G26010	lignin	misc.peroxidases	_	IS	OS	L	_	AA2	CW
Lus10006748	AT4G27270	_	lipid metabolism.exotics(steroids, squalene etc)	S	_	_	_	_	AA6	_
Lus10007612	AT4G27270	_	lipid metabolism.exotics(steroids, squalene etc)	S	IS	_	_	_	AA6	CW
Lus10043042	AT4G29270	_	misc.acid and other phosphatases	S	_	_	_	_	_	CW
Lus10024535	AT4G31140	Hemicellulose (xyloglucan)	misc.beta 1,3 glucan hydrolases	S	_	_	_	_	CBM43	_
Lus10020166	AT4G31340	_	cell.organisation.cytoskeleton.Myosin	_	IS	OS	_	R	_	CW
Lus10026967	AT4G31340	_	cell.organisation.cytoskeleton.Myosin	_	IS	OS	_	R	_	CW
Lus10033784	AT4G33820	Xylan	not assigned.no ontology	S	_	_	_	_	GH10	_
Lus10001508	AT4G34180	_	DNA.unspecified	S	IS	OS	_	R	_	CW
Lus10031453	AT4G34180	_	DNA.unspecified	_	IS	_	_	_	_	CW
Lus10031454	AT4G34180	_	DNA.unspecified	S	_	_	L	_	_	CW
Lus10019822	AT4G35220	_	DNA.unspecified	_	_	_	_	R	_	CW
Lus10028365	AT4G36195	_	protein.degradation.serine protease	_	IS	OS	_	R	_	CW
Lus10011079	AT4G37530	lignin	misc.peroxidases	_	_	OS	L	R	AA2	_
Lus10008635	AT5G03810	_	misc.GDSL-motif lipase	S	_	_	_	_	_	CW
Lus10005706	AT5G04885	Xylan	misc.gluco-, galacto- and mannosidases	S	_	_	_	_	GH3	_
Lus10033393	AT5G04885	Xylan	misc.gluco-, galacto- and mannosidases	S	_	_	_	_	GH3	CW
Lus10013376	AT5G06570	_	not assigned.no ontology	S	_	_	_	_	CE10	_
Lus10036168	AT5G06570	_	not assigned.no ontology	S	_	_	_	_	CE10	_
Lus10008173	AT5G06730	lignin	misc.peroxidases	_	_	_	_	R	AA2	CW
Lus10008174	AT5G06730	lignin	misc.peroxidases	S	_	_	_	_	AA2	CW
Lus10026171	AT5G07030	_	RNA.regulation of transcription	_	IS	OS	L	R	_	CW

Lus10038732	AT5G10770	_	RNA.regulation of transcription.unclassified	S	_	_	_	_	_	CW
Lus10042426	AT5G10770	_	RNA.regulation of transcription.unclassified	S	IS	OS	L	R	_	CW
Lus10005474	AT5G11420	_	not assigned.unknown	S	_	_	_	_	_	CW
Lus10013112	AT5G11420	_	not assigned.unknown	_	_	_	L	_	_	CW
Lus10031819	AT5G12890	_	misc.UDP glucosyl and glucoronyl transferases	S	_	_	_	_	GT1	_
Lus10031258	AT5G12950	_	not assigned.unknown	_	_	OS	_	_	_	CW
Lus10012577	AT5G13980	N-glycosilation of proteins !!	misc.gluco-, galacto- and mannosidases	_	_	OS	_	_	GH38	_
Lus10041516	AT5G13980	N-glycosilation of proteins !!	misc.gluco-, galacto- and mannosidases	_	IS	OS	_	R	GH38	_
Lus10020218	AT5G16380	_	not assigned.unknown	S	_	_	_	_	_	CW
Lus10017665	AT5G20860	pectin	cell wall.pectin*esterases.PME	S	_	_	_	_	CE8	CW
Lus10008432	AT5G20950	Hemicellulose (xyloglucan)	cell wall.degradation.cellulases and beta -1,4-glucanases	S	_	_	_	_	GH3	CW
Lus10025538	AT5G21105	_	redox.ascorbate and glutathione.ascorbate	S	_	_	_	_	AA1	_
Lus10026753	AT5G21105	_	redox.ascorbate and glutathione.ascorbate	S	_	_	_	_	AA1	_
Lus10027509	AT5G24010	_	signalling.receptor kinases.Catharanthus roseus-like RLK1	_	IS	_	_	_	_	CW
Lus10006538	AT5G39120	_	stress.abiotic.unspecified	_	_	_	_	R	_	CW
Lus10003116	AT5G39150	_	stress.abiotic.unspecified	_	_	_	_	R	_	CW
Lus10006543	AT5G39150	_	stress.abiotic.unspecified	_	_	_	_	R	_	CW
Lus10039445	AT5G40150	lignin	misc.peroxidases	S	_	_	_	_	AA2	CW
Lus10028304	AT5G44310	_	not assigned.no ontology.late embryogenesis abundant domain-containing protein	S	_	_	_	_	_	CW
Lus10033301	AT5G45280	pectin	cell wall.pectin*esterases.acetyl esterase	_	_	_	_	R	CE13	CW
Lus10034762	AT5G45280	pectin	cell wall.pectin*esterases.acetyl esterase	_	_	OS	_	_	CE13	CW
Lus10007795	AT5G47380	_	not assigned.unknown	S	_	_	_	_	_	CW
Lus10003517	AT5G49900	Hemicellulose (xyloglucan)	not assigned.unknown	_	_	OS	_	_	GH116	_
Lus10001321	AT5G57655	sugar precursor	minor CHO metabolism.others.Xylose isomerase	S	IS	OS	_	R	_	CW
Lus10006983	AT5G57655	sugar precursor	minor CHO metabolism.others.Xylose isomerase	_	IS	OS	L	_	_	CW
Lus10002964	AT5G59100	_	protein.degradation.subtilases	S	_	_	_	_	_	CW
Lus10042555	AT5G59100	_	protein.degradation.subtilases	S	_	_	_	_	_	CW
Lus10022071	AT5G61750	_	stress.abiotic.unspecified	_	IS	_	_	_	_	CW

Lus10011585	AT5G65700.1	_	signalling.receptor kinases.leucine rich repeat XI	_	_	OS	_	_	_	CW
Lus10041784	AT5G66390	lignin	misc.peroxidases	_	_	OS	_	R	AA2	CW
Lus10009365	AT5G67090	_	protein.degradation.subtilases	_	_	_	L	_	_	CW
Lus10019811	ATCAD4	lignin	secondary metabolism.lignin biosynthesis.CAD	S	_	_	_	_	_	_
Lus10027864	ATCAD4	lignin	secondary metabolism.lignin biosynthesis.CAD	S	_	_	_	_	_	_
Lus10041041	ATSPS4F	_	major CHO metabolism.synthesis.sucrose.SPS	S	_	_	_	_	GT4	_
Lus10006510	AUD1	sugar precursor	cell wall.precursor synthesis.UXS	_	IS	OS	_	_	_	_
Lus10018491	BAM5	_	major CHO metabolism.degradation	S	_	_	_	_	GH14	_
Lus10039701	BAM5	_	major CHO metabolism.degradation	S	_	_	_	_	GH14	_
Lus10014108	BG1	Callose, MLG	misc.beta 1,3 glucan hydrolases.	_	_	_	_	R	GH17	CW
Lus10019801	BG1	Callose, MLG	misc.beta 1,3 glucan hydrolases.	S	_	_	L	R	GH17	CW
Lus10002807	BG3	Callose, MLG	misc.beta 1,3 glucan hydrolases.	_	_	_	L	R	GH17	CW
Lus10025108	BGAL10	galactan	misc.gluco-, galacto- and mannosidases.beta-galactosidase	S	_	_	_	_	GH35	CW
Lus10008974	BGAL12	Hemicellulose/galactan	misc.gluco-, galacto- and mannosidases.beta-galactosidase	_	IS	OS	_	_	GH35	CW
Lus10018138	BGAL16	pectin	misc.gluco-, galacto- and mannosidases.beta-galactosidase	S	_	_	_	_	GH35	CW
Lus10000271	BGAL8	pectin	misc.gluco-, galacto- and mannosidases.beta-galactosidase	_	_	OS	_	_	GH35	_
Lus10016655	BGAL9	pectin	misc.gluco-, galacto- and mannosidases.beta-galactosidase	_	IS	OS	_	R	GH35	CW
Lus10007871	BGLU11	Hemicellulose (Xyloglucans, xylan, MLG)	misc.gluco-, galacto- and mannosidases	S	_	_	_	_	GH1	CW
Lus10007872	BGLU11	Hemicellulose (Xyloglucans, xylan, MLG)	misc.gluco-, galacto- and mannosidases	S	_	_	_	_	GH1	_
Lus10031235	BGLU15	Hemicellulose (xyloglucan)	misc.gluco-, galacto- and mannosidases	S	_	_	_	_	GH1	CW
Lus10022883	BGLU17	Hemicellulose (xyloglucan)	misc.gluco-, galacto- and mannosidases	S	IS	_	_	_	GH1	_
Lus10030576	BGLU17	Hemicellulose (xyloglucan)	misc.gluco-, galacto- and mannosidases	_	IS	OS	_	R	GH1	CW
Lus10039541	BGLU17	Hemicellulose (xyloglucan)	misc.gluco-, galacto- and mannosidases	_	IS	_	_	_	GH1	CW
Lus10012869	BGLU32	Hemicellulose (xyloglucan)	misc.gluco-, galacto- and mannosidases	_	_	_	_	R	GH1	CW
Lus10012871	BGLU32	Hemicellulose (xyloglucan)	misc.gluco-, galacto- and mannosidases	_	IS	OS	L	_	GH1	_
Lus10030577	BGLU32	Hemicellulose (xyloglucan)	misc.gluco-, galacto- and mannosidases	_	IS	OS	_	_	GH1	CW
Lus10020232	BGLU42	Hemicellulose (xyloglucan)	misc.gluco-, galacto- and mannosidases	S	_	_	_	_	GH1	_

Lus10030515	BGLU44	Hemicellulose (xyloglucan)	misc.gluco-, galacto- and mannosidases	S	_	_	L	_	GH1	CW
Lus10035679	BMY2	_	major CHO metabolism.degradation	S	_	_	_	_	GH14	_
Lus10016858	BXL1	Hemicellulose (xylan)	cell wall.degradation.mannan-xylose-arabinose-fucose	_	IS	OS	L	R	GH3	_
Lus10010015	BXL2	Hemicellulose (Xyloglucans, xylan, MLG)	cell wall.degradation.mannan-xylose-arabinose-fucose	S	_	_	_	_	GH3	CW
Lus10027598	C4H	lignin	secondary metabolism..lignin biosynthesis.C4H	_	_	_	_	R	_	_
Lus10034449	C4H	lignin	secondary metabolism.lignin biosynthesis.C4H	_	IS	OS	_	R	_	_
Lus10035011	C4H	lignin	secondary metabolism.lignin biosynthesis.C4H	S	IS	_	_	R	_	_
Lus10002089	CAD9	lignin	secondary metabolism.lignin biosynthesis.CAD	S	_	_	_	_	_	_
Lus10027888	CCoAOMT1	lignin	secondary metabolism.lignin biosynthesis.CCoAOMT	_	IS	_	_	_	_	_
Lus10018083	CEP1	remodelling	protein.degradation.cysteine protease	_	IS	_	_	_	_	CW
Lus10042078	CEP1	remodelling	protein.degradation.cysteine protease	S	_	_	_	_	_	CW
Lus10008226	CESA4	cellulose	cell wall.cellulose synthesis.cellulose synthase	_	IS	_	_	_	GT2_proc	_
Lus10039607	CEV1	cellulose	cell wall.cellulose synthesis.cellulose synthase	_	IS	_	_	_	GT2_proc	_
Lus10019060	ChiC	_	stress.biotic	S	_	_	_	_	GH18	CW
Lus10036310	ChiC	_	stress.biotic	_	_	_	_	R	GH18	CW
Lus10004076	CLH1	_	stress.biotic	S	_	OS	L	_	_	CW
Lus10032521	CNX1	_	signalling.calcium	S	IS	OS	L	R	_	CW
Lus10010341	CRA1	_	development.storage proteins	S	_	_	_	_	_	CW
Lus10011817	CRA1	_	development.storage proteins	S	_	_	_	_	_	CW
Lus10021179	CRA1	_	development.storage proteins	S	_	_	_	_	_	CW
Lus10021180	CRA1	_	development.storage proteins	S	_	_	_	_	_	CW
Lus10022927	CRA1	_	development.storage proteins	S	_	_	_	_	_	CW
Lus10022929	CRA1	_	development.storage proteins	S	_	_	_	_	_	CW
Lus10024889	CRA1	_	development.storage proteins	S	_	_	_	_	_	CW
Lus10024891	CRA1	_	development.storage proteins	S	_	_	_	_	_	CW
Lus10020222	CRT3	_	signalling.calcium	_	IS	OS	L	R	_	CW
Lus10024087	CXE17	_	Biodegradation of Xenobiotics	S	_	_	_	_	CE10	_
Lus10038315	CYP5	_	cell.cycle.peptidylprolyl isomerase	S	_	_	_	_	_	CW
Lus10033524	CYP98A3	lignin	secondary metabolism.lignin biosynthesis.C3H	S	IS	OS	_	_	_	_

Lus10026117	CYSB	_	protein.degradation.cysteine protease	S	_	OS	_	_	_	CW
Lus10004699	CYT1	sugar precursor	cell wall.precursor synthesis.NDP sugar pyrophosphorylase	S	_	_	_	_	_	_
Lus10014330	DGL1	_	protein.glycosylation	S	IS	OS	_	R	_	CW
Lus10026039	DGL1	_	protein.glycosylation	_	IS	_	_	_	_	CW
Lus10015062	DIN9	sugar precursor	cell wall.precursor synthesis.phosphomannose isomerase	S	_	_	_	_	_	_
Lus10035408	DPE2	_	major CHO metabolism.degradation.starch.D enzyme	S	_	_	_	_	CBM20	_
Lus10015475	EBS1	_	protein.glycosylation	_	IS	_	_	_	GT24	_
Lus10019946	EBS1	_	protein.glycosylation	_	IS	OS	_	_	GT24	CW
Lus10009680	EBS5	_	not assigned.no ontology	_	IS	_	_	_	_	CW
Lus10007526	ENGase85B	_	not assigned.no ontology	S	_	_	_	_	GH85	_
Lus10003226	EP3	_	stress.biotic	_	_	OS	_	R	GH19	CW
Lus10024366	EP3	_	stress.biotic	_	_	_	L	_	GH19	CW
Lus10035625	EP3	_	stress.biotic	_	_	_	_	R	GH19	CW
Lus10010352	EXL2	_	signalling.in sugar and nutrient physiology	_	_	OS	L	_	_	CW
Lus10036484	EXL2	_	signalling.in sugar and nutrient physiology	_	_	OS	L	_	_	CW
Lus10031130	EXL3	_	not assigned.no ontology	_	_	OS	_	_	_	CW
Lus10003336	EXLA1	Polysaccharide (remodelling)	cell wall.modification	S	_	_	L	_	EXPN	CW
Lus10003822	EXPA13	Polysaccharide polymer (remodelling)	cell wall.modification	S	_	_	_	_	EXPN	CW
Lus10020312	F20D22.6	_	protein.synthesis.initiation	S	_	_	_	_	AA2	CW
Lus10010149	F28P22.13	lignin	secondary metabolism.lignin biosynthesis.CAD	S	_	_	_	_	_	_
Lus10002795	F28P22.20	_	stress.abiotic.unspecified	_	_	_	L	_	_	CW
Lus10010948	FCLY	_	not assigned.unknown	_	IS	_	_	_	_	CW
Lus10014185	FER	_	signalling.receptor kinases.Catharanthus roseus-like RLK1	_	IS	OS	_	_	_	CW
Lus10036609	Fh5	_	formin homology 2 domain-containing protein	S	_	_	_	_	_	CW
Lus10002984	FLA11	glycoproteins	cell wall.cell wall proteins.AGPs.AGP	_	_	OS	_	_	_	CW
Lus10036112	FLA11	glycoproteins	cell wall.cell wall proteins.AGPs.AGP	_	IS	OS	_	_	_	CW
Lus10036114	FLA11	glycoproteins	cell wall.cell wall proteins.AGPs.AGP	_	_	OS	_	_	_	CW
Lus10006391	FLA2	glycoproteins	cell wall.cell wall proteins.AGPs.AGP	_	_	OS	_	_	_	CW

Lus10017696	FLA6	glycoproteins	cell wall.cell wall proteins.AGPs.AGP	_	IS	_	_	_	_	CW
Lus10009235	FLA8	glycoproteins	cell wall.cell wall proteins.AGPs.AGP	S	_	OS	L	_	_	CW
Lus10020979	FUC95A	Hemicellulose (xyloglucan)	not assigned.unknown	S	_	_	_	_	GH95	_
Lus10012591	GAPA-2	_	PS.calvin cycle.GAP	_	IS	OS	L	_	_	CW
Lus10012162	GH9A1	cellulose	cell wall.cellulose synthesis	_	IS	OS	_	_	GH9	_
Lus10038544	GHR1	_	signalling.receptor kinases.leucine rich repeat III	_	_	OS	_	_	_	CW
Lus10034959	GLOX1	_	Biodegradation of Xenobiotics	S	_	_	_	_	AA5	CW
Lus10010726	GLP7	_	stress.abiotic.unspecified	_	IS	_	_	_	_	CW
Lus10029214	GLP7	_	stress.abiotic.unspecified	_	IS	_	_	_	_	CW
Lus10015337	GMII	N-glycosilation of proteins !!	protein.glycosylation	_	IS	_	_	_	GH38	_
Lus10033689	GSL05	callose	minor CHO metabolism.callose	S	_	_	_	_	GT48	_
Lus10001906	GT72B1	_	secondary metabolism.flavonoids.dihydroflavonols	S	_	_	_	_	GT1	_
Lus10008823	GUS3	glucuronidase 3	cell wall.degradation	_	_	OS	_	_	GH79	CW
Lus10043326	GUT2	Hemicellulose (xylan)	misc.UDP glucosyl and glucoronyl transferases	_	IS	OS	_	_	GT47	CW
Lus10005829	GXM2	Hemicellulose (xylan)	not assigned.unknown	_	IS	OS	_	_	_	CW
Lus10041830	HCHIB	_	stress.biotic	_	_	_	L	R	GH19	CW
Lus10002321	HCT	lignin	secondary metabolism.lignin biosynthesis.HCT	S	_	_	_	_	_	_
Lus10026097	HCT	lignin	secondary metabolism.lignin biosynthesis.HCT	_	IS	_	_	_	_	_
Lus10028140	HERK1	_	signalling.receptor kinases.Catharanthus roseus-like RLK1	_	IS	OS	_	_	_	CW
Lus10030323	HEXO1	_	protein.glycosylation.	_	_	OS	_	_	GH20	CW
Lus10006864	IAR3	_	hormone metabolism.auxin.synthesis-degradation	_	IS	OS	_	R	_	CW
Lus10014236	IMK2	_	signalling.receptor kinases.leucine rich repeat III	_	_	OS	_	_	_	CW
Lus10029245	IRX1	cellulose	cell wall.cellulose synthesis.cellulose synthase	_	IS	_	_	_	GT2_proc	_
Lus10017863	IRX6	cellulose	cell wall.cellulose synthesis.COBRA	_	IS	OS	_	_	_	_
Lus10036474	ISA3	_	major CHO metabolism.degradation.starch.ISA3	S	_	_	_	_	GH13	_
Lus10029429	KTI1	_	stress.biotic.PR-proteins	_	_	_	L	R	_	CW
Lus10010850	LAC17	lignin	secondary metabolism.simple phenols	S	_	_	_	_	AA1	_
Lus10019227	LCR68	_	stress.biotic	_	IS	OS	_	_	_	CW
Lus10028097	LRX2	glycoproteins	cell wall.cell wall proteins.LRR	S	_	_	_	_	_	CW

Lus10011358	LTPG1	_	lipid metabolism.lipid transfer proteins etc	_	_	_	L	_	_	CW
Lus10025643	LYM1	_	not assigned.no ontology	_	_	OS	_	_	_	CW
Lus10007211	MAN1	Hemicellulose (mannan)	misc.gluco-, galacto- and mannosidases	S	_	_	_	_	GH5	_
Lus10023361	MEE23	_	misc.nitrilases, *nitrile lyases, berberine bridge enzymes	_	IS	OS	_	_	AA7	CW
Lus10033502	MUM2	Pectin (mucilage)	misc.gluco-, galacto- and mannosidases.beta-galactosidase	_	_	OS	_	_	GH35	CW
Lus10018678	NPGR1	_	signalling.calcium	S	_	_	_	_	_	CW
Lus10006719	NRS/ER	sugar precursor	cell wall.precursor synthesis.UER	S	_	_	_	_	_	_
Lus10014147	NRS/ER	sugar precursor	cell wall.precursor synthesis.UER	S	_	_	_	_	_	_
Lus10015576	OMT1	lignin	secondary metabolism.lignin biosynthesis.COMT	S	_	_	_	_	_	_
Lus10032929	OMT1	lignin	secondary metabolism.lignin biosynthesis.COMT	S	IS	OS	_	_	_	_
Lus10006302	OSM34	_	stress.abiotic	S	_	_	_	R	_	CW
Lus10024511	OSM34	_	stress.abiotic	_	_	_	_	R	_	CW
Lus10004163	PA2	lignin	misc.peroxidases	_	_	_	_	R	AA2	CW
Lus10008167	PA2	lignin	misc.peroxidases	_	_	OS	L	R	AA2	CW
Lus10027989	PA2	lignin	misc.peroxidases	S	_	_	_	R	AA2	CW
Lus10023531	PAL1	lignin	secondary metabolism.lignin biosynthesis.PAL	S	IS	_	_	_	_	_
Lus10026518	PAL1	lignin	secondary metabolism.lignin biosynthesis.PAL	S	_	_	_	_	_	_
Lus10040416	PAL1	lignin	secondary metabolism.lignin biosynthesis.PAL	_	IS	_	_	_	_	_
Lus10019725	PAO1	_	polyamine metabolism.degradation.polyamin oxidase	_	_	OS	_	_	_	CW
Lus10022070	PAP85	_	development.storage proteins	S	_	_	_	_	_	CW
Lus10042615	PAP85	_	development.storage proteins	S	_	_	_	_	_	CW
Lus10042617	PAP85	_	development.storage proteins	S	_	_	_	_	_	CW
Lus10018156	PDIL1-1	_	redox.thioredoxin.PDIL	S	IS	OS	L	R	_	CW
Lus10021478	PDIL1-4	_	redox.thioredoxin.PDIL	_	IS	OS	L	R	_	CW
Lus10022581	PDIL1-4	_	redox.thioredoxin.PDIL	_	IS	OS	L	R	_	CW
Lus10037403	PDIL1-4	_	redox.thioredoxin.PDIL	_	IS	OS	L	R	_	CW
Lus10035870	PDIL1-5	_	redox.thioredoxin.PDIL	S	_	_	_	_	_	CW
Lus10015160	PDIL2-2	_	redox.thioredoxin.PDIL	_	IS	OS	_	_	_	CW
Lus10031522	PDIL2-2	_	redox.thioredoxin.PDIL	S	IS	OS	_	R	_	CW

Lus10016798	PDIL5-2	_	redox.thioredoxin.PDIL	_	IS	OS	_	_	_	CW
Lus10021028	PGIP1	pectin	cell wall.degradation.pectate lyases and polygalacturonases	S	_	_	_	_	_	CW
Lus10021029	PGIP1	pectin	cell wall.degradation.pectate lyases and polygalacturonases	S	_	OS	_	R	_	CW
Lus10023824	PGIP1	pectin	cell wall.degradation.pectate lyases and polygalacturonases	S	_	_	L	R	_	CW
Lus10001200	PHS2	_	major CHO metabolism.degradation.starch.starch phosphorylase	S	_	OS	_	_	GH3	_
Lus10006497	PLAT2	_	not assigned.no ontology	_	_	OS	_	_	_	CW
Lus10013344	PME3	pectin	cell wall.pectin*esterases.PME	_	_	OS	L	_	CE8	CW
Lus10039314	PME3	pectin	cell wall.pectin*esterases.PME	_	_	_	_	R	CE8	CW
Lus10040934	PMM	sugar precursor	cell wall.precursor synthesis.phosphomannomutase	_	_	OS	_	_	_	_
Lus10020480	PR1	_	stress.biotic	_	_	_	_	R	_	CW
Lus10020491	PR1	_	stress.biotic	_	_	_	L	_	_	CW
Lus10013165	PRK6	_	signalling.receptor kinases.leucine rich repeat III	S	_	_	_	_	_	CW
Lus10003573	PRX52	lignin	misc.peroxidases	_	_	_	_	R	AA2	_
Lus10009932	PRX52	lignin	misc.peroxidases	_	_	OS	_	_	AA2	CW
Lus10009933	PRX52	lignin	misc.peroxidases	_	_	_	_	R	AA2	_
Lus10030148	PRX52	lignin	misc.peroxidases	_	IS	OS	_	R	AA2	_
Lus10034207	PRX52	lignin	misc.peroxidases	_	_	OS	_	R	AA2	CW
Lus10026817	PYD2	_	nucleotide metabolism.degradation.pyrimidine.dihydropyrimidinase	S	_	_	L	_	_	CW
Lus10011834	QUA1	pectin	cell wall.pectin synthesis	_	IS	OS	_	R	GT8	_
Lus10043276	RAN3	_	signalling.G-proteins	_	IS	_	L	R	_	CW
Lus10032926	RCI3	lignin	misc.peroxidases	_	_	_	_	R	AA2	CW
Lus10027164	RCI3	lignin	stress.abiotic.cold	_	_	_	_	R	AA2	CW
Lus10031548	RCI3	lignin	stress.abiotic.cold	_	_	OS	_	_	AA2	CW
Lus10027877	RD21B	_	protein.degradation.cysteine protease	S	_	_	_	_	_	CW
Lus10004640	RGP3	Polysaccharide polymer (synthesis)	cell wall.cell wall proteins.RGP	S	_	_	_	_	GT75	_
Lus10026676	RGP3	Polysaccharide polymer (synthesis)	cell wall.cell wall proteins.RGP	S	_	_	_	_	GT75	_
Lus10010942	RHM1	sugar precursor	cell wall.precursor synthesis.RHM	S	_	_	_	_	_	_

Lus10038146	RHM1	sugar precursor	cell wall.precursor synthesis.RHM	S	_	_	_	_	_	_
Lus10042497	RHM1	sugar precursor	cell wall.precursor synthesis.RHM	S	_	_	_	_	_	_
Lus10032035	RHS19	lignin	misc.peroxidases	_	_	_	_	R	AA2	CW
Lus10004040	RKL1	_	signalling.receptor kinases.leucine rich repeat III	_	IS	OS	_	_	_	CW
Lus10017996	RKL1	_	signalling.receptor kinases.leucine rich repeat III	_	_	OS	_	_	_	CW
Lus10003110	RNS3	_	RNA.processing.ribonucleases	_	_	_	_	R	_	CW
Lus10040666	ROC7	_	cell.cycle.peptidylprolyl isomerase	_	IS	OS	_	_	_	CW
Lus10026327	RRA3	AGP	not assigned.unknown	_	IS	_	_	_	GT77	_
Lus10020887	RSW3	Hemicellulose (xyloglucan)	misc.gluco-, galacto- and mannosidases.alpha-galactosidase	S	IS	OS	L	R	GH31	CW
Lus10033490	RSW3	Hemicellulose (xyloglucan)	misc.gluco-, galacto- and mannosidases.alpha-galactosidase	S	IS	OS	_	R	GH31	CW
Lus10014382	SBE2.1	_	major CHO metabolism.synthesis.starch.starch branching	S	_	_	_	_	GH13	_
Lus10038240	scpl20	_	protein.degradation.serine protease	S	_	_	_	_	_	CW
Lus10007732	scpl3	_	protein.degradation.serine protease	S	_	_	_	_	_	CW
Lus10018924	scpl42	_	protein.degradation.serine protease	S	_	_	_	_	_	CW
Lus10028625	scpl42	_	protein.degradation.serine protease	S	_	_	_	_	_	CW
Lus10038691	SCPL49	_	protein.degradation.serine protease	S	_	_	_	_	_	CW
Lus10004958	SERK1	_	signalling.receptor kinases.leucine rich repeat II	S	_	_	_	_	_	CW
Lus10040396	SESA5	_	not assigned.unknown	S	_	_	_	_	_	CW
Lus10029118	SEX1	_	major CHO metabolism.degradation	S	_	_	_	_	CBM45	_
Lus10031051	SFGH	_	C1-metabolism	S	_	_	_	_	CE1	_
Lus10011882	SGT	_	lipid metabolism.exotics (steroids, squalene etc)	_	IS	_	_	_	GT1	_
Lus10011883	SGT	_	lipid metabolism.exotics (steroids, squalene etc)	S	_	_	_	_	GT1	_
Lus10033135	SHD	_	stress.abiotic.heat	S	IS	OS	_	R	_	CW
Lus10034535	SHD	_	stress.abiotic.heat	_	IS	OS	_	_	_	CW
Lus10014903	SHV3	cellulose	lipid metabolism.glycerophosphodiester phosphodiesterase	_	_	OS	_	_	_	_
Lus10016584	SHV3	cellulose	lipid metabolism.glycerophosphodiester phosphodiesterase	_	_	OS	L	_	_	CW
Lus10032576	SHV3	cellulose	lipid metabolism.glycerophosphodiester phosphodiesterase	_	_	OS	_	_	_	CW

Lus10005649	sks5	_	not assigned.no ontology	_	_	OS	_	_	AA1	CW
Lus10021238	sks5	_	not assigned.no ontology	S	_	_	_	_	AA1	CW
Lus10021832	sks5	_	not assigned.no ontology	_	_	OS	_	_	AA1	CW
Lus10004339	SKU5	lignin	development.unspecified	_	IS	OS	_	_	AA1	CW
Lus10024604	SKU5	lignin	development.unspecified	S	IS	OS	_	_	AA1	CW
Lus10028921	SKU5	lignin	development.unspecified	_	IS	OS	_	_	AA1	CW
Lus10032241	SKU5	lignin	development.unspecified	S	_	_	_	_	AA1	CW
Lus10034647	SPPA	_	protein.degradation	S	_	_	_	_	_	CW
Lus10008056	SPS1F	_	major CHO metabolism.synthesis.sucrose.SPS	S	_	_	_	_	GT4	_
Lus10038119	SPS1F	_	major CHO metabolism.synthesis.sucrose.SPS	S	_	_	_	_	GT4	_
Lus10018116	SS1	_	major CHO metabolism.synthesis.starch.starch synthase	S	_	_	_	_	GT5	_
Lus10008451	SSL3	_	secondary metabolism.N misc.alkaloid-like	S	IS	_	_	_	_	CW
Lus10013370	SSL3	_	secondary metabolism.N misc.alkaloid-like	_	IS	OS	_	_	_	CW
Lus10010308	SUS3	cellulose	major CHO metabolism.degradation.sucrose.Susy	S	_	_	_	_	GT4	_
Lus10013417	SUS3	cellulose	major CHO metabolism.degradation.sucrose.Susy	_	IS	_	_	_	GT4	_
Lus10012454	SUS4	cellulose	major CHO metabolism.degradation.sucrose.Susy	_	IS	OS	_	_	GT4	_
Lus10020506	SUS4	cellulose	major CHO metabolism.degradation.sucrose.Susy	S	IS	OS	_	R	GT4	_
Lus10022211	UAH	_	protein.degradation	_	IS	_	_	_	_	CW
Lus10018057	UFD1	_	protein.degradation.ubiquitin	S	_	_	_	_	_	CW
Lus10004656	UGD1	sugar precursor	cell wall.precursor synthesis.UDP-Glc dehydrogenase (UGD)	S	_	_	_	_	_	_
Lus10026656	UGD1	sugar precursor	cell wall.precursor synthesis.UDP-Glc dehydrogenase (UGD)	S	_	_	_	_	_	_
Lus10037096	UGD1	sugar precursor	cell wall.precursor synthesis.UDP-Glc dehydrogenase (UGD)	_	IS	OS	_	_	_	_
Lus10001822	UGE5	sugar precursor	cell wall.precursor synthesis.UGE	_	_	OS	_	_	_	_
Lus10007370	UGP2	callose	glycolysis.cytosolic branch.UGPase	_	_	OS	_	_	_	_
Lus10010957	UGP2	callose	glycolysis.cytosolic branch.UGPase	_	_	OS	_	_	_	_
Lus10020788	UGP2	callose	glycolysis.cytosolic branch.UGPase	S	IS	OS	L	R	_	_
Lus10001905	UGT72B3	_	secondary metabolism.flavonoids.dihydroflavonols	S	_	_	_	_	GT1	_
Lus10017825	UGT74F1	_	hormone metabolism.salicylic acid.synthesis-degradation	S	_	_	_	_	GT1	_

Lus10025854	UGT78D2	_	secondary metabolism.flavonol-3-O-rhamnosyltransferase	S	_	_	_	_	GT1	_
Lus10007421	UGT85A1	_	hormone metabolism.cytokinin.synthesis-degradation	S	_	_	_	_	GT1	_
Lus10036967	UGT85A4	_	misc.UDP glucosyl and glucoronyl transferases	S	_	_	_	_	GT1	_
Lus10010274	UNE5	_	redox.thioredoxin.PDIL	S	IS	_	L	R	_	CW
Lus10020691	UNE5	_	redox.thioredoxin.PDIL	S	_	_	_	_	_	CW
Lus10036337	UNE5	_	redox.thioredoxin.PDIL	S	_	_	_	_	_	CW
Lus10028658	USP	sugar precursor	cell wall.precursor synthesis.NDP sugar pyrophosphorylase	S	_	_	_	_	_	_
Lus10025293	UXS1	sugar precursor	cell wall.precursor synthesis.UXS	_	IS	_	_	_	_	_
Lus10003605	UXS4	sugar precursor	cell wall.precursor synthesis.UXS	_	IS	OS	_	R	_	_
Lus10005450	UXS5	sugar precursor	cell wall.precursor synthesis.UXS	S	_	_	_	_	_	_
Lus10040847	UXS5	sugar precursor	cell wall.precursor synthesis.UXS	S	IS	OS	_	_	_	_
Lus10001707	UXS6	sugar precursor	cell wall.precursor synthesis.UXS	S	_	_	_	_	_	_
Lus10005155	UXS6	sugar precursor	cell wall.precursor synthesis.UXS	_	IS	OS	_	_	_	_
Lus10014176	VSR1	_	protein.targeting.secretory pathway.vacuole	_	IS	OS	L	R	_	CW
Lus10030722	XCP2	_	protein.degradation.cysteine protease	_	IS	_	_	_	_	CW
Lus10009313	XSP1	_	protein.degradation.subtilases	_	IS	_	_	_	_	CW
Lus10023009	XTH32	Hemicellulose (remodelling)	cell wall.modification	_	_	OS	_	_	GH16	CW
Lus10030923	XTH5	Hemicellulose (XTH)	cell wall.modification	_	_	_	_	R	GH16	CW
Lus10035019	XYL1	Hemicellulose (xylan, xyloglucan)	misc.gluco-, galacto- and mannosidases.alpha-galactosidase	S	_	_	_	_	GH31	CW
Lus10041457	XYL1	Hemicellulose (xylan, xyloglucan)	misc.gluco-, galacto- and mannosidases.alpha-galactosidase	S	_	_	_	_	GH31	CW
Lus10020966	XYL4	Hemicellulose (xyloglucan)	cell wall.degradation.mannan-xylose-arabinose-fucose	_	_	_	_	R	GH3	CW

Partie 2:Mise en évidence de la variabilité génotypique associée aux différences de métabolisme pariétal chez le lin

Genotype- and environmentally-dependent differences in gene accumulation and protein abundance in flax (*Linum usitatisimum* L).

The results presented in this manuscript will be integrated with those of other partners from the FIBRAGEN consortium (article in preparation). For space reasons, only the supplementary data 3 are presented in the printed version of this paper.

Chabi M[a]; Lucau A[*] ; Arribat S[a]; Goulas E[a] Blervacq AS[a], Renaud J[b] ; Hawkins S[a].

[a.]Unité de Glycobiologie Structurale et Fonctionnelle, UMR8576 CNRS-Université Lille 1, 59655 Villeneuve d'Ascq Cedex, France

[b]Centre de Recherche Public-Gabriel Lippmann, L-4422 Belvaux, Luxembourg

Abstract

A comparative transcriptomic and proteomic analyses was undertaken on the outer stem tissues of 7 flax varieties (4 spring fiber varieties, 2 winter fiber varieties and 1 oil winter variety) grown in the field over 2 consecutive years. 659 genes are differentially expressed at the variety level and 1571 genes are differentially expressed between years in at least one variety. The winter oil variety seemed to be the most different from the spring fiber varieties and functional classification indicated that some of these genes were related to cell wall metabolism (COBRA, PMEI, PEA) and might therefore play a role in defining fiber quality Proteomic analysis highlighted the differential abundance of a flax protein belonging to the XTH IIIA family between a spring fiber and winter oil varieties. Detailed analyses of unassigned peptides spectra allowed us to define a new gene model for a flax expansin. Altogether, these results represent an important contribution to our understanding of the molecular processes potentially underlying fiber quality in flax

Introduction

Global warming and diminishing fossil fuel reserves are driving the transition towards to a greener economy. In addition to alternative energy sources, it is also vital to develop new environmentally friendly materials. Natural fibers from plant species such as flax and hemp are known to possess interesting mechanical properties and represent an interesting alternative to the use of glass fibers in composites for the transport and construction industries(Summerscales et *al.,* 2010; Bourmaud et *al.,* 2013). However, the development of an efficient natural fiber based industry depends on a better understanding of the different factors affecting both fiber and material quality (Lefeuvre et *al.*, 2013)

Fiber quality itself is the result of the architecture and structural composition of the fiber composite cell wall consisting of the middle lamella, primary cell wall and thick secondary cell wall characterized by high amounts of crystalline cellulose and very low lignin levels (Morvan et *al.*, 2003b; Day et *al.*, 2005c; Bourmaud et *al.*, 2013). Although some broad correlations can currently be made between cell wall structure and fiber quality, (e.g. negative relationship between lignin content and textile fiber quality), it is likely that many other more or less subtle variations in cell wall structure and metabolism could contribute to differences in fiber quality. In order to obtain some preliminary information about the genetic variability of cell wall metabolism in flax we performed a comparative whole genome transcriptomics analyses on the fiber-bearing tissues from seven different varieties grown under field conditions. Current flax varieties consist of 2 main types – i) fiber varieties and ii) oil varieties that have been selected for in different breeding programmes. These varieties are planted in France and other European countries in the spring and are harvested at the end of the summer in the same year. Other varieties have been selected for their resistance to low temperature conditions enabling them to be planted in the winter and harvested the following year.

Whole genome transcriptomics have been previously used to compare climacteric and non-climacteric varieties of developing melons (Saladié et *al.*, 2015), to analyse postharvest network senescence processes in 4 citrus varieties (Ding et *al.*, 2015), to determine the factors affecting fruit length in two cucumber varieties (Jiang et *al.*, 2015) and to compare chilling tolerant rice varieties during early chilling stress (Lindlöf et *al.*, 2015). This type of strategy is generally considered to be an unbiased analytical approach for comparative studies (van Dijk et *al.*, 2014). In flax, this approach has been used to analyse drought stress (Dash et *al.*, 2014), to study paleopolyploidy events within the flax genus (Sveinsson et *al.*, 2014), to investigate the natural

hypolignification of flax fibers (Huis et *al.*, 2012), and to study spatiotemporal formation and deposition of polymeric lignans and plant defense cyanogenic glucosides (Dalisay et *al.*, 2015).

In addition to transcriptomics, we also compared the proteomes of a flax spring fiber variety and a winter oil variety. Proteomics has been previously used to study the responses of two citrus rootstocks to water deficit (Oliveira et *al.*, 2015) and the response of two wheat varieties to nitrogen induced stress (Chandna and Ahmad, 2015). Differential proteomic profiling was also used to compare 2 GM bean varieties with non-GM varieties (Valentim-Neto et *al.*, 2015), rice varieties (Yang et *al.*, 2014) and grape varieties (Fraige et *al.*, 2015). This method was also used to identify proteins enriched in best flax fiber (Hotte and Deyholos, 2008). Proteomic analysis were recently developed for flax (Hotte and Deyholos, 2008; Day et *al.*, 2013), that allow us to use their extraction protocol, and improve it.

Contrasted growing conditions and a wide range of abiotic stress are also known to impact on cell wall metabolism and structure. For example in responses to drought stress, both pectin and particularly xyloglucan composition, as well as arabinogalactan proteins were modified in resurrection plants (Moore et *al.*, 2006; Leucci et *al.*, 2008). Interestingly the response to the same stress varied between the different varieties studied indicating a combined effect of both genotype and environment (Piro et *al.*, 2003; Konno et *al.*, 2008). Similar observations were reported for soybean roots that accumulate pectin under salt stress, as well as for three tolerant wheat cultivars that show an increase in arabinoxylan in dietry fibers under heat and drought stress (An et *al.*, 2014)(Rakszegi et *al.*, 2014). Cold stress was also reported to have an impact on the cell wall structure in different frost tolerant Miscanthus genotypes that accumulated pectin, lignin and some cell wall proteins (Domon et *al.*, 2013).

In order to evaluate the potential environmental impact on flax fiber cell wall metabolism we performed both transcriptomics and proteomics on samples grown from 2 consecutive years.

Material and Methods

Plant growth and materiel

Seven Flax cultivars: Diane, Hermes, Drakkars, Belinka, Adelie, Violin, and Oliver were grown under field conditions by LINEA, Grandvillier France during 2012 and 2013. Stem outer tissues (containing bast fibers) were harvested from flax plants prior to the flowering stage and plunged in liquid nitrogen before storing at -80° C for transcriptomics and proteomics.

Histological analysis

Ethanol-fixed samples of the median region of flax whole stems were dehydrated using ethanol series. Three plants were used and for each plant, three hand-cross sections were used to perform the measurements. Toluidine-Blue O (TBO) coloration was used to allow the visual separation between the inner and outer tissues and facilitate fiber bundle analysis using imageJ software. Significant differences were evaluated using ANOVA.

Transcriptomics

RNA extraction and quality verification

Total RNA was extracted from three pools of outer stem tissues for each of the 7 varieties using the TriReagent kit (Molecular Research Center) method. RNA integrity and concentration were evaluated with RNA StdSens Chips using the Experion automated electrophoresis system (Bio-Rad).

Hybridization and data analyses

RNA processing, CY3-labeling and hybridization were made following manufacture's instruction for One-Color Microarray base gene Expression Analysis (Agilent Technologies). Hybridization was performed on Agilent microarrays Agilent-045382 UGSF flax 45K v1.0 array based upon flax genome coding sequence (Wang et *al.*, 2012) available at Phytozome (http://phytozome.org). The 4 plex array contains 45,220 60-mer in situ oligonucleotides per block. Hybridization and washing were performed following Agilent manufacture's instruction, and slides were immediately scanned at 5 mm pixel-1 resolution using an Axon GenePix 4000B scanner (Molecular Devices Corporation, Sunnyvale, CA, USA) piloted by GenePix Pro 6.0 software (Molecular DevicesCorporation, Sunnyvale, CA, USA).

Scanned images (TIFF format) were then submitted to grid alignment and expression data analyses. For each slide a global lowess followed by a print-tip median normalization was performed using R packages (http://cran.r-project.org/) as implemented in CLC bio software (http://www.clcbio.com), followed by an inter-slide normalization.

For each condition, three repetitions were performed. To identify genes displaying a significant change in expression over the repetitions, a script utilizing library functions in R with a Bonferroni-corrected P value of less than 5% and a cut-off of +/- 1.5 log2 ratio was used for all experimental conditions. Only genes with smooth expression profiles were retained. Differential analyses were performed for each variety *vs* the other 6 varieties giving a total of 21 possible combinations.

Proteomics:

Protein extraction

Four g frozen outer stem materiel were ground in liquid nitrogen to a fine powder, followed by 5 mins grinding in 10 ml of Tris HCl buffer 50 mM, 0,06% IPC, pH 7.5, before centrifuging 10min at 4°C /16000g. The supernatant was incubated 15min at room temperature with protamine sulfate under a low agitation then centrifuged 10 min at18000g. The pellet was discarded, and the proteins were precipitated with 10% TCA, 1 hour at -20°C. The pellet was washed 1 time with cold acetone, then dried 5min at room temperature.

Dry protein pellets were resolved in 2D-DIGE labeling buffer [7 M urea, 2 M thiourea, 4% (w/v) CHAPS and 30 mM Tris] at room temperature with agitation for 4 hours. The pH of the solution was adjusted to 8.5, and protein concentration was determined by 2D Quant kit (GE Healthcare). The protein samples were labeled with CyDye minimal dyes (GE Healthcare) according to manufacturer's instructions. Thirty µg of proteins were labeled by adding 240 pmol of fluorochromes (Cy3 or Cy5). A pooled internal standard was performed by mixing 15 µg of each sample and labeled with Cy2 dye and included in all gel runs.

For each 2-D gel a mixture with a final volume of 120 µl containing two samples labeled cy3 and cy5, internal standard labeled with cy2, and labeling buffer [7 M urea, 2 M thiourea, 4% (w/v) CHAPS, 6 µl.ml−1 DeStreak reagent (GE Healthcare)] and 1,5% (v/v) Biolyte, were loaded onto a 24 cm IPG non-linear strip with a pI range 3−10 (Bio-Rad) and separated using the IPGphor3 system with a rapid voltage slope to reach the maximum of 10 000 V and a total of 72 000 Vh. Second dimension resolution was carried out by SDS-PAGE on a 12.5% (v/v) resolving gel (EttanDALTsix, GE Healthcare) according to 2D HPE Large Gels manufacturer instructions.

Image analysis and protein identifications

Scanning of the 2D-DIGE gels was carried out using a Typhoon Imager 9400 (GE Healthcare) at three different wavelengths corresponding to the different CyDyes.The generated gel images were analyzed with DeCyder 7.0 software, using the Cy2 channel as a standard.

Gene and Protein annotation:

The differentially expressed genes and proteins were annotated using MapMan database (Lohse et al 2013).

Results

Varieties show differences in fiber bundle area and fiber number

In this work we compared 4 spring fiber varieties (Diane, Drakkar, Hermes, Belinka), 2 winter fiber varieties (Adelie, Violin) and 1 winter oil variety (Oliver). Microscopic analyses of freehand cross-sections made from the middle region of flax stems revealed that there are significant differences in fiber bundle area and fiber number between certain varieties (Fig.1). Generally, the 4 spring fibers varieties had significantly greater (ANOVA, $P < 0.05$) bundle surface areas and number of fibers per bundle than both winter fiber and winter oil varieties in both the 2012 and 2013 samples. The one exception to this trend concerned the 2012 fiber bundle area data where no significant differences could be observed between spring- and winter-fiber varieties. Although no significant differences were detected between spring fibers varieties in 2012 samples, both Belinka and Diane showed greater bundle surface area and fiber number per bundle than Hermes and Drakkar in the 2013 samples.

Whole genome transcriptomics reveals important inter-year and inter-variety differences in gene expression

Comparative whole transcriptomics of the fiber-bearing outer-stem tissues from the 7 varieties harvested in 2012 and 2013 allowed us to identify 2,451 genes that showed significant differences in transcript accumulation (Supplementary Data 1).

Principal component analyses performed on the differentially expressed genes (DEG) show a strong impact of the years on the data repartition (Fig. 2), with more than 43% (PC1) of the difference shown being explained by the years. In addition, the PCA also allows a clear separation between the oil variety (Oliver) and both spring- and winter-fiber varieties. Within the fiber varieties, the winter fiber variety (Violin) is also clearly separated from the spring fiber varieties.

DEGs were then separated at two levels according to the method used by Ding et *al.* 2015: i) differences between the seven varieties for each year (= 'variety level'), and ii) differences between the 2 years for each variety (= 'environmental' level). For the variety level, 42 comparisons were made (21 comparisons per year) and for the environmental level 7 comparisons were made.

At the variety level, we considered only those genes showing the same or similar values in a pairwise comparison between 2 varieties for the two years. This approach allowed us to select a total of 659 DEGs (Fig. 3A). Our results show that 4 of the 5 pairwise comparisons with the highest numbers of differentially expressed genes corresponded to winter oil (Oliver) *vs.* spring

fiber (Belinka, Drakkar, Hermes, Diane) varieties. Functional classification of these genes using MapMan (Lohse et al 2013) indicated that 25 of the 35 possible functional categories were represented (Fig.3B) (Supplementary Data 2). Highly represented classes included 35 (not assigned), 26 (miscellaneous) and 20 (stress). Cell wall related genes (class 10) represented between 0 and 6 % of differentially-expressed genes depending upon the variety couple. Principal component analyses and hierarchical clustering of this subset of genes (Fig. 4) confirmed the separation of varieties previously observed with the total gene set. PCA once again separated the winter oil variety (Oliver) (and the winter fiber variety, Violin) from the spring fiber varieties with the genotype effect explaining more than 30% of the observed difference (PCA1). Despite selecting only those genes showing the same or similar values in a pairwise comparison between 2 varieties for the two years our results show that 24 % of the variation can still be attributed to differences between the 2 years (PC2). Examination of PC3 (Fig. 4B) showed that the spring, winter fiber varieties Belinka and Adelie respectively could be separated from the other fiber varieties (Fig. 4B). The separation of the flax varieties by PCA was confirmed by hierarchical analysis of the gene subset (Fig. 4C**).** Both Violin and Oliver are separated from the other varieties with 2012/2013 data being grouped together for these 2 varieties. In contrast, the fiber varieties show more variability and only Belinka and Diane have grouped 2012/2013 data. These results would suggest that the environment has a greater/lesser impact on gene expression depending upon the variety considered.

At the environmental level the comportment of the same variety is compared for two years, in this way 7 comparisons were performed (Fig. 5). Our results show that the number of genes differentially expressed between the years varies according to the variety considered. The variety Hermes showed the highest number of DEG (798) genes), followed by Drakkar (704), Oliver (674), Adelie (639), Belinka (606), Diane (462) and Violin (402) (Fig. 5a). For each variety comparison the DEGs could be divided into 3 groups: i) DEGs that are common to all 7 varieties, ii) DEGS that are only found in one variety and iii) DEGs that are found in between 2 and 6 varieties. Functional classification of the class 1 DEGs indicated that the most represented bin is stress (29 %) including genes involved in biotic and abiotic stress responses (Fig 5b). Functional classification (MapMan) of genes in the second class (Fig. 5c) provided an overview of variety-specific responses to the environment. Although the not-assigned bin was the most represented category in all varieties, no other similarities between the different varieties could be detected. In contrast, when all spring fiber variety DEGs and winter fiber variety DEGs from the third class were grouped together we were able to identify 63 DEGs specific to the first group (spring fiber) and 101 DEGs specific to the second group (winter fiber). Functional

classification of these genes (Fig. 5d) indicated that for the winter fiber varieties the bin 26 (miscellaneous) was the most represented. For spring fiber varieties, 14 out of the 35 possible bin classes were represented with the bins 35 (unknown, 20 %) and 27 RNA, 19 %) being the most represented.

Cell wall related genes

The goal of this study is to try and link differential gene expression to cell wall quality and/or quantity and so we therefore selected 87 DEGs related to cell wall metabolism (= MapMan cell wall bin plus expert selection of other cell wall DEGs not included in the MapMan bin) (Supplementary Data 3). Twenty-six of the 87 cell wall related DEGs belong to the variety level DEGs, 61 belong to the environmental level and 18 are common to both groups (Fig. 6). The most highly represented DEGs are associated with pectin metabolism (degradation/modification) with 31 genes, followed by the hemicellulose related genes (synthesis/ modification) with 15 DEGs. We also identified genes involved in cell wall modifications (expansins), lignin biosynthesis (peroxidases, laccases) and related to cellulose (*COBRA*). Interestenly *COBRA* genes are highly expressed in 3 spring fiber varieties compared to the winter oil variety.

Comparative proteomics reveals important inter-year and inter-variety differences

Both histological analyses and transcriptomics indicated that the oil variety Oliver could be distinguished from the spring fiber varieties and we therefore selected the varieties Diane (Spring fiber variety) and Oliver (Winter oil) for differential proteomic analyses by 2D DIGE gel electrophoresis. Proteins were extracted from the same 2012/2013 material used for transcriptomics analysis.

278 protein spots showed significant differential abundance (at least 1.5 fold change, $p \leq 0.05$) between the varieties and years. Subsequent excision, trypsin digestion and mass spectrometry allowed thesuccessful identification of 129 proteins representing 59 non-redundant proteins (Table 1). As for transcriptomics, differential proteins were separated into i) a variety level group and ii) an environmental level group (Fig. 7). Our results show that 16 proteins weredifferentially accumulated at the variety level and 31 at theenvironment level. 8 were in common between both.

Identified proteins were classified into different functional categories using MapMan (Fig. 7). 18 of the 35 possible functional categories were represented with the class ‘photosynthesis’ (bin

1) being the most highly represented (21 % of total proteins) (12 proteins), followed by 'miscellaneous', 'redox' and 'not assigned' (bin 26, 21 and 35) with 8, 77% (5 proteins). The category 'cell wall' (bin 10), together with 'amino acid metabolism' (bin 13) and 'protein' (bin 29), constituted the 3rd most represented classes with7%(4 proteins). Altogether we were able to identify 5 proteins related to cell wall metabolism including 4 proteins from the cell wall bin (XTH6, XTH32, XTH8 and a beta-glucosidase) and a UDP-glucose/galactose epimerase.

New gene model for a putative expansin/EG45-like domain containing protein.

We identified flax proteins by using Mascot software based on the flax genome sequence (*Linum usitassinum* genome *v.*1) present in the Phytozome database (http://phytozome.jgi.doe.gov/pz/port*al.*html). Despite the efficiency of this procedure 4 high quality spectra/peptide sequences (ion scores 37, 96, 101, 115) were not identified. Subsequent investigation using NCBI public databases identified a *Linum usitatissimum*EST corresponding to an Expansin protein but not integrated into the first version of the flax genome. A proposed model of the flax expansin protein based on homologues sequences found in other plant genomes is shown in Fig. 8). In Oliver, but not Diane, the expansin protein was up-regulated in 2013 compared to 2012.

Transcriptomic and proteomics cross-link

Both transcriptomics and proteomics were performed on the same material and we were therefore able to compare differentially expressed genes and proteins. Our results (Table 2) allowed us to identify 6 genes/proteins differentially expressed/ accumulated and including 4 genes/proteins involved in cell wall metabolism (3 Xyloglucan endotransglycosylases/hydrolases : XTH and 1 beta glucosidase 32), as well as 1 cyclase family protein and 1 unknown protein. Of the 3 XTH proteins identified, only the flax gene encoding an XTH32 ortholog shows a significant difference between Oliver and Diane for 2012 and 2013. A phylogenetic tree of the flax XTH IIIA gene family (Fig. 9) reveals the strong paralogy of this family in the flax genome as previously reported (Chabi et *al.*, submitted, Flax Proteome Atlas).

Discussion

The quality of cultivated flax varieties is related to the fiber content of the stems and the bundle surface area (Brutch et *al.*, 2008; Charlet et *al.*, 2009; Charlet et *al.*, 2010). Our results indicated significant differences between the varieties in terms of the fiber bundle surface area and fibers

numbers. The difference is observed mostly between the spring fiber variety (SF) and the winter fiber variety (WF) and is in agreement with previous observations (Brutch *et al.*, 2008; Diederichsen and Ulrich, 2009). Significant differences were also observed between the two tested years with the difference being the most marked for the SF as previously reported (Brutch et *al.*, 2011). Genotype variability in *Linum usitatisimum* L is reported to be low (Terre de lin, Linea flax farmers, personal comments). Nevertheless sufficient genetic variability exists and has allowed the use of quantitative genetic approaches in linseed (Soto-Cerda et *al.*, 2014; Kumar et *al.*, 2015). The use of flax microarrays has also demonstrated the existence of variety-dependent differences in gene expression (Fenart et *al.*, 2010).

Variety Level DEGs

At the variety level we observed differences in the number of DEGs obtained for each variety comparison. Overall these differences are related to the 'closeness' between the varieties with, for example, Drakkar and Hermes being the 'closet' varieties. This observation correlates well with the hierarchical dendrogram of microarray expression data, and the histological analysis showing that there is no significant difference between these 2 varieties in terms of bundle surface area and fiber number. The most important difference in gene number was observed when SF varieties were compared to Oliver (WO). Violin (WF) is the closest to Oliver whereas Adelie (WF) shows the same difference to Oliver as the other SF varieties also in agreement with the histological data. Among the DEGs between the WF and SF varieties it is interesting to note that the gene encoding AGAMOUS-like 67, related to flowering inititiation, is down regulated in SF varieties compared to WF varieties, as is a gene belonging to the UDP-Glycosyltransferase superfamily known to play a key role in anthocyanin production and correlated with flowering. These differences are presumable related to the differences in the life cycle between Winter and spring flax varieties (Alister D. Muir, 2003).

Other DEGs of interest that are up-regulated in FS varieties compared to WV include a member of the CYP76C family associated with increased pesticide resistance in *A thaliana*, as well as a gene encoding a TIR-NBS-LRR resistance gene (Höfer et *al.*, 2014) .

A gene encoding an F-box protein SLY1 is up-regulated in SF varieties compared to the WO variety and is known to play an important role in growth as mutants have a dwarf phenotype and low germination (Ariizumi et *al.*, 2011). Interestingly, a gene encoding jmjC, which is known to be involved in gibberellin associated plant growth is also up-regulated in SF varieties compared to the WO variety. It is possible that the differential expression of this gene might contribute to

the observed differences between the size of fiber varieties and oil varieties in addition to , the effect of agricultural practices (e.g. seed density) (Diederichsen and Richards, 2003).

Environmental Level DEGs

At the environment level, 1571 genes are differentially expressed between the years for at least one variety. Among these genes, 107 are differentially expressed in all 7 varieties. Twenty of these genes encoded heat shock proteins and were down regulated in 2012 compared to 2013. In contrast 5 genes encoding WCOR413-like cold acclimation proteins are up regulated in 2012 *vs* 2013. Interestingly, analyses of themeterological data for both years indicated that the average temperatures during plant growth were higher in 2013 (mean max: 16.8 °C; mean min: 9.1 °C) than in 2012 (mean max: 14.5 °C; mean min: 6.7 °C).

Cell wall related genes

Most of the genes belonging to this category were detected at the environmental level. This observation is in agreement with data showing that fiber production and quality fluctuates between the years and can be related to contrasted weather conditions (Brutch et *al.*, 2011). In this study the extent to which environmental conditions affected quality depended upon the variety studied.. Among the DEGs observed in our study we identified 7 genes encoding FLA proteins: 2 orthologs of FLA11, 2 orthologs of FLA12, 1 ortholog of FLA7 and 2 orthologs of FLA2. Of these FLAs, the FLA 12 orthologs are up-regulated in Diane, Drakkar, and Hermes compared to Oliver. This gene is reported to expressed specifically in stems, and stem tensile strength and stiffness are affected in the corresponding *Arabidopsis* mutant(MacMillan et *al.*, 2010). A correlation has also been demonstrated between FLA12 transcript accumulation and cellulose microfibril orientation and wood properties, in *Eucalyptus* (Qiu et *al.*, 2008).

Two flax orthologs of the *COBRA* gene that is known to play an important role in determining cellulose crystallinity are also up-regulated in Diane, Drakkar and Hermes compared to Oliver (Sorek et *al.*, 2014). In contrast a gene encoding a fucosyltransferase 1 was up-regulated in Belinka and Violin compared to Diane and Adelie. This protein regulates side-chain fucosylation in pectins and mutations in Arabidopsis are associated with problems in cell wall biosynthesis, structural integrity and a dwarf phenotype (Perrin et *al.*, 1999)(Vanzin et *al.*, 2002). The observation that the fucosyltransferase 1 gene was up-regulated in 1 spring fiber variety and 1 winter fiber variety in comparison with 2 other spring fiber varieties illustrates the subtle differences in cell wall DGE between the different varieties. This fact was underlined by the observation that a gene encoding a desulfoglucosinolate sulfotransferase that has been

shown to play a role in fiber elongation in cotton is up-regulated in flax SF varieties compared to WF varieties (Fang et *al.*, 2014),.

Generally, the most represented cell wall DEGs are those implicated in the degradation/ modification of pectin reflecting the importance of this polymer in different biological processes including cell expansion and response to stress (Willats et *al.*, 2001). Two flax genes encoding PMEIs (Lus10008203,Lus10001467) are up-regulated in three varieties (Diane, Oliver and Violin) compared to the others varieties, while another PMEI (Lus10031138) abundant in flax fiber-enriched tissue and expressed during intrusive fiber growth in stem peel tissues (Pinzon-Latorre and Deyholos, 2014) is up-regulated in Violin compared to Drakkar. Similarly the pectate lyase related genes PEA(Lus10037945) is up-regulated in Oliver compared to Belika, and up regulated in Diane comparing to Adelie, whereas Lus10024398 is down-regulated in Oliver compared to other fiber varieties. Most of the identified genes can potentially be related to fiber quality.

Proteomics allowed us to identify 59 non-redundant, differentially-accumulated proteins of which 6 could be related to cell wall metabolism (a UDP-glactose/glucose epimerase, XTH6, XTH32, XTH8, a beta-glucosidase and an expansin). The identification of 3 XTH proteins out of a total of 6 is remarkable and strongly suggests that these proteins are involved in flax cell wall metabolism. XTH enzymes cut (hydrolase activity) xyloglucan hemicellulose polymers linking cellulose microfibrils and are generally believed to play a key role in cell expansion during cell growth even though recent work has cast doubt upon hypothesis (Cosgrove, 2005; EKlöf and Brumer, 2010; Parks and Cosgrove, 2012, 2015). Subsequently, the XTH relinks (endotransglycosylation activity) the cut end of the xyloglucan to another xyloglucan polymer thereby relinking cellulose microfibrils and contributing to stopping cell expansion (Eklöf et *al.*, 2013; Nishikubo et *al.* 2011). Both XTH activity and gene expression have previously been correlated with G-fiber formation in tension wood in poplar and, given the presence of xyloglucans in the flax fiber cell wall (Chabi et *al.*, submitted, Flax Proteome Atlas), it is not illogical to find XTH proteins in fiber-containing tissues in flax (Nishikubo et *al.*, 2007, 2011; Baba et *al.*, 2009; Mellerowicz et Gorshkova, 2012). Of the 3 XTH proteins identified, only XTH32 shows a significant difference between Oliver and Diane for 2012 and 2013. As previously reported (Chabi et *al.*, submitted, Flax Proteome Atlas) this protein belongs to the IIIA XTH family that contains one loop contributing to predominant hydrolysis activity (Eklöf et *al.*, 2013). The observed paralogy of this protein is particularly interesting and could be related to flax fiber formation and quality (Nishikubo et *al.* 2011).

Acknowledgements

References

Adolphe JL, Whiting SJ, Juurlink BHJ, Thorpe LU, Alcorn J (2009) Health effects with consumption of the flax lignan secoisolariciresinol diglucoside. Br J Nutr **103**: 929

Agarwal UP (2006) Raman imaging to investigate ultrastructure and composition of plant cell walls: Distribution of lignin and cellulose in black spruce wood (Picea mariana). Planta **224**: 1141–1153

Ageeva M V., Petrovská B, Kieft H, Sal'nikov V V., Snegireva a. V., Van Dam JEG, Van Veenendaal WLH, Emons a. MC, Gorshkova T a., Van Lammeren a. a M (2005) Intrusive growth of flax phloem fibers is of intercalary type. Planta **222**: 565–574

Alister D. Muir NDW (2003) Flax: The genus Linum.

An P, Li X, Zheng Y, Matsuura A, Abe J, Eneji AE, Tanimoto E, Inanaga S (2014) Effects of NaCl on Root Growth and Cell Wall Composition of Two Soya bean Cultivars with Contrasting Salt Tolerance. J Agron Crop Sci **200**: 212–218

Anders N, Wilkinson MD, Lovegrove A, Freeman J, Tryfona T, Pellny TK, Weimar T, Mortimer JC, Stott K, Baker JM, et al (2012) Glycosyl transferases in family 61 mediate arabinofuranosyl transfer onto xylan in grasses. Proc Natl Acad Sci U S A **109**: 989–93

Appenzeller L, Doblin M, Barreiro R, Wang H, Niu X, Kollipara K, Carrigan L, Tomes D, Chapman M, Dhugga KS (2004) Cellulose synthesis in maize: isolation and expression analysis of the cellulose synthase (CesA) gene family. Cellulose **11**: 287–299

Ariizumi T, Lawrence PK, Steber CM (2011) The role of two f-box proteins, SLEEPY1 and SNEEZY, in Arabidopsis gibberellin signaling. Plant Physiol **155**: 765–75

Atmodjo MA, Hao Z, Mohnen D (2013) Evolving views of pectin biosynthesis. Annu Rev Plant Biol **64**: 747–79

Baba K, Park YW, Kaku T, Kaida R, Takeuchi M, Yoshida M, Hosoo Y, Ojio Y, Okuyama T, Taniguchi T, et al (2009) Xyloglucan for generating tensile stress to bend tree stem. Mol Plant **2**: 893–903

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

Barnett JR, Bonham VA (2004) Cellulose microfibril angle in the cell wall of wood fibres. Biol Rev Camb Philos Soc **79**: 461–472

Bar-Peled M, Urbanowicz BR, O'Neill MA (2012) The Synthesis and Origin of the Pectic Polysaccharide Rhamnogalacturonan II - Insights from Nucleotide Sugar Formation and Diversity. Front Plant Sci **3**: 92

Barvkar VT, Pardeshi VC, Kale SM, Kadoo NY, Giri AP, Gupta VS (2012) Proteome profiling of flax (Linum usitatissimum) seed: Characterization of functional metabolic pathways operating during seed development. J Proteome Res **11**: 6264–6276

Battaglia M, Solórzano RM, Hernández M, Cuéllar-Ortiz S, García-Gómez B, Márquez J, Covarrubias AA (2006) Proline-rich cell wall proteins accumulate in growing regions and phloem tissue in response to water deficit in common bean seedlings. Planta **225**: 1121–1133

Baumann MJ, Eklöf JM, Michel G, Kallas AM, Teeri TT, Czjzek M, Brumer H (2007) Structural evidence for the evolution of xyloglucanase activity from xyloglucan endo-transglycosylases: biological implications for cell wall metabolism. Plant Cell **19**: 1947–63

Becker B, Marin B (2009) Streptophyte algae and the origin of embryophytes. Ann Bot **103**: 999–1004

Bernhardt C, Tierney ML (2000) Expression of AtPRP3, a proline-rich structural cell wall protein from Arabidopsis, is regulated by cell-type-specific developmental pathways involved in root hair formation. Plant Physiol **122**: 705–14

Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. Annu Rev Plant Biol **54**: 519–46

Bonello P, Blodgett JT (2003) Pinus nigra–Sphaeropsis sapinea as a model pathosystem to investigate local and systemic effects of fungal infection of pines. Physiol Mol Plant Pathol **63**: 249–261

Bosch M, Hepler PK (2005) Pectin methylesterases and pectin dynamics in pollen tubes. Plant Cell **17**: 3219–3226

Bourmaud A, Ausias G, Lebrun G, Tachon M-L, Baley C (2013) Observation of the structure of a composite polypropylene/flax and damage mechanisms under stress. Ind Crops Prod **43**: 225–236

Bray EA (2004) Genes commonly regulated by water-deficit stress in Arabidopsis thaliana. J Exp Bot **55**: 2331–41

Brown DM, Goubet F, Wong VW, Goodacre R, Stephens E, Dupree P, Turner SR (2007) Comparison of five xylan synthesis mutants reveals new insight into the mechanisms of xylan synthesis. Plant J **52**: 1154–1168

Brown DM, Zeef LAH, Ellis J, Goodacre R, Turner SR (2005) Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. Plant Cell **17**: 2281–95

Brutch NB, Sharov IY, Pavlov a. V., Porokhovinova E a. (2011) Diversity of flax characters associated with fibre formation and environmental influence on their formation. Russ J Genet Appl Res **1**: 361–370

Brutch NB, Soret-Morvan O, Porokhovinova E a., Sharov IY, Morvan C (2008) Characters of Fibre Quality in Lines of Flax Genetic Collection. J Nat Fibers **5**: 95–126

Buendgen MR, Coors JG, Grombacher AW, Russell WA (1990) European Corn Borer Resistance and Cell Wall Composition of Three Maize Populations. Crop Sci **30**: 505

Buffetto F, Ropartz D, Zhang XJ, Gilbert HJ, Guillon F, Ralet M-C (2014) Recovery and fine structure variability of RGII sub-domains in wine (Vitis vinifera Merlot). Ann Bot **114**: 1327–1337

Burn JE, Hurley UA, Birch RJ, Arioli T, Cork A, Williamson RE (2002) The cellulose-deficient Arabidopsis mutant rsw3 is defective in a gene encoding a putative glucosidase II, an enzyme processing N-glycans during ER quality control. Plant J **32**: 949–60

Burton R a, Gidley MJ, Fincher GB (2010) Heterogeneity in the chemistry, structure and function of plant cell walls. Nat Chem Biol **6**: 724–732

Caffall KH, Mohnen D (2009) The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. Carbohydr Res **344**: 1879–900

Campargue C, Lafitte C, Esquerré-Tugayé MT, Mazau D (1998) Analysis of hydroxyproline and hydroxyproline-arabinosides of plant origin by high-performance anion-exchange chromatography-pulsed amperometric detection. Anal Biochem **257**: 20–5

Cannon MC, Terneus K, Hall Q, Tan L, Wang Y, Wegenhart BL, Chen L, Lamport DTA, Chen Y, Kieliszewski MJ (2008) Self-assembly of the plant cell wall requires an extensin scaffold. Proc Natl Acad Sci U S A **105**: 2226–31

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res **37**: D233–8

Carpita N, Tierney M, Campbell M (2001) Molecular biology of the plant cell wall: searching for the genes that define structure, architecture and dynamics. Plant Mol Biol **47**: 1–5

Carpita NC (1996) STRUCTURE AND BIOGENESIS OF THE CELL WALLS OF GRASSES. Annu Rev Plant Physiol Plant Mol Biol **47**: 445–476

Carroll A, Mansoori N, Li S, Lei L, Vernhettes S, Visser RGF, Somerville C, Gu Y, Trindade LM (2012) Complexes with mixed primary and secondary cellulose synthases are functional in Arabidopsis plants. Plant Physiol **160**: 726–37

Cassab GI, Varner JE (1988) Cell w all proteins. Ann Rev Plant Physiol Plant Mol Biol **39**: 321–353

Chandna R, Ahmad A (2015) Nitrogen stress-induced alterations in the leaf proteome of two wheat varieties grown at different nitrogen levels. Physiol Mol Biol Plants **21**: 19–33

Chantreau M, Chabbert B, Billiard S, Hawkins S, Neutelings G (2015) Functional analyses of cellulose synthase genes in flax (Linum usitatissimum) by virus-induced gene silencing. Plant Biotechnol J. doi: 10.1111/pbi.12350

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 (2014a) 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 **26**: 4462–82

Chantreau M, Portelette A, Dauwe R, Kiyoto S, Crônier D, Morreel K, Arribat S, Neutelings G, Chabi M, Boerjan W, et al (2014b) 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 **26**: 4462–82

Charlet K, Eve S, Jernot JP, Gomina M, Breard J (2009) Tensile deformation of a flax fiber. Procedia Eng **1**: 233–236

Charlet K, Jernot JP, Eve S, Gomina M, Bréard J (2010) Multi-scale morphological characterisation of flax: From the stem to the fibrils. Carbohydr Polym **82**: 54–61

Chemikosova SB, Pavlencheva N V., Gur'yanov OP, Gorshkova TA (2006) The effect of soil drought on the phloem fiber development in long-fiber flax. Russ J Plant Physiol **53**: 656–662

Cocuron J-C, Lerouxel O, Drakakaki G, Alonso AP, Liepman AH, Keegstra K, Raikhel N, Wilkerson CG (2007) A gene from the cellulose synthase-like C family encodes a -1,4 glucan synthase. Proc Natl Acad Sci **104**: 8550–8555

Cosgrove DJ (1999) Enzymes and other agents that enhance cell wall extensibility. Annu Rev Plant Physiol Plant Mol Biol **50**: 391–417

Cosgrove DJ (2014) ScienceDirect Re-constructing our models of cellulose and primary cell wall assembly. Curr Opin Plant Biol **22**: 122–131

Cosgrove DJ (2005) Growth of the plant cell wall. Nat Rev Mol Cell Biol **6**: 850–61

Criscuolo A, Gribaldo S (2010) BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol Biol **10**: 210

Cui S, Hu J, Guo S, Wang J, Cheng Y, Dang X, Wu L, He Y (2012) Proteome analysis of Physcomitrella patens exposed to progressive dehydration and rehydration. J Exp Bot **63**: 711–26

Dalisay DS, Kim KW, Lee C, Yang H, Rübel O, Bowen BP, Davin LB, Lewis NG (2015) Dirigent Protein-Mediated Lignan and Cyanogenic Glucoside Formation in Flax Seed: Integrated Omics and MALDI Mass Spectrometry Imaging. J Nat Prod **78**: 1231–42

Dash PK, Cao Y, Jailani AK, Gupta P, Venglat P, Xiang D, Rai R, Sharma R (2014) Genome-wide analysis of drought induced gene expression changes in flax (Linum usitatissimum). **5**: 37–41

Davis EL, Hussey RS, Baum TJ, Bakker J, Schots A, Rosso M-N, Abad P (2000) N EMATODE P ARASITISM G ENES. Annu Rev Phytopathol **38**: 365–396

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

Day A, Fénart S, Neutelings G, Hawkins S, Rolando C, Tokarski C (2013) Identification of cell wall proteins in the flax (Linum usitatissimum) stem. Proteomics **13**: 812–825

Day A, Ruel K, Neutelings G, Crônier D, David H, Hawkins S, Chabbert B (2005b) Lignification in the flax stem: evidence for an unusual lignin in bast fibers. Planta **222**: 234–45

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

Delmer DP (1999) CELLULOSE BIOSYNTHESIS: Exciting Times for A Difficult Field of Study. Annu Rev Plant Physiol Plant Mol Biol **50**: 245–276

DeMartini JD, Pattathil S, Miller JS, Li H, Hahn MG, Wyman CE (2013) Investigating plant cell wall components that affect biomass recalcitrance in poplar and switchgrass. Energy Environ Sci **6**: 898

Desprez T., Juraniec M. b, Crowell EF. F, Jouy H., Pochylova Z., Parcy F., Höfte H., Gonneau M., Vernhettes S. (2007) Organization of cellulose synthase complexes involved in primary cell wall synthesis in Arabidopsis thaliana. Proc Natl Acad Sci U S A **104**: 15572–15577

Diederichsen A, Ulrich A (2009) Variability in stem fibre content and its association with other characteristics in 1177 flax (Linum usitatissimum L.) genebank accessions. Ind Crops Prod **30**: 33–39

Van Dijk JP, de Mello CS, Voorhuijzen MM, Hutten RCB, Arisi ACM, Jansen JJ, Buydens LMC, van der Voet H, Kok EJ (2014) Safety assessment of plant varieties using transcriptomics profiling and a one-class classifier. Regul Toxicol Pharmacol **70**: 297–303

Dillman AC, Brinsmade JCJ (1938) Effect of spacing on the development of the flax plant. J. Am. Soc. Agron.

Ding Y, Chang J, Ma Q, Chen L, Liu S, Jin S, Han J, Xu R, Zhu A, Guo J, et al (2015) Network Analysis of Postharvest Senescence Process in Citrus Fruits Revealed by Transcriptomic and Metabolomic Profiling. Plant Physiol **168**: pp.114.255711

Doblin MS, Pettolino F, Bacic A (2010) Evans Review: Plant cell walls: The skeleton of the plant world. Funct Plant Biol **37**: 357–381

Domon J-M, Baldwin L, Acket S, Caudeville E, Arnoult S, Zub H, Gillet F, Lejeune-Hénaut I, Brancourt-Hulmel M, Pelloux J, et al (2013) Cell wall compositional modifications of Miscanthus ecotypes in response to cold acclimation. Phytochemistry **85**: 51–61

Donaldson L, Hague J, Snell R (2001) Lignin Distribution in Coppice Poplar, Linseed and Wheat Straw. Holzforschung **55**: 379–385

Dong J, Jiang Y, Chen R, Xu Z, Gao X (2011) Isolation of a novel xyloglucan endotransglucosylase (OsXET9) gene from rice and analysis of the response of this gene to abiotic stresses. African J Biotechnol **10**: 17424–17434

Doré C, Varoquaux F (2006) Histoire et amélioration de cinquante plantes cultivées.

Dumont M, Lehner A, Bouton S, Kiefer-Meyer MC, Voxeur A, Pelloux J, Lerouge P, Mollet J-C (2014) The cell wall pectic polymer rhamnogalacturonan-II is required for proper pollen tube elongation: implications of a putative sialyltransferase-like protein. Ann Bot 1177–1188

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32**: 1792–7

Egelund J, Petersen BL, Motawia MS, Damager I, Faik A, Olsen CE, Ishii T, Clausen H, Ulvskov P, Geshi N (2006) Arabidopsis thaliana RGXT1 and RGXT2 encode Golgi-localized (1,3)-alpha-D-xylosyltransferases involved in the synthesis of pectic rhamnogalacturonan-II. Plant Cell **18**: 2593–607

Eklöf JM, Brumer H (2010) The XTH gene family: an update on enzyme structure, function, and phylogeny in xyloglucan remodeling. Plant Physiol **153**: 456–66

Eklöf JM, Shojania S, Okon M, McIntosh LP, Brumer H (2013) Structure-function analysis of a broad specificity Populus trichocarpa endo-β-glucanase reveals an evolutionary link between bacterial licheninases and plant XTH gene products. J Biol Chem **288**: 15786–15799

Esau K (1943) Vascular Differentiation in the Vegetative Shoot of Linum. III. The Origin of the Bast Fibers on JSTOR. Am J Bot **30**: 579–586

Estévez JM, Kieliszewski MJ, Khitrov N, Somerville C (2006) Characterization of synthetic hydroxyproline-rich proteoglycans with arabinogalactan protein and extensin motifs in Arabidopsis. Plant Physiol **142**: 458–70

Fan L, Linker R, Gepstein S, Tanimoto E, Yamamoto R, Neumann PM (2006) Progressive inhibition by water deficit of cell wall extensibility and growth along the elongation zone of maize roots is related to increased lignin metabolism and progressive stelar accumulation of wall phenolics. Plant Physiol **140**: 603–612

Fang L, Tian R, Li X, Chen J, Wang S, Wang P, Zhang T (2014) Cotton fiber elongation network revealed by expression profiling of longer fiber lines introgressed with different Gossypium barbadense chromosome segments. BMC Genomics **15**: 838

Farooq M, Wahid A, Kobayashi N, Fujita D, Basra S (2009) Plant drought stress: effects, mechanisms and management. Agron Sustain Dev **29**: 185–212

Feiz L, Irshad M, Pont-lezica RF, Canut H, Jamet E (2006) Evaluation of cell wall preparations for proteomics : a new procedure for purifying cell walls from Arabidopsis hypocotyls. **13**: 1–13

Fenart S, Chabi M, Gallina S, Huis R, Neutelings G, Riviere N, Thomasset B, Hawkins S, Lucau-Danila A (2013) Intra-platform comparison of 25-mer and 60-mer oligonucleotide Nimblegen DNA microarrays. BMC Res Notes **6**: 43

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

Festucci-Buselli R a., Otoni WC, Joshi CP (2007) Structure, organization, and functions of cellulose synthase complexes in higher plants. Brazilian J Plant Physiol **19**: 1–13

Foster R, POONI HS, MACKAY IJ (1997) Quantitative evaluation of Linum usitatissimum varieties for dual-purpose traits. J Agric Sci **129**: 179–185

Fraige K, González-Fernández R, Carrilho E, Jorrín-Novo J V (2015) Metabolite and proteome changes during the ripening of Syrah and Cabernet Sauvignon grape varieties cultured in a nontraditional wine region in Brazil. J Proteomics **113**: 206–25

Franková L, Fry SC (2013) Biochemistry and physiological roles of enzymes that "cut and paste" plant cell-wall polysaccharides. J Exp Bot **64**: 3519–3550

Fry SC (1988) The growing plant cell wall: chemical and metabolic analysis. New York **203**: 333

Fry SC (2004) Primary cell wall metabolism: tracking the careers of wall polymers in living plant cells. New Phytol **161**: 641–675

Fu J, Huang B (2001) Involvement of antioxidants and lipid peroxidation in the adaptation of two cool-season grasses to localized drought stress. Environ Exp Bot **45**: 105–114

Gierlinger N, Schwanninger M (2006) Chemical Imaging of Poplar Wood Cell Walls by Confocal Raman Microscopy. **140**: 1246–1254

Gille S, Pauly M (2012) O-acetylation of plant cell wall polysaccharides. Front Plant Sci **3**: 12

Girault R, Bert F, Rihouey C, Jauneau A, Morvan C, Jarvis M (1997) Galactans and cellulose in flax fibres: putative contributions to the tensile strength. Int J Biol Macromol **21**: 179–188

Gorshkova T a., Wyatt SE, Salnikov V V., Gibeaut DM, Ibragimov MR, Lozovaya V V., Carpita NC (1996) Cell-Wall Polysaccharides of Developing Flax Plants. Plant Physiol **110**: 721–729

Gorshkova T, Mokshina N, Chernova T, Ibragimova N, Salnikov V, Mikshina P, Tryfona T, Banasiak A, Immerzeel P, Dupree P, et al (2015) Aspen tension wood fibers contain β-(1→4)-galactans and acidic arabinogalactans retained by cellulose microfibrils in gelatinous walls. Plant Physiol. doi: 10.1104/pp.15.00690

Gorshkova T, Morvan C (2006) Secondary cell-wall assembly in flax phloem fibres: Role of galactans. Planta **223**: 149–158

Gorshkova TA, Sal'nikov V V., Chemikosova SB, Ageeva M V., Pavlencheva N V., van Dam JEG (2003) The snap point: a transition point in Linum usitatissimum bast fiber development. Ind Crops Prod **18**: 213–221

Gothandam KM, Nalini E, Karthikeyan S, Shin JS (2009) OsPRP3, a flower specific proline-rich protein of rice, determines extracellular matrix structure of floral organs and its overexpression confers cold-tolerance. Plant Mol Biol **72**: 125–135

Gou J-Y, Miller LM, Hou G, Yu X-H, Chen X-Y, Liu C-J (2012) Acetylesterase-mediated deacetylation of pectin impairs cell elongation, pollen germination, and plant reproduction. Plant Cell **24**: 50–65

Goubet F, Bourlard T, Girault R, Alexandre C, Vandevelde M-C, Morvan C (1995) Structural features of galactans from flax fibres. Carbohydr Polym **27**: 221–227

Grabber JH (2005) How Do Lignin Composition, Structure, and Cross-Linking Affect Degradability? A Review of Cell Wall Model Studies. Crop Sci **45**: 820

Gunl M, Neumetzler L, Kraemer F, de Souza a., Schultink a., Pena M, York WS, Pauly M (2011) AXY8 Encodes an -Fucosidase, Underscoring the Importance of Apoplastic Metabolism on the Fine Structure of Arabidopsis Cell Wall Polysaccharides. Plant Cell **23**: 4025–4040

Guo J, Shi G, Guo X, Zhang L, Xu W, Wang Y, Su Z, Hua J (2015) Transcriptome analysis reveals that distinct metabolic pathways operate in salt-tolerant and salt-sensitive upland cotton varieties subjected to salinity stress. Plant Sci **238**: 33–45

Guo R, Hao W, Gong D (2012) Effects of Water Stress on Germination and Growth of Linseed Seedlings (Linum usitatissimum L), Photosynthetic Efficiency and Accumulation of Metabolites. J Agric Sci **4**: p253

Gurjanov OP. P, Ibragimova NN. N, Gnezdilov OI. I, Gorshkova TA. A (2008) Polysaccharides, tightly bound to cellulose in cell wall of flax bast fibre: Isolation and identification. Carbohydr Polym **72**: 719–729

Hall Q, Cannon MC (2002) The cell wall hydroxyproline-rich glycoprotein RSH is essential for normal embryo development in Arabidopsis. Plant Cell **14**: 1161–72

Han Y yang, Li A xiu, Li F, Zhao M rong, Wang W (2012) Characterization of a wheat (Triticum aestivum L.) expansin gene, TaEXPB23, involved in the abiotic stress response and phytohormone regulation. Plant Physiol Biochem **54**: 49–58

Harholt J, Suttangkakul A, Vibe Scheller H (2010) Biosynthesis of pectin. Plant Physiol **153**: 384–95

Harper JF, Manney L, DeWitt ND, Yoo MH, Sussman MR (1990) The Arabidopsis thaliana plasma membrane H(+)-ATPase multigene family. Genomic sequence and expression of a third isoform. J Biol Chem **265**: 13601–8

Van Hazendonk JM, Reinerik EJM, de Waard P, van Dam JEG (1996) Structural analysis of acetylated hemicellulose polysaccharides from fibre flax (Linum usitatissimum L.). Carbohydr Res **291**: 141–154

Van Hengel AJ, Roberts K (2003) AtAGP30, an arabinogalactan-protein in the cell walls of the primary root, plays a role in root regeneration and seed germination. Plant J **36**: 256–70

Hijazi M, Velasquez SM, Jamet E, Estevez JM, Albenne C (2014) An update on post-translational modifications of hydroxyproline-rich glycoproteins : toward a model highlighting their contribution to plant cell wall architecture. **5**: 1–10

Höfer R, Boachon B, Renault H, Gavira C, Miesch L, Iglesias J, Ginglinger J-F, Allouche L, Miesch M, Grec S, et al (2014) Dual function of the cytochrome P450 CYP76 family from Arabidopsis thaliana in the metabolism of monoterpenols and phenylurea herbicides. Plant Physiol **166**: 1149–61

Holland N (2000) A Comparative Analysis of the Plant Cellulose Synthase (CesA) Gene Family. PLANT Physiol **123**: 1313–1324

Hotte NSC, Deyholos MK (2008) A flax fibre proteome: identification of proteins enriched in bast fibres. BMC Plant Biol **8**: 52

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

Ibragimova NN, Mokshina NE, Gorshkova TA (2012) [Cell wall proteins of plant fibers]. Bioorg Khim **38**: 139–48

Immerzeel P, Eppink MM, De Vries SC, Schols H a., Voragen AGJ (2006) Carrot arabinogalactan proteins are interlinked with pectins. Physiol Plant **128**: 18–28

Ingram J, Bartels D (1996) THE MOLECULAR BASIS OF DEHYDRATION TOLERANCE IN PLANTS. Annu Rev Plant Physiol Plant Mol Biol **47**: 377–403

Ishiguro S, Watanabe Y, Ito N, Nonaka H, Takeda N, Sakai T, Kanaya H, Okada K (2002) SHEPHERD is the Arabidopsis GRP94 responsible for the formation of functional CLAVATA proteins. EMBO J **21**: 898–908

Iwai H, Masaoka N, Ishii T, Satoh S (2002) A pectin glucuronyltransferase gene is essential for intercellular attachment in the plant meristem. Proc Natl Acad Sci U S A **99**: 16319–24

Jamet E, Boudart G, Borderies G, Charmont S, Lafitte C, Rossignol M, Canut H, Pont-Lezica R (2008) Isolation of plant cell wall proteins. Methods Mol Biol **425**: 187–201

Jamet E, Canut H, Boudart G, Pont-Lezica RF (2006) Cell wall proteins: a new insight through proteomics. Trends Plant Sci **11**: 33–9

Jauh G, Lord E (1996) Localization of pectins and arabinogalactan-proteins in lily (Lilium longiflorum L.) pollen tube and style, and their possible roles in pollination. Planta. doi: 10.1007/BF00196566

Jiang L, Yan S, Yang W, Li Y, Xia M, Chen Z, Wang Q, Yan L, Song X, Liu R, et al (2015) Transcriptomic analysis reveals the roles of microtubule-related genes and transcription factors in fruit length regulation in cucumber (Cucumis sativus L.). Sci Rep **5**: 8031

Joshi S. V, Drzal L. T, Mohanty a. . K, Arora S (2004) Are natural fiber composites environmentally superior to glass fiber reinforced composites? Compos Part A Appl Sci Manuf **35**: 371–376

Kaewthai N, Gendre D, Eklöf JM, Ibatullin FM, Ezcurra I, Bhalerao RP, Brumer H (2013) Group III-A XTH genes of Arabidopsis encode predominant xyloglucan endohydrolases that are dispensable for normal growth. Plant Physiol **161**: 440–54

Kantar M, Lucas SJ, Budak H (2011) Drought Stress. Molecular Genetics and Genomics Approaches. Adv Bot Res. doi: 10.1016/B978-0-12-387692-8.00013-8

Kaya MD, Okçu G, Atak M, Çıkılı Y, Kolsarıcı Ö (2006) Seed treatments to overcome salt and drought stress during germination in sunflower (Helianthus annuus L.). Eur J Agron **24**: 291–295

Keckes J, Burgert I, Frühmann K, Müller M, Kölln K, Hamilton M, Burghammer M, Roth S V, Stanzl-Tschegg S, Fratzl P (2003) Cell-wall recovery after irreversible deformation of wood. Nat Mater **2**: 810–4

Keegstra K, Raikhel N (2001) Plant glycosyltransferases. Curr Opin Plant Biol **4**: 219–24

Keegstra K, Talmadge KW, Bauer WD, Albersheim P (1973) The Structure of Plant Cell Walls: III. A Model of the Walls of Suspension-cultured Sycamore Cells Based on the Interconnections of the Macromolecular Components. Plant Physiol **51**: 188–97

Keller B, Schmid J, Lamb CJ (1989) Vascular expression of a bean cell wall glycine-rich protein-beta-glucuronidase gene fusion in transgenic tobacco. EMBO J **8**: 1309–14

Kenrick P, Crane PR (1997) The origin and early evolution of plants on land. **389**: 33–39

Kieliszewski MJ (2001) The latest hype on Hyp-O-glycosylation codes. Phytochemistry **57**: 319–23

Kimura S (1999) Immunogold Labeling of Rosette Terminal Cellulose-Synthesizing Complexes in the Vascular Plant Vigna angularis. PLANT CELL ONLINE **11**: 2075–2086

Kitazawa K, Tryfona T, Yoshimi Y, Hayashi Y, Kawauchi S, Antonov L, Tanaka H, Takahashi T, Kaneko S, Dupree P, et al (2013) β-galactosyl Yariv reagent binds to the β-1,3-galactan of arabinogalactan proteins. Plant Physiol **161**: 1117–26

Kitts DD, Yuan Y V, Wijewickreme AN, Thompson LU (1999) Antioxidant activity of the flaxseed lignan secoisolariciresinol diglycoside and its mammalian lignan metabolites enterodiol and enterolactone. Mol Cell Biochem **202**: 91–100

Konno H, Yamasaki Y, Sugimoto M, Takeda K (2008) Differential changes in cell wall matrix polysaccharides and glycoside-hydrolyzing enzymes in developing wheat seedlings differing in drought tolerance. J Plant Physiol **165**: 745–54

Kumar S, You FM, Duguid S, Booker H, Rowland G, Cloutier S (2015) QTL for fatty acid composition and yield in linseed (Linum usitatissimum L.). Theor Appl Genet **128**: 965–984

Kymäläinen H-R, Sjöberg A-M (2008) Flax and hemp fibres as raw materials for thermal insulations. Build Environ **43**: 1261–1269

Lafarguette F, Leplé J-C, Déjardin A, Laurans F, Costa G, Lesage-Descauses M-C, Pilate G (2004) Poplar genes encoding fasciclin-like arabinogalactan proteins are highly expressed in tension wood. New Phytol **164**: 107–121

Lamport DTA, Tan L, Kieliszewski MJ (2011) Structural proteins of the primary cell wall: extraction, purification, and analysis. Methods Mol Biol **715**: 209–19

Lao NT, Long D, Kiang S, Coupland G, Shoue DA, Carpita NC, Kavanagh TA (2003) Mutation of a family 8 glycosyltransferase gene alters cell wall carbohydrate composition and causes a humidity-sensitive semi-sterile dwarf phenotype in Arabidopsis. Plant Mol Biol **53**: 647–661

Le SQ, Gascuel O (2008) An improved general amino acid replacement matrix. Mol Biol Evol **25**: 1307–20

Lee B-R, Kim K-Y, Jung W-J, Avice J-C, Ourry A, Kim T-H (2007) Peroxidases and lignification in relation to the intensity of water-deficit stress in white clover (Trifolium repens L.). J Exp Bot **58**: 1271–9

Lee C, Teng Q, Huang W, Zhong R, Ye Z-H (2010) The Arabidopsis family GT43 glycosyltransferases form two functionally nonredundant groups essential for the elongation of glucuronoxylan backbone. Plant Physiol **153**: 526–41

Lee C, Teng Q, Zhong R, Ye Z-H (2011) The four Arabidopsis reduced wall acetylation genes are expressed in secondary wall-containing cells and required for the acetylation of xylan. Plant Cell Physiol **52**: 1289–301

Lefeuvre A, Bourmaud A, Lebrun L, Morvan C, Baley C (2013) A study of the yearly reproducibility of flax fiber tensile properties. Ind Crops Prod **50**: 400–407

Leonard R, Petersen BO, Himly M, Kaar W, Wopfner N, Kolarich D, van Ree R, Ebner C, Duus JO, Ferreira F, et al (2005) Two Novel Types of O-Glycans on the Mugwort Pollen Allergen Art v 1 and Their Role in Antibody Binding. J Biol Chem **280**: 7932–7940

Léonard R, Wopfner N, Pabst M, Stadlmann J, Petersen BO, Duus JØ, Himly M, Radauer C, Gadermaier G, Razzazi-Fazeli E, et al (2010) A new allergen from ragweed (Ambrosia artemisiifolia) with homology to art v 1 from mugwort. J Biol Chem **285**: 27192–200

Lerouxel O, Cavalier DM, Liepman AH, Keegstra K (2006) Biosynthesis of plant cell wall polysaccharides - a complex process. Curr Opin Plant Biol **9**: 621–30

Leucci MR, Lenucci MS, Piro G, Dalessandro G (2008) Water stress and cell wall polysaccharides in the apical root zone of wheat cultivars varying in drought tolerance. J Plant Physiol **165**: 1168–1180

Lewis NG, Yamamoto E (1990) Lignin: occurrence, biogenesis and biodegradation. Annu Rev Plant Physiol Plant Mol Biol **41**: 455–96

Lindlöf A, Chawade A, Sikora P, Olsson O (2015) Comparative Transcriptomics of Sijung and Jumli Marshi Rice during Early Chilling Stress Imply Multiple Protective Mechanisms. PLoS One **10**: e0125385

Liu C, Cai Y, Zhang X, Gou M, Yang H (2014) Tailoring lignin biosynthesis for efficient and sustainable biofuel production. Plant Biotechnol J **12**: 1154–62

Liu X-L, Liu L, Niu Q-K, Xia C, Yang K-Z, Li R, Chen L-Q, Zhang X-Q, Zhou Y, Ye D (2011) MALE GAMETOPHYTE DEFECTIVE 4 encodes a rhamnogalacturonan II xylosyltransferase and is important for growth of pollen tubes and roots in Arabidopsis. Plant J **65**: 647–660

Liwanag AJM, Ebert B, Verhertbruggen Y, Rennie EA, Rautengarten C, Oikawa A, Andersen MCF, Clausen MH, Scheller H V. (2012) Pectin Biosynthesis: GALS1 in Arabidopsis thaliana Is a -1,4-Galactan -1,4-Galactosyltransferase. Plant Cell **24**: 5024–5036

Lohse M, Nagel A, Herter T, May P, Schroda M, Zrenner R, Tohge T, Fernie AR, Stitt M, Usadel B (2014) Mercator: a fast and simple web server for genome scale functional annotation of plant sequence data. Plant Cell Environ **37**: 1250–8

Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. **42**: 490–495

Lowry B, Lee D, Hébant C (1980) The Origin of Land Plants: A New Look at an Old Problem on JSTOR. Int Assoc Plant Taxon **29**: 183–197

Lygin A V., UPTON J, DOHLEMAN FG, JUVIK J, ZABOTINA OA, WIDHOLM JM, LOZOVAYA V V. (2011) Composition of cell wall phenolics and polysaccharides of the potential bioenergy crop -Miscanthus. GCB Bioenergy **3**: 333–345

M. Sfiligoj Smole SHKSK and TK (2013) Advances in Agrophysical Research. doi: 10.5772/3341

MacMillan CP, Mansfield SD, Stachurski ZH, Evans R, Southerton SG (2010) Fasciclin-like arabinogalactan proteins: specialization for stem biomechanics and cell wall architecture in Arabidopsis and Eucalyptus. Plant J **62**: 689–703

Manabe Y, Nafisi M, Verhertbruggen Y, Orfila C, Gille S, Rautengarten C, Cherk C, Marcus SE, Somerville S, Pauly M, et al (2011) Loss-of-Function Mutation of REDUCED WALL ACETYLATION2 in Arabidopsis Leads to Reduced Cell Wall Acetylation and Increased Resistance to Botrytis cinerea. PLANT Physiol **155**: 1068–1078

Mangeon A, Junqueira RM, Sachetto-Martins G (2010) Functional diversity of the plant glycine-rich proteins superfamily. Plant Signal Behav **5**: 99–104

Manickavelu A, Nadarajan N, Ganesh SK, Gnanamalar RP, Chandra Babu R (2006) Drought tolerance in rice: morphological and molecular genetic consideration. Plant Growth Regul **50**: 121–138

Matsunaga T, Ishii T, Matsumoto S, Higuchi M, Darvill A, Albersheim P, O'Neill MA (2004) Occurrence of the primary cell wall polysaccharide rhamnogalacturonan II in pteridophytes, lycophytes, and bryophytes. Implications for the evolution of vascular plants. Plant Physiol **134**: 339–51

McCourt RM, Delwiche CF, Karol KG (2004) Charophyte algae and land plant origins. Trends Ecol Evol **19**: 661–6

McDill J, Repplinger M, Simpson BB, Kadereit JW (2009) The Phylogeny of <I>Linum</I> and Linaceae Subfamily Linoideae, with Implications for Their Systematics, Biogeography, and Evolution of Heterostyly. Syst Bot **34**: 386–405

McDougall GJ (1993) Isolation and partial characterisation of the non-cellulosic polysaccharides of flax fibre. Carbohydr Res **241**: 227–236

Mcfarlane HE, Anett D (2014) The Cell Biology of Cellulose Synthesis. doi: 10.1146/annurev-arplant-050213-040240

Mellerowicz E, Sundberg B (2008) Wood cell walls: biosynthesis, developmental dynamics and their implications for wood properties. Curr Opin Plant Biol **11**: 293–300

Mellerowicz EJ, Gorshkova TA (2012) Tensional stress generation in gelatinous fibres: a review and possible mechanism based on cell-wall structure and composition. J Exp Bot **63**: 551–65

Michel SAAX, Vogels RRM, Bouvy ND, Knetsch MLW, van den Akker NMS, Gijbels MJJ, van der Marel C, Vermeersch J, Molin DGM, Koole LH (2014) Utilization of flax fibers for biomedical applications. J Biomed Mater Res B Appl Biomater **102**: 477–87

Miedes E, Vanholme R, Boerjan W, Molina A (2014) The role of the secondary cell wall in plant resistance to pathogens. Front Plant Sci **5**: 1–13

Minic Z (2008) Physiological roles of plant glycoside hydrolases. Planta **227**: 723–40

Minic Z, Jamet E, Negroni L, Arsene der Garabedian P, Zivy M, Jouanin L (2007) A sub-proteome of Arabidopsis thaliana mature stems trapped on Concanavalin A is enriched in cell wall glycoside hydrolases. J Exp Bot **58**: 2503–2512

Mohnen D (2008) Pectin structure and biosynthesis. Curr Opin Plant Biol **11**: 266–277

Mokshina N, Gorshkova T, Deyholos MK (2014) Chitinase-like (CTL) and cellulose synthase (CESA) gene expression in gelatinous-type cellulosic walls of flax (Linum usitatissimum L.) bast fibers. PLoS One **9**: e97949

Moore JP, Farrant JM, Driouich A (2008) A role for pectin-associated arabinans in maintaining the flexibility of the plant cell wall during water deficit stress. Plant Signal Behav **3**: 102–104

Moore JP, Nguema-Ona E, Chevalier L, Lindsey GG, Brandt WF, Lerouge P, Farrant JM, Driouich A (2006) Response of the leaf cell wall to desiccation in the resurrection plant Myrothamnus flabellifolius. Plant Physiol **141**: 651–662

Morvan C, Andème-Onzighi C, Girault R, Himmelsbach DS, Driouich A, Akin DE (2003a) Building flax fibres: more than one brick in the walls. Plant Physiol Biochem **41**: 935–944

Morvan C, Andème-Onzighi C, Girault R, Himmelsbach DS, Driouich A, Akin DE (2003b) Building flax fibres: More than one brick in the walls. Plant Physiol Biochem **41**: 935–944

Motose H, Sugiyama M, Fukuda H (2004) A proteoglycan mediates inductive interaction during plant vascular development. Nature **429**: 873–878

Mousavi A, Hotta Y (2005) Glycine-rich proteins: a class of novel proteins. Appl Biochem Biotechnol **120**: 169–74

Munns R, James RA, Sirault XRR, Furbank RT, Jones HG (2010) New phenotyping methods for screening wheat and barley for beneficial responses to water deficit. J Exp Bot **61**: 3499–3507

Murashige T, Skoog F (1962) A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. Physiol Plant **15**: 473–497

Muravenko O V, Amosova A V, Samatadze TE, Semenova OI, Nosova I V, Popov K V, Shostak NG, Zoshchuk SA, Zelenin A V (2004) [Chromosomal localization of 5S and 45S ribosomal DNA in species of Linum L. section Linum (syn=Protolinum and Adenolinum)]. Genetika **40**: 256–60

Muravenko O V, Lemesh VA, Samatadze TE, Amosova A V, Grushetskaia ZE, Popov K V, Semenova OI, Khotyleva L V, Zelenin A V (2003) [Genome comparisons of three closely related flax species and their hybrids with chromosomal and molecular markers]. Genetika **39**: 510–8

Muravenko V, Samatadze TE, Popov K V, Amosova A V, Zelenin A V (2001) Polymorphism of heterochromatic regions of flax chromosomes. Membr Cell Biol **14**: 743–8

Nakashima J, Chen F, Jackson L, Shadle G, Dixon RA (2008) Multi-site genetic modification of monolignol biosynthesis in alfalfa (Medicago sativa): effects on lignin composition in specific cell types. New Phytol **179**: 738–750

Neumann PM (2011) Plant Responses to Drought and Salinity Stress - Developments in a Post-Genomic Era. Adv Bot Res. doi: 10.1016/B978-0-12-387692-8.00002-3

Le Ngoc Huyen T, Rémond C, Dheilly RM, Chabbert B (2010) Effect of harvesting date on the composition and saccharification of Miscanthus x giganteus. Bioresour Technol **101**: 8224–31

Nguema-Ona E, Vicré-Gibouin M, Cannesan M-A, Driouich A (2013) Arabinogalactan proteins in root–microbe interactions. Trends Plant Sci **18**: 440–449

Nguema-Ona E, Vicré-Gibouin M, Gotté M, Plancot B, Lerouge P, Bardor M, Driouich A (2014) Cell wall O-glycoproteins and N-glycoproteins: aspects of biosynthesis and function. Front Plant Sci **5**: 499

Nicotra AB, Atkin OK, Bonser SP, Davidson AM, Finnegan EJ, Mathesius U, Poot P, Purugganan MD, Richards CL, Valladares F, et al (2010) Plant phenotypic plasticity in a changing climate. Trends Plant Sci **15**: 684–92

Niklas KJ, Kutschera U (2010) The evolution of the land plant life cycle. New Phytol **185**: 27–41

Nishikubo N, Awano T, Banasiak A, Bourquin V, Ibatullin F, Funada R, Brumer H, Teeri TT, Hayashi T, Sundberg B, et al (2007) Xyloglucan endo-transglycosylase (XET) functions in gelatinous layers of tension wood fibers in poplar--a glimpse into the mechanism of the balancing act of trees. Plant Cell Physiol **48**: 843–55

Nishikubo N, Takahashi J, Roos A a, Derba-Maceluch M, Piens K, Brumer H, Teeri TT, Stålbrand H, Mellerowicz EJ (2011) Xyloglucan endo-transglycosylase-mediated xyloglucan rearrangements in developing wood of hybrid aspen. Plant Physiol **155**: 399–413

O'Neill MA, Ishii T, Albersheim, P. A, Darvill AG (2004) Rhamnogalacturonan II: structure and function of a borate cross-linked cell wall pectic polysaccharide. Annu Rev Plant Biol **55**: 109–39

Obel N, Erben V, Schwarz T, Kühnel S, Fodor A, Pauly M (2009) Microanalysis of plant cell wall polysaccharides. Mol Plant **2**: 922–32

Obembe OO, Jacobsen E, Timmers J, Gilbert H, Blake AW, Knox JP, Visser RGF, Vincken J-P (2007) Promiscuous, non-catalytic, tandem carbohydrate-binding modules modulate the cell-wall structure and development of transgenic tobacco (Nicotiana tabacum) plants. J Plant Res **120**: 605–617

Ogawa-Ohnishi M, Matsushita W, Matsubayashi Y (2013) Identification of three hydroxyproline O-arabinosyltransferases in Arabidopsis thaliana. Nat Chem Biol **9**: 726–30

Oliveira TM, da Silva FR, Bonatto D, Neves DM, Morillon R, Maserti BE, Filho MAC, Costa MGC, Pirovani CP, Gesteira AS (2015) Comparative study of the protein profiles of Sunki mandarin and Rangpur lime plants in response to water deficit. BMC Plant Biol **15**: 69

Paës G (2014) Fluorescent Probes for Exploring Plant Cell Wall Deconstruction: A Review. 9380–9402

Park YB, Cosgrove DJ (2012) A revised architecture of primary cell walls based on biomechanical changes induced by substrate-specific endoglucanases. Plant Physiol **158**: 1933–43

Park YB, Cosgrove DJ (2015) Xyloglucan and its interactions with other components of the growing cell wall. Plant Cell Physiol **56**: 180–94

Passardi F, Penel C, Dunand C (2004) Performing the paradoxical: how plant peroxidases modify the cell wall. Trends Plant Sci **9**: 534–40

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

Pauly M, Keegstra K (2008) Cell-wall carbohydrates and their modification as a resource for biofuels. Plant J **54**: 559–68

Penna S (2003) Building stress tolerance through over-producing trehalose in transgenic plants. Trends Plant Sci **8**: 355–357

Pérez S, Rodríguez-Carvajal MA, Doco T (2003) A complex plant cell wall polysaccharide: rhamnogalacturonan II. A structure in quest of a function. Biochimie **85**: 109–21

Perrin RM, DeRocher AE, Bar-Peled M, Zeng W, Norambuena L, Orellana A, Raikhel N V, Keegstra K (1999) Xyloglucan fucosyltransferase, an enzyme involved in plant cell wall biosynthesis. Science **284**: 1976–9

Petersen BL, Egelund J, Damager I, Faber K, Krüger Jensen J, Yang Z, Bennett EP, Scheller HV, Ulvskov P (2009) Assay and heterologous expression in Pichia pastoris of plant cell wall type-II membrane anchored glycosyltransferases. Glycoconj J **26**: 1235–1246

Pilling E, Höfte H (2003) Feedback from the wall. Curr Opin Plant Biol **6**: 611–6

Pinzon-Latorre D, Deyholos MK (2014) Pectinmethylesterases (PME) and Pectinmethylesterase Inhibitors (PMEI) Enriched during Phloem Fiber Development in Flax (Linum usitatissimum). PLoS One **9**: e105386

Piro G, Leucci MR, Waldron K, Dalessandro G (2003) Exposure to water stress causes changes in the biosynthesis of cell wall polysaccharides in roots of wheat cultivars varying in drought tolerance. Plant Sci **165**: 559–569

Popper Z (2008) Evolution and diversity of green plant cell walls. Curr Opin Plant Biol **11**: 286–292

Popper ZA, Michel G, Hervé C, Domozych DS, Willats WGT, Tuohy MG, Kloareg B, Stengel DB (2011) Evolution and diversity of plant cell walls: from algae to flowering plants. Annu Rev Plant Biol **62**: 567–590

Prasad K (1997) Hydroxyl radical-scavenging property of secoisolariciresinol diglucoside (SDG) isolated from flax-seed. Mol Cell Biochem **168**: 117–23

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics **155**: 945–959

Qi X, Behrens BX, West PR, Mort AJ (1995) Solubilization and partial characterization of extensin fragments from cell walls of cotton suspension cultures. Evidence for a covalent cross-link between extensin and pectin. Plant Physiol **108**: 1691–701

Qiu D, Wilson IW, Gan S, Washusen R, Moran GF, Southerton SG (2008) Gene expression in Eucalyptus branch wood with marked variation in cellulose microfibril orientation and lacking G-layers. New Phytol **179**: 94–103

Ragauskas AJ, Ragauskas AJ, Beckham GT, Biddy MJ, Chandra R, Chen F, Naskar AK, Saddler JN, Tschaplinski TJ, Tuskan GA, et al (2014) Lignin Valorization : Improving Lignin Processing in the Biorefi nery. doi: 10.1126/science.1246843

Rakszegi M, Lovegrove A, Balla K, Láng L, Bedő Z, Veisz O, Shewry PR (2014) Effect of heat and drought stress on the structure and composition of arabinoxylan and β-glucan in wheat grain. Carbohydr Polym **102**: 557–65

Ralph J, Lundquist K, Brunow PG, Lu F, Kim H, Schatz P, Marita J, Hatfield R, Ralph S, Christensen J, et al (2004) Lignins: Natural polymer from oxidative coupling of 4-hydroxyphenylpropanoids. Phytochem Rev **3**: 29–60

Ramachandra Reddy A, Chaitanya KV, Vivekanandan M (2004) Drought-induced responses of photosynthesis and antioxidant metabolism in higher plants. J Plant Physiol **161**: 1189–202

Ray S, Paynel F, Morvan C, Lerouge P, Driouich A, Ray B (2013) Characterization of mucilage polysaccharides, arabinogalactanproteins and cell-wall hemicellulosic polysaccharides isolated from flax seed meal: A wealth of structural moieties. Carbohydr Polym **93**: 651–60

Rennie EA, Hansen SF, Baidoo EEK, Hadi MZ, Keasling JD, Scheller HV (2012) Three members of the Arabidopsis glycosyltransferase family 8 are xylan glucuronosyltransferases. Plant Physiol **159**: 1408–17

Rich MK, Schorderet M, Reinhardt D (2014) The role of the cell wall compartment in mutualistic symbioses of plants. Front Plant Sci **5**: 238

Richmond T (2000) Higher plant cellulose synthases. Genome Biol **1**: REVIEWS3001

Ridley BL, O'Neill MA, Mohnen D (2001) Pectins: structure, biosynthesis, and oligogalacturonide-related signaling. Phytochemistry **57**: 929–67

Ringli C, Keller B, Ryser U (2001) Glycine-rich proteins as structural components of plant cell walls. Cell Mol Life Sci **58**: 1430–1441

Del Río JC, Rencoret J, Gutiérrez A, Nieto L, Jiménez-Barbero J, Martínez ÁT (2011) Structural characterization of guaiacyl-rich lignins in flax (Linum usitatissimum) fibers and shives. J Agric Food Chem **59**: 11088–11099

Roach MJ, Deyholos MK (2007) Microarray analysis of flax (Linum usitatissimum L.) stems identifies transcripts enriched in fibre-bearing phloem tissues. Mol Genet Genomics **278**: 149–165

Roach MJ, Mokshina NY, Badhan A, Snegireva A V., Hobson N, Deyholos MK, Gorshkova TA (2011) Development of cellulosic secondary walls in flax fibers requires beta-galactosidase. Plant Physiol **156**: 1351–1363

Rossignol M, Peltier J-B, Mock H-P, Matros A, Maldonado AM, Jorrín J V (2006) Plant proteome analysis: a 2004-2006 update. Proteomics **6**: 5529–48

Sachetto-Martins G, Franco LO, de Oliveira DE (2000) Plant glycine-rich proteins: a family or just proteins with a common motif? Biochim Biophys Acta - Gene Struct Expr **1492**: 1–14

Saito F, Suyama A, Oka T, Yoko-O T, Matsuoka K, Jigami Y, Shimma Y-I (2014) Identification of Novel Peptidyl Serine α-Galactosyltransferase Gene Family in Plants. J Biol Chem **289**: 20405–20420

Saladié M, Cañizares J, Phillips MA, Rodriguez-Concepcion M, Larrigaudière C, Gibon Y, Stitt M, Lunn JE, Garcia-Mas J (2015) Comparative transcriptional profiling analysis of developing melon (Cucumis melo L.) fruit from climacteric and non-climacteric varieties. BMC Genomics **16**: 440

Sampedro J, Cosgrove DJ (2005) The expansin superfamily. Genome Biol **6**: 242

San Clemente H, Jamet E (2015) WallProtDB, a database resource for plant cell wall proteomics. Plant Methods **11**: 2

Scarwell H-J, Roussel I (2010) Le changement climatique : Quand le climat nous pousse à changer d'ère. Développement durable Territ. 1:

Scheller HV, Ulvskov P (2010) Hemicelluloses. Annu Rev Plant Biol **61**: 263–289

Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Schölkopf B, Weigel D, Lohmann JU (2005) A gene expression map of Arabidopsis thaliana development. Nat Genet **37**: 501–6

Schultz CJ (2004) Post-translational Modifications of Arabinogalactan-peptides of Arabidopsis thaliana: ENDOPLASMIC RETICULUM AND GLYCOSYLPHOSPHATIDYLINOSITOL-ANCHOR SIGNAL CLEAVAGE SITES AND HYDROXYLATION OF PROLINE. J Biol Chem **279**: 45503–45511

Seifert GJ, Roberts K (2007) The Biology of Arabinogalactan Proteins. Annu Rev Plant Biol **58**: 137–161

Sharma HSS, Faughey G, Lyons G (1999) Comparison of physical, chemical, and thermal characteristics of water-, dew-, and enzyme-retted flax fibers. J Appl Polym Sci **74**: 139–143

Sharma HSS, Faughey G, McCall D (1996) Effect of Sample Preparation and Heating Rate on the Differential Thermogravimetric Analysis of Flax Fibres. J Text Inst **87**: 249–257

Shi YZ, Zhu XF, Miller JG, Gregson T, Zheng SJ, Fry SC (2014) Distinct catalytic capacities of two aluminium-repressed Arabidopsis thaliana xyloglucan endotransglucosylase/hydrolases, XTH15 and XTH31, heterologously produced in Pichia. Phytochemistry **112**: 160–169

Showalter AM (1993) Structure and Function of Plant Cell Wall Proteins. PLANT CELL ONLINE **5**: 9–23

Singh KK, Mridula D, Rehal J, Barnwal P (2011) Flaxseed: a potential source of food, feed and fiber. Crit Rev Food Sci Nutr **51**: 210–22

Smith BG, Harris PJ (1999) The polysaccharide composition of Poales cell walls. Biochem Syst Ecol **27**: 33–53

Smith E, Roberts K, Hutchings A, Galfre G (1984) Monoclonal antibodies to the major structural glycoprotein of the Chlamydomonas cell wall. Planta **161**: 330–8

Snegireva A V., Ageeva M V., Amenitskii SI, Chernova TE, Ebskamp M, Gorshkova TA (2010) Intrusive growth of sclerenchyma fibers. Russ J Plant Physiol **57**: 342–355

Somerville C (2006) Cellulose synthesis in higher plants. Annu Rev Cell Dev Biol **22**: 53–78

Song D, Xi W, Shen J, Bi T, Li L (2011) Characterization of the plasma membrane proteins and receptor-like kinases associated with secondary vascular differentiation in poplar. Plant Mol Biol **76**: 97–115

Sorek N, Sorek H, Kijac A, Szemenyei HJ, Bauer S, Hématy K, Wemmer DE, Somerville CR (2014) The Arabidopsis COBRA protein facilitates cellulose crystallization at the plasma membrane. J Biol Chem **289**: 34911–20

Sørensen I, Domozych D, Willats WGT (2010) How Have Plant Cell Walls Evolved ? 1. **153**: 366–372

Soto-Cerda BJ, Duguid S, Booker H, Rowland G, Diederichsen A, Cloutier S (2014) Association mapping of seed quality traits using the Canadian flax (Linum usitatissimum L.) core collection. Theor Appl Genet **127**: 881–96

Souza AP De, Kamei CLA, Torres AF, Pattathil S, Hahn MG, Trindade LM, Buckeridge MS (2015) How cell wall complexity influences saccharification efficiency in Miscanthus sinensis. doi: 10.1093/jxb/erv183

Strasser G (2001) Pectic substances from red beet (Beta vulgaris conditiva). Part I. Structural analysis of rhamnogalacturonan I using enzymic degradation and methylation analysis. Carbohydr Polym **44**: 63–70

Summerscales J, Dissanayake NPJ, Virk AS, Hall W (2010) Composites : Part A A review of bast fibres and their composites . Part 1 – Fibres as reinforcements. Compos Part A **41**: 1329–1335

Sveinsson S, McDill J, Wong GKS, Li J, Li X, Deyholos MK, Cronk QCB (2014) Phylogenetic pinpointing of a paleopolyploidy event within the flax genus (Linum) using transcriptomics. Ann Bot **113**: 753–761

Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, et al (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res **43**: D447–52

Tan L, Qiu F, Lamport DTA, Kieliszewski MJ (2004) Structure of a hydroxyproline (Hyp)-arabinogalactan polysaccharide from repetitive Ala-Hyp expressed in transgenic Nicotiana tabacum. J Biol Chem **279**: 13156–65

Tan L, Showalter AM, Egelund J, Hernandez-Sanchez A, Doblin MS, Bacic A (2012) Arabinogalactan-proteins and the research challenges for these enigmatic plant cell surface proteoglycans. Front Plant Sci **3**: 140

Tan L, Varnai P, Lamport DTA, Yuan C, Xu J, Qiu F, Kieliszewski MJ (2010) Plant O-hydroxyproline arabinogalactans are composed of repeating trigalactosyl subunits with short bifurcated side chains. J Biol Chem **285**: 24575–83

Taylor NG (2008) Cellulose biosynthesis and deposition in higher plants. New Phytol **178**: 239–52

Tenhaken R (2015) Cell wall remodeling under abiotic stress. Front Plant Sci **5**: 1–9

Thompson JE, Fry SC (2001) Restructuring of wall-bound xyloglucan by transglycosylation in living plant cells. Plant J **26**: 23–34

Tobimatsu Y, Wagner A, Donaldson L, Mitra P, Niculaes C, Dima O, Kim JI, Anderson N, Loque D, Boerjan W, et al (2013) Visualization of plant cell wall lignification using fluorescence-tagged monolignols. Plant J **76**: 357–366

Touré A, Xueming X (2010) Flaxseed Lignans: Source, Biosynthesis, Metabolism, Antioxidant Activity, Bio-Active Components, and Health Benefits. Compr Rev Food Sci Food Saf **9**: 261–269

Trethewey JAK, Campbell LM, Harris PJ (2005) (1->3),(1->4)-{beta}-d-Glucans in the cell walls of the Poales (sensu lato): an immunogold labeling study using a monoclonal antibody. Am J Bot **92**: 1660–74

Tripathy JN, Zhang J, Robin S, Nguyen TT, Nguyen HT (2000) QTLs for cell-membrane stability mapped in rice (Oryza sativa L.) under drought stress. TAG Theor Appl Genet **100**: 1197–1202

Tsekos I (1999) THE SITES OF CELLULOSE SYNTHESIS IN ALGAE: DIVERSITY AND EVOLUTION OF CELLULOSE-SYNTHESIZING ENZYME COMPLEXES. J Phycol **35**: 635–655

Uddin MN, Hanstein S, Leubner R, Schubert S (2013) Leaf Cell-Wall Components as Influenced in the First Phase of Salt Stress in Three Maize (Zea mays L.) Hybrids Differing in Salt Resistance. J Agron Crop Sci **199**: 405–415

Valentim-Neto PA, Rossi GB, Anacleto KB, de Mello CS, Balsamo GM, Arisi ACM (2015) Leaf proteome comparison of two GM common bean varieties and their non-GM counterparts by principal component analysis. J Sci Food Agric. doi: 10.1002/jsfa.7166

Vanholme R, Demedts B, Morreel K, Ralph J, Boerjan W (2010) Lignin biosynthesis and structure. Plant Physiol **153**: 895–905

Vanholme R, Morreel K, Darrah C, Oyarce P, Grabber JH, Ralph J, Boerjan W (2012) Metabolic engineering of novel lignin in biomass crops. New Phytol **196**: 978–1000

Vanzin GF, Madson M, Carpita NC, Raikhel N V, Keegstra K, Reiter W-D (2002) The mur2 mutant of Arabidopsis thaliana lacks fucosylated xyloglucan because of a lesion in fucosyltransferase AtFUT1. Proc Natl Acad Sci U S A **99**: 3340–5

Velasquez M, Salter JS, Dorosz JG, Petersen BL, Estevez JM (2012) Recent Advances on the Posttranslational Modifications of EXTs and Their Roles in Plant Cell Walls. Front Plant Sci **3**: 93

Velasquez SM, Ricardi MM, Dorosz JG, Fernandez P V, Nadra AD, Pol-Fachin L, Egelund J, Gille S, Harholt J, Ciancia M, et al (2011) O-glycosylated cell wall proteins are essential in root hair growth. Science **332**: 1401–3

Vidal S, Doco T, Williams P, Pellerin P, York WS, O'Neill MA, Glushka J, Darvill AG, Albersheim P (2000) Structural characterization of the pectic polysaccharide rhamnogalacturonan II: evidence for the backbone location of the aceric acid-containing oligoglycosyl side chain. Carbohydr Res **326**: 277–94

Vincken J-P (2003) If Homogalacturonan Were a Side Chain of Rhamnogalacturonan I. Implications for Cell Wall Architecture. PLANT Physiol **132**: 1781–1789

Vogel J (2008) Unique aspects of the grass cell wall. 301–307

Voragen AGJ, Coenen G-J, Verhoef RP, Schols HA (2009) Pectin, a versatile polysaccharide present in plant cell walls. Struct Chem **20**: 263–275

Vorwerk S, Somerville S, Somerville C (2004) The role of plant cell wall polysaccharide composition in disease resistance. Trends Plant Sci **9**: 203–9

Wang Z, Hobson N, Galindo L, Zhu S, Shi D, McDill J, Yang L, Hawkins S, Neutelings G, Datla R, et al (2012) The genome of flax (Linum usitatissimum) assembled de novo from short shotgun sequence reads. Plant J **72**: no–no

Whitcombe AJ, O'Neill MA, Steffan W, Albersheim P, Darvill AG (1995) Structural characterization of the pectic polysaccharide, rhamnogalacturonan-II. Carbohydr Res **271**: 15–29

Willats WG, Knox JP (1996) A role for arabinogalactan-proteins in plant cell expansion: evidence from studies on the interaction of beta-glucosyl Yariv reagent with seedlings of Arabidopsis thaliana. Plant J **9**: 919–25

Willats WG, McCartney L, Mackie W, Knox JP (2001) Pectin: cell biology and prospects for functional analysis. Plant Mol Biol **47**: 9–27

Wodniok S, Brinkmann H, Glöckner G, Heidel AJ, Philippe H, Melkonian M (2011) Origin of land plants : Do conjugating green algae hold the key ? BMC Evol Biol **11**: 104

Wolf S, Rausch T, Greiner S (2009) The N-terminal pro region mediates retention of unprocessed type-I PME in the Golgi apparatus. Plant J **58**: 361–75

Wróbel-Kwiatkowska M, Czemplik M, Kulma A, Zuk M, Kaczmar J, Dymińska L, Hanuza J, Ptak M, Szopa J (2012) New biocomposites based on bioplastic flax fibers and biodegradable polymers. Biotechnol Prog **28**: 1336–46

Wu Y, Cosgrove DJ (2000) Adaptation of roots to low water potentials by changes in cell wall extensibility and cell wall proteins. J Exp Bot **51**: 1543–53

Xu H, Ke K, Zhang Z, Luo X, Jin X, Liu SS-Y, Fan Z (2013) Effects of platelet-rich plasma and recombinant human bone morphogenetic protein-2 on suture distraction osteogenesis. J Craniofac Surg **24**: 645–50

Yang L, Wang CC, Guo WD, Li XB, Lu M, Yu CL (2006) Differential expression of cell wall related genes in the elongation zone of rice roots under water deficit. Russ J Plant Physiol **53**: 390–395

Yang Y, Zhu K, Xia H, Chen L, Chen K (2014) Comparative proteomic analysis of indica and japonica rice varieties. Genet Mol Biol **37**: 652–61

Yi H, Hwang KT, Regenstein JM, Shin SW (2014) Fatty Acid Composition and Sensory Characteristics of Eggs Obtained from Hens Fed Flaxseed Oil, Dried Whitebait and/or Fructo-oligosaccharide. Asian-Australasian J Anim Sci **27**: 1026–34

Zabotina OA (2012) Xyloglucan and Its Biosynthesis. Front Plant Sci **3**: 134

Zeid IM, Shedeed ZA (2006) Response of alfalfa to putrescine treatment under drought stress. Biol Plant **50**: 635–640

Zeng W, Jiang N, Nadella R, Killen TL, Nadella V, Faik A (2010) A glucurono(arabino)xylan synthase complex from wheat contains members of the GT43, GT47, and GT75 families and functions cooperatively. Plant Physiol **154**: 78–97

Zhong R, Peña MJ, Zhou G-K, Nairn CJ, Wood-Jones A, Richardson EA, Morrison WH, Darvill AG, York WS, Ye Z-H (2005) Arabidopsis fragile fiber8, which encodes a putative glucuronyltransferase, is essential for normal secondary wall synthesis. Plant Cell **17**: 3390–408

Zhong R, Ye Z (2014) Secondary Cell Walls: Biosynthesis, Patterned Deposition and Transcriptional Regulation. Plant Cell Physiol **0**: 1–20

Zini E, Scandola M, Gatenholm P (2003) Heterogeneous acylation of flax fibers. Reaction kinetics and surface properties. Biomacromolecules **4**: 821–7

Figure legends

Figure 1: Comparison of fiber bundle parameters in flax stems of different varieties. **A**. Photomicrograph of freehand cross-section of typical flax stem stained with Touidine Blue O showing fiber bundles. **B**. Zoom of fiber bundle showing elementary fibers (arrows). **C**. Total fiber bundle surface area and elementary fiber number per bundle in 2012 samples. **D**. Total

fiber bundle surface area and elementary fiber number per bundle in 2013 samples. A: Adelie; B: Belinka; D: Diane; H: Hermes; DR: Drakkar; V: Violin; O: Oliver. Spring fiber varieties are represented by green columns, winter fiber varieties are represented by grey columns and the winter oil variety is represented by a red column. Statistically significant differences (ANOVA, $p < 0.05$) are represented by different letters.

Figure 2: Principle component analysis representing the repartition of the seven varieties used according to their genotype and years based on differentially-accumulated transcripts. The axe 1 (PC1) is explained by the differences between years and the axe 2 (PC2) is explained by the differences between the varieties. FS = fiber spring, FW = fiber winter, OW = oil winter.

Figure 3: Variety level differential gene expression in pairwise comparisons of different flax varieties. Number of differentially expressed genes (DEGs) (**A**) and MapMan-based functional classification of differentially expressed genes (**B**). Spring fiber varieties (SF) *vs* Spring fiber varieties (Diane : D; Belinka : B; Drakkar : DR; Hermes : H) green; Spring fiber varieties (SF) *vs* Winter fiber varieties (WF) (Adelie : A; Violin : V) grey; Winter fiber varieties (WF) *vs* Winter fiber varieties (WF) blue; Winter oil variety (Oliver : O) *vs* Spring fiber varieties orange; Winter oil variety (Oliver : O) *vs* Winter fiber varieties (WF) yellow. Functional classes : photosynthesis (1), major CHO metabolism (2) minor CHO metabolism (3), OPP (7), mitochondrial electron transport/ATP synthesis (9), cell wall (10), lipid metabolism (11), N-metabolism (12), amino acid metabolism, (13), metal handling (15), secondary metabolism (16), hormones metabolism (17), stress (20), redox (21), polyamine metabolism (22), nucleotide metabolism (23), miscellaneous (26), RNA (27), DNA (28), proteins (29), signaling (30), cell (31), development (33), transport (34), not assigned (35).

Figure 4: Principal component analysis of DEGs at the variety level and environmental level (**A**) and variety level (**B**). Hierarchical representation of the varieties and years with those replicas (**C**).

Figure 5: Number of genes differentially expressed at the environment level. (**A**) Genes are divided into three classes: i) genes differentially expressed between the years for all varieties (blue); ii) genes influenced by the environment-variety specific (orange) and iii) genes influenced by the environment common to 2 or more varieties (gray). (**B**) Functional classification of class i genes. (**C**) Functional classification of class ii genes. (**D**) Functional classification of class iii genes.Functional classes : photosynthesis (1), major CHO metabolism (2) minor CHO metabolism (3), OPP (7), mitochondrial electron transport/ATP synthesis (9),

cell wall (10), lipid metabolism (11), N-metabolism (12), amino acid metabolism, (13), metal handling (15), secondary metabolism (16), hormones metabolism (17), stress (20), redox (21), polyamine metabolism (22), nucleotide metabolism (23), miscellaneous (26), RNA (27), DNA (28), proteins (29), signaling (30), cell (31), development (33), transport (34), not assigned (35).

Figure 6: Classification of cell wall related DEGS according to their variety level and environmental level behavior The LusID of the genes are represented with different colors according to the cell wall polymer/function with which they are associated. Varieties are represented with the following different colors: blue: Adelie; green: Belinka; red: Diane; gray: Drakkar; black: Hermes; turquoise: Violin, pink: Oliver. Genes that are differentially expressed at both the variety- and environmental-levels are represented by their LusID. The first line indicates the varieties that show differential expression of the gene, the second line shows the differential expression (>/<) between the varieties.

Figure 7: Differentially-accumulated proteins between Diane (SF) and Oliver (WO) varieties. Venn diagram showing the distribution numbers of proteins identified betweenthe 4 biological relevant ratios investigated in this analysis. B) Functional classification of all the proteins differentially abundant between Diane and Oliver in two successive years.

Figure 8: New gene model for a putative expansin/EG45-like domain containing protein. Scaffold indicates where the gene was found, the transcript and its relative coordinates are given. For the transcript: the simple line represents genome sequences not found in the transcript sequence, white rectangles are parts of the sequences found in the transcript but not coding the protein, while colored rectangles are transcripts that are translated into protein. The protein sequence is shown, and peptides found by protein sequencing are underlined.

Figure 9: Phylogenetic tree of the xyloglucan endo-transglycosylases/hydrolases IIIA, found in all plant species present in the phytozome database v10, performed with Bayesian analyses. The support value for each branch indicated is the posterior probability and the bootstraps values obtained with 1000 bootstrap repetitions. The proteins sequences found in *Linum usitassinum*, *Populus thiocharpa*, and *Arabidopsis thaliana* are colored respectively in red, blue and purple. Red arrow indicates flax XTH gene found, in this study, blue arrow indicates XTH gene identified previously (Chabi et *al.*, submitted, Flax Proteome Atlas).

Tables

Table 1: List of non-redundant differentially abundant flax proteins

Table 2: List of genes/ proteins differentially accumulated and identified by both transcriptomics and proteomics

Supplementary Data

For space reasons, only the supplementary data 3 are presented in the printed version of this paper.

Supplementary data 1. List of all differentially expressed genes (DEGs).

Supplementary data 2. List of Variety level DEGs

Supplementary data 3. List of cell wall related DEGs

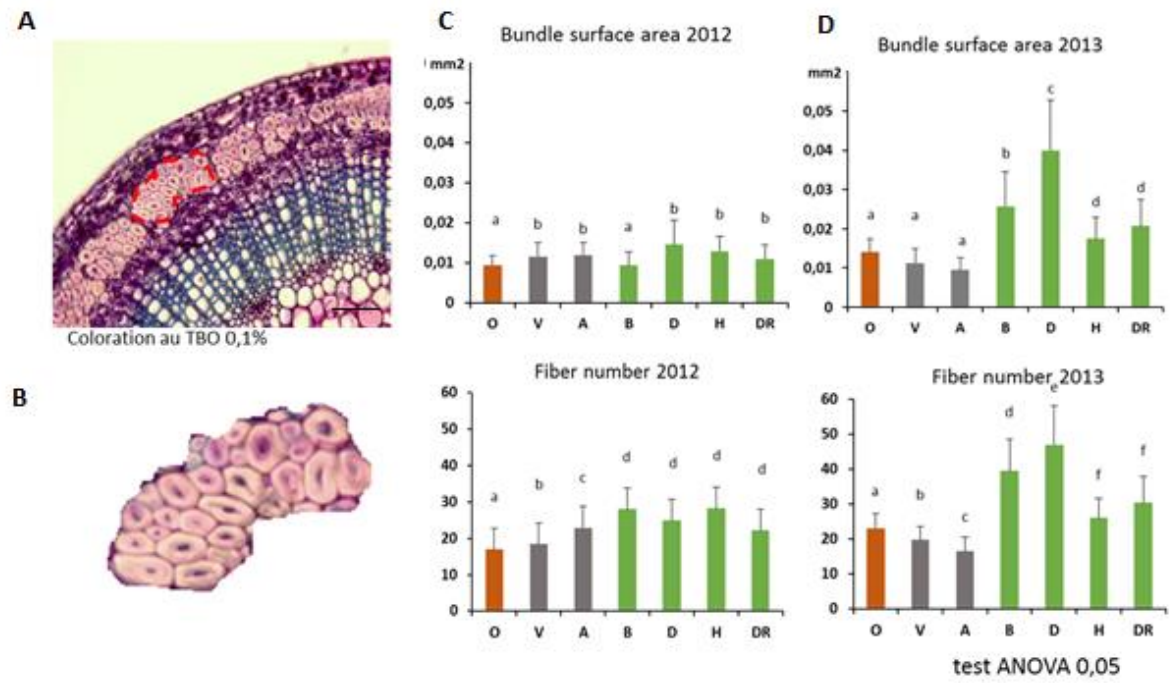


Figure 1: Comparison of fiber bundle parameters in flax stems of different varieties. A. Photomicrograph of freehand cross-section of typical flax stem stained with Touidine Blue O showing fiber bundles. B . Zoom of fiber bundle showing elementary fibers (arrows). C . Total fiber bundle surface area and elementary fiber number per bundle in 2012 samples. D. Total fiber bundle surface area and elementary fiber number per bundle in 2013 samples. A: Adelie; B: Belinka; D: Diane; H: Hermes; DR: Drakkar; V: Violin; O: Oliver. Spring fiber varieties are represented by green columns, winter fiber varieties are represented by grey columns and the winter oil variety is represented by a red column. Statistically significant differences (ANOVA, $p < 0.05$) are represented by different letters.

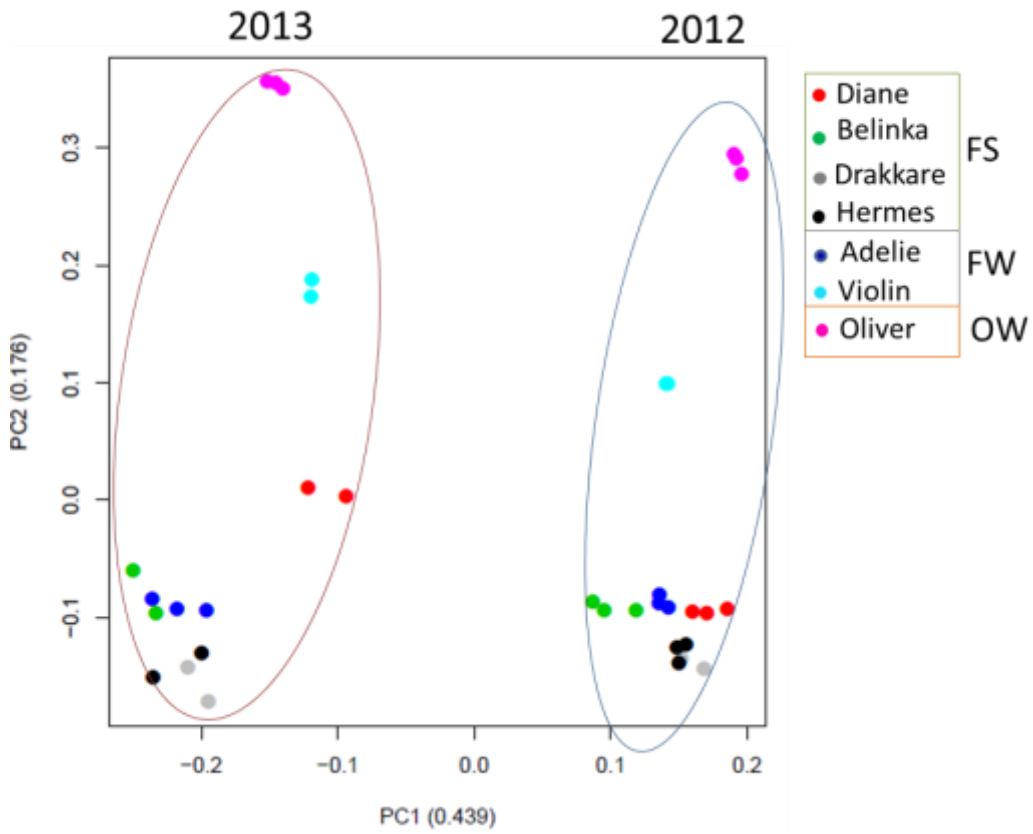


Figure 2: Principle component analysis representing the repartition of the seven varieties used according to their genotype and years based on differentially-accumulated transcripts. The axe 1 (PC1) is explained by the differences between years and the axe 2 (PC2) is explained by the differences between the varieties. FS = fiber spring, FW = fiber winter, OW = oil winter.

.

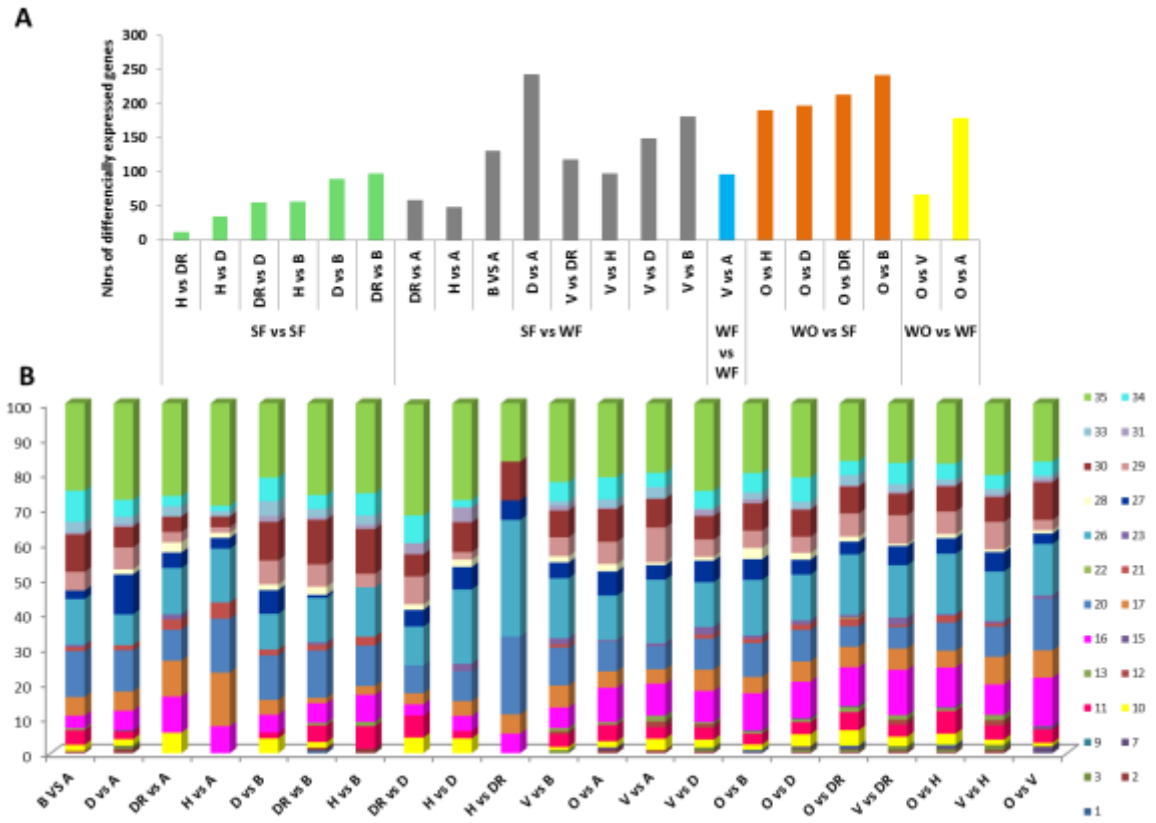


Figure 3: Variety level differential gene expression in pairwise comparisons of different flax varieties. Number of differentially expressed genes (DEGs) (**A**) and MapMan-based functional classification of differentially expressed genes (**B**). Spring fiber varieties (SF) *vs* Spring fiber varieties (Diane : D; Belinka : B; Drakkar : DR; Hermes : H) green; Spring fiber varieties (SF) *vs* Winter fiber varieties (WF) (Adelie : A; Violin : V) grey; Winter fiber varieties (WF) *vs* Winter fiber varieties (WF) blue; Winter oil variety (Oliver : O) *vs* Spring fiber varieties orange; Winter oil variety (Oliver : O) *vs* Winter fiber varieties (WF) yellow. Functional classes : photosynthesis (1), major CHO metabolism (2) minor CHO metabolism (3), OPP (7), mitochondrial electron transport/ATP synthesis (9), cell wall (10), lipid metabolism (11), N-metabolism (12), amino acid metabolism, (13), metal handling (15), secondary metabolism (16), hormones metabolism (17), stress (20), redox (21), polyamine metabolism (22), nucleotide metabolism (23), miscellaneous (26), RNA (27), DNA (28), proteins (29), signaling (30), cell (31), development (33), transport (34), not assigned (35).

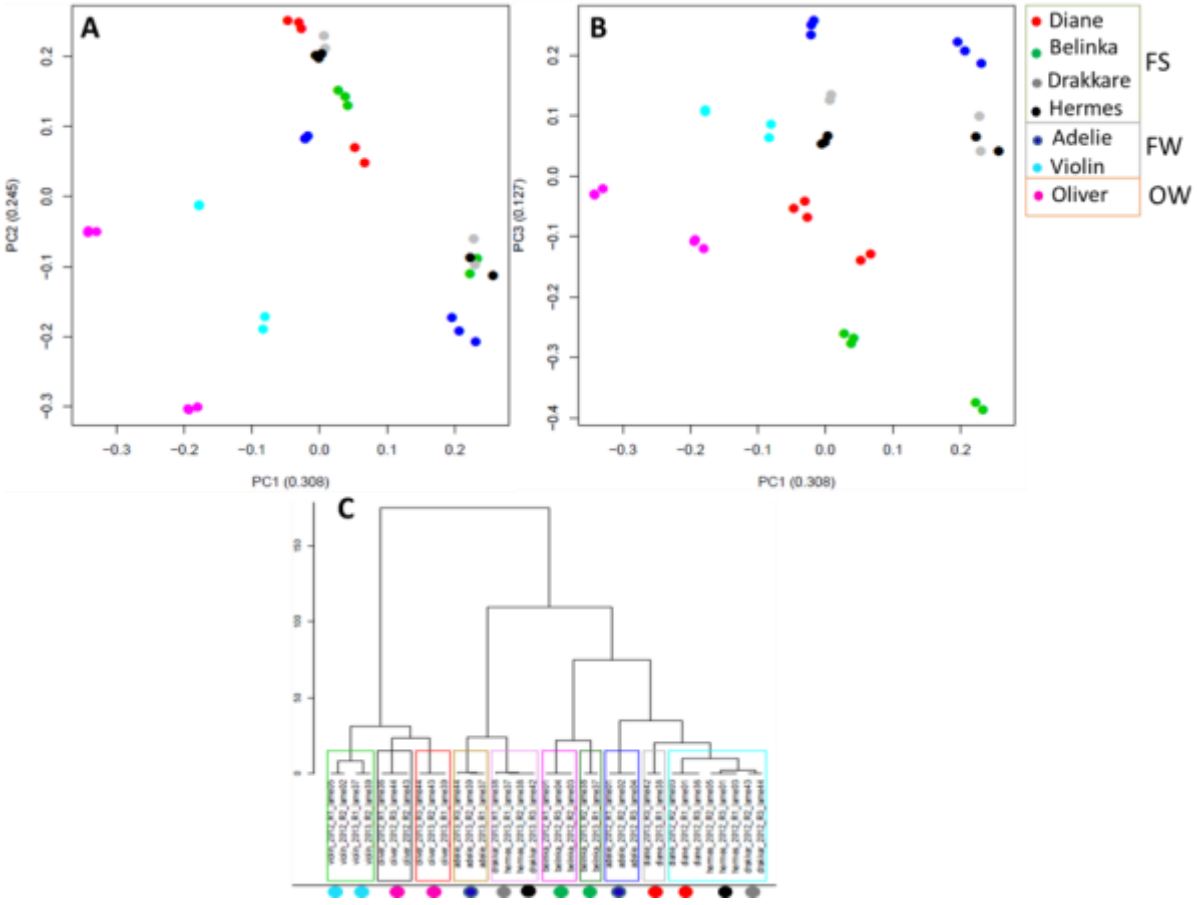


Figure 4: Principal component analysis of DEGs at the variety level and environmental level (**A**) and variety level (**B**). Hierarchical representation of the varieties and years with those replicas (**C**).

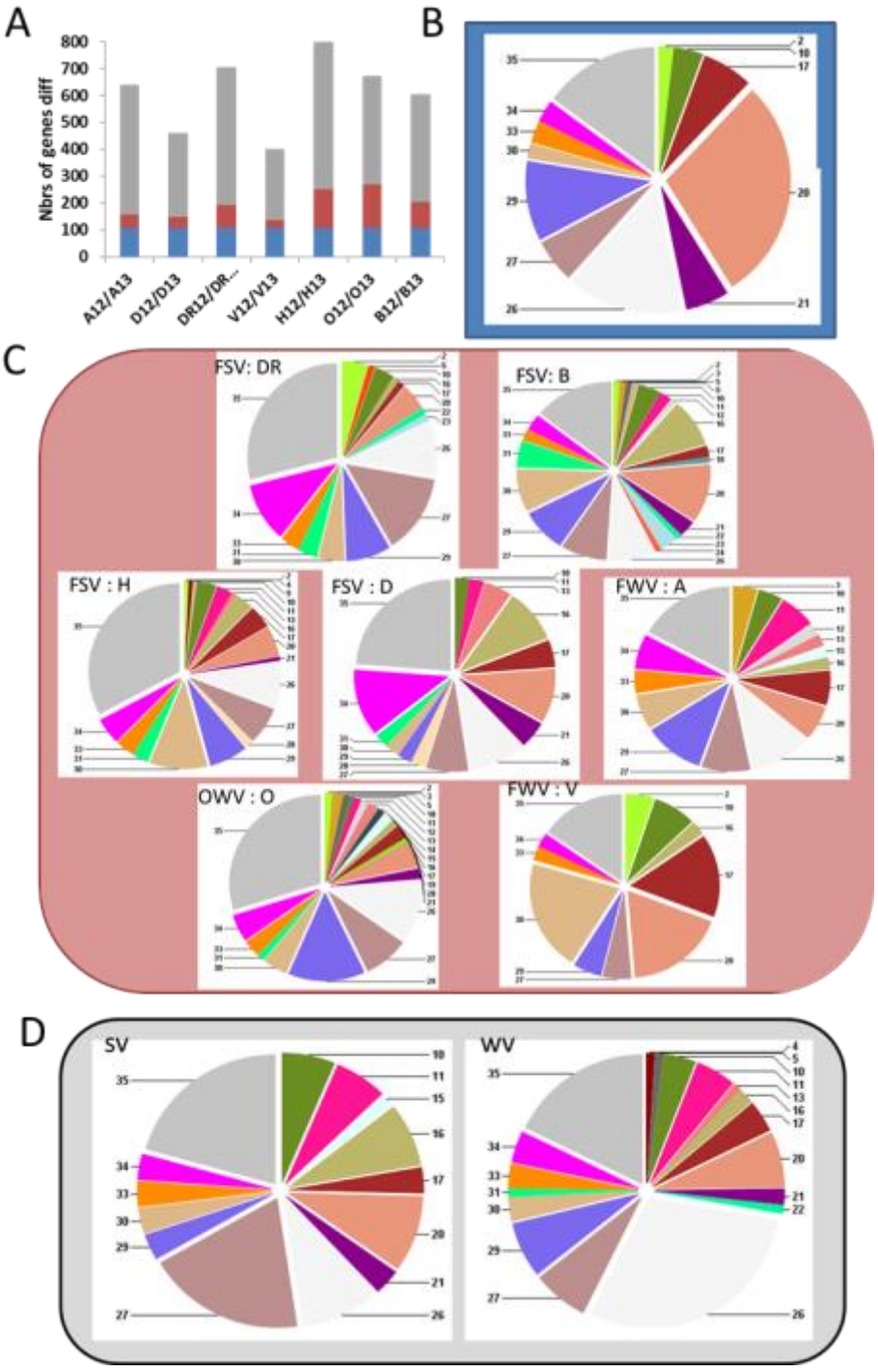


Figure 5: Number of genes differentially expressed at the environment level. (**A**) Genes are divided into three classes: i) genes differentially expressed between the years for all varieties (blue); ii) genes influenced by the environment-variety specific (orange) and iii) genes influenced by the environment common to 2 or more varieties (gray). (**B**) Functional classification of class i genes. (**C**) Functional classification of class ii genes. (**D**) Functional classification of class iii genes.Functional classes : photosynthesis (1), major CHO metabolism (2) minor CHO metabolism (3), OPP (7), mitochondrial electron transport/ATP synthesis (9), cell wall (10), lipid metabolism (11), N-metabolism (12), amino acid metabolism, (13), metal handling (15), secondary metabolism (16), hormones metabolism (17), stress (20), redox (21), polyamine metabolism (22), nucleotide metabolism (23), miscellaneous (26), RNA (27), DNA (28), proteins (29), signaling (30), cell (31), development (33), transport (34), not assigned (35).

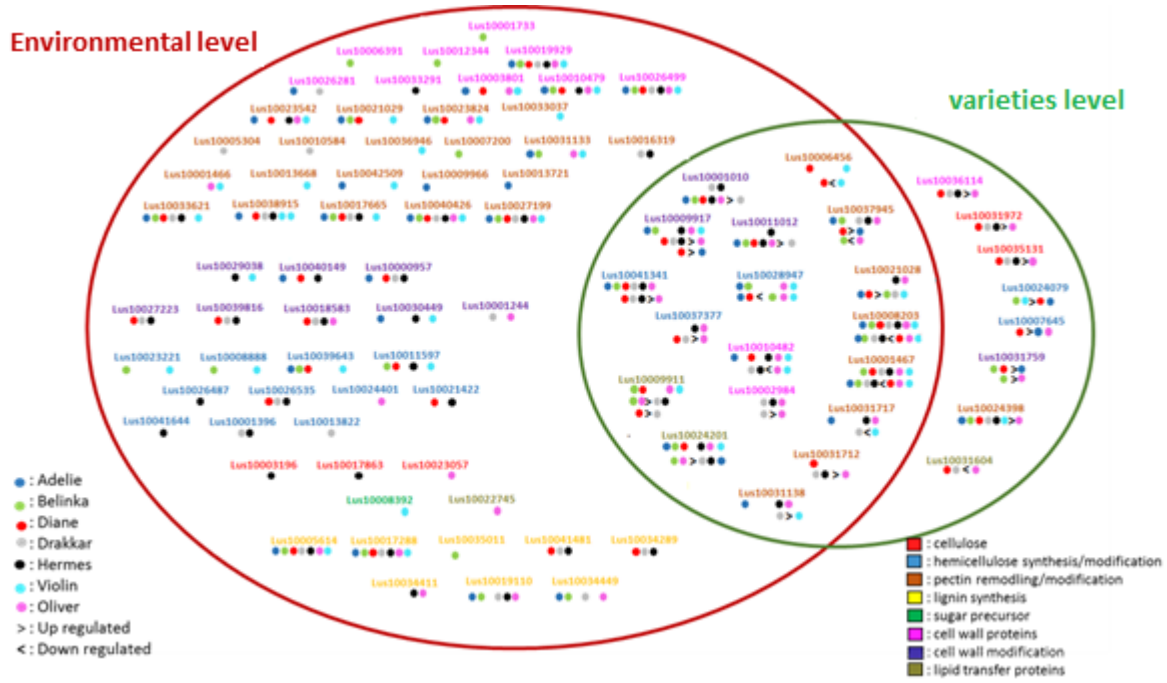


Figure 6: Classification of cell wall related DEGS according to their variety level and environmental level behavior The LusID of the genes are represented with different colors according to the cell wall polymer/function with which they are associated. Varieties are represented with the following different colors: blue: Adelie; green: Belinka; red: Diane; gray: Drakkar; black: Hermes; turquoise: Violin, pink: Oliver. Genes that are differentially expressed at both the variety- and environmental-levels are represented by their LusID. The first line indicates the varieties that show differential expression of the gene, the second line shows the differential expression (>/<) between the varieties.

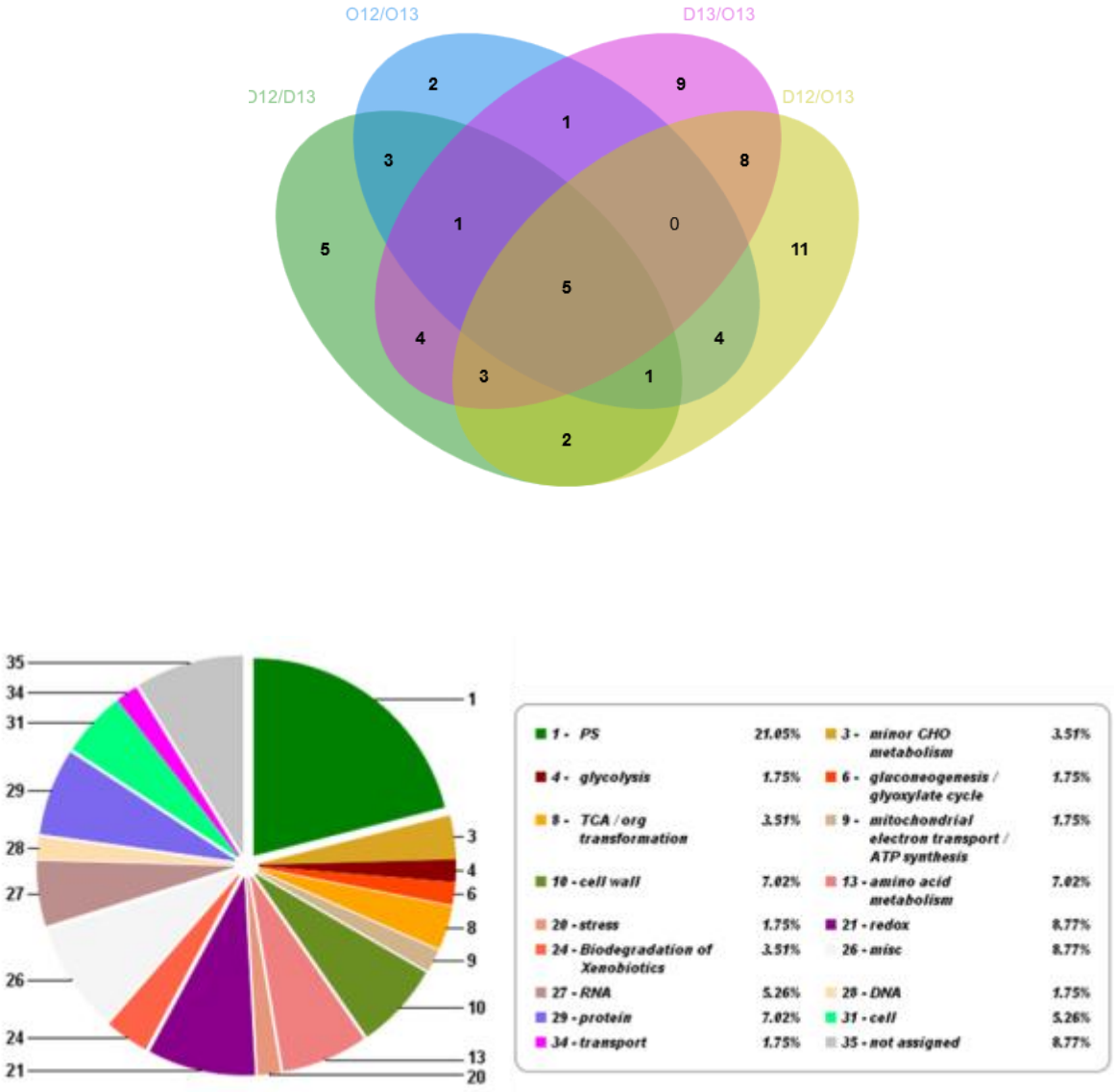


Figure 7: Differentially-accumulated proteins between Diane (SF) and Oliver (WO) varieties. Venn diagram showing the distribution numbers of proteins identified betweenthe 4 biological relevant ratios investigated in this analysis. B) Functional classification of all the proteins differentially abundant between Diane and Oliver in two successive years.

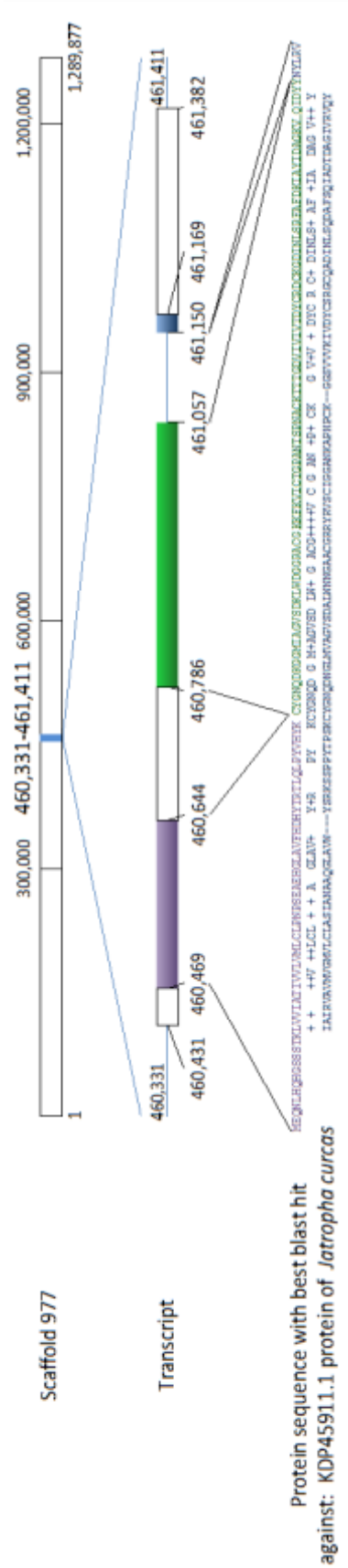


Figure 8: New gene model for a putative expansin/EG45-like domain containing protein. Scaffold indicates where the gene was found, the transcript and its relative coordinates are given. For the transcript: the simple line represents genome sequences not found in the transcript sequence, white rectangles are parts of the sequences found in the transcript but not coding the protein, while colored rectangles are transcripts that are translated into protein. The protein sequence is shown, and peptides found by protein sequencing are underlined.

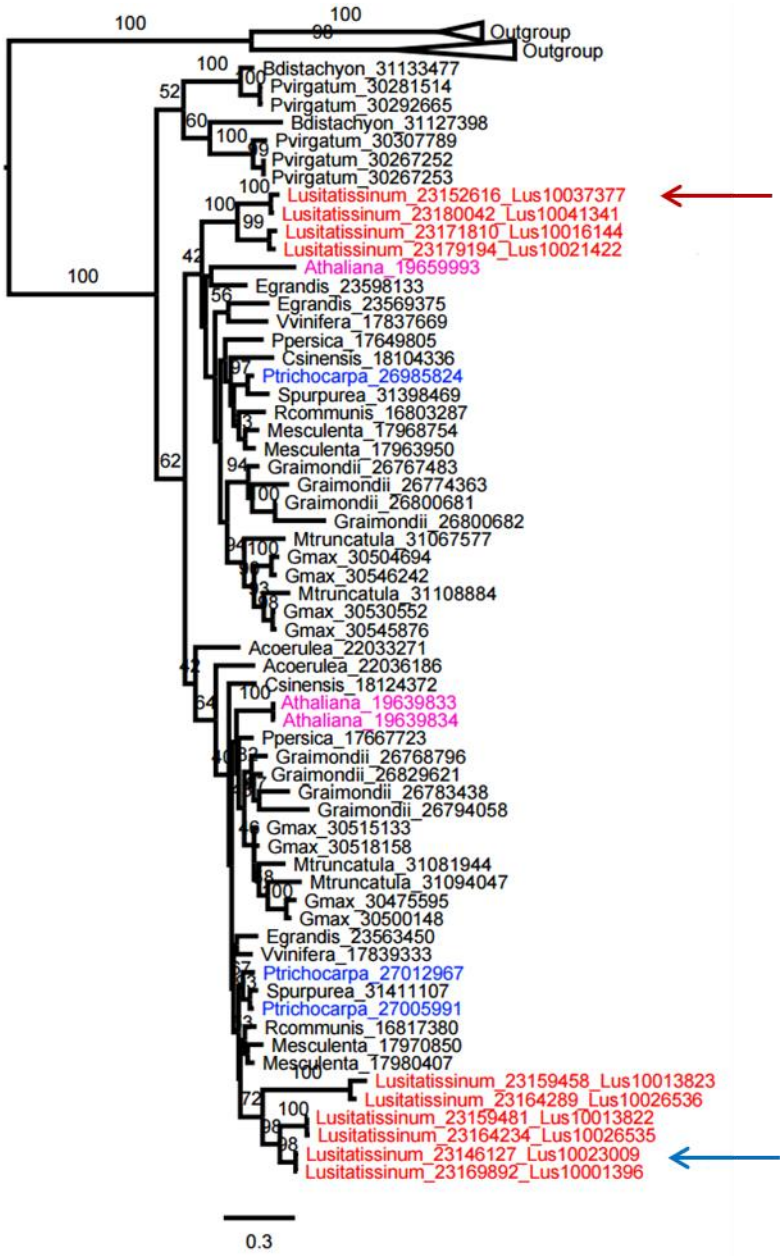


Figure 9: Phylogenetic tree of the xyloglucan endo-transglycosylases/hydrolases IIIA, found in all plant species present in the phytozome database v10, performed with Bayesian analyses. The support value for each branch indicated is the posterior probability and the bootstraps values obtained with 1000 bootstrap repetitions. The proteins sequences found in *Linum usitassinum*, *Populus thiocharpa*, and *Arabidopsis thaliana* are colored respectively in red, blue and purple. Red arrow indicates flax XTH gene found, in this study, blue arrow indicates XTH gene identified previously (Chabi et *al.*, submitted, Flax Proteome Atlas).

Supplementary data 3: cell wall related genes differentially expressed between the years and varieties. Up: up regulated in 2012 vs 2013; Down : down regulated on 2012 vs 2013; D diane (sfv); B belinka (sfv); Dr: Dreakkar (sfv); H: Hermes (sfv); A: Adelie (wfv); V: Violin (wfv); O: Oliver (wov).

Feature ID	Discription	At ID	Enverenment level		variety level	
Lus10026499	cell wall proteins.AGP	FLA11	yes	all up	No	_
Lus10019929	cell wall proteins.AGP	FLA11	yes	all up	No	_
Lus10036114	cell wall proteins.AGP	FLA12	No	_	yes	D,Dr ,H> O
Lus10002984	cell wall proteins.AGP	FLA12	yes	Up Dr, H, O	yes	D,Dr> O
Lus10001733	cell wall proteins.AGP	Fla7	yes	B up	No	_
Lus10006391	cell wall proteins.AGP	Fla2	yes	_	No	_
Lus10012344	cell wall proteins.AGP	Fla2	yes	_	No	_
Lus10003801	cell wall proteins.proline rich proteins	(LTP) family protein	yes	up O, V, A, D	No	_
Lus10010482	cell wall proteins.proline rich proteins	(LTP) family protein	yes	up O, V, A, D, H	yes	O, V>Dr, H
Lus10010479	cell wall proteins.proline rich proteins	(LTP) family protein	yes	up O, V, A, D, H, B	No	_
Lus10026281	cell wall proteins.proline rich proteins	proline-rich protein 2	yes	up A, DR	No	_
Lus10040149	cell wall.modification (EXP)	expansin-like A2	yes	up A D, H	No	_
Lus10000957	cell wall.modification (EXP)	expansin-like A2	yes	up A D, Dr, H	No	_
Lus10039816	cell wall.modification (EXP)	expansin B1	yes	down D, Dr, H	No	_
Lus10018583	cell wall.modification (EXP)	expansin B1	yes	down D, Dr, H, O	No	_
Lus10027223	cell wall.modification (EXP)	expansin B4	yes	down D, Dr, H	No	_
Lus10031759	cell wall.modification (EXP)	plant natriuretic peptide A	no	_	yes	D>A; B>O
Lus10001010	cell wall.modification (EXP)	expansin-like A2	yes	down DR; UP H	yes	A, O, V, B>Dr;
Lus10011012	cell wall.modification (EXP)	expansin-like A3	yes	UP H	yes	A, O, V, B, D>Dr
Lus10029038	cell wall.modification (EXP)	expansin A2	yes	up H, V	no	_
Lus10001244	cell wall.modification (EXP)	expansin-like A3	yes	up Dr, O	no	_
Lus10030449	cell wall.modification (EXP)	expansin A1	yes	up A,H, V	No	_

Lus10009917	cell wall.modification (EXP)	expansin A2	yes	up A, B, H,O, V	yes	H, Dr, D> O; D>A
Lus10023057	cellulose	cellulose synthase like G1	yes	down O	No	_
Lus10003196	cellulose	cellulose synthase like G1	yes	down O	No	_
Lus10035131	cellulose	COBRA-like	No	_	yes	D, Dr, H>O
Lus10031972	cellulose	COBRA-like	No	_	yes	D, Dr, H>O
Lus10017863	cellulose	IRX6	yes	up H	no	_
Lus10041644	hemicellulose biosynthesis	fucosyltransferase 7	yes	up H	no	_
Lus10024079	hemicellulose biosynthesis	fucosyltransferase 1	No	_	yes	B>D, A; V>D
Lus10026487	hemicellulose biosynthesis	IRX9	yes	up H	no	_
Lus10041341	hemicellulose modification	XTR8	yes	up A,B,D,Dr,O	yes	D, Dr, H> O
Lus10011597	hemicellulose modification	XTH7	yes	up B,D, H, V	No	_
Lus10039643	hemicellulose modification	XTH7	yes	up A, B, D, V	No	_
Lus10021422	hemicellulose modification	XTH32	yes	down D, H	No	_
Lus10026535	hemicellulose modification	XTH32	yes	down D, Dr, H	No	_
Lus10013822	hemicellulose modification	XTH32	yes	down Dr	No	_
Lus10037377	hemicellulose modification	XTH32	yes	up H, O	yes	D, Dr> O
Lus10007645	hemicellulose modification	XTH8	No	_	yes	D>A, O
Lus10001396	hemicellulose modification	XTR8	yes	up Dr, H	No	_
Lus10023221	hemicellulose modification	XTH16	yes	up B, V	No	_
Lus10008888	hemicellulose modification	XTH15	yes	up B, V	No	_
Lus10028947	hemicellulose modification	XTH16	yes	up A, B, O, V	yes	B, Dr, O, V>A; B, O V> D
Lus10041481	lignin biosynthesis	laccase 10	yes	up D, Dr, H	No	_
Lus10034289	lignin biosynthesis	laccase 10	yes	up D, Dr, H	No	_
Lus10005614	lignin biosynthesis	peroxidases	yes	up all	No	_
Lus10017288	lignin biosynthesis	peroxidases	yes	up all	No	_
Lus10019110	lignin biosynthesis	C4H	yes	down A, B, Dr, H, O	No	_
Lus10034449	lignin biosynthesis	C4H	yes	down A, B, Dr, O	No	_
Lus10035011	lignin biosynthesis	C4H	yes	dawn B	No	_
Lus10024201	lipid transfer proteins	LTP1	yes	dawn A, B, D, H, O, V	yes	B, O >A, Dr, H

Lus10009911	lipid transfer proteins	LTP3	yes	dawnB, D, O, V	yes	B, O >A, Dr, H ; D>Dr
Lus10022745	lipid transfer proteins	LTP2	yes	dawn O	No	_
Lus10031604	lipid transfer proteins	SHAVEN 3	no	_	yes	A, B, D, Dr, H< O, V
Lus10006456	pecine degradation	Pectin lyase	yes	up D, V	yes	V>D
Lus10023542	pecine degradation	Pectin lyase	yes	up A, D, H, O, V	No	_
Lus10021029	pecine degradation	PGIPs	yes	up A, B, D, V	No	_
Lus10023824	pecine degradation	PGIPs	yes	up A, B, D, V, O	No	_
Lus10040426	pecine degradation	Pectin lyase	yes	up all	No	_
Lus10005304	pecine degradation	Pectin lyase	yes	dawn Dr	No	_
Lus10010584	pecine degradation	Pectin lyase	yes	dawn Dr	No	_
Lus10036946	pecine degradation	Pectin lyase	yes	up V	No	_
Lus10033037	pecine degradation	Pectin lyase	yes	up V	No	_
Lus10021028	pecine degradation	PGIPs	yes	up H, O	yes	A, D> B, Dr, V
Lus10013668	pecine degradation	Pectin lyase	yes	up V	No	_
Lus10042509	pecine degradation	Pectin lyase	yes	up A, V	No	_
Lus10037945	pecine degradation	Pectin lyase	yes	up A, B, Dr, H, O	yes	D>A; O>B
Lus10031712	pectin modification	PMEI	yes	up D	yes	Dr, H >O
Lus10038915	pectin modification	PMEI	yes	up A, D, Dr, H, V, O	No	_
Lus10027199	pectin modification	PMEI	yes	ap all	No	_
Lus10008203	pectin modification	PMEI	yes	up all	yes	D, O,V>A, B,Dr, H
Lus10033621	pectin modification	PMEI	yes	up A, B, D, H, V	No	_
Lus10001467	pectin modification	PMEI	yes	up B, D, Dr, H, O,V	yes	D, O,V>A,B, Dr, H
Lus10017665	pectin modification	PME	yes	up A, B, D, Dr, H, V	No	_
Lus10002739	pectin modification	Cell-wall inhibitor of beta fructosidase	yes	dawn Dr, H	No	_
Lus10016319	pectin modification	Cell-wall inhibitor of beta fructosidase	yes	dawn Dr, H	No	_
Lus10024398	pectin modification	PAE	no	_	yes	A, B, D, Dr, H, V> O
Lus10001466	pectin modification	PME	yes	up O, V	No	_

Lus10031717	pectin modification	PMEI	yes	up A, H, Dr	yes	V>Dr
Lus10031138	pectin modification	PMEI	yes	up A, H, Dr	yes	V>Dr
Lus10013721	pectin modification	PME	yes	up A	No	_
Lus10009966	pectin modification	PAE	yes	up A	No	_
Lus10007200	pectin modification	PAE	yes	up B	No	_
Lus10031133	pectin modification	PMEI	yes	up A, B, O, V	No	_
Lus10031713	pectin modification	PMEI	yes	up A, B, O, V	No	_
Lus10008392	suger precursor	Beta-fructofuranosidase	yes	Up V	No	_
Lus10024401	suger precursor	alpha-galactosidase 2	yes	up O	No	_

Table 1: Proteins differentially accumulated between the years and varieties.

			Environment level				Varieties level			
			D2012 / DI2013		**O 2012 / O2013**		**D 2013 / O2013**		**D 2012 / O 2012**	
spot n°	**LusID**	**Discription**	**T-test**	**Av. Ratio**	**T-test**	**Av. Ratio**	**T-test**	**Av. Ratio**	**T-test**	**Av. Ratio**
1419	gi|171462119	Expansin-related protein 3 precursor	0,33	1,1	0,0048	-1,5	0,00082	-1,72	0,74	-1,05
1535	gi|324742904	FLAXSQ25_RP_005_D08_30NOV2007_058 LUSST1MD Linum usitatissimum cDNA, mRNA sequence	4,70E-05	3,03	0,0044	1,76	1,50E-05	-3,3	0,0031	-1,91
321	Lus10000692	Inositol-3-phosphate synthase	6,00E-05	-1,53	0,043	-1,15	0,00014	1,41	0,33	1,06
878	Lus10001508	Cyclase family protein	0,017	-1,51	0,00044	-1,62	6,10E-06	-3,63	4,60E-05	-3,39
646	Lus10001822	UDP-D-glucose 4-epimerase activity	0,00051	-1,49	0,064	-1,28	0,0052	1,48	0,019	1,27
69	Lus10002620	cytosolic cobalamin-independent methionine synthase	7,40E-01	-1,04	0,13	1,44	0,0071	-1,5	0,0055	-2,24
774	Lus10002943	glyoxalase I homolog ATGLX1	0,07	-1,21	0,43	1,09	0,032	-1,19	0,0061	-1,57
1457	Lus10004155	recombination and DNA-damage resistance protein	0,61	-1,04	0,062	-1,29	0,017	1,3	0,0091	1,62
1325	Lus10007627	photosystem II subunit Q-2	0,39	-1,14	0,0061	-1,53	0,021	1,37	0,02	1,83
1269	Lus10008932	MLP-like protein 328 (MLP328)	0,0017	-1,77	0,85	1,03	2,60E-05	4,07	0,006	2,25
940	Lus10009205	methyl esterase 7	0,00022	-1,91	0,0044	-1,41	2,90E-05	2,4	0,00046	1,76
639	Lus10010340	methyl esterase 7	0,00072	-1,21	0,37	-1,05	4,80E-06	1,58	0,0014	1,37
890	Lus10010846	NAD(P)-binding Rossmann-fold superfamily protein	0,0031	-1,59	0,029	-1,24	0,00064	2,09	0,0019	1,63
232	Lus10011125	NADP-dependent malic enzyme	0,79	-1,02	0,33	-1,06	0,00054	-1,51	0,0088	-1,46
1232	Lus10011330	Peptidyl-prolyl cis-trans isomerase	0,16	1,18	0,35	-1,1	0,034	1,22	0,0072	1,58
839	Lus10011597	xyloglucan endotransglucosylase/hydrolase 6	1,80E-05	2,03	0,0058	1,43	0,056	-1,24	0,04	1,14
1227	Lus10012167	Peptidyl-prolyl cis-trans isomerase	0,016	-1,77	0,0024	-2,61	2,60E-05	-4,52	0,0063	-3,06

779	Lus10012459	Malate dehydrogenase	0,24	1,21	0,019	1,52	0,77	-1,06	0,068	-1,32
1526	Lus10013078	50S ribosomal protein L12	5,50E-06	3,75	0,0019	3,26	0,0049	-1,58	0,25	-1,37
1046	Lus10013537	L-ascorbate peroxidase	0,00014	-1,99	0,039	-1,34	0,0047	1,55	0,66	1,04
1449	Lus10014003	unknown	0,00025	1,85	0,013	-1,53	5,00E-06	-5,81	0,00014	-2,05
690	Lus10014669	protein with ribonuclease activity	0,069	-1,16	0,84	1,01	0,0058	-1,3	0,00016	-1,52
941	Lus10015049	Glutathione transferase L3	0,00072	-1,32	0,0052	-1,22	2,80E-06	1,99	1,90E-05	1,84
834	Lus10016783	Oxygen-evolving enhancer protein 1	0,084	-1,19	0,00023	-1,9	0,0059	1,46	2,00E-05	2,34
1166	Lus10016969	2-Cys peroxiredoxin	0,013	1,66	0,51	1,06	0,74	1,06	0,0019	1,67
1354	Lus10017158	Lactoylglutathione lyase	0,038	-1,41	0,68	-1,04	0,013	1,53	0,35	1,13
291	Lus10018156	Protein disulfide-isomerase precursor	0,71	-1,03	0,0092	-1,28	0,0001	-1,5	0,17	-1,2
77	Lus10020273	cobalamin-independent methionine synthase	1,30E-02	-1,21	0,047	1,28	0,012	-1,29	0,00019	-1,99
1429	Lus10020555	heat stable protein	0,0067	-1,63	0,026	-1,5	0,3	-1,19	0,053	-1,29
1165	Lus10020717	Oxygen-evolving enhancer protein 2-1	0,00069	1,54	0,00049	2,08	0,016	-1,51	5,10E-06	-2,03
1525	Lus10022261	50S ribosomal protein L12	4,50E-05	3,17	0,00078	2,94	0,088	-1,51	0,0082	-1,4
519	Lus10022424	peroxisomal photorespiratory enzyme	0,0025	-1,57	0,0025	-1,58	0,11	1,12	0,41	1,12
837	Lus10022479	Oxygen-evolving enhancer protein 1	0,26	-1,08	0,00033	-1,59	0,0061	1,36	9,50E-06	2
1307	Lus10023180	thioredoxin-dependent peroxidase	0,83	-1,01	0,033	1,09	3,60E-06	1,67	0,011	1,52
1350	Lus10023569	Glycine-rich RNA-binding protein GRP1A	0,12	-1,21	0,31	1,17	0,0015	2,03	0,017	1,43
755	Lus10023674	potassium channel beta subunit 1	0,11	-1,2	0,0047	1,14	0,00052	-1,36	0,00071	-1,88
870	Lus10024383	NAD(P)-binding Rossmann-fold superfamily protein	0,55	-1,09	0,2	1,23	0,032	1,5	0,3	1,12
1415	Lus10024959	PSII polypeptide subunits	4,00E-05	-2,04	0,00062	-1,32	4,70E-05	1,38	0,18	-1,11
1391	Lus10025889	peptidylprolyl isomerase	0,95	1	0,0008	1,19	0,0053	-1,6	0,00035	-1,9

1030	Lus10026117	cysteine protease	0,014	1,72	0,0098	3,77	0,11	-2,59	2,80E-05	-5,66
797	Lus10026241	phosphoglycolate phosphatase	0,0046	-1,69	0,12	-1,21	0,0072	1,63	0,17	1,17
215	Lus10026370	phosphoglycerate mutase	0,029	1,12	0,19	1,13	0,00022	-1,49	0,0014	-1,51
1112	Lus10026643	glutathione S transferases	0,0076	1,31	0,1	1,16	0,00035	-1,49	0,016	-1,32
1474	Lus10026779	cysteine protease	0,016	1,36	0,31	-1,1	0,69	1,03	0,0061	1,54
731	Lus10027056	Cysteine synthase	0,5	-1,13	0,095	-1,34	0,21	1,31	0,005	1,55
609	Lus10028261	eduheptulose bisphosphatase	0,5	1,19	0,023	1,49		-1,44	0,0054	-1,8
814	Lus10028451	ferredoxin reductase	0,026	-1,54	0,071	-1,16	0,96	1,01	0,0049	-1,32
198	Lus10030577	beta glucosidase 32	0,65	1,06	0,054	-1,17	0,00015	-4,03	2,70E-05	-3,27
1278	Lus10031428	DUF538	0,019	1,64	0,0045	1,43	0,96	1,01	0,34	1,16
187	Lus10031708	transketolase	0,012	-1,27	0,23	-1,09	2,30E-05	1,5	0,14	1,29
648	Lus10032956	Aldolase superfamily protein	0,48	-1,05	0,069	1,33	0,61	-1,1	0,00053	-1,55
402	Lus10035263	ATP synthase subunit beta	0,00029	1,5	0,014	1,37	0,03	-1,52	0,002	-1,39
612	Lus10035593	Monodehydroascorbate reductase	0,48	-1,19	0,017	1,7		-1,11	0,0049	-2,25
859	Lus10037377	xyloglucan endotransglucosylase/hydrolase 32	0,96	-1,07	0,021	1,4	0,012	2,38	0,0013	1,59
860	Lus10041341	xyloglucan endo-transglycosylase-related 8	1,20E-05	3,15	0,43	-1,35	0,75	-1,18	2,00E-05	3,62
1125	Lus10041602	Plastid-specific 30S ribosomal protein 2	0,83	-1,01	0,94	1	0,011	1,75	0,11	1,72
1078	Lus10042468	Probable glutathione S-transferase	0,052	-1,13	0,22	-1,07	4,70E-06	-1,66	9,00E-05	-1,76
249	Lus10043025	NADP-dependent malic enzyme	0,34	-1,08	0,37	-1,08	0,00066	-1,77	0,0006	-1,77
441	Lus10043157	serine hydroxymethyltransferase	0,0041	1,23	0,0013	1,55	0,14	1,1	0,097	-1,15

Table 2: genes/ proteins differencially accumulated, identified with both transcriptomics and proteomics.

Lus ID	Transcriptomics				Proteomics							
	variety level	environment level			environment level				variety level			
	O/D	D12/13 O 12/13			O12 /O13		O12 /O13		D13 / O13		D12 / O12	
					T-test	Av. Ratio	T-test	Av. Ratio	T-test	Av. Ratio	T-test	Av. Ratio
Lus10001508	yes	no	no	Cyclase family protein	0,017	-1,51	0,00044	-1,62	6,10E-06	-3,63	4,60E-05	-3,39
Lus10011597	no	yes	no	XTH6	1,80E-05	2,03	0,0058	1,43	0,056	-1,24	0,04	1,14
Lus10014003	yes	yes	yes	unknwon	0,00025	1,85	0,013	-1,53	5,00E-06	-5,81	0,00014	-2,05
Lus10030577	yes	no	no	BGLU32	0,65	1,06	0,054	-1,17	0,00015	-4,03	2,70E-05	-3,27
Lus10037377	yes	no	no	XTH32	0,96	-1,07	0,021	1,4	0,012	2,38	0,0013	1,59
Lus10041341	yes	yes	yes	XTR8	1,20E-05	3,15	0,43	-1,35	0,75	-1,18	2,00E-05	3,62

Chapitre 4 : Etude de l'impact d'un stress hydrique chez le lin

Chapitre 4 : Etude de l'impact d'un stress hydrique chez le lin

Avant-propos

La production du lin est dépendante des conditions climatiques car tant la quantité que la qualité des fibres de lin sont étroitement liées aux différents facteurs environnementaux. L'un des facteurs majeurs qui impact toutes les cultures agricoles, y compris le lin, est le stress hydrique. Ce stress résultant d'un manque d'eau dans les tissus de la plante, est associé aux conditions de sécheresse provoquées par une hausse de température et une diminution de précipitations, ainsi que par une sur-exploitation par l'homme des réserves naturelles du sol. Les impacts négatifs de ce stress sur l'agriculture sont estimés en croissance continue dans la situation actuelle où le changement climatique se confirme dans de nombreuses régions du globe. Chez les producteurs de lin, la sécheresse est associée à la production de fibres « plates », de mauvaise qualité (com. pers. Linea, Terre de Lin), mais également à la production de fibre plus courtes (Chemikosova et *al.*, 2006).

Cependant, notre compréhension de l'impact d'un stress hydrique sur le métabolisme pariétal du lin reste extrêmement limitée et il était donc primordial d'apporter de nouvelles connaissances sur ce sujet, ce qui permettra de comprendre, que ce soit au niveau fondamentale ou appliqué, le comportement de cette plante vis-à-vis d'un tel stress. Le projet ANR NoStressWall (ANR-12-ADAP-0011; 09/2012 – 08/2015) a été conçu dans le but de d'identifier les réponses de deux espèces, le lin et *Brachypodium distachyon*, à un stress hydrique avec un focus particulier sur le métabolisme pariét*al.* Ce projet, impliquant plusieurs partenaires, prévoyait une approche basée sur l'utilisation de plusieurs technologies de type « omics » (transcriptomiques, protéomiques et métabolomiques) ainsi que des analyses histologiques, physiologiques et biochimiques afin de fournir un vue globale de l'impact d'un stress hydrique.

La première phase de ce projeta concernéla mise en place d'une culture hydroponique du lin (Fig. 1), étape que j'ai réalisée et qui m'a permis de produire du matériel végétal pour les différentes études « omics » prévues au cours de ce projet. En effet, nous avons décidé d'analyser séparément les feuilles, tiges et racines des plantes et le système d'hydroponie nous permettait de récolter rapidement ces dernières. Pour mimer le stress hydrique nous avons utilisé du polyethylene glycol (PEG 6000) car ce produit a été utilisé précédemment dans des études semblables (Munns *et al.*, 2010 ; Guo *et al.*, 2012).

Figure 3 : Culture en hydroponie chez le lin. A gauche le système de culture de plantes de lin avec une couverture empêchant l'entrée de la lumière au niveau racinaire. A droite l'aspect des racines, tiges et feuilles au moment de la récolte des échantillons de lin. (M. Chabi)

Dans le cadre du projet NoStressWall, j'ai effectué les mesures physiologiques, les analyses histologiques ainsi que l'étude transcriptomique, et celles-ci sur les 3 organes végétales: feuilles, tiges et racines, récoltés après 1, 4 et 10 jours de stress (Fig. 2).

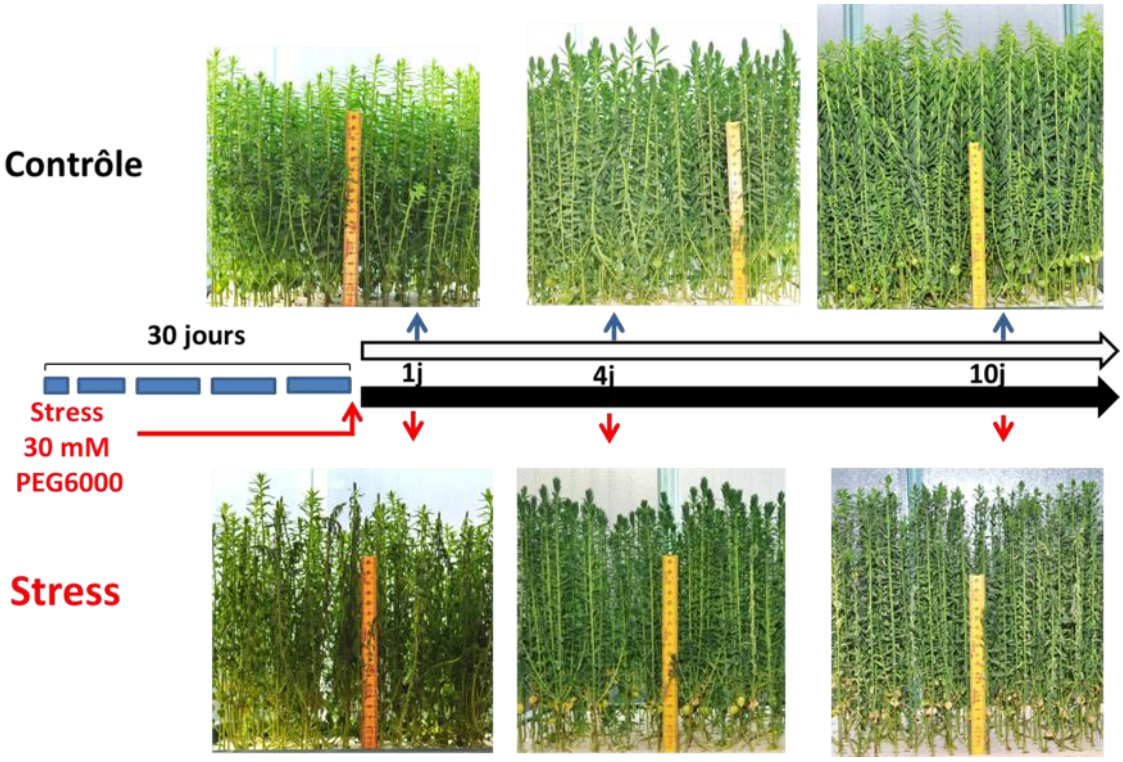


Figure 4 : Aspect des plantes de lin récoltées pour les analyses. Le stress est appliqué sur des plantes âgées de 30 jours par l'addition du PEG 6000 à 30mM au milieu de culture. Les plantes sont récoltées un, quatre et dix jours après l'application du stress.

Ainsi les mesures phénotypiques (voir ci-dessous) montrent une diminution de la croissance des parties aériennes de la plante dès le 4ème jour après l'application du stress. Ce ralentissement est associé aux déformations globales de la structure des feuilles, des racines, et des fibres périphloèmiennes de la tige. Au niveau de l'étude transcriptomique, nous avons observé des

différences significatives (*p* value < 0,01 et log2 ratio <-3 et > +3) dans l'accumulation de transcrits correspondant à 769 gènes.L'analyse globale de ces données est actuellement en cours et nous avons déjà identifié des profils d'expression communs aux 3 organes, ainsi que des profils spécifiques à un seul organe. Des différences significatives sont observées dès le premier jour après l'application du stress et plusieurs catégories fonctionnelles sont représentées dans les profils. Des résultats préliminaires sont présentés ci-dessous et serontintégrés dans une publication globale qui regroupera les résultats des analyses protéomiques (obtenus par le laboratoire MSAP, CNRS USR 3290, Université Lille 1) et les résultats des analyses métabolomiques et biochimiques (obtenus par le laboratoire BioPI, EA3900, Université de Picardie Jules Vernes).

Preliminary Results

Morphological and physiological changes of flax plants in response to drought

Water content (WC) was measured in flax stems and roots during stress (Figure 1) and showed a significant loss in stems after 4 and 10 days of drought. On the other hand, root WC showed no difference between stressed and control plants. Drought stress significantly reduced stem height, but not root length (except for day 10) resulting in an overall increase in the root/stem length ratio (Figure 2). Drought stress resulted in a significant reduction in leaf surface area at day 10, however the efficiency of photosystem II was affected at all 3 time points (Figure 2).

Microscopic analyses indicated that flax leaves undergo a very early disorganisation of palisade mesophyll tissues associated with a drastic decrease of intercellular spaces. A progressive plasmolysis occurs and is visible after 4 days of stress, accompanied by a densification of the cell content. After 10 days of drought stress, leaf cells are less plasmolysed but are characterized by the presence of a condensed nucleus and are more densely packed (1.5 - 2.7 x) than in control leaves (Additional file 1 A). A similar decrease of intercellular spaces is found in stems (Additional file 1 B). Roots seem to be less affected by drought and the general histological organisation is conserved even though the volume of cortex tissue appears to increase in stressed plants at day 10 (Additional file 1 C).

As flax is an industrial plant used for its cellulose-rich bast fibers, we focused on the changes in the morphology of these cells during drought stress (Figure 3). After 1 day of stress fibers

became deformed and then then flattened at D4 and D10 (Figure 3 A), with a larger lumen (Figure 3B) and thinner cell wall (Figure 3C).

Transcriptome analyses of flax plants under drought conditions

A total of 769 unigenes were differentially expressed between PEG-treated plants and control plants ($-3 \leq \log2ratio \geq 3$, p-value ≤ 0.01), at one or more of the three time points (day 1, 4 and 10), and in one or more of the three tissues (roots, stems, leaves) (Additional file 2).

Venn diagrams representing the number of differentially-expressed genes (DEGs) in each tissue after 1, 4 and 10 days of PEG-treatment in flax is represented in the Figure 4A. The results indicate that the different organs show both a common response and an organ-specific response to the stress. While roots show a continued increase in the number of DEGs between day 1 and day 10, the number of DEGS in both leaves and roots decreased between days 1 and 4, before increasing at day 10 (Figure 4b).

Cell wall related DEGs

A preliminary examination of cell wall related DEGs (additional file 2) clearly shows that an important number of transcripts related to cell wall metabolism are differentially accumulated between stressed- and control plants in the 3 different organs. DEGs related to polysaccharide biosynthesis (cellulose: CesA, hemicelluloses: GT47 and pectin: GT8), can be observed, as well as an important number of DEGs involved in cell wall remodeling events (XTHs, expansins, beta-mannosidases, endo-beta-glucanases, beta-glucosidases etc.). Interestingly, only one DEG (C4H: cinnamate-4-hydroxylase) associated with lignin monomer biosynthesis was observed. In contrast there was a differential accumulation of several phenoloxidases including 3 orthologs of Lac17 suggesting that cell wall lignin formation might be affected at the oxidative polymerization step.

Nous continuons à analyser ces données transcriptomiques pour préciser l'impact du stress sur le métabolisme pariétal, mais également sur autres processus physiologiques. Les données seront confronté avec des données protéomiques (MSAP, CNRS USR 3290, Université Lille 1) et métabolomique pour fournir une vision plus globale de l'impact de ce stress chez le lin.

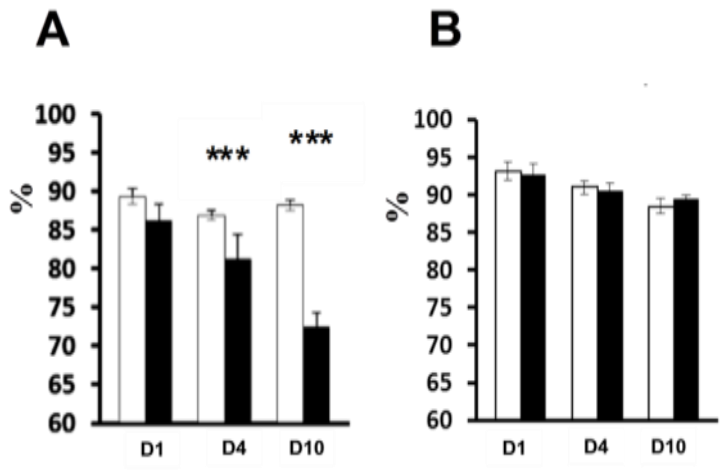


Figure 1. **The water content (WC) in stems (A) and roots (B) of flax plants submitted to a drought stress.**

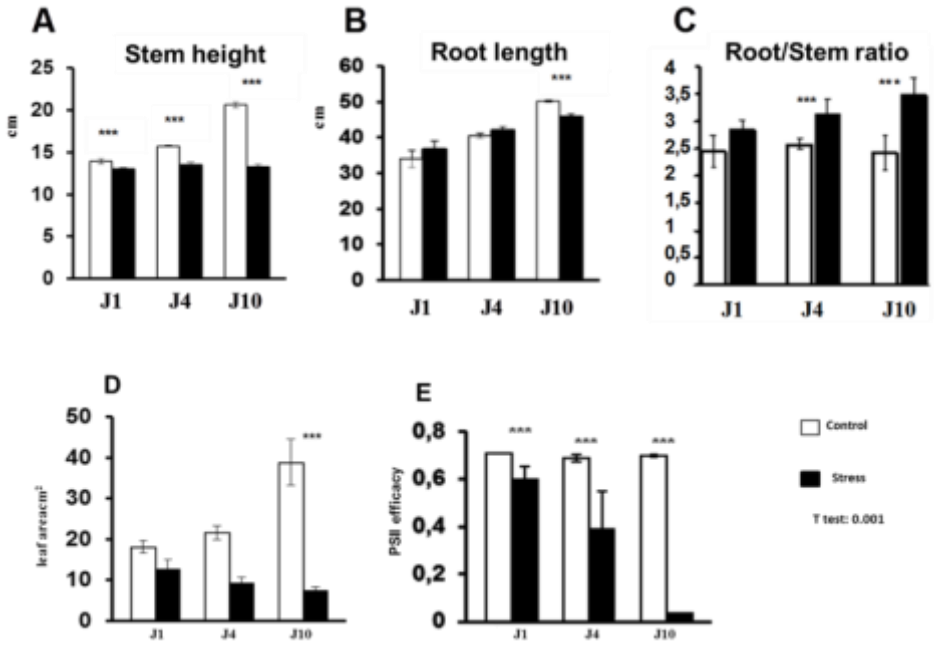


Figure 2. Changes in plant morphology and leaves activity during drought stress in flax plants. A: stem height; B: root length; C: root/stem length ratios; D: Leaf surface; E: PSII efficacy.

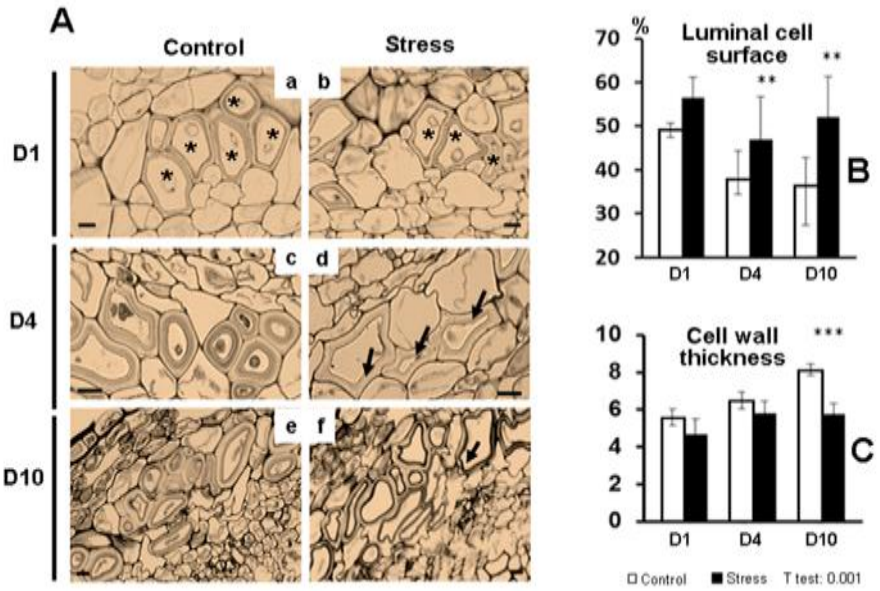


Figure 3. Histological analysis of flax stem fibers. A: microscopic observations in transversal sections; B: Fiber luminal surface (ratio internal cell surface *vs*. external cell surface, %); C: Fiber cell wall thickness (ratio external cell diameter *vs*. 2*internal cell diameter). Phloem fibers (*)became deformed (arrows d, f), with a larger lumen (B) and thinner cell wall (C). a, c, e – controls, b, d, f – PEG stressed plants. D1, D4, D10 = 1, 4, 10 days of PEG treatment. Bar = 20 µm.

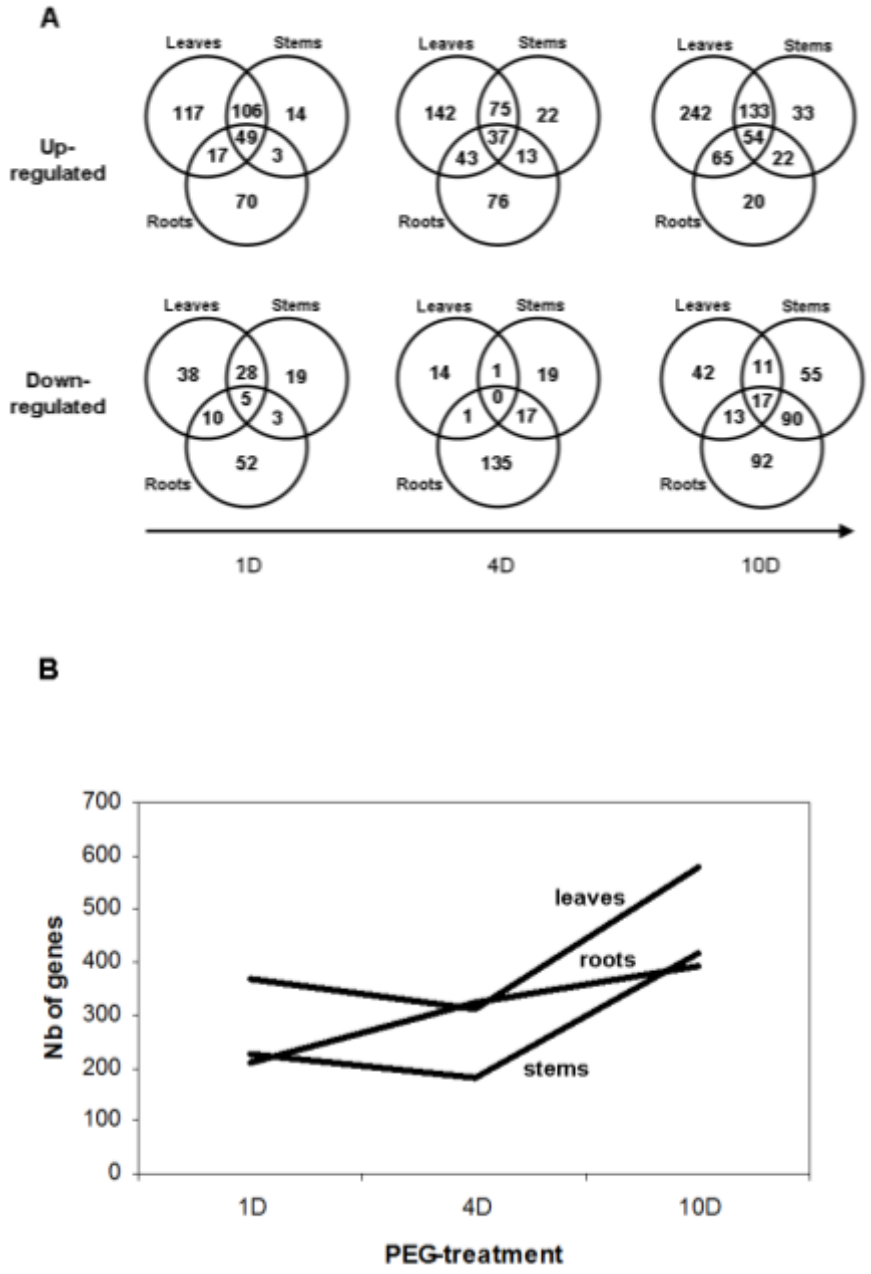


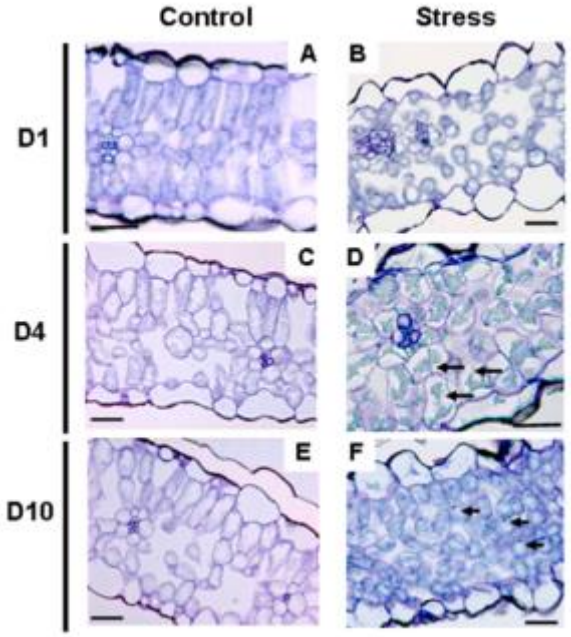
Figure 4. Number of differentially expressed genes (DEGs) during drought-induced stress in leaves, stems and roots of flax plants.

A: Venn diagram (http://bioinformatics.psb.ugent.be/webtools/Venn/) representing the number of up- and down-regulated genes after 1, 4 and 10 days (1D, 4D, 10D) of PEG-treatment. B: Graphical representation of stress-related DEG numbers in leaves, stems and roots.

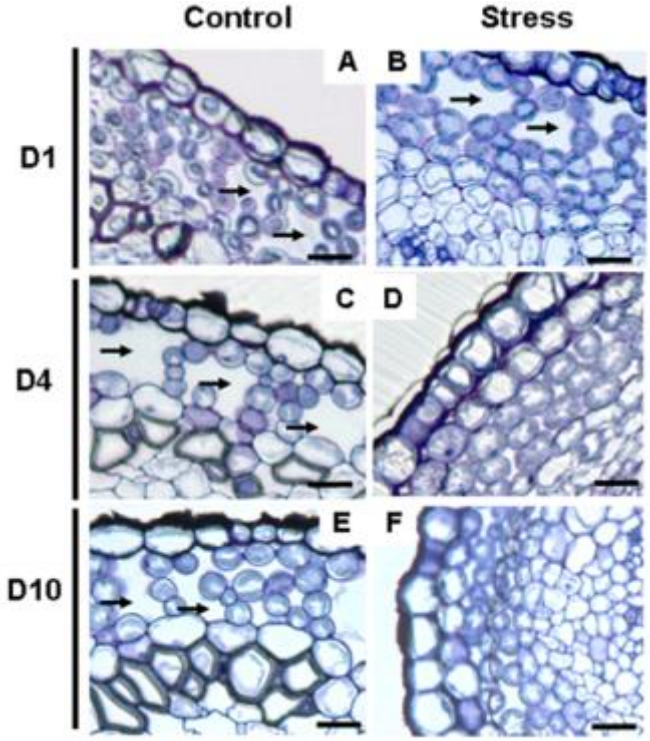
Additional files

Additional file 1. Histological analysis of flax organs under drought.

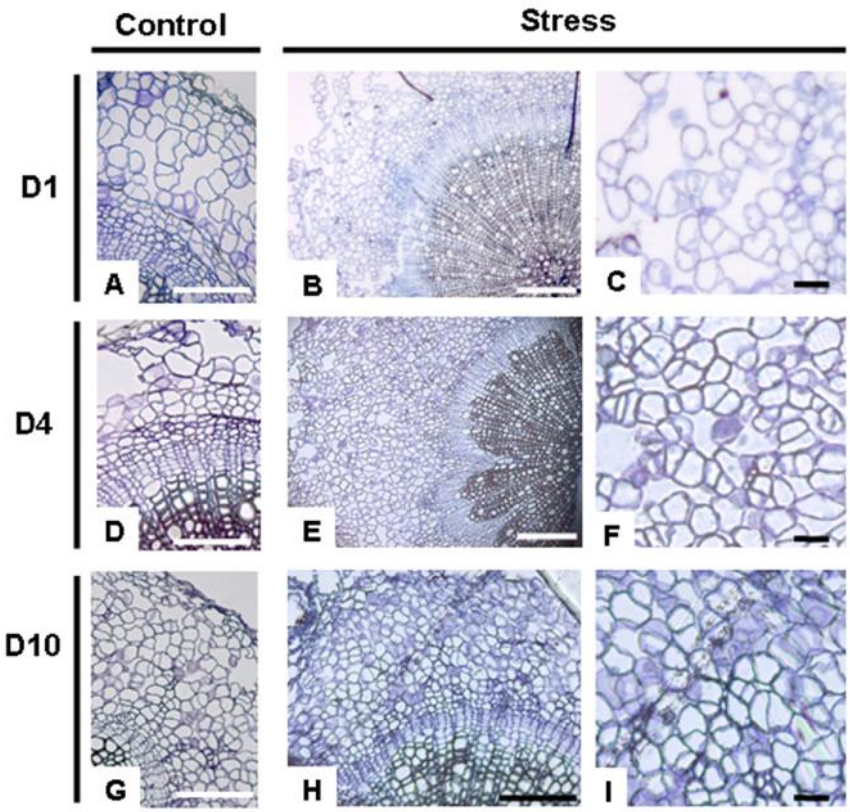
Additional file 1A. Histological analysis of flax leaves. A, C, E – controls, B, D, F – PEG stressed plants. Bar = 30 mm.



Additional file 1B. Histological analysis of flax stems.A, C, E – controls, B, D, F – PEG stressed plants. Bar = 30 mm.



Additional file 1C. Histological analysis of flax roots. A, D, G – controls, B, C, E, F, H, I – PEG stressed plants. Bar = 30 mm.



Additional file 2. List of cell wall related DEGs in flax leaves, stems and roots at 1, 4 and 10 days of PEG-treatment. Log2 ratios were calculated for PEG-treated plants vs. control and significant gene expression results ($-3 \leq log2ratio \geq 3$, p-value ≤ 0.001) were filtered for at least one time course point.

Gene ID		Leaves	Leaves	Leaves	Stems	Stems	Stems	Roots	Roots	Roots	Gene name
Time course of PEG treatment (days)		D1	D4	D10	D1	D4	D10	D1	D4	D10	
genolin_c48486	Lus10007372	3,32	4,16	3,66	0,47	0,45	1,75	-0,27	-0,86	-0,91	sucrose synthase 6
genolin_c30425	Lus10008206	3,04	3,13	1,52	0,00	0,31	1,84	-0,05	-0,17	0,32	sucrose synthase 3
genolin_c30425	Lus10008206	3,04	3,13	1,52	0,00	0,31	1,84	-0,05	-0,17	0,32	sucrose synthase 3
genolin_c1436	Lus10007296	1,99	0,01	0,89	-0,16	-0,50	-3,62	-0,05	-2,29	-4,01	cellulose synthase family protein
genolin_c2228	Lus10029245	1,52	0,21	0,92	-0,24	-0,33	-2,11	-0,28	-2,67	-4,48	cellulose synthase family protein
genolin_c1532	Lus10007296	1,86	0,29	0,52	-0,12	-0,35	-2,65	-0,39	-2,82	-4,99	cellulose synthase family protein
genolin_c15644	Lus10008226	1,78	-0,01	0,63	0,37	-0,43	-2,19	0,21	-1,85	-3,42	cellulose synthase A4
genolin_c12565	Lus10008225	1,12	-0,38	0,60	-0,13	-0,93	-3,17	-0,34	-1,87	-3,52	cellulose synthase A4
genolin_c4754	Lus10034670	1,41	0,52	0,71	-0,53	-0,55	-2,76	-0,78	-1,99	-4,50	COBRA-like extracellular glycosyl-phosphatidyl inositol-anchored protein family
genolin_c41437	Lus10017863	1,11	0,18	0,65	-0,03	-0,50	-2,80	-0,75	-2,01	-3,87	COBRA-like extracellular glycosyl-phosphatidyl inositol-anchored protein family
genolin_c7984	Lus10023968	-0,38	1,71	1,32	-0,36	0,59	-1,52	-1,87	-1,37	-4,56	expansin-like A1
genolin_c20545	Lus10025116	-0,49	1,75	1,92	-0,76	0,56	-1,36	-1,64	-1,64	-3,85	expansin-like A2
genolin_c1226	Lus10024615	-0,41	-0,08	3,58	-0,77	-0,75	-0,56	-0,82	-1,61	-2,25	unknown
genolin_c8840	Lus10033672	-0,13	-0,79	1,76	-0,20	-1,69	-4,55	-0,41	-1,60	-2,63	Leucine-rich repeat (LRR) family protein
genolin_c23152	Lus10001321	1,54	-0,40	-1,01	1,59	1,54	2,67	2,13	1,45	3,71	xylose isomerase family protein
genolin_c43928	Lus10002919	0,44	-1,02	3,88	0,77	-0,52	2,07	2,07	1,47	2,14	Core-2/I-branching beta-1,6-N-acetylglucosaminyltransferase family protein
genolin_c4378	Lus10030484	3,66	4,09	3,36	4,75	2,44	3,85	1,38	0,67	0,84	xyloglucan endotransglycosylase 6
genolin_c7864	Lus10030484	2,54	3,36	3,04	3,73	2,30	3,77	1,41	1,21	0,47	xyloglucan endotransglycosylase 6
genolin_c35302	Lus10013000	-3,34	-0,38	-0,82	-2,34	-0,98	0,32	-3,35	-1,68	-2,82	xyloglucan:xyloglucosyl transferase 33
genolin_c39270	Lus10033755	-2,07	-0,01	1,69	-1,22	-1,62	-1,84	-1,55	-1,92	-3,32	xyloglucan endotransglucosylase/hydrolase 9
genolin_c24685	Lus10031098	0,76	0,45	3,47	-0,06	-1,63	-2,73	-0,90	-3,04	-2,46	O-Glycosyl hydrolases family 17 protein
genolin_c21380	Lus10002930	0,52	-0,04	1,73	0,02	0,49	-1,97	-0,88	-3,07	-0,91	cellulase 3
genolin_c12851	Lus10002930	0,26	-0,58	3,27	-0,11	1,24	-1,82	-0,99	-2,93	-1,29	cellulase 3
genolin_c34877	Lus10029071	-0,78	0,30	2,32	-0,43	-1,30	-2,05	-0,35	-1,77	-4,00	glycosyl hydrolase 9B1
genolin_c3841	Lus10039541	0,12	-0,67	1,87	0,54	0,72	-3,82	-0,60	-1,33	-2,39	beta glucosidase 17
genolin_c26638	Lus10001151	-0,76	-0,53	2,48	-0,49	-0,52	-0,58	-2,21	-2,77	-3,33	Glycosyl hydrolase family protein
genolin_c56327	Lus10029917	0,08	-0,34	4,26	0,32	0,05	-0,27	-0,27	-0,31	0,78	glycoside hydrolase family 2 protein

genolin_c2876	Lus10014184	1,41	1,03	2,15	0,21	-0,49	-3,34	0,07	-1,27	-2,63	Glycosyl hydrolase superfamily protein
genolin_c19367	Lus10007243	0,10	-0,40	3,22	-0,49	-0,50	-2,82	0,18	-1,17	-0,46	glycosyl hydrolase family 10 protein / carbohydrate-binding domain-containing pro
genolin_c13554	Lus10028243	-0,49	0,34	2,65	0,03	-0,57	-3,69	0,17	-0,79	-0,26	glycosyl hydrolase family 10 protein / carbohydrate-binding domain-containing pro
genolin_c11942	Lus10034829	0,61	0,12	4,00	-0,34	-0,77	-1,65	-0,57	-1,04	-1,15	Core-2/I-branching beta-1,6-N-acetylglucosaminyltransferase family protein
genolin_c25080	Lus10019917	1,20	0,57	2,04	-0,44	-1,01	-4,31	-1,02	-2,30	-3,68	Nucleotide-diphospho-sugar transferases superfamily protein
genolin_c3061	Lus10043326	1,02	0,43	2,72	-0,17	-0,93	-3,90	-0,16	-2,31	-2,53	Exostosin family protein
genolin_c59305	Lus10019479	2,12	0,67	3,37	-0,10	-0,74	-3,87	-0,21	-2,22	-2,84	Exostosin family protein
genolin_c25760	Lus10020890	1,24	0,48	3,09	0,22	-1,39	-2,69	0,22	-1,89	-1,49	plant glycogenin-like starch initiation protein 1
genolin_c55644	Lus10033485	-0,13	0,12	2,35	-0,34	-1,25	-3,58	0,48	-1,55	-1,54	plant glycogenin-like starch initiation protein 1
genolin_c32505	Lus10035836	1,62	0,69	3,54	0,70	-0,53	-4,75	-1,65	-2,53	-5,30	galacturonosyltransferase 12
genolin_c6820	Lus10035836	1,81	0,64	1,86	0,03	-0,67	-4,32	-1,19	-2,33	-4,19	galacturonosyltransferase 12
genolin_c7526	Lus10035836	1,52	0,84	1,33	-0,17	-0,80	-4,03	-1,16	-2,30	-4,44	galacturonosyltransferase 12
genolin_c10867	Lus10024583	4,10	2,43	0,62	0,31	0,50	1,18	0,85	0,45	0,29	UDP-glucosyl transferase 85A3
genolin_c25686	Lus10024583	3,74	2,51	0,66	0,54	1,09	1,37	0,84	0,74	0,77	UDP-glucosyl transferase 85A3
genolin_c13889	Lus10039277	5,89	5,12	4,41	3,15	2,04	3,87	1,62	0,86	0,86	UDP-Glycosyltransferase superfamily protein
genolin_c10886	Lus10039277	5,15	4,79	3,82	1,82	1,62	3,90	1,07	0,64	0,43	UDP-Glycosyltransferase superfamily protein
genolin_c49925	Lus10039277	6,02	4,88	2,59	1,69	2,04	3,04	1,69	1,03	0,82	UDP-Glycosyltransferase superfamily protein
genolin_c41285	Lus10039277	6,09	3,82	2,15	2,28	1,28	2,57	1,06	0,72	1,78	UDP-Glycosyltransferase superfamily protein
genolin_c29976	Lus10031388	0,49	0,19	-0,20	1,23	1,35	3,68	3,10	4,48	0,32	UDP-Glycosyltransferase superfamily protein
genolin_c32116	Lus10031388	-0,06	0,32	0,93	0,18	0,76	2,93	2,10	3,32	0,14	UDP-Glycosyltransferase superfamily protein
genolin_c34700	Lus10011885	-1,27	-0,93	-0,70	-0,47	-3,10	-2,26	-0,36	-1,02	0,23	Pectin lyase-like superfamily protein
genolin_c22007	Lus10033305	1,36	2,50	3,20	0,56	-1,36	-1,95	-0,73	-1,42	-2,78	Pectinacetylesterase family protein
genolin_c36319	Lus10039314	-0,66	0,10	5,37	-2,72	-1,89	-2,79	-1,27	-1,54	-0,82	pectin methylesterase 3
genolin_c37408	Lus10043035	-0,84	-0,42	3,98	0,05	-0,20	0,78	-0,05	-1,98	-3,81	Pectin lyase-like superfamily protein
genolin_c50322	Lus10002124	0,64	-0,11	5,71	-0,14	0,35	-0,26	-0,67	-0,32	0,80	Pectin lyase-like superfamily protein
genolin_c36377	Lus10007922	-0,77	0,08	1,15	-0,60	-1,46	-0,81	-0,06	-1,27	-3,26	Pectin lyase-like superfamily protein
genolin_c11945	Lus10030886	-0,99	-0,33	3,33	-0,34	-0,55	-0,70	0,09	-1,00	-4,72	Pectin lyase-like superfamily protein
genolin_c50219	Lus10007922	-0,41	0,47	1,91	-0,40	-1,33	-0,34	-0,19	-1,40	-3,17	Pectin lyase-like superfamily protein
genolin_c1115	Lus10008646	-0,92	0,73	0,92	-0,78	-0,67	-1,21	-1,04	-1,90	-3,38	Nucleotide-diphospho-sugar transferases superfamily protein
genolin_c40683	Lus10010024	-0,91	0,25	0,95	-0,89	-0,71	-1,69	-1,42	-3,22	-1,43	cellulose synthase-like D5
genolin_c10515	Lus10015625	2,31	0,13	0,35	1,52	0,24	2,64	1,52	2,97	4,74	beta galactosidase 1
genolin_c26437	Lus10037644	2,15	-0,25	-0,84	1,74	0,49	2,69	2,03	3,30	5,12	beta galactosidase 1

genolin_c21480	Lus10015625	2,70	0,12	0,55	1,66	0,11	2,51	1,57	3,26	5,25	beta galactosidase 1
genolin_c51183	Lus10002984	1,19	0,98	5,46	-0,22	0,13	-1,05	0,90	0,76	0,61	FASCICLIN-like arabinogalactan-protein 11
genolin_c2859	Lus10036113	1,27	0,95	4,09	-0,22	-0,06	-1,00	0,65	0,46	0,77	FASCICLIN-like arabinogalactan-protein 11
genolin_c57242	Lus10036111	0,18	-1,03	2,58	-0,87	-1,86	-4,11	-0,70	-2,88	-3,13	FASCICLIN-like arabinogalactan-protein 11
genolin_c3396	Lus10019929	0,78	0,20	1,49	-0,59	-0,87	-3,71	-0,64	-2,27	-3,44	FASCICLIN-like arabinogalactan-protein 11
genolin_c13982	Lus10012351	-0,17	0,93	2,61	-0,77	-0,28	-3,04	-1,58	-2,03	-2,50	FASCICLIN-like arabinoogalactan 7
genolin_c19397	Lus10029353	-0,44	-0,83	1,49	-0,34	-0,30	-1,26	-1,57	-1,40	-4,17	FASCICLIN-like arabinogalactan protein 17 precursor
genolin_c11608	Lus10006399	0,39	0,62	2,48	-0,48	-0,49	-3,19	-1,45	-1,93	-2,94	FASCICLIN-like arabinoogalactan 7
genolin_c1905	Lus10025231	1,28	0,82	3,34	3,31	0,71	3,99	0,99	2,14	1,62	lipid transfer protein 3
genolin_c39078	Lus10025230	0,88	2,34	4,26	1,62	0,09	1,64	-0,22	1,30	0,63	lipid transfer protein 6
genolin_c353	Lus10025234	4,74	3,92	3,88	1,26	0,53	-0,15	-0,21	-0,20	-5,70	lipid transfer protein 2
genolin_c10558	Lus10025234	3,74	3,14	3,47	1,13	0,32	-0,17	0,21	-0,06	-3,83	lipid transfer protein 2
genolin_c39726	Lus10015278	0,06	-0,17	0,07	-0,18	-0,10	-0,70	-1,76	-2,35	-2,98	lipid transfer protein 4
genolin_c52667	Lus10041830	0,14	1,08	1,75	1,94	1,04	3,57	0,32	-0,45	-0,41	basic chitinase
genolin_c29349	Lus10040420	2,30	2,00	4,10	1,17	0,38	0,87	0,49	0,03	0,69	chitinase A
genolin_c438	Lus10010861	1,97	2,46	5,22	1,61	-0,08	0,90	0,04	0,20	0,61	Chitinase family protein
genolin_c53940	Lus10000193	0,74	1,33	2,31	2,66	1,51	4,25	0,52	-0,19	-0,25	basic chitinase
genolin_c2271	Lus10006691	-2,85	-0,99	-3,94	-0,91	0,82	-1,88	-2,08	-2,79	-3,65	O-methyltransferase family protein
genolin_c11079	Lus10027598	-0,82	0,92	3,72	0,01	-0,34	-2,31	0,73	-0,05	-1,86	cinnamate-4-hydroxylase
genolin_c3993	Lus10007599	0,85	1,97	4,40	-0,23	-0,77	-2,43	-0,01	0,35	-2,12	Domain of unknown function (DUF1995)
genolin_c1251	Lus10012143	0,24	-0,08	2,56	-0,55	-1,01	-2,90	-0,70	-1,12	-3,14	pinoresinol reductase 1
genolin_c14103	Lus10007599	-0,56	1,24	3,41	0,04	-0,16	-1,49	-0,68	0,65	-2,15	Domain of unknown function (DUF1995)
genolin_c6075	Lus10041481	0,60	-0,04	3,55	0,12	-0,96	-4,12	-0,57	-2,69	-2,49	Laccase/Diphenol oxidase family protein
genolin_c3323	Lus10028263	1,91	0,78	2,96	0,31	0,35	-5,04	0,36	-1,34	-2,11	Laccase/Diphenol oxidase family protein
genolin_c25502	Lus10041481	0,52	0,24	2,20	-0,02	-1,19	-4,82	-0,63	-3,37	-4,10	Laccase/Diphenol oxidase family protein
genolin_c8754	Lus10010850	0,83	-0,73	2,95	-0,39	-1,59	-4,04	-0,10	-2,89	-1,60	laccase 17
genolin_c3505	Lus10035517	0,67	-0,48	2,61	-0,43	-1,40	-3,47	0,09	-2,32	-1,91	Laccase/Diphenol oxidase family protein
genolin_c17032	Lus10018208	1,96	0,24	3,65	0,43	-1,35	-2,30	0,44	-1,29	-1,15	Laccase/Diphenol oxidase family protein
genolin_c18244	Lus10019122	0,56	0,64	2,64	0,79	-1,06	-3,26	0,59	-1,42	-1,04	laccase 17
genolin_c14063	Lus10010850	0,79	-0,48	1,74	-0,11	-1,49	-3,75	0,07	-1,91	-1,12	laccase 17
genolin_c16680	Lus10032894	0,52	-0,19	1,17	0,15	-0,97	-3,63	-0,16	-1,13	-2,44	Laccase/Diphenol oxidase family protein
genolin_c53315	Lus10017426	0,41	0,06	4,69	-0,79	-0,85	-2,02	0,15	-1,73	-1,37	laccase 5
genolin_c41672	Lus10021133	0,39	0,60	4,63	-0,30	-0,16	-1,68	-1,22	-1,42	-1,80	Laccase/Diphenol oxidase family protein
genolin_c5106	Lus10035517	0,65	-0,43	3,26	-1,00	-1,31	-2,53	-0,86	-1,80	-1,32	Laccase/Diphenol oxidase family protein

genolin_c9414	Lus10017426	1,57	0,77	4,21	-0,96	-0,93	-3,12	0,08	-1,28	-1,58	laccase 5
genolin_c19304	Lus10024121	-2,38	0,16	4,53	0,43	0,20	0,57	-1,39	-1,89	0,38	unknown
genolin_c6397	Lus10024121	-1,77	0,18	4,28	0,04	0,04	-0,33	-2,17	-2,65	-0,24	unknown
genolin_c12961	Lus10005393	-0,71	0,05	3,72	-0,06	-0,39	-0,35	-2,43	-2,36	-0,41	unknown
genolin_c6899	Lus10024121	-1,67	0,81	3,18	0,00	0,19	-0,02	-1,52	-2,20	0,08	unknown
genolin_c54381	Lus10009898	0,49	1,48	4,74	-0,06	0,14	1,24	-0,80	0,49	-3,00	Peroxidase superfamily protein
genolin_c39688	Lus10001442	3,76	3,93	4,84	0,38	0,65	1,60	-0,47	-0,77	-1,19	Peroxidase family protein
genolin_c27447	Lus10007638	0,01	-0,27	4,47	0,28	-0,62	-0,36	-1,21	1,23	0,22	Peroxidase superfamily protein
genolin_c8134	Lus10018374	-0,36	0,29	5,66	0,21	-0,01	-0,47	-0,82	0,51	0,06	Peroxidase superfamily protein
genolin_c59350	Lus10005614	-0,50	-0,07	3,69	-0,78	-1,12	-4,69	-0,24	-0,05	-0,93	Peroxidase superfamily protein
genolin_c6312	Lus10017288	0,85	0,01	3,06	-0,42	-1,51	-4,77	-0,29	-0,84	-1,56	Peroxidase superfamily protein
genolin_c30365	Lus10032926	2,10	2,11	4,73	-1,46	-0,14	-1,26	-0,53	-0,70	-1,06	Peroxidase superfamily protein
genolin_c13126	Lus10015127	1,83	2,47	1,77	-0,72	-1,39	-3,47	-1,05	-0,74	-1,04	Peroxidase superfamily protein

Conclusions et perspectives

Le lin est une espèce qui présente un intérêt économique tout autant pour ses graines oléagineuses riches en omega 3 et pour ses fibres cellulosiques périphloèmiennes. Ces fibres naturelles sont utilisées depuis longtempsdans l'industrie textile et elles intéressent de plus en plus les secteurs du bâtiment et de transport pour la production de matériaux biocomposites. Ces différentes industries ont des exigences variées en termes de la « qualité » de fibres. Bien qu'il soit difficile de trouver une simple définition de ce concept, il est généralement accepté que la « qualité » de fibres dépend, en partie, de ses propriétés mécaniqueset chimiques conditionnées elles-mêmes par la structure et l'architecture de la paroi cellulaire. Les industries textiles recherchent par exemple des fibres les plus fines, qui soient les plus longues possible, avec une très faible teneur en lignine, conférant ainsi une flexibilité importante (Foster et *al.*, 1997), tandis que les industries des matériaux composites vont,à l'inverse, rechercher des fibres avec plus de rigidité et une interaction compatible avec les différentes matrices hydrophobes utilisées dans ce domaine.

Du fait de ce rapport intime entre la qualité de fibres de lin et la paroi cellulaire, une valorisation optimalede cette espèce nécessite une meilleure compréhension des mécanismes fondamentaux responsables de la mise en place de la paroi de ses fibres périphloèmiennes. Des recherches spécifiques sont d'autant plus nécessaires que la structure de l'épaisse paroi secondaire est atypique avec une hypolignification importante et des teneurs élevées en cellulose cristalline.

Au sein de notre laboratoire plusieurs travaux ont déjà étaient menés avec l'objectif d'améliorer nos connaissances de mécanismes responsables de la formation de la parois des fibres de lin (Huis et *al.*, 2012), (Chantreau et *al.*, 2014b)(Chantreau et *al.*, 2015). Mes travaux de thèse se situent dans cette démarche globale en continuant à mettre en place des approches haut débit (transcriptomique, protéomique) pour mieux comprendre les origines de la variabilité de la structure – et donc « qualité » des fibres de lin.

Projet 1 : Développement des outils de transcriptomique (comparaison et validation des microarrays NimbleGen 60 mers)

Les nouvelles technologies de transcriptomique de type « génome entier » (whole genome transcriptomics) ont joué un rôle majeur dans les travaux de notre laboratoire visant une

meilleure compréhension de mécanismes responsables de la mise en place de la paroi cellulaire des fibres de lin. C'est dans le cadre du projet ANR GENOLIN que notre laboratoire a contribué au développement du premier support transcriptomique – des microarrays « NimbleGen » (société Roche NimbleGen) – couvrant le « génome entier » du lin (Fenart et *al.*, 2010) . Le génome du lin n'étant pas disponible avant 2012 (Wang et *al.*, 2012), ces premiers microarrays étaient basées sur un design à partir des ESTs et une assemblage de séquences courtes obtenues par la technologie 454 (Fenart et *al.*, 2010). Les puces étaient basées dans un premier temps sur des sondes avec une longueur d'environ 25 oligonucléotides (25-mer). Par la suite, un deuxième design a été conçu par Biogemma (partenaire du consortium GENOLIN) sur la base de sondes plus longuesavec une taille d'environ60 nucléotides (60 mer). Cette modification technique a été décidée dans le but d'obtenir des sondes plus spécifiques. La validation de ces microarrays a fait l'objet de la publication présentée dans le premier chapitre de ce mémoire (Fenart et *al.*, 2013). Ce travail a permis de valider le passage des sondes 25 à 60 mer, et d'utiliser des microarrays NimbleGen basés sur les sondes longues dans le cadre du projet ANR NoStressWall.

Cependant, et malgré les résultats intéressants que nous avons pu obtenir avec les microarrays NimbleGen, la société Roche a décidé abruptement d'arrêter la commercialisation de ses microarrays en 2012 face à l'essor d'autres technologies (RNAseq). Confronté à cette situation le laboratoire a décidé, en attendant de mettre en place une approche de RNAseq, de continuer à favoriser une technologie de microarray en utilisant un support Agilent. Ces puces sont largement répandues dans le domaine des études transcriptomique, mais si le principe général reste le même, des différences techniques existent (différence de température d'hybridation, marquage des ARN et non pas des ADNc, nombre de sondes par gène…). Autre différence majeure est que le design des microarrays Agilent est basé sur le génome du lin (Wang et *al.*, 2012) et non pas des ESTs.

Ce dont il faut tenir compte aussi, est que les technologies utilisées pour le transcriptomique évoluent très rapidement, nécessitant un effort important d'adaptation par les laboratoires. Actuellement la technologie majeure qui a émergé ces dernières années et qui devient la norme (pour le moment) est le RNAseq. Cette technique fait partie des séquençages de seconde génération, de plus en plus utilisée, au profit des puces à ADN classique.

En effet cette technologie offre quelques avantages par rapport aux puces à ADN classique (spécificité, sensibilité et précision plus importantes que dans le cas de puce à ADN). Autre avantage pour la technique de RNAseq est que cette technologie ne restreint pas le profil des

données d'expression à un contenu et à une annotation spécifique. Malgré ces avantages, le RNAseq présente quelques inconvénients par rapport aux puces à ADN (accès à un séquenceur, compétences en bio-informatique). En comparaison, les puces à ADN restent relativement moins couteûses et plus faciles d'utilisation. Cependant, l'équipe est en train d'acquérir les compétences en bio-informatique nécessaires dans le cadre d'une autre thèse et les séquençages seront réalisés par prestation pour les études futures de transcriptomique.

Projet 2 : « Proteome Atlas » de lin

Dans le contexte du développement des outils, bien que la transcriptomique nous permet d'analyser l'expression des gènes, les processus de régulation post-transcriptionnelle ne sont pas négligeables et il est nécessaire d'aller au niveau protéomique pour s'assurer de la présence réelle du produit du gène. En effet, entre le transcrit d'un gène et l'action finale de celui-ci, il existe plusieurs étapes, un transcrit peut être abondamment présent mais peut être dégradé par la suite induisant des biais dans les résultats et des approximations dans l'analyse scientifique. Ce qui veut dire que techniquement une étude transcriptomique nous permet d'orienter notre recherche sur certaines voies mais pour prouver une hypothèse il faut investir en utilisant soit d'autres échelles, soit d'autres méthodes. Ainsi, il nous a semblé indispensable de développer une méthode complémentaire, et nous nous sommes tournés vers le protéomique. Dans ce contexte une première publication a été réalisé au sein de notre laboratoire (Day et *al.*, 2013). Cette étude s'est focalisée sur la tige entière du lin, et a porté un intérêt particulier aux protéines pariétales. Cependant, la tige de lin contient 2 populations de cellules ayant des structures pariétales fortement contrastées (parois secondaires du xylème fortement lignifiées *vs* parois cellulosiques des fibres très faiblement lignifiées). Pour aller plus loin dans ce domaine nous avons opté pour la réalisation d'un « atlas des protéines chez le lin ». En effet de cette manière nous avons pu identifier des protéines présentes dans les différents organes de la plante ; feuilles, racine et tige (tissus internes et tissus externes). Du fait que le lin est également cultivé pour ses graines, nous avons décidé de récupérer des données protéomiques sur cet organe et disponibles dans le domaine publique (Barvkar et *al.*, 2012) pour créer une ressource qui sera utile pour tous les scientifiques travaillant sur cette espèce.

Toutefois, les résultats obtenus restent qualitatifs et afin d'affiner les informations obtenues, une approche de protéomique quantitative serait une amélioration non négligeable de ce travail réalisé. En effet, la différence entre organes peut ne pas être aussi extrême, c'est-à-dire, de l'ordre de la présence ou de l'absence d'une protéine, de plus les différences bien que petites au niveau de la régulation peuvent aboutir à des variations majeurs au niveau du phénotype. Une

telle précision dans la quantité de protéine pourrait ainsinous aider à mieux comprendre, pas seulement les différences morphologiques entre les organes, mais également l'hétérogénéité de la paroi cellulaire.

Le cycle de vie de la plante du lin consiste en cinq stades de développement : la levée, le stade végétatif, la floraison, le stade capsules vertes, et enfin le stade capsules brunes.

Les travaux effectués sur les protéines du lin et qui représente le chapitre 2 de mes travaux de thèse, ont été réalisés sur des plantes âgées de 70 jours, ce qui correspond au stade végétatif le plus tardif juste avant l'apparition des premiers bourgeons floraux. Afin d'enrichir ce travail, il serait intéressant de réaliser une étude protéomique sur les 5 principaux stades de développement précédemment cités, ou bien encore, sur d'autres organes tels que les bourgeons floraux. Les précédentes études transcriptomique et histologique réalisée chez le lin (Ageeva et *al.*, 2005; Huis et *al.*, 2012) ont montré qu'il existe une variabilité importante entre les différentes régions de la tige de lincorrespondant à un gradient de développement. En effet, la tige peut être séparée en région apicale (la plus « jeune »), en région basale (la plus « matures ») et en région médiane (développement intermédiaire). Le développement de fibres de lin suit cette progression – les parois des fibres élémentaires en région apicale ne sont pas remplies de manière homogène, mais se remplissent au fur et à mesure qu'on descend le long de la tige. Ainsi, dans le contexte d'une étude plus approfondie,il serait d'un grand intérêt d'appliquer une approche de protéomique pour disséquer ce dynamisme pariétale.

Bien que nous avons pu identifier 1242 protéines non-redondantes, plusieurs spectres de « bonne qualité » n'étaient pas assignées à une protéine. Ceci peut être expliqué parl'existence de protéinesnon recensées dans la première version du génome.Une autre raison est peut-être liée au fait que le génome de référence utilisé est celui de CDC Bethune (variété canadienne, huile printemps) tandis que la variété utilisée pour l'étude protéomique était Diane (une variété fibre printemps). Ainsi des différences induites par des SNP pourraient aussi être à l'origine d'une non-identification de protéine. Il serait nécessaire de construire et utiliser une base de données plus élargie comme il a été faitavec le protéome de l'homme. Finalement la protéomique peut être une bonne méthode qui peut non seulement corriger, mais également compléter les annotations des génomes.

Par ailleurs pour comprendre la composition et la structure complexe de la paroi cellulaire, une des méthodes « non-destructives » qui est utilisée pour la visualisation et la localisation de ses polymères,est l'immunolocalisation. Nous avons utilisé plusieurs anticorps pour caractériser les

hémicelluloses des parois de fibres et mis en évidence la présence des xyloglucanes et, en moindre quantité, des xylanes, pas seulement au niveau de la paroi primaire, mais très probablement au niveau de la couche S1 de la paroi secondaire. Une approche de microscopie électronique permettra de préciser ce point (Gorshkova et *al.*, 2015).
L'utilisation de la technique d'empreinte enzymatique a permis de mettre en évidence plusieurs motifs d'oligo-xyloglucanes/xylanes dans la fraction enrichie en hémicelluloses. Il serait intéressant d'utiliser la même approche pour analyser la fraction pectique de la paroi des tissus internes et externes de la tige de lin.

Pour ce travail nous avons aussi réalisé une étude phylogénétique à grand échelle, c'est-à-dire, que pour chacune des protéines pariétales identifiées, nous avons réalisé sa phylogénie en utilisant 21 autres espèces ayant un génome complet ; de la plus proche espèce du lin (*P. trichocarpa*, *M. truncatula*, *R. communis*), mais aussi avec des groupes plus externes comme des monocotylédones (*B. distachyon*), des bryophytes (*P. patens*)et même des algues vertes unicellulaires (*C. reinhardtii*, *O. lucimarinus*). Ceci nous a permis d'apporter une vision plus précise de certaines fonctions et de leurs évolutions spécifiques chez le lin comparé aux autres espèces. D'une manière très intéressante nous avons constaté que pour certaines protéinespariétales, nous retrouvons des ultra-paralogies (Jourda et *al.* 2014) chez le lin par rapport aux autres espèces.C'est le cas, par exemple, des chitinases (GH18),des expansines (GH45) et des xyloglucan endo-transglycosylases/hydrolases (GH16). La paralogie des chitinases a déjà été reportée et été expliqué par la duplication du génome du lin lors de son évolution (Sviensson et *al.*, 2013 ; Mokshina et *al.*, 2014).Il est possible que cette paralogie soit en rapport avec la structure pariétale particulière du lin et mérite plus d'attention.Autre point positif de l'approche phylogénétique est qu'elle peut permettre une annotation précise des protéines identifiées. En effet pour chacune des arbres phylogénétiques nous avons généré une image correspondant au domaine fonctionnel des protéines ce qui permet de détecter des protéines tronqués, ou avec différents domaines pouvant induire différentes fonctions. L'utilisation de la banque de mutants EMS du lin, ou bien l'exploitation du système VIGS, dans une démarche de génétique inverse permettra d'aborder une approche fonctionnelle pour évaluer le(s) rôle(s) des différentes protéines mise en évidence dans l'étude protéomique (Chantreau et *al.*, 2013 ; 2015).

Projet 3 : Mise en évidence de la variabilité génotypique associée aux différences de métabolisme pariétal chez le lin

Depuis environ une vingtaine d'années la culture du lin a repris un nouvel élan dans le monde économique, et ce grâce à l'émergence de nouveaux secteurs de valorisation qui sont les biocomposites (Brutch et *al.*, 2008). En effet, le lin classiquement cultivé pour l'industrie textile, se voit propulsé dans un autre secteur potentiellement plus porteur mais qui a des exigences différentes en termes de matière première, de celles recherchées dans l'industrie textile. Ce dédoublement de l'intérêt économique pour le lin associé à une demande de « qualité » différente, pousse à reconsidérer les origines de la variabilité de la structure de la fibre de lin et, à terme, la question de la sélection variétale. De façon simple, les différentes variétés de lin sur le marché se classent dans 2 groupes principaux : les variétés à fibres – cultivées pour les fibres cellulosiques à usage textile et les variétés à huile – cultivées pour les graines et les huiles. A ces 2 groupes viennent s'ajouter les variétés « printemps », et les variétés de « hiver », ce dernier permettant de récolter les fibres/graines plus tôt.

Plusieurs études (Brutch et *al.*, 2008; Diederichsen and Ulrich, 2009; Brutch et *al.*, 2011) ont été réalisées sur différentes variétés de lin pour établir les liens qui existent entre les différents caractères de la plante, afin de faciliter les croisements et la création de nouvelles variétés. Plusieurs caractères d'intérêt économique tels quela qualité des fibres (flexibilité, force, longueur), la quantité des fibres produites et la résistance aux maladies dépendent d'un déterminisme génétique, et ce malgré certaines études préalables qui rapportent la faible diversité génétique du lin (Doré and Varoquaux, 2006). Des études récentes réalisées avec des technologies « haut débit » (« omics ») chez autres espèces ont permis d'avancer les connaissances concernant les rapport entre la diversité génétique et la variabilité de la résistance aux différents stress (a)biotiques (Rakszegi et *al.*, 2014; Chandna and Ahmad, 2015; Fraige et *al.*, 2015; Guo et *al.*, 2015). De même ces approches ont permis d'établir des liens entre la diversité génétique et la biomasse/composition pariétale (Souza et *al.*, 2015).

Avec le projet FIBRAGEN (*Flax for Improved Biomaterials thRough Applied GENomics* - KBBE 2011-2013) auquel j'ai pu participer, nous avons démontré qu'il existe des différences significatives entre plusieurs variétés de lin au niveau de l'expression des gènes et de l'abondance de protéines. Ce projet a eu comme objectif ledéveloppementde connaissances biologiques et techniques nécessaires pour améliorer et/ou sélectionner des génotypes de lin optimisés pour l'utilisation dans l'industrie des biomatériaux. En ce qui concerne les travaux de doctorat réalisés dans le contexte de ce projet nous pouvons souligner des avancées réalisées sur 2 volets majeurs: 1) l'analyse à grande échelle (« omics ») des profils d'expression des gènes (7 variétés) et des protéines (2 variétés) ; 2) la mise en évidence d'un rôle potentiel de la protéine

XTH (xyloglucan endotransglycosylase/hydrolase) dans la structuration de la paroi de la fibre de lin.

Concernant le premier point, les analyses transcriptomiques ont permis de mettre en évidence un total de 659 gènes exprimés différentiellement entre les variétés dont un certain nombre impliqué dans le métabolisme pariétale. De façon intéressante, l'approche protéomique nous a permis de détecter plusieurs peptides de bonne qualité mais qui n'étaient pas associés à une protéine prédite dans le génome du lin (Wang et *al.*, 2012). Une recherche approfondie a pu associer ces peptides à une expansine permettant ainsi d'améliorer l'annotation du génome et soulignant l'intérêt de l'approche protéomique. Pour complémenter ces analyses « omics », des études de la composition (chimie, FT-IR spectroscopie), ainsi que des tests micromécaniques (élasticité, résistance à la traction, etc.) sur les faisceaux de fibres issus de mêmes variétés de lin ont été réalisés par autres partenaires du consortium Fibragen (UMR FARE, INRA Reims et l'Université de Brême). Le défi majeur est maintenant d'essayer de croiser les données issues de ces analyses et, à terme de les confronter avec les résultats réalises sur des matériaux composites des tests faits par les partenaires du projet (UMR FARE, INRA Reims, l'Université de Brême et l'institut technique AIMPLAS).

Les approches transcriptomique et protéomique ont mis en évidence la présence de transcrits/protéines correspondant à 3 XTHs (xyloglucan endo-transglycosylase hydrolases). Les XTHs font partie de la famille CAZy GH16 et sont proposées de jouer un rôle important dans les processus de l'expansion cellulaire via leur action sur les xyloglucanes pariétaux. Etant donné la longueur très importante de la fibre de lin – ce qui implique une expansion cellulaire conséquente, ainsi que sa similarité à une « fibre G » du bois de tension et dans laquelle les xyloglucanes sont proposés de jouer un rôle structurant, il n'est pas illogique de penser que les XTHs pourraient aussi représenter un acteur majeur du métabolisme pariétal chez le lin. Cette hypothèse est également appuyée par l'étude phylogénétique montrant une duplication importante au sein de la famille XTH IIIA.

Pour vérifier cette hypothèse il serait intéressantd'entreprendre la purification de la protéine et/ou la production de protéines recombinantes pour tester l'activité des XTH de lin. Des approches d'enzymologie (pH, température, Km, Vmax) pour mieux caractériser les protéines et leurs comportements vis à vis de xyloglucanes apporteront des informations clés. En complément de l'analyse enzymologique, l'hybridation *in situ* et/ou l'immunolocalisation seraient souhaitables pour localiser les transcrits/protéines des XTHs au niveau des

tissus/cellules. Dans ce cas, il serait également intéressant de comparer la localisation tissulaire/cellulaire avec les résultats de tests *in situ* de l'activité XTH. Concernant les approches *in situ*, il serait intéressant de réaliser des sections longitudinales en plus des sections transversales plus classiques. Comme évoqué plus haut, l'utilisation de mutants EMS, une approche VIGS ou transformation génétique permettra une caractérisation fonctionnelle.

Le projet FIBRAGENdont le but était de mettre en avant des différences éventuelles au niveau des différentes variétés, a égalementdémontré l'effet de l'environnement surles expressions de gènes chez les différentes variétés. En effet, le fait d'avoir réalisé les analyses de manière successive sur deux années différentes, nous a permis de révéler le comportement des variétés étudiées dans des conditions climatiques différentes.Ceci permet d'avoir une image de la stabilité/labilité de ces variétés en réponse aux conditions environnementales. Dans le projet ANR NoStressWall, dont le but était d'analyser l'impact d'un stress hydrique sur le lin avec un focus particulier sur le métabolisme pariétal, nous avons utilisé la variété à fibres printemps « Diane ». Il serait intéressant d'élargir ce type d'étude à d'autres variétés de lin. Si dans le projet FIBRAGEN les expériences ont été menées en champ (donc en conditions environnementales peu contrôlées), le projet sur le stress hydrique a été réalisé en conditions contrôlées. Une analyse approfondie des gènes differentiellement exprimés (ou des métabolismes pariétaux concernés) dans ces deux projets apportera des informations nouvelles sur la contribution respective des facteurs environnementaux et génétiques à la structure, et donc « qualité » de la fibre de lin.

Bibliographie

Adolphe JL, Whiting SJ, Juurlink BHJ, Thorpe LU, Alcorn J (2009) Health effects with consumption of the flax lignan secoisolariciresinol diglucoside. Br J Nutr **103**: 929

Agarwal UP (2006) Raman imaging to investigate ultrastructure and composition of plant cell walls: Distribution of lignin and cellulose in black spruce wood (Picea mariana). Planta **224**: 1141–1153

Ageeva M V., Petrovská B, Kieft H, Sal'nikov V V., Snegireva a. V., Van Dam JEG, Van Veenendaal WLH, Emons a. MC, Gorshkova T a., Van Lammeren a. a M (2005) Intrusive growth of flax phloem fibers is of intercalary type. Planta **222**: 565–574

Alister D. Muir NDW (2003) Flax: The genus Linum.

An P, Li X, Zheng Y, Matsuura A, Abe J, Eneji AE, Tanimoto E, Inanaga S (2014) Effects of NaCl on Root Growth and Cell Wall Composition of Two Soya bean Cultivars with Contrasting Salt Tolerance. J Agron Crop Sci **200**: 212–218

Anders N, Wilkinson MD, Lovegrove A, Freeman J, Tryfona T, Pellny TK, Weimar T, Mortimer JC, Stott K, Baker JM, et al (2012) Glycosyl transferases in family 61 mediate arabinofuranosyl transfer onto xylan in grasses. Proc Natl Acad Sci U S A **109**: 989–93

Appenzeller L, Doblin M, Barreiro R, Wang H, Niu X, Kollipara K, Carrigan L, Tomes D, Chapman M, Dhugga KS (2004) Cellulose synthesis in maize: isolation and expression analysis of the cellulose synthase (CesA) gene family. Cellulose **11**: 287–299

Ariizumi T, Lawrence PK, Steber CM (2011) The role of two f-box proteins, SLEEPY1 and SNEEZY, in Arabidopsis gibberellin signaling. Plant Physiol **155**: 765–75

Atmodjo MA, Hao Z, Mohnen D (2013) Evolving views of pectin biosynthesis. Annu Rev Plant Biol **64**: 747–79

Baba K, Park YW, Kaku T, Kaida R, Takeuchi M, Yoshida M, Hosoo Y, Ojio Y, Okuyama T, Taniguchi T, et al (2009) Xyloglucan for generating tensile stress to bend tree stem. Mol Plant **2**: 893–903

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

Barnett JR, Bonham VA (2004) Cellulose microfibril angle in the cell wall of wood fibres. Biol Rev Camb Philos Soc **79**: 461–472

Bar-Peled M, Urbanowicz BR, O'Neill MA (2012) The Synthesis and Origin of the Pectic Polysaccharide Rhamnogalacturonan II - Insights from Nucleotide Sugar Formation and Diversity. Front Plant Sci **3**: 92

Barvkar VT, Pardeshi VC, Kale SM, Kadoo NY, Giri AP, Gupta VS (2012) Proteome profiling of flax (Linum usitatissimum) seed: Characterization of functional metabolic pathways operating during seed development. J Proteome Res **11**: 6264–6276

Battaglia M, Solórzano RM, Hernández M, Cuéllar-Ortiz S, García-Gómez B, Márquez J, Covarrubias AA (2006) Proline-rich cell wall proteins accumulate in growing regions and phloem tissue in response to water deficit in common bean seedlings. Planta **225**: 1121–1133

Baumann MJ, Eklöf JM, Michel G, Kallas AM, Teeri TT, Czjzek M, Brumer H (2007) Structural evidence for the evolution of xyloglucanase activity from xyloglucan endo-transglycosylases: biological implications for cell wall metabolism. Plant Cell **19**: 1947–63

Becker B, Marin B (2009) Streptophyte algae and the origin of embryophytes. Ann Bot **103**: 999–1004

Bernhardt C, Tierney ML (2000) Expression of AtPRP3, a proline-rich structural cell wall protein from Arabidopsis, is regulated by cell-type-specific developmental pathways involved in root hair formation. Plant Physiol **122**: 705–14

Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. Annu Rev Plant Biol **54**: 519–46

Bonello P, Blodgett JT (2003) Pinus nigra–Sphaeropsis sapinea as a model pathosystem to investigate local and systemic effects of fungal infection of pines. Physiol Mol Plant Pathol **63**: 249–261

Bosch M, Hepler PK (2005) Pectin methylesterases and pectin dynamics in pollen tubes. Plant Cell **17**: 3219–3226

Bourmaud A, Ausias G, Lebrun G, Tachon M-L, Baley C (2013) Observation of the structure of a composite polypropylene/flax and damage mechanisms under stress. Ind Crops Prod **43**: 225–236

Bray EA (2004) Genes commonly regulated by water-deficit stress in Arabidopsis thaliana. J Exp Bot **55**: 2331–41

Brown DM, Goubet F, Wong VW, Goodacre R, Stephens E, Dupree P, Turner SR (2007) Comparison of five xylan synthesis mutants reveals new insight into the mechanisms of xylan synthesis. Plant J **52**: 1154–1168

Brown DM, Zeef LAH, Ellis J, Goodacre R, Turner SR (2005) Identification of novel genes in Arabidopsis involved in secondary cell wall formation using expression profiling and reverse genetics. Plant Cell **17**: 2281–95

Brutch NB, Sharov IY, Pavlov a. V., Porokhovinova E a. (2011) Diversity of flax characters associated with fibre formation and environmental influence on their formation. Russ J Genet Appl Res **1**: 361–370

Brutch NB, Soret-Morvan O, Porokhovinova E a., Sharov IY, Morvan C (2008) Characters of Fibre Quality in Lines of Flax Genetic Collection. J Nat Fibers **5**: 95–126

Buendgen MR, Coors JG, Grombacher AW, Russell WA (1990) European Corn Borer Resistance and Cell Wall Composition of Three Maize Populations. Crop Sci **30**: 505

Buffetto F, Ropartz D, Zhang XJ, Gilbert HJ, Guillon F, Ralet M-C (2014) Recovery and fine structure variability of RGII sub-domains in wine (Vitis vinifera Merlot). Ann Bot **114**: 1327–1337

Burn JE, Hurley UA, Birch RJ, Arioli T, Cork A, Williamson RE (2002) The cellulose-deficient Arabidopsis mutant rsw3 is defective in a gene encoding a putative glucosidase II, an enzyme processing N-glycans during ER quality control. Plant J **32**: 949–60

Burton R a, Gidley MJ, Fincher GB (2010) Heterogeneity in the chemistry, structure and function of plant cell walls. Nat Chem Biol **6**: 724–732

Caffall KH, Mohnen D (2009) The structure, function, and biosynthesis of plant cell wall pectic polysaccharides. Carbohydr Res **344**: 1879–900

Campargue C, Lafitte C, Esquerré-Tugayé MT, Mazau D (1998) Analysis of hydroxyproline and hydroxyproline-arabinosides of plant origin by high-performance anion-exchange chromatography-pulsed amperometric detection. Anal Biochem **257**: 20–5

Cannon MC, Terneus K, Hall Q, Tan L, Wang Y, Wegenhart BL, Chen L, Lamport DTA, Chen Y, Kieliszewski MJ (2008) Self-assembly of the plant cell wall requires an extensin scaffold. Proc Natl Acad Sci U S A **105**: 2226–31

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res **37**: D233–8

Carpita N, Tierney M, Campbell M (2001) Molecular biology of the plant cell wall: searching for the genes that define structure, architecture and dynamics. Plant Mol Biol **47**: 1–5

Carpita NC (1996) STRUCTURE AND BIOGENESIS OF THE CELL WALLS OF GRASSES. Annu Rev Plant Physiol Plant Mol Biol **47**: 445–476

Carroll A, Mansoori N, Li S, Lei L, Vernhettes S, Visser RGF, Somerville C, Gu Y, Trindade LM (2012) Complexes with mixed primary and secondary cellulose synthases are functional in Arabidopsis plants. Plant Physiol **160**: 726–37

Cassab GI, Varner JE (1988) Cell w all proteins. Ann Rev Plant Physiol Plant Mol Biol **39**: 321–353

Chandna R, Ahmad A (2015) Nitrogen stress-induced alterations in the leaf proteome of two wheat varieties grown at different nitrogen levels. Physiol Mol Biol Plants **21**: 19–33

Chantreau M, Chabbert B, Billiard S, Hawkins S, Neutelings G (2015) Functional analyses of cellulose synthase genes in flax (Linum usitatissimum) by virus-induced gene silencing. Plant Biotechnol J. doi: 10.1111/pbi.12350

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 (2014a) 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 **26**: 4462–82

Chantreau M, Portelette A, Dauwe R, Kiyoto S, Crônier D, Morreel K, Arribat S, Neutelings G, Chabi M, Boerjan W, et al (2014b) 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 **26**: 4462–82

Charlet K, Eve S, Jernot JP, Gomina M, Breard J (2009) Tensile deformation of a flax fiber. Procedia Eng **1**: 233–236

Charlet K, Jernot JP, Eve S, Gomina M, Bréard J (2010) Multi-scale morphological characterisation of flax: From the stem to the fibrils. Carbohydr Polym **82**: 54–61

Chemikosova SB, Pavlencheva N V., Gur'yanov OP, Gorshkova TA (2006) The effect of soil drought on the phloem fiber development in long-fiber flax. Russ J Plant Physiol **53**: 656–662

Cocuron J-C, Lerouxel O, Drakakaki G, Alonso AP, Liepman AH, Keegstra K, Raikhel N, Wilkerson CG (2007) A gene from the cellulose synthase-like C family encodes a -1,4 glucan synthase. Proc Natl Acad Sci **104**: 8550–8555

Cosgrove DJ (1999) Enzymes and other agents that enhance cell wall extensibility. Annu Rev Plant Physiol Plant Mol Biol **50**: 391–417

Cosgrove DJ (2014) ScienceDirect Re-constructing our models of cellulose and primary cell wall assembly. Curr Opin Plant Biol **22**: 122–131

Cosgrove DJ (2005) Growth of the plant cell wall. Nat Rev Mol Cell Biol **6**: 850–61

Criscuolo A, Gribaldo S (2010) BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evol Biol **10**: 210

Cui S, Hu J, Guo S, Wang J, Cheng Y, Dang X, Wu L, He Y (2012) Proteome analysis of Physcomitrella patens exposed to progressive dehydration and rehydration. J Exp Bot **63**: 711–26

Dalisay DS, Kim KW, Lee C, Yang H, Rübel O, Bowen BP, Davin LB, Lewis NG (2015) Dirigent Protein-Mediated Lignan and Cyanogenic Glucoside Formation in Flax Seed: Integrated Omics and MALDI Mass Spectrometry Imaging. J Nat Prod **78**: 1231–42

Dash PK, Cao Y, Jailani AK, Gupta P, Venglat P, Xiang D, Rai R, Sharma R (2014) Genome-wide analysis of drought induced gene expression changes in flax (Linum usitatissimum). **5**: 37–41

Davis EL, Hussey RS, Baum TJ, Bakker J, Schots A, Rosso M-N, Abad P (2000) N EMATODE P ARASITISM G ENES. Annu Rev Phytopathol **38**: 365–396

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

Day A, Fénart S, Neutelings G, Hawkins S, Rolando C, Tokarski C (2013) Identification of cell wall proteins in the flax (Linum usitatissimum) stem. Proteomics **13**: 812–825

Day A, Ruel K, Neutelings G, Crônier D, David H, Hawkins S, Chabbert B (2005b) Lignification in the flax stem: evidence for an unusual lignin in bast fibers. Planta **222**: 234–45

Delmer DP (1999) CELLULOSE BIOSYNTHESIS: Exciting Times for A Difficult Field of Study. Annu Rev Plant Physiol Plant Mol Biol **50**: 245–276

DeMartini JD, Pattathil S, Miller JS, Li H, Hahn MG, Wyman CE (2013) Investigating plant cell wall components that affect biomass recalcitrance in poplar and switchgrass. Energy Environ Sci **6**: 898

Desprez T., Juraniec M. b, Crowell EF. F, Jouy H., Pochylova Z., Parcy F., Höfte H., Gonneau M., Vernhettes S. (2007) Organization of cellulose synthase complexes involved in primary cell wall synthesis in Arabidopsis thaliana. Proc Natl Acad Sci U S A **104**: 15572–15577

Diederichsen A, Ulrich A (2009) Variability in stem fibre content and its association with other characteristics in 1177 flax (Linum usitatissimum L.) genebank accessions. Ind Crops Prod **30**: 33–39

Van Dijk JP, de Mello CS, Voorhuijzen MM, Hutten RCB, Arisi ACM, Jansen JJ, Buydens LMC, van der Voet H, Kok EJ (2014) Safety assessment of plant varieties using transcriptomics profiling and a one-class classifier. Regul Toxicol Pharmacol **70**: 297–303

Dillman AC, Brinsmade JCJ (1938) Effect of spacing on the development of the flax plant. J. Am. Soc. Agron.

Ding Y, Chang J, Ma Q, Chen L, Liu S, Jin S, Han J, Xu R, Zhu A, Guo J, et al (2015) Network Analysis of Postharvest Senescence Process in Citrus Fruits Revealed by Transcriptomic and Metabolomic Profiling. Plant Physiol **168**: pp.114.255711

Doblin MS, Pettolino F, Bacic A (2010) Evans Review: Plant cell walls: The skeleton of the plant world. Funct Plant Biol **37**: 357–381

Domon J-M, Baldwin L, Acket S, Caudeville E, Arnoult S, Zub H, Gillet F, Lejeune-Hénaut I, Brancourt-Hulmel M, Pelloux J, et al (2013) Cell wall compositional modifications of Miscanthus ecotypes in response to cold acclimation. Phytochemistry **85**: 51–61

Donaldson L, Hague J, Snell R (2001) Lignin Distribution in Coppice Poplar, Linseed and Wheat Straw. Holzforschung **55**: 379–385

Dong J, Jiang Y, Chen R, Xu Z, Gao X (2011) Isolation of a novel xyloglucan endotransglucosylase (OsXET9) gene from rice and analysis of the response of this gene to abiotic stresses. African J Biotechnol **10**: 17424–17434

Doré C, Varoquaux F (2006) Histoire et amélioration de cinquante plantes cultivées.

Dumont M, Lehner A, Bouton S, Kiefer-Meyer MC, Voxeur A, Pelloux J, Lerouge P, Mollet J-C (2014) The cell wall pectic polymer rhamnogalacturonan-II is required for proper pollen tube elongation: implications of a putative sialyltransferase-like protein. Ann Bot 1177–1188

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32**: 1792–7

Egelund J, Petersen BL, Motawia MS, Damager I, Faik A, Olsen CE, Ishii T, Clausen H, Ulvskov P, Geshi N (2006) Arabidopsis thaliana RGXT1 and RGXT2 encode Golgi-localized (1,3)-alpha-D-xylosyltransferases involved in the synthesis of pectic rhamnogalacturonan-II. Plant Cell **18**: 2593–607

Eklöf JM, Brumer H (2010) The XTH gene family: an update on enzyme structure, function, and phylogeny in xyloglucan remodeling. Plant Physiol **153**: 456–66

Eklöf JM, Shojania S, Okon M, McIntosh LP, Brumer H (2013) Structure-function analysis of a broad specificity Populus trichocarpa endo-β-glucanase reveals an evolutionary link between bacterial licheninases and plant XTH gene products. J Biol Chem **288**: 15786–15799

Esau K (1943) Vascular Differentiation in the Vegetative Shoot of Linum. III. The Origin of the Bast Fibers on JSTOR. Am J Bot **30**: 579–586

Estévez JM, Kieliszewski MJ, Khitrov N, Somerville C (2006) Characterization of synthetic hydroxyproline-rich proteoglycans with arabinogalactan protein and extensin motifs in Arabidopsis. Plant Physiol **142**: 458–70

Fan L, Linker R, Gepstein S, Tanimoto E, Yamamoto R, Neumann PM (2006) Progressive inhibition by water deficit of cell wall extensibility and growth along the elongation zone of maize roots is related to increased lignin metabolism and progressive stelar accumulation of wall phenolics. Plant Physiol **140**: 603–612

Fang L, Tian R, Li X, Chen J, Wang S, Wang P, Zhang T (2014) Cotton fiber elongation network revealed by expression profiling of longer fiber lines introgressed with different Gossypium barbadense chromosome segments. BMC Genomics **15**: 838

Farooq M, Wahid A, Kobayashi N, Fujita D, Basra S (2009) Plant drought stress: effects, mechanisms and management. Agron Sustain Dev **29**: 185–212

Feiz L, Irshad M, Pont-lezica RF, Canut H, Jamet E (2006) Evaluation of cell wall preparations for proteomics : a new procedure for purifying cell walls from Arabidopsis hypocotyls. **13**: 1–13

Fenart S, Chabi M, Gallina S, Huis R, Neutelings G, Riviere N, Thomasset B, Hawkins S, Lucau-Danila A (2013) Intra-platform comparison of 25-mer and 60-mer oligonucleotide Nimblegen DNA microarrays. BMC Res Notes **6**: 43

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

Festucci-Buselli R a., Otoni WC, Joshi CP (2007) Structure, organization, and functions of cellulose synthase complexes in higher plants. Brazilian J Plant Physiol **19**: 1–13

Foster R, POONI HS, MACKAY IJ (1997) Quantitative evaluation of Linum usitatissimum varieties for dual-purpose traits. J Agric Sci **129**: 179–185

Fraige K, González-Fernández R, Carrilho E, Jorrín-Novo J V (2015) Metabolite and proteome changes during the ripening of Syrah and Cabernet Sauvignon grape varieties cultured in a nontraditional wine region in Brazil. J Proteomics **113**: 206–25

Franková L, Fry SC (2013) Biochemistry and physiological roles of enzymes that "cut and paste" plant cell-wall polysaccharides. J Exp Bot **64**: 3519–3550

Fry SC (1988) The growing plant cell wall: chemical and metabolic analysis. New York **203**: 333

Fry SC (2004) Primary cell wall metabolism: tracking the careers of wall polymers in living plant cells. New Phytol **161**: 641–675

Fu J, Huang B (2001) Involvement of antioxidants and lipid peroxidation in the adaptation of two cool-season grasses to localized drought stress. Environ Exp Bot **45**: 105–114

Gierlinger N, Schwanninger M (2006) Chemical Imaging of Poplar Wood Cell Walls by Confocal Raman Microscopy. **140**: 1246–1254

Gille S, Pauly M (2012) O-acetylation of plant cell wall polysaccharides. Front Plant Sci **3**: 12

Girault R, Bert F, Rihouey C, Jauneau A, Morvan C, Jarvis M (1997) Galactans and cellulose in flax fibres: putative contributions to the tensile strength. Int J Biol Macromol **21**: 179–188

Gorshkova T a., Wyatt SE, Salnikov V V., Gibeaut DM, Ibragimov MR, Lozovaya V V., Carpita NC (1996) Cell-Wall Polysaccharides of Developing Flax Plants. Plant Physiol **110**: 721–729

Gorshkova T, Mokshina N, Chernova T, Ibragimova N, Salnikov V, Mikshina P, Tryfona T, Banasiak A, Immerzeel P, Dupree P, et al (2015) Aspen tension wood fibers contain β-(1→4)-galactans and acidic arabinogalactans retained by cellulose microfibrils in gelatinous walls. Plant Physiol. doi: 10.1104/pp.15.00690

Gorshkova T, Morvan C (2006) Secondary cell-wall assembly in flax phloem fibres: Role of galactans. Planta **223**: 149–158

Gorshkova TA, Sal'nikov V V., Chemikosova SB, Ageeva M V., Pavlencheva N V., van Dam JEG (2003) The snap point: a transition point in Linum usitatissimum bast fiber development. Ind Crops Prod **18**: 213–221

Gothandam KM, Nalini E, Karthikeyan S, Shin JS (2009) OsPRP3, a flower specific proline-rich protein of rice, determines extracellular matrix structure of floral organs and its overexpression confers cold-tolerance. Plant Mol Biol **72**: 125–135

Gou J-Y, Miller LM, Hou G, Yu X-H, Chen X-Y, Liu C-J (2012) Acetylesterase-mediated deacetylation of pectin impairs cell elongation, pollen germination, and plant reproduction. Plant Cell **24**: 50–65

Goubet F, Bourlard T, Girault R, Alexandre C, Vandevelde M-C, Morvan C (1995) Structural features of galactans from flax fibres. Carbohydr Polym **27**: 221–227

Grabber JH (2005) How Do Lignin Composition, Structure, and Cross-Linking Affect Degradability? A Review of Cell Wall Model Studies. Crop Sci **45**: 820

Gunl M, Neumetzler L, Kraemer F, de Souza a., Schultink a., Pena M, York WS, Pauly M (2011) AXY8 Encodes an -Fucosidase, Underscoring the Importance of Apoplastic Metabolism on the Fine Structure of Arabidopsis Cell Wall Polysaccharides. Plant Cell **23**: 4025–4040

Guo J, Shi G, Guo X, Zhang L, Xu W, Wang Y, Su Z, Hua J (2015) Transcriptome analysis reveals that distinct metabolic pathways operate in salt-tolerant and salt-sensitive upland cotton varieties subjected to salinity stress. Plant Sci **238**: 33–45

Guo R, Hao W, Gong D (2012) Effects of Water Stress on Germination and Growth of Linseed Seedlings (Linum usitatissimum L), Photosynthetic Efficiency and Accumulation of Metabolites. J Agric Sci **4**: p253

Gurjanov OP. P, Ibragimova NN. N, Gnezdilov OI. I, Gorshkova TA. A (2008) Polysaccharides, tightly bound to cellulose in cell wall of flax bast fibre: Isolation and identification. Carbohydr Polym **72**: 719–729

Hall Q, Cannon MC (2002) The cell wall hydroxyproline-rich glycoprotein RSH is essential for normal embryo development in Arabidopsis. Plant Cell **14**: 1161–72

Han Y yang, Li A xiu, Li F, Zhao M rong, Wang W (2012) Characterization of a wheat (Triticum aestivum L.) expansin gene, TaEXPB23, involved in the abiotic stress response and phytohormone regulation. Plant Physiol Biochem **54**: 49–58

Harholt J, Suttangkakul A, Vibe Scheller H (2010) Biosynthesis of pectin. Plant Physiol **153**: 384–95

Harper JF, Manney L, DeWitt ND, Yoo MH, Sussman MR (1990) The Arabidopsis thaliana plasma membrane H(+)-ATPase multigene family. Genomic sequence and expression of a third isoform. J Biol Chem **265**: 13601–8

Van Hazendonk JM, Reinerik EJM, de Waard P, van Dam JEG (1996) Structural analysis of acetylated hemicellulose polysaccharides from fibre flax (Linum usitatissimum L.). Carbohydr Res **291**: 141–154

Van Hengel AJ, Roberts K (2003) AtAGP30, an arabinogalactan-protein in the cell walls of the primary root, plays a role in root regeneration and seed germination. Plant J **36**: 256–70

Hijazi M, Velasquez SM, Jamet E, Estevez JM, Albenne C (2014) An update on post-translational modifications of hydroxyproline-rich glycoproteins : toward a model highlighting their contribution to plant cell wall architecture. **5**: 1–10

Höfer R, Boachon B, Renault H, Gavira C, Miesch L, Iglesias J, Ginglinger J-F, Allouche L, Miesch M, Grec S, et al (2014) Dual function of the cytochrome P450 CYP76 family from Arabidopsis thaliana in the metabolism of monoterpenols and phenylurea herbicides. Plant Physiol **166**: 1149–61

Holland N (2000) A Comparative Analysis of the Plant Cellulose Synthase (CesA) Gene Family. PLANT Physiol **123**: 1313–1324

Hotte NSC, Deyholos MK (2008) A flax fibre proteome: identification of proteins enriched in bast fibres. BMC Plant Biol **8**: 52

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

Ibragimova NN, Mokshina NE, Gorshkova TA (2012) [Cell wall proteins of plant fibers]. Bioorg Khim **38**: 139–48

Immerzeel P, Eppink MM, De Vries SC, Schols H a., Voragen AGJ (2006) Carrot arabinogalactan proteins are interlinked with pectins. Physiol Plant **128**: 18–28

Ingram J, Bartels D (1996) THE MOLECULAR BASIS OF DEHYDRATION TOLERANCE IN PLANTS. Annu Rev Plant Physiol Plant Mol Biol **47**: 377–403

Ishiguro S, Watanabe Y, Ito N, Nonaka H, Takeda N, Sakai T, Kanaya H, Okada K (2002) SHEPHERD is the Arabidopsis GRP94 responsible for the formation of functional CLAVATA proteins. EMBO J **21**: 898–908

Iwai H, Masaoka N, Ishii T, Satoh S (2002) A pectin glucuronyltransferase gene is essential for intercellular attachment in the plant meristem. Proc Natl Acad Sci U S A **99**: 16319–24

Jamet E, Boudart G, Borderies G, Charmont S, Lafitte C, Rossignol M, Canut H, Pont-Lezica R (2008) Isolation of plant cell wall proteins. Methods Mol Biol **425**: 187–201

Jamet E, Canut H, Boudart G, Pont-Lezica RF (2006) Cell wall proteins: a new insight through proteomics. Trends Plant Sci **11**: 33–9

Jauh G, Lord E (1996) Localization of pectins and arabinogalactan-proteins in lily (Lilium longiflorum L.) pollen tube and style, and their possible roles in pollination. Planta. doi: 10.1007/BF00196566

Jiang L, Yan S, Yang W, Li Y, Xia M, Chen Z, Wang Q, Yan L, Song X, Liu R, et al (2015) Transcriptomic analysis reveals the roles of microtubule-related genes and transcription factors in fruit length regulation in cucumber (Cucumis sativus L.). Sci Rep **5**: 8031

Joshi S. V, Drzal L. T, Mohanty a. . K, Arora S (2004) Are natural fiber composites environmentally superior to glass fiber reinforced composites? Compos Part A Appl Sci Manuf **35**: 371–376

Kaewthai N, Gendre D, Eklöf JM, Ibatullin FM, Ezcurra I, Bhalerao RP, Brumer H (2013) Group III-A XTH genes of Arabidopsis encode predominant xyloglucan endohydrolases that are dispensable for normal growth. Plant Physiol **161**: 440–54

Kantar M, Lucas SJ, Budak H (2011) Drought Stress. Molecular Genetics and Genomics Approaches. Adv Bot Res. doi: 10.1016/B978-0-12-387692-8.00013-8

Kaya MD, Okçu G, Atak M, Çıkılı Y, Kolsarıcı Ö (2006) Seed treatments to overcome salt and drought stress during germination in sunflower (Helianthus annuus L.). Eur J Agron **24**: 291–295

Keckes J, Burgert I, Frühmann K, Müller M, Kölln K, Hamilton M, Burghammer M, Roth S V, Stanzl-Tschegg S, Fratzl P (2003) Cell-wall recovery after irreversible deformation of wood. Nat Mater **2**: 810–4

Keegstra K, Raikhel N (2001) Plant glycosyltransferases. Curr Opin Plant Biol **4**: 219–24

Keegstra K, Talmadge KW, Bauer WD, Albersheim P (1973) The Structure of Plant Cell Walls: III. A Model of the Walls of Suspension-cultured Sycamore Cells Based on the Interconnections of the Macromolecular Components. Plant Physiol **51**: 188–97

Keller B, Schmid J, Lamb CJ (1989) Vascular expression of a bean cell wall glycine-rich protein-beta-glucuronidase gene fusion in transgenic tobacco. EMBO J **8**: 1309–14

Kenrick P, Crane PR (1997) The origin and early evolution of plants on land. **389**: 33–39

Kieliszewski MJ (2001) The latest hype on Hyp-O-glycosylation codes. Phytochemistry **57**: 319–23

Kimura S (1999) Immunogold Labeling of Rosette Terminal Cellulose-Synthesizing Complexes in the Vascular Plant Vigna angularis. PLANT CELL ONLINE **11**: 2075–2086

Kitazawa K, Tryfona T, Yoshimi Y, Hayashi Y, Kawauchi S, Antonov L, Tanaka H, Takahashi T, Kaneko S, Dupree P, et al (2013) β-galactosyl Yariv reagent binds to the β-1,3-galactan of arabinogalactan proteins. Plant Physiol **161**: 1117–26

Kitts DD, Yuan Y V, Wijewickreme AN, Thompson LU (1999) Antioxidant activity of the flaxseed lignan secoisolariciresinol diglycoside and its mammalian lignan metabolites enterodiol and enterolactone. Mol Cell Biochem **202**: 91–100

Konno H, Yamasaki Y, Sugimoto M, Takeda K (2008) Differential changes in cell wall matrix polysaccharides and glycoside-hydrolyzing enzymes in developing wheat seedlings differing in drought tolerance. J Plant Physiol **165**: 745–54

Kumar S, You FM, Duguid S, Booker H, Rowland G, Cloutier S (2015) QTL for fatty acid composition and yield in linseed (Linum usitatissimum L.). Theor Appl Genet **128**: 965–984

Kymäläinen H-R, Sjöberg A-M (2008) Flax and hemp fibres as raw materials for thermal insulations. Build Environ **43**: 1261–1269

Lafarguette F, Leplé J-C, Déjardin A, Laurans F, Costa G, Lesage-Descauses M-C, Pilate G (2004) Poplar genes encoding fasciclin-like arabinogalactan proteins are highly expressed in tension wood. New Phytol **164**: 107–121

Lamport DTA, Tan L, Kieliszewski MJ (2011) Structural proteins of the primary cell wall: extraction, purification, and analysis. Methods Mol Biol **715**: 209–19

Lao NT, Long D, Kiang S, Coupland G, Shoue DA, Carpita NC, Kavanagh TA (2003) Mutation of a family 8 glycosyltransferase gene alters cell wall carbohydrate composition and causes a humidity-sensitive semi-sterile dwarf phenotype in Arabidopsis. Plant Mol Biol **53**: 647–661

Le SQ, Gascuel O (2008) An improved general amino acid replacement matrix. Mol Biol Evol **25**: 1307–20

Lee B-R, Kim K-Y, Jung W-J, Avice J-C, Ourry A, Kim T-H (2007) Peroxidases and lignification in relation to the intensity of water-deficit stress in white clover (Trifolium repens L.). J Exp Bot **58**: 1271–9

Lee C, Teng Q, Huang W, Zhong R, Ye Z-H (2010) The Arabidopsis family GT43 glycosyltransferases form two functionally nonredundant groups essential for the elongation of glucuronoxylan backbone. Plant Physiol **153**: 526–41

Lee C, Teng Q, Zhong R, Ye Z-H (2011) The four Arabidopsis reduced wall acetylation genes are expressed in secondary wall-containing cells and required for the acetylation of xylan. Plant Cell Physiol **52**: 1289–301

Lefeuvre A, Bourmaud A, Lebrun L, Morvan C, Baley C (2013) A study of the yearly reproducibility of flax fiber tensile properties. Ind Crops Prod **50**: 400–407

Leonard R, Petersen BO, Himly M, Kaar W, Wopfner N, Kolarich D, van Ree R, Ebner C, Duus JO, Ferreira F, et al (2005) Two Novel Types of O-Glycans on the Mugwort Pollen Allergen Art v 1 and Their Role in Antibody Binding. J Biol Chem **280**: 7932–7940

Léonard R, Wopfner N, Pabst M, Stadlmann J, Petersen BO, Duus JØ, Himly M, Radauer C, Gadermaier G, Razzazi-Fazeli E, et al (2010) A new allergen from ragweed (Ambrosia artemisiifolia) with homology to art v 1 from mugwort. J Biol Chem **285**: 27192–200

Lerouxel O, Cavalier DM, Liepman AH, Keegstra K (2006) Biosynthesis of plant cell wall polysaccharides - a complex process. Curr Opin Plant Biol **9**: 621–30

Leucci MR, Lenucci MS, Piro G, Dalessandro G (2008) Water stress and cell wall polysaccharides in the apical root zone of wheat cultivars varying in drought tolerance. J Plant Physiol **165**: 1168–1180

Lewis NG, Yamamoto E (1990) Lignin: occurrence, biogenesis and biodegradation. Annu Rev Plant Physiol Plant Mol Biol **41**: 455–96

Lindlöf A, Chawade A, Sikora P, Olsson O (2015) Comparative Transcriptomics of Sijung and Jumli Marshi Rice during Early Chilling Stress Imply Multiple Protective Mechanisms. PLoS One **10**: e0125385

Liu C, Cai Y, Zhang X, Gou M, Yang H (2014) Tailoring lignin biosynthesis for efficient and sustainable biofuel production. Plant Biotechnol J **12**: 1154–62

Liu X-L, Liu L, Niu Q-K, Xia C, Yang K-Z, Li R, Chen L-Q, Zhang X-Q, Zhou Y, Ye D (2011) MALE GAMETOPHYTE DEFECTIVE 4 encodes a rhamnogalacturonan II xylosyltransferase and is important for growth of pollen tubes and roots in Arabidopsis. Plant J **65**: 647–660

Liwanag AJM, Ebert B, Verhertbruggen Y, Rennie EA, Rautengarten C, Oikawa A, Andersen MCF, Clausen MH, Scheller H V. (2012) Pectin Biosynthesis: GALS1 in Arabidopsis thaliana Is a -1,4-Galactan -1,4-Galactosyltransferase. Plant Cell **24**: 5024–5036

Lohse M, Nagel A, Herter T, May P, Schroda M, Zrenner R, Tohge T, Fernie AR, Stitt M, Usadel B (2014) Mercator: a fast and simple web server for genome scale functional annotation of plant sequence data. Plant Cell Environ **37**: 1250–8

Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. **42**: 490–495

Lowry B, Lee D, Hébant C (1980) The Origin of Land Plants: A New Look at an Old Problem on JSTOR. Int Assoc Plant Taxon **29**: 183–197

Lygin A V., UPTON J, DOHLEMAN FG, JUVIK J, ZABOTINA OA, WIDHOLM JM, LOZOVAYA V V. (2011) Composition of cell wall phenolics and polysaccharides of the potential bioenergy crop -Miscanthus. GCB Bioenergy **3**: 333–345

M. Sfiligoj Smole SHKSK and TK (2013) Advances in Agrophysical Research. doi: 10.5772/3341

MacMillan CP, Mansfield SD, Stachurski ZH, Evans R, Southerton SG (2010) Fasciclin-like arabinogalactan proteins: specialization for stem biomechanics and cell wall architecture in Arabidopsis and Eucalyptus. Plant J **62**: 689–703

Manabe Y, Nafisi M, Verhertbruggen Y, Orfila C, Gille S, Rautengarten C, Cherk C, Marcus SE, Somerville S, Pauly M, et al (2011) Loss-of-Function Mutation of REDUCED WALL ACETYLATION2 in Arabidopsis Leads to Reduced Cell Wall Acetylation and Increased Resistance to Botrytis cinerea. PLANT Physiol **155**: 1068–1078

Mangeon A, Junqueira RM, Sachetto-Martins G (2010) Functional diversity of the plant glycine-rich proteins superfamily. Plant Signal Behav **5**: 99–104

Manickavelu A, Nadarajan N, Ganesh SK, Gnanamalar RP, Chandra Babu R (2006) Drought tolerance in rice: morphological and molecular genetic consideration. Plant Growth Regul **50**: 121–138

Matsunaga T, Ishii T, Matsumoto S, Higuchi M, Darvill A, Albersheim P, O'Neill MA (2004) Occurrence of the primary cell wall polysaccharide rhamnogalacturonan II in pteridophytes, lycophytes, and bryophytes. Implications for the evolution of vascular plants. Plant Physiol **134**: 339–51

McCourt RM, Delwiche CF, Karol KG (2004) Charophyte algae and land plant origins. Trends Ecol Evol **19**: 661–6

McDill J, Repplinger M, Simpson BB, Kadereit JW (2009) The Phylogeny of <I>Linum</I> and Linaceae Subfamily Linoideae, with Implications for Their Systematics, Biogeography, and Evolution of Heterostyly. Syst Bot **34**: 386–405

McDougall GJ (1993) Isolation and partial characterisation of the non-cellulosic polysaccharides of flax fibre. Carbohydr Res **241**: 227–236

Mcfarlane HE, Anett D (2014) The Cell Biology of Cellulose Synthesis. doi: 10.1146/annurev-arplant-050213-040240

Mellerowicz E, Sundberg B (2008) Wood cell walls: biosynthesis, developmental dynamics and their implications for wood properties. Curr Opin Plant Biol **11**: 293–300

Mellerowicz EJ, Gorshkova TA (2012) Tensional stress generation in gelatinous fibres: a review and possible mechanism based on cell-wall structure and composition. J Exp Bot **63**: 551–65

Michel SAAX, Vogels RRM, Bouvy ND, Knetsch MLW, van den Akker NMS, Gijbels MJJ, van der Marel C, Vermeersch J, Molin DGM, Koole LH (2014) Utilization of flax fibers for biomedical applications. J Biomed Mater Res B Appl Biomater **102**: 477–87

Miedes E, Vanholme R, Boerjan W, Molina A (2014) The role of the secondary cell wall in plant resistance to pathogens. Front Plant Sci **5**: 1–13

Minic Z (2008) Physiological roles of plant glycoside hydrolases. Planta **227**: 723–40

Minic Z, Jamet E, Negroni L, Arsene der Garabedian P, Zivy M, Jouanin L (2007) A sub-proteome of Arabidopsis thaliana mature stems trapped on Concanavalin A is enriched in cell wall glycoside hydrolases. J Exp Bot **58**: 2503–2512

Mohnen D (2008) Pectin structure and biosynthesis. Curr Opin Plant Biol **11**: 266–277

Mokshina N, Gorshkova T, Deyholos MK (2014) Chitinase-like (CTL) and cellulose synthase (CESA) gene expression in gelatinous-type cellulosic walls of flax (Linum usitatissimum L.) bast fibers. PLoS One **9**: e97949

Moore JP, Farrant JM, Driouich A (2008) A role for pectin-associated arabinans in maintaining the flexibility of the plant cell wall during water deficit stress. Plant Signal Behav **3**: 102–104

Moore JP, Nguema-Ona E, Chevalier L, Lindsey GG, Brandt WF, Lerouge P, Farrant JM, Driouich A (2006) Response of the leaf cell wall to desiccation in the resurrection plant Myrothamnus flabellifolius. Plant Physiol **141**: 651–662

Morvan C, Andème-Onzighi C, Girault R, Himmelsbach DS, Driouich A, Akin DE (2003a) Building flax fibres: more than one brick in the walls. Plant Physiol Biochem **41**: 935–944

Morvan C, Andème-Onzighi C, Girault R, Himmelsbach DS, Driouich A, Akin DE (2003b) Building flax fibres: More than one brick in the walls. Plant Physiol Biochem **41**: 935–944

Motose H, Sugiyama M, Fukuda H (2004) A proteoglycan mediates inductive interaction during plant vascular development. Nature **429**: 873–878

Mousavi A, Hotta Y (2005) Glycine-rich proteins: a class of novel proteins. Appl Biochem Biotechnol **120**: 169–74

Munns R, James RA, Sirault XRR, Furbank RT, Jones HG (2010) New phenotyping methods for screening wheat and barley for beneficial responses to water deficit. J Exp Bot **61**: 3499–3507

Murashige T, Skoog F (1962) A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures. Physiol Plant **15**: 473–497

Muravenko O V, Amosova A V, Samatadze TE, Semenova OI, Nosova I V, Popov K V, Shostak NG, Zoshchuk SA, Zelenin A V (2004) [Chromosomal localization of 5S and 45S ribosomal DNA in species of Linum L. section Linum (syn=Protolinum and Adenolinum)]. Genetika **40**: 256–60

Muravenko O V, Lemesh VA, Samatadze TE, Amosova A V, Grushetskaia ZE, Popov K V, Semenova OI, Khotyleva L V, Zelenin A V (2003) [Genome comparisons of three closely related flax species and their hybrids with chromosomal and molecular markers]. Genetika **39**: 510–8

Muravenko V, Samatadze TE, Popov K V, Amosova A V, Zelenin A V (2001) Polymorphism of heterochromatic regions of flax chromosomes. Membr Cell Biol **14**: 743–8

Nakashima J, Chen F, Jackson L, Shadle G, Dixon RA (2008) Multi-site genetic modification of monolignol biosynthesis in alfalfa (Medicago sativa): effects on lignin composition in specific cell types. New Phytol **179**: 738–750

Neumann PM (2011) Plant Responses to Drought and Salinity Stress - Developments in a Post-Genomic Era. Adv Bot Res. doi: 10.1016/B978-0-12-387692-8.00002-3

Le Ngoc Huyen T, Rémond C, Dheilly RM, Chabbert B (2010) Effect of harvesting date on the composition and saccharification of Miscanthus x giganteus. Bioresour Technol **101**: 8224–31

Nguema-Ona E, Vicré-Gibouin M, Cannesan M-A, Driouich A (2013) Arabinogalactan proteins in root–microbe interactions. Trends Plant Sci **18**: 440–449

Nguema-Ona E, Vicré-Gibouin M, Gotté M, Plancot B, Lerouge P, Bardor M, Driouich A (2014) Cell wall O-glycoproteins and N-glycoproteins: aspects of biosynthesis and function. Front Plant Sci **5**: 499

Nicotra AB, Atkin OK, Bonser SP, Davidson AM, Finnegan EJ, Mathesius U, Poot P, Purugganan MD, Richards CL, Valladares F, et al (2010) Plant phenotypic plasticity in a changing climate. Trends Plant Sci **15**: 684–92

Niklas KJ, Kutschera U (2010) The evolution of the land plant life cycle. New Phytol **185**: 27–41

Nishikubo N, Awano T, Banasiak A, Bourquin V, Ibatullin F, Funada R, Brumer H, Teeri TT, Hayashi T, Sundberg B, et al (2007) Xyloglucan endo-transglycosylase (XET) functions in gelatinous layers of tension wood fibers in poplar--a glimpse into the mechanism of the balancing act of trees. Plant Cell Physiol **48**: 843–55

Nishikubo N, Takahashi J, Roos A a, Derba-Maceluch M, Piens K, Brumer H, Teeri TT, Stålbrand H, Mellerowicz EJ (2011) Xyloglucan endo-transglycosylase-mediated xyloglucan rearrangements in developing wood of hybrid aspen. Plant Physiol **155**: 399–413

O'Neill MA, Ishii T, Albersheim, P. A, Darvill AG (2004) Rhamnogalacturonan II: structure and function of a borate cross-linked cell wall pectic polysaccharide. Annu Rev Plant Biol **55**: 109–39

Obel N, Erben V, Schwarz T, Kühnel S, Fodor A, Pauly M (2009) Microanalysis of plant cell wall polysaccharides. Mol Plant **2**: 922–32

Obembe OO, Jacobsen E, Timmers J, Gilbert H, Blake AW, Knox JP, Visser RGF, Vincken J-P (2007) Promiscuous, non-catalytic, tandem carbohydrate-binding modules modulate the cell-wall structure and development of transgenic tobacco (Nicotiana tabacum) plants. J Plant Res **120**: 605–617

Ogawa-Ohnishi M, Matsushita W, Matsubayashi Y (2013) Identification of three hydroxyproline O-arabinosyltransferases in Arabidopsis thaliana. Nat Chem Biol **9**: 726–30

Oliveira TM, da Silva FR, Bonatto D, Neves DM, Morillon R, Maserti BE, Filho MAC, Costa MGC, Pirovani CP, Gesteira AS (2015) Comparative study of the protein profiles of Sunki mandarin and Rangpur lime plants in response to water deficit. BMC Plant Biol **15**: 69

Paës G (2014) Fluorescent Probes for Exploring Plant Cell Wall Deconstruction: A Review. 9380–9402

Park YB, Cosgrove DJ (2012) A revised architecture of primary cell walls based on biomechanical changes induced by substrate-specific endoglucanases. Plant Physiol **158**: 1933–43

Park YB, Cosgrove DJ (2015) Xyloglucan and its interactions with other components of the growing cell wall. Plant Cell Physiol **56**: 180–94

Passardi F, Penel C, Dunand C (2004) Performing the paradoxical: how plant peroxidases modify the cell wall. Trends Plant Sci **9**: 534–40

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

Pauly M, Keegstra K (2008) Cell-wall carbohydrates and their modification as a resource for biofuels. Plant J **54**: 559–68

Penna S (2003) Building stress tolerance through over-producing trehalose in transgenic plants. Trends Plant Sci **8**: 355–357

Pérez S, Rodríguez-Carvajal MA, Doco T (2003) A complex plant cell wall polysaccharide: rhamnogalacturonan II. A structure in quest of a function. Biochimie **85**: 109–21

Perrin RM, DeRocher AE, Bar-Peled M, Zeng W, Norambuena L, Orellana A, Raikhel N V, Keegstra K (1999) Xyloglucan fucosyltransferase, an enzyme involved in plant cell wall biosynthesis. Science **284**: 1976–9

Petersen BL, Egelund J, Damager I, Faber K, Krüger Jensen J, Yang Z, Bennett EP, Scheller HV, Ulvskov P (2009) Assay and heterologous expression in Pichia pastoris of plant cell wall type-II membrane anchored glycosyltransferases. Glycoconj J **26**: 1235–1246

Pilling E, Höfte H (2003) Feedback from the wall. Curr Opin Plant Biol **6**: 611–6

Pinzon-Latorre D, Deyholos MK (2014) Pectinmethylesterases (PME) and Pectinmethylesterase Inhibitors (PMEI) Enriched during Phloem Fiber Development in Flax (Linum usitatissimum). PLoS One **9**: e105386

Piro G, Leucci MR, Waldron K, Dalessandro G (2003) Exposure to water stress causes changes in the biosynthesis of cell wall polysaccharides in roots of wheat cultivars varying in drought tolerance. Plant Sci **165**: 559–569

Popper Z (2008) Evolution and diversity of green plant cell walls. Curr Opin Plant Biol **11**: 286–292

Popper ZA, Michel G, Hervé C, Domozych DS, Willats WGT, Tuohy MG, Kloareg B, Stengel DB (2011) Evolution and diversity of plant cell walls: from algae to flowering plants. Annu Rev Plant Biol **62**: 567–590

Prasad K (1997) Hydroxyl radical-scavenging property of secoisolariciresinol diglucoside (SDG) isolated from flax-seed. Mol Cell Biochem **168**: 117–23

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics **155**: 945–959

Qi X, Behrens BX, West PR, Mort AJ (1995) Solubilization and partial characterization of extensin fragments from cell walls of cotton suspension cultures. Evidence for a covalent cross-link between extensin and pectin. Plant Physiol **108**: 1691–701

Qiu D, Wilson IW, Gan S, Washusen R, Moran GF, Southerton SG (2008) Gene expression in Eucalyptus branch wood with marked variation in cellulose microfibril orientation and lacking G-layers. New Phytol **179**: 94–103

Ragauskas AJ, Ragauskas AJ, Beckham GT, Biddy MJ, Chandra R, Chen F, Naskar AK, Saddler JN, Tschaplinski TJ, Tuskan GA, et al (2014) Lignin Valorization : Improving Lignin Processing in the Biorefi nery. doi: 10.1126/science.1246843

Rakszegi M, Lovegrove A, Balla K, Láng L, Bedő Z, Veisz O, Shewry PR (2014) Effect of heat and drought stress on the structure and composition of arabinoxylan and β-glucan in wheat grain. Carbohydr Polym **102**: 557–65

Ralph J, Lundquist K, Brunow PG, Lu F, Kim H, Schatz P, Marita J, Hatfield R, Ralph S, Christensen J, et al (2004) Lignins: Natural polymer from oxidative coupling of 4-hydroxyphenylpropanoids. Phytochem Rev **3**: 29–60

Ramachandra Reddy A, Chaitanya KV, Vivekanandan M (2004) Drought-induced responses of photosynthesis and antioxidant metabolism in higher plants. J Plant Physiol **161**: 1189–202

Ray S, Paynel F, Morvan C, Lerouge P, Driouich A, Ray B (2013) Characterization of mucilage polysaccharides, arabinogalactanproteins and cell-wall hemicellulosic polysaccharides isolated from flax seed meal: A wealth of structural moieties. Carbohydr Polym **93**: 651–60

Rennie EA, Hansen SF, Baidoo EEK, Hadi MZ, Keasling JD, Scheller HV (2012) Three members of the Arabidopsis glycosyltransferase family 8 are xylan glucuronosyltransferases. Plant Physiol **159**: 1408–17

Rich MK, Schorderet M, Reinhardt D (2014) The role of the cell wall compartment in mutualistic symbioses of plants. Front Plant Sci **5**: 238

Richmond T (2000) Higher plant cellulose synthases. Genome Biol **1**: REVIEWS3001

Ridley BL, O'Neill MA, Mohnen D (2001) Pectins: structure, biosynthesis, and oligogalacturonide-related signaling. Phytochemistry **57**: 929–67

Ringli C, Keller B, Ryser U (2001) Glycine-rich proteins as structural components of plant cell walls. Cell Mol Life Sci **58**: 1430–1441

Del Río JC, Rencoret J, Gutiérrez A, Nieto L, Jiménez-Barbero J, Martínez ÁT (2011) Structural characterization of guaiacyl-rich lignins in flax (Linum usitatissimum) fibers and shives. J Agric Food Chem **59**: 11088–11099

Roach MJ, Deyholos MK (2007) Microarray analysis of flax (Linum usitatissimum L.) stems identifies transcripts enriched in fibre-bearing phloem tissues. Mol Genet Genomics **278**: 149–165

Roach MJ, Mokshina NY, Badhan A, Snegireva A V., Hobson N, Deyholos MK, Gorshkova TA (2011) Development of cellulosic secondary walls in flax fibers requires beta-galactosidase. Plant Physiol **156**: 1351–1363

Rossignol M, Peltier J-B, Mock H-P, Matros A, Maldonado AM, Jorrín J V (2006) Plant proteome analysis: a 2004-2006 update. Proteomics **6**: 5529–48

Sachetto-Martins G, Franco LO, de Oliveira DE (2000) Plant glycine-rich proteins: a family or just proteins with a common motif? Biochim Biophys Acta - Gene Struct Expr **1492**: 1–14

Saito F, Suyama A, Oka T, Yoko-O T, Matsuoka K, Jigami Y, Shimma Y-I (2014) Identification of Novel Peptidyl Serine α-Galactosyltransferase Gene Family in Plants. J Biol Chem **289**: 20405–20420

Saladié M, Cañizares J, Phillips MA, Rodriguez-Concepcion M, Larrigaudière C, Gibon Y, Stitt M, Lunn JE, Garcia-Mas J (2015) Comparative transcriptional profiling analysis of developing melon (Cucumis melo L.) fruit from climacteric and non-climacteric varieties. BMC Genomics **16**: 440

Sampedro J, Cosgrove DJ (2005) The expansin superfamily. Genome Biol **6**: 242

San Clemente H, Jamet E (2015) WallProtDB, a database resource for plant cell wall proteomics. Plant Methods **11**: 2

Scarwell H-J, Roussel I (2010) Le changement climatique : Quand le climat nous pousse à changer d'ère. Développement durable Territ. 1:

Scheller HV, Ulvskov P (2010) Hemicelluloses. Annu Rev Plant Biol **61**: 263–289

Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Schölkopf B, Weigel D, Lohmann JU (2005) A gene expression map of Arabidopsis thaliana development. Nat Genet **37**: 501–6

Schultz CJ (2004) Post-translational Modifications of Arabinogalactan-peptides of Arabidopsis thaliana: ENDOPLASMIC RETICULUM AND GLYCOSYLPHOSPHATIDYLINOSITOL-ANCHOR SIGNAL CLEAVAGE SITES AND HYDROXYLATION OF PROLINE. J Biol Chem **279**: 45503–45511

Seifert GJ, Roberts K (2007) The Biology of Arabinogalactan Proteins. Annu Rev Plant Biol **58**: 137–161

Sharma HSS, Faughey G, Lyons G (1999) Comparison of physical, chemical, and thermal characteristics of water-, dew-, and enzyme-retted flax fibers. J Appl Polym Sci **74**: 139–143

Sharma HSS, Faughey G, McCall D (1996) Effect of Sample Preparation and Heating Rate on the Differential Thermogravimetric Analysis of Flax Fibres. J Text Inst **87**: 249–257

Shi YZ, Zhu XF, Miller JG, Gregson T, Zheng SJ, Fry SC (2014) Distinct catalytic capacities of two aluminium-repressed Arabidopsis thaliana xyloglucan endotransglucosylase/hydrolases, XTH15 and XTH31, heterologously produced in Pichia. Phytochemistry **112**: 160–169

Showalter AM (1993) Structure and Function of Plant Cell Wall Proteins. PLANT CELL ONLINE **5**: 9–23

Singh KK, Mridula D, Rehal J, Barnwal P (2011) Flaxseed: a potential source of food, feed and fiber. Crit Rev Food Sci Nutr **51**: 210–22

Smith BG, Harris PJ (1999) The polysaccharide composition of Poales cell walls. Biochem Syst Ecol **27**: 33–53

Smith E, Roberts K, Hutchings A, Galfre G (1984) Monoclonal antibodies to the major structural glycoprotein of the Chlamydomonas cell wall. Planta **161**: 330–8

Snegireva A V., Ageeva M V., Amenitskii SI, Chernova TE, Ebskamp M, Gorshkova TA (2010) Intrusive growth of sclerenchyma fibers. Russ J Plant Physiol **57**: 342–355

Somerville C (2006) Cellulose synthesis in higher plants. Annu Rev Cell Dev Biol **22**: 53–78

Song D, Xi W, Shen J, Bi T, Li L (2011) Characterization of the plasma membrane proteins and receptor-like kinases associated with secondary vascular differentiation in poplar. Plant Mol Biol **76**: 97–115

Sorek N, Sorek H, Kijac A, Szemenyei HJ, Bauer S, Hématy K, Wemmer DE, Somerville CR (2014) The Arabidopsis COBRA protein facilitates cellulose crystallization at the plasma membrane. J Biol Chem **289**: 34911–20

Sørensen I, Domozych D, Willats WGT (2010) How Have Plant Cell Walls Evolved ? 1. **153**: 366–372

Soto-Cerda BJ, Duguid S, Booker H, Rowland G, Diederichsen A, Cloutier S (2014) Association mapping of seed quality traits using the Canadian flax (Linum usitatissimum L.) core collection. Theor Appl Genet **127**: 881–96

Souza AP De, Kamei CLA, Torres AF, Pattathil S, Hahn MG, Trindade LM, Buckeridge MS (2015) How cell wall complexity influences saccharification efficiency in Miscanthus sinensis. doi: 10.1093/jxb/erv183

Strasser G (2001) Pectic substances from red beet (Beta vulgaris conditiva). Part I. Structural analysis of rhamnogalacturonan I using enzymic degradation and methylation analysis. Carbohydr Polym **44**: 63–70

Summerscales J, Dissanayake NPJ, Virk AS, Hall W (2010) Composites : Part A A review of bast fibres and their composites . Part 1 – Fibres as reinforcements. Compos Part A **41**: 1329–1335

Sveinsson S, McDill J, Wong GKS, Li J, Li X, Deyholos MK, Cronk QCB (2014) Phylogenetic pinpointing of a paleopolyploidy event within the flax genus (Linum) using transcriptomics. Ann Bot **113**: 753–761

Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, et al (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res **43**: D447–52

Tan L, Qiu F, Lamport DTA, Kieliszewski MJ (2004) Structure of a hydroxyproline (Hyp)-arabinogalactan polysaccharide from repetitive Ala-Hyp expressed in transgenic Nicotiana tabacum. J Biol Chem **279**: 13156–65

Tan L, Showalter AM, Egelund J, Hernandez-Sanchez A, Doblin MS, Bacic A (2012) Arabinogalactan-proteins and the research challenges for these enigmatic plant cell surface proteoglycans. Front Plant Sci **3**: 140

Tan L, Varnai P, Lamport DTA, Yuan C, Xu J, Qiu F, Kieliszewski MJ (2010) Plant O-hydroxyproline arabinogalactans are composed of repeating trigalactosyl subunits with short bifurcated side chains. J Biol Chem **285**: 24575–83

Taylor NG (2008) Cellulose biosynthesis and deposition in higher plants. New Phytol **178**: 239–52

Tenhaken R (2015) Cell wall remodeling under abiotic stress. Front Plant Sci **5**: 1–9

Thompson JE, Fry SC (2001) Restructuring of wall-bound xyloglucan by transglycosylation in living plant cells. Plant J **26**: 23–34

Tobimatsu Y, Wagner A, Donaldson L, Mitra P, Niculaes C, Dima O, Kim JI, Anderson N, Loque D, Boerjan W, et al (2013) Visualization of plant cell wall lignification using fluorescence-tagged monolignols. Plant J **76**: 357–366

Touré A, Xueming X (2010) Flaxseed Lignans: Source, Biosynthesis, Metabolism, Antioxidant Activity, Bio-Active Components, and Health Benefits. Compr Rev Food Sci Food Saf **9**: 261–269

Trethewey JAK, Campbell LM, Harris PJ (2005) (1->3),(1->4)-{beta}-d-Glucans in the cell walls of the Poales (sensu lato): an immunogold labeling study using a monoclonal antibody. Am J Bot **92**: 1660–74

Tripathy JN, Zhang J, Robin S, Nguyen TT, Nguyen HT (2000) QTLs for cell-membrane stability mapped in rice (Oryza sativa L.) under drought stress. TAG Theor Appl Genet **100**: 1197–1202

Tsekos I (1999) THE SITES OF CELLULOSE SYNTHESIS IN ALGAE: DIVERSITY AND EVOLUTION OF CELLULOSE-SYNTHESIZING ENZYME COMPLEXES. J Phycol **35**: 635–655

Uddin MN, Hanstein S, Leubner R, Schubert S (2013) Leaf Cell-Wall Components as Influenced in the First Phase of Salt Stress in Three Maize (Zea mays L.) Hybrids Differing in Salt Resistance. J Agron Crop Sci **199**: 405–415

Valentim-Neto PA, Rossi GB, Anacleto KB, de Mello CS, Balsamo GM, Arisi ACM (2015) Leaf proteome comparison of two GM common bean varieties and their non-GM counterparts by principal component analysis. J Sci Food Agric. doi: 10.1002/jsfa.7166

Vanholme R, Demedts B, Morreel K, Ralph J, Boerjan W (2010) Lignin biosynthesis and structure. Plant Physiol **153**: 895–905

Vanholme R, Morreel K, Darrah C, Oyarce P, Grabber JH, Ralph J, Boerjan W (2012) Metabolic engineering of novel lignin in biomass crops. New Phytol **196**: 978–1000

Vanzin GF, Madson M, Carpita NC, Raikhel N V, Keegstra K, Reiter W-D (2002) The mur2 mutant of Arabidopsis thaliana lacks fucosylated xyloglucan because of a lesion in fucosyltransferase AtFUT1. Proc Natl Acad Sci U S A **99**: 3340–5

Velasquez M, Salter JS, Dorosz JG, Petersen BL, Estevez JM (2012) Recent Advances on the Posttranslational Modifications of EXTs and Their Roles in Plant Cell Walls. Front Plant Sci **3**: 93

Velasquez SM, Ricardi MM, Dorosz JG, Fernandez P V, Nadra AD, Pol-Fachin L, Egelund J, Gille S, Harholt J, Ciancia M, et al (2011) O-glycosylated cell wall proteins are essential in root hair growth. Science **332**: 1401–3

Vidal S, Doco T, Williams P, Pellerin P, York WS, O'Neill MA, Glushka J, Darvill AG, Albersheim P (2000) Structural characterization of the pectic polysaccharide rhamnogalacturonan II: evidence for the backbone location of the aceric acid-containing oligoglycosyl side chain. Carbohydr Res **326**: 277–94

Vincken J-P (2003) If Homogalacturonan Were a Side Chain of Rhamnogalacturonan I. Implications for Cell Wall Architecture. PLANT Physiol **132**: 1781–1789

Vogel J (2008) Unique aspects of the grass cell wall. 301–307

Voragen AGJ, Coenen G-J, Verhoef RP, Schols HA (2009) Pectin, a versatile polysaccharide present in plant cell walls. Struct Chem **20**: 263–275

Vorwerk S, Somerville S, Somerville C (2004) The role of plant cell wall polysaccharide composition in disease resistance. Trends Plant Sci **9**: 203–9

Wang Z, Hobson N, Galindo L, Zhu S, Shi D, McDill J, Yang L, Hawkins S, Neutelings G, Datla R, et al (2012) The genome of flax (Linum usitatissimum) assembled de novo from short shotgun sequence reads. Plant J **72**: no–no

Whitcombe AJ, O'Neill MA, Steffan W, Albersheim P, Darvill AG (1995) Structural characterization of the pectic polysaccharide, rhamnogalacturonan-II. Carbohydr Res **271**: 15–29

Willats WG, Knox JP (1996) A role for arabinogalactan-proteins in plant cell expansion: evidence from studies on the interaction of beta-glucosyl Yariv reagent with seedlings of Arabidopsis thaliana. Plant J **9**: 919–25

Willats WG, McCartney L, Mackie W, Knox JP (2001) Pectin: cell biology and prospects for functional analysis. Plant Mol Biol **47**: 9–27

Wodniok S, Brinkmann H, Glöckner G, Heidel AJ, Philippe H, Melkonian M (2011) Origin of land plants : Do conjugating green algae hold the key ? BMC Evol Biol **11**: 104

Wolf S, Rausch T, Greiner S (2009) The N-terminal pro region mediates retention of unprocessed type-I PME in the Golgi apparatus. Plant J **58**: 361–75

Wróbel-Kwiatkowska M, Czemplik M, Kulma A, Zuk M, Kaczmar J, Dymińska L, Hanuza J, Ptak M, Szopa J (2012) New biocomposites based on bioplastic flax fibers and biodegradable polymers. Biotechnol Prog **28**: 1336–46

Wu Y, Cosgrove DJ (2000) Adaptation of roots to low water potentials by changes in cell wall extensibility and cell wall proteins. J Exp Bot **51**: 1543–53

Xu H, Ke K, Zhang Z, Luo X, Jin X, Liu SS-Y, Fan Z (2013) Effects of platelet-rich plasma and recombinant human bone morphogenetic protein-2 on suture distraction osteogenesis. J Craniofac Surg **24**: 645–50

Yang L, Wang CC, Guo WD, Li XB, Lu M, Yu CL (2006) Differential expression of cell wall related genes in the elongation zone of rice roots under water deficit. Russ J Plant Physiol **53**: 390–395

Yang Y, Zhu K, Xia H, Chen L, Chen K (2014) Comparative proteomic analysis of indica and japonica rice varieties. Genet Mol Biol **37**: 652–61

Yi H, Hwang KT, Regenstein JM, Shin SW (2014) Fatty Acid Composition and Sensory Characteristics of Eggs Obtained from Hens Fed Flaxseed Oil, Dried Whitebait and/or Fructo-oligosaccharide. Asian-Australasian J Anim Sci **27**: 1026–34

Zabotina OA (2012) Xyloglucan and Its Biosynthesis. Front Plant Sci **3**: 134

Zeid IM, Shedeed ZA (2006) Response of alfalfa to putrescine treatment under drought stress. Biol Plant **50**: 635–640

Zeng W, Jiang N, Nadella R, Killen TL, Nadella V, Faik A (2010) A glucurono(arabino)xylan synthase complex from wheat contains members of the GT43, GT47, and GT75 families and functions cooperatively. Plant Physiol **154**: 78–97

Zhong R, Peña MJ, Zhou G-K, Nairn CJ, Wood-Jones A, Richardson EA, Morrison WH, Darvill AG, York WS, Ye Z-H (2005) Arabidopsis fragile fiber8, which encodes a putative glucuronyltransferase, is essential for normal secondary wall synthesis. Plant Cell **17**: 3390–408

Zhong R, Ye Z (2014) Secondary Cell Walls: Biosynthesis, Patterned Deposition and Transcriptional Regulation. Plant Cell Physiol **0**: 1–20

Zini E, Scandola M, Gatenholm P (2003) Heterogeneous acylation of flax fibers. Reaction kinetics and surface properties. Biomacromolecules **4**: 821–7

Résumé

Le lin (*Linum usitatissimum* L.) est cultivé pour ses fibres riches en cellulose utilisées dans l'industrie textile et pour l'élaboration des matériaux composites. La « qualité » des fibres dépend, en partie, de la structure de la paroi cellulaire et nous avons donc essayé de mieux comprendre les différents facteurs pouvant impacter sur la composition des parois cellulaires chez le lin. Dans un premier temps nous avons validé un nouvel support de microarray de type Nimblegen en passant d'un système à base d'oligonucléotides courts (25-mer) à une version avec oligonucléootides longs (60-mer) pour des analyses de transcriptomique. Ensuite une étude protéomique sur plusieurs organes végétatif nous a permis d'identifier 1242 protéines non-redondantes dont 410 sont associées au métabolisme pariét*al.* En parallèle nous avons démontré la présence des hémicellulose de type xyloglucane dans les parois des fibres de lin et mis en évidence une importante paralogie de la famille IIIA des XTHs (xyloglucan endo-transglycosylase/hydrolase) potentiellement impliqué dans la formation/structuration de la paroi des fibres de lin. Puis, une comparaison transcriptomique et protéomique entre différentes variétés de lin (fibre printemps, fibre hiver, huile hiver) cultivées au champ sur 2 années consécutives a permis d'identifier 659 gènes différentiellement exprimés (DEGs) au niveau variétale, et 1571 DEGs au niveau environnement*al.* Un nombre non-négligeable de ces gènes est impliqué dans le métabolisme pariétal permettant ainsi de fournir les premières indices expliquant le lien entre variété et qualité. Cette dernière étude a également souligné l'importance potentielle de la protéine XTH dans le métabolisme pariétal du lin. Le rôle de l'environnement sur le métabolisme pariétale était explorée d'avantage dans une étude visant à disséquer l'impact d'un stress hydrique sur les transcriptomes de 3 organes végétatifs (tige, feuille, racine). Les analyses préliminaires ont identifié un nombre important de DEGs impliqués dans la biosynthèse et le remodelage de plusieurs polymères pariétaux

Abstract

Flax (Linum usitatissimum L.) is grown for its cellulose-rich bast fibers used in the textile industry and for reinforcing composite materials. Fiber “quality” depends partly on the structure of the cell wall and we have therefore tried to obtain a better understanding of the different factors that can influence the structure of flax cell walls. We firstly confirmed the use of a new Nimblegen microarray changing from a system based on short (25-mer) oligonucleotides to a system based on long oligonucleotides (60 mers). A proteomics approach was then used and allowed us to identify 1,242 non-redundant proteins of which 410 could be related to cell wall metabolism. In parallel we demonstrated the presence of xyloglucan hemicelluloses in flax fiber cell walls and identified an important paralogy in the IIIA XTH (xyloglucan endo-transglycosylase/hydrolase) family potentially implicated in the formation/structuration of the flax fiber cell wall. Then a transcriptomic and proteomic comparison between different field-grown flax varieties (spring fiber, winter fiber, winter oil) over 2 consecutive years allowed us to identify 659 differentially-expressed genes (DEGs) at the variety level, and 1,571 genes at the environmental level. A non-negligible number of these genes is involved in cell wall metabolism thereby providing some initial clues allowing a link to be made between variety and quality. This study also underlined the potential importance of the XTH protein in flax cell wall metabolism. The role of the environment on cell wall metabolism was further explored in a study aiming to dissect the impact of drought stress on the transcriptomes of 3 vegetative organs (stem, leaf, root). Preliminary analyses identified an important number of DEGs involved in the biosynthesis and remodeling of several cell wall polymers.

The Plant Cell, Vol. 25: 3961–3975, October 2013, www.plantcell.org

Convergent Evolution of Polysaccharide Debranching Defines a Common Mechanism for Starch Accumulation in Cyanobacteria and Plants[W]

Ugo Cenci,[a] Malika Chabi,[a,1] Mathieu Ducatez,[a,1] Catherine Tirtiaux,[a] Jennifer Nirmal-Raj,[a] Yoshinori Utsumi,[b] Daiki Kobayashi,[b] Satoshi Sasaki,[b] Eiji Suzuki,[b] Yasunori Nakamura,[b] Jean-Luc Putaux,[c] Xavier Roussel,[a] Amandine Durand-Terrasson,[c] Debashish Bhattacharya,[d] Anne-Sophie Vercoutter-Edouart,[a] Emmanuel Maes,[a] Maria Cecilia Arias,[a] Monica Palcic,[e] Lyann Sim,[e] Steven G. Ball,[a] and Christophe Colleoni[a,2]

[a] Université des Sciences et Technologies de Lille, Unité de Glycobiologie Structurale et Fonctionnelle, Unité Mixte de Recherche 8576, Centre National de la Recherche Scientifique, Cité Scientifique, 59655 Villeneuve d'Ascq cedex, France

[b] Department of Biological Production, Akita Prefectural University, Akita 010-0195, Japan

[c] Centre de Recherches sur Les Macromolécules Végétales (Centre National de la Recherche Scientifique), F-38041 Grenoble cedex 9, France (affiliated with Université Joseph Fourier and Member of Institut de Chimie Moléculaire de Grenoble and Insitut Carnot PolyNat)

[d] Department of Ecology, Evolution, and Natural Resources, Institute of Marine and Coastal Sciences, Rutgers University, New Brunswick, New Jersey 08901

[e] Carlsberg Laboratory, Copenhagen V DK-1799, Denmark

ORCID IDs: 0000-0003-4941-4266 (M.D.); 0000-0002-3981-7806 (J.N.-R.); 0000-0002-9196-7123 (Y.U.); 0000-0002-0153-7497 (E.S.); 0000-0002-9760-5369 (J.-L.P.); 0000-0003-2923-4559 (X.R.); 0000-0003-0611-1273 (D.B.); 0000-0002-4193-1309 (M.C.A.); 0000-0002-5221-089X (M.P.); 0000-0003-1629-1650 (S.B.); 0000-0002-6534-6120 (C.C.).

Starch, unlike hydrosoluble glycogen particles, aggregates into insoluble, semicrystalline granules. In photosynthetic eukaryotes, the transition to starch accumulation occurred after plastid endosymbiosis from a preexisting cytosolic host glycogen metabolism network. This involved the recruitment of a debranching enzyme of chlamydial pathogen origin. The latter is thought to be responsible for removing misplaced branches that would otherwise yield a water-soluble polysaccharide. We now report the implication of starch debranching enzyme in the aggregation of semicrystalline granules of single-cell cyanobacteria that accumulate both glycogen and starch-like polymers. We show that an enzyme of analogous nature to the plant debranching enzyme but of a different bacterial origin was recruited for the same purpose in these organisms. Remarkably, both the plant and cyanobacterial enzymes have evolved through convergent evolution, showing novel yet identical substrate specificities from a preexisting enzyme that originally displayed the much narrower substrate preferences required for glycogen catabolism.

INTRODUCTION

Bacteria, Archaea, and eukaryotes often store Glc in the form of glycogen. This storage material consists of small hydrosoluble particles composed of α-1,4–linked and α-1,6–branched glucan chains (Shearer and Graham, 2002). Because current structure models envision that each α-1,4–linked chain supports on average two novel branched chains, mathematical modeling predicts that chain density will increase with particle size up to a maximal possible diameter of 42 nm. Such glycogen particles contain up to 55,000 Glc residues with over 36% resting in the outer particle chains (Meléndez et al., 1999). These are thus readily accessible to cell metabolism without the need for polysaccharide debranching. Hence, in glycogen, Glc residues remain rapidly available to cellular enzymes as if they were in the soluble phase but remain much less active osmotically. Archaeplastida (also called Kingdom Plantae) consisting of the Chloroplastida (green algae and all land plants), the Rhodophyceae (red algae), and the Glaucophyta (glaucophytes) store starch granules of potentially unlimited size and no glycogen (Ball et al., 2011). Starch always contains a glycogen-like polymer named amylopectin occasionally mixed with a very moderately branched amylose polysaccharide. Amylopectin aggregates into semicrystalline granules of potentially unlimited diameter. This organization results from an asymmetrical distribution of branches, allowing the formation of double helical structures that align and crystallize into two different allomorphs (the so-called A and B) or a mixture of both (Buléon et al., 1998). This aggregation directly affects the properties of most food sources, including their digestibility as well as the functional properties in all nonfood uses of the polymer.

Starch is osmotically inert, allowing for the accumulation of very large quantities of Glc (60 to 90% of the dry weight) in the storage organs of plants. Until recently, the distribution of starch

[1] These authors contributed equally to this work.

[2] Address correspondence to christophe.colleoni@univ-lille1.fr.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Christophe Colleoni (christophe.colleoni@univ-lille1.fr).

[W] Online version contains Web-only data.

www.plantcell.org/cgi/doi/10.1105/tpc.113.118174

seemed restricted to photosynthetic eukaryotes, including several secondary endosymbiosis lineages derived from Archaeplastida, such as the cryptophytes (Deschamps et al., 2006), the dinoflagellates (Dauvillée et al., 2009), and some apicomplexa parasites (Coppin et al., 2005). However, more recent studies revealed the existence of starch-like structures in unicellular diazotrophic cyanobacteria belonging to the order Chroococcales (Nakamura et al., 2005; Deschamps et al., 2008; Suzuki et al., 2013). The presence of anomalous glycogen particles had already been identified previously in this clade, while it is only very recently that this material was recognized as starch-like and the term "semi-amylopectin" was coined to describe the major amylopectin-like fraction within these granules (Schneegurt et al., 1994; Suzuki et al., 2013). Four out of the six reported starch-accumulating cyanobacteria strains accumulate only this polysaccharide fraction. However, two different strains also synthesize amylose using an enzyme phylogenetically related to the archaeplastidial granule-bound starch synthase (GBSS), an enzyme known to be selectively responsible for the synthesis of this fraction in plants (Delrue et al., 1992; Deschamps et al., 2008).

Clues as to the nature of the biochemical mechanism distinguishing starch from glycogen synthesis came from the study of glycogen-accumulating mutants of *Chlamydomonas* and cereals that proved defective for the same GH13-type (for Glycosyl Hydrolase family 13, according to the carbohydrate active enzyme (CAZy) classification) of debranching enzyme (DBE; James et al., 1995; Mouille et al., 1996; Kubo et al., 1999). In the green algae, the substitution of starch by glycogen was complete, thereby hinting that the absence of DBE prevented amylopectin synthesis altogether. This enzyme thus catalyzed an essential previously unrecognized step distinguishing starch from glycogen synthesis. From these observations, several groups proposed that DBE was selectively responsible for removal of misplaced branches, preventing polysaccharide aggregation in an otherwise hydrosoluble precursor (Ball et al., 1996). This model, known as the preamylopectin trimming model, is well sustained by the evidence gathered through mutant analysis in all plant systems examined to date, including *Arabidopsis thaliana* (Wattebled et al., 2005). However, direct proof of this model will only be obtained by achieving in vitro synthesis of starch granules. The DBEs missing in the glycogen-accumulating plants are known as isoamylases and belong to a family of direct DBEs that apart from Archaeplastida are only distributed in bacteria. In bacteria, such enzymes, generally named GlgX (by reference to the *Escherichia coli* glycogen metabolism locus coding it), are known to be involved in glycogen catabolism by debranching only those external chains that have been first recessed by the catabolic enzyme glycogen phosphorylase. The *E. coli* enzyme is thus known to be very selective for chains of 3-4 Glc residues left over by glycogen/starch phosphorylase action on external chains and displays very little residual activity on longer chains (Jeanningros et al., 1976; Dauvillée et al., 2005). This restricted enzyme selectivity prevents futile cycles during polysaccharide synthesis since such activities are unable to debranch directly the products of branching enzyme activity that branches a minimum of six Glc residues on an acceptor chain. In line with their supposed catabolic function, the *E. coli* GlgX defective mutants overproduce glycogen with short external chains (Dauvillée et al., 2005). Interestingly, analogous results have also been obtained in glycogen-accumulating cyanobacteria (Suzuki et al., 2007).

We now report the existence of both glycogen and starch in a recently axenized marine *Cyanobacterium* (strain CLg1) related isolate. We show through the isolation of an allelic series of 14 distinct mutants that alteration, decrease, or disappearance of a particular GH13 GlgX type of DBE correlates with the disappearance of the starch fraction and a large increase in the accumulation of glycogen. Remarkably, the wild-type enzyme defective in the mutants has evolved an isoamylase-type of

Table 1. Polysaccharide Composition in Wild-Type and Class A Mutant Strains

	Starch Granules		WSP		% of Each Polysaccharide	
Strain	mg/mg of Protein	%	mg/mg of Protein	%	Starch	WSP
Wild type	2.80 ± 0.36	100	0.50 ± 0.03	100	84.7	15.3
174H3	0.08 ± 0.02	3	1.89 ± 0.57	375	4.3	95.7
107D8	0.04 ± 0.01	1.5	0.87 ± 0.23	173	4.6	95.4
91G1	0.09 ± 0.04	3.3	1.57 ± 0.4	312	5.5	94.5
153H12	0.22 ± 0.20	7.7	1.47 ± 0.38	292	12.8	87.2
123B3	0.15 ± 0.05	5.5	1.38 ± 0.14	273	10.1	89.9
99A7	0.09 ± 0.03	3.3	2.69 ± 1.10	533	3.3	96.7
21'C6	0.07 ± 0.02	2.6	2.04 ± 0.42	403	3.4	96.6
118H9	0.05 ± 0.03	1.7	1.64 ± 0.83	324	2.9	97.1
91D6	0.07 ± 0.03	2.45	1.99 ± 0.35	395	3.3	96.7
175E12	0.12 ± 0.06	4.2	2.19 ± 0.72	435	5	95.0
174B6	0.17 ± 0.1	6	1.64 ± 0.51	326	9.3	90.7
134A4	0.06 ± 0.03	2.1	1.80 ± 0.70	354	3.2	96.8
81G3	0.07 ± 0.02	2.6	2.37 ± 0.1	469	3	97.0
80D5	0.07 ± 0.04	2.5	2.13 ± 0.2	422	3.2	96.8

The measurements are means ± SD of three independent extractions.

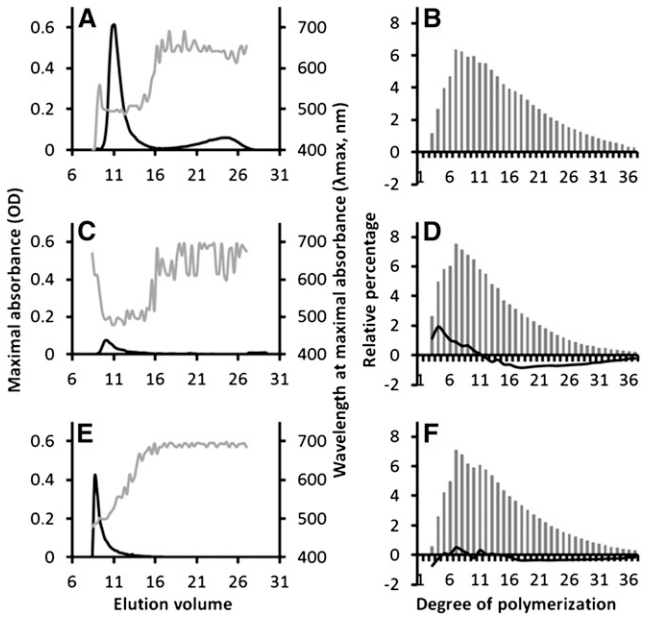


Figure 1. Structural Analyses of Starch-Like Granules Accumulated by the Wild-Type Strain and Class A Mutants.

Insoluble water polysaccharides purified from the wild type **(A)**. A severe mutant 99A7 **(C)** and an intermediate mutant 153H12 **(E)** were subjected to size-exclusion chromatography analysis (CL2B sepharose). The wavelength (λ_{max}, gray lines) at the maximal absorbance (black lines) of the iodine-polysaccharide complex was determined for each fraction. As previously described, starch-like granules of the wild type are composed of semiamylopectin (fractions 9 to 16) and amylose-like polysaccharides (fractions 17 to 27). Both starch-like granules of class A mutants are composed exclusively of abnormal semiamylopectin fraction (λ_{max} values ranging between 490 and 500 nm instead of 507 nm for the wild-type semiamylopectin). Fractions containing semiamylopectin material were pooled and subjected to chain-length distribution analysis. After complete digestion with commercial isoamylase, chains of Glc were separated according to their DP by HPAEC-PAD. The relative abundance for each DP (gray bars) was determined for the wild type **(B)**, the 99A7 mutant **(D)**, and the 153H12 mutant **(F)** from three independent extractions. Subtractive analyses (percentage of each DP in the mutant's semiamylopectin minus percentage of each DP in the wild type's semiamylopectin), depicted as black lines in **(D)** and **(F)**, reveal an increase of short chains (DP 3 to 10) and a decrease in long-chain content (DP 12 to 35) in the severe mutant 99A7 (in **[D]**) and no significant difference in the intermediate 153H12 mutant (in **[F]**).

activity very similar to that present in plants. This activity evolved, as was the case for the plant enzyme from preexisting GlgX-like proteins. However, phylogenetic analysis proves that the phylogenetic origin of this gene is independent from the enzyme recruited by the Archaeplastida. These results point to the independent acquisition of starch metabolism in cyanobacteria and plants and suggest that polysaccharide debranching by an isoamylase-like enzyme defines a common mechanism for the synthesis and aggregation of starch polymers from enzymes of a preexisting glycogen metabolism network.

RESULTS

Cyanobacterium sp CLg1 Accumulates Both Glycogen and Starch from a Gene-Rich Suite of Enzymes of Bacterial Glycogen Metabolism

Unlike Archaeplastida, which only contain starch like material, the wild-type axenized strain CLg1 always accumulates two

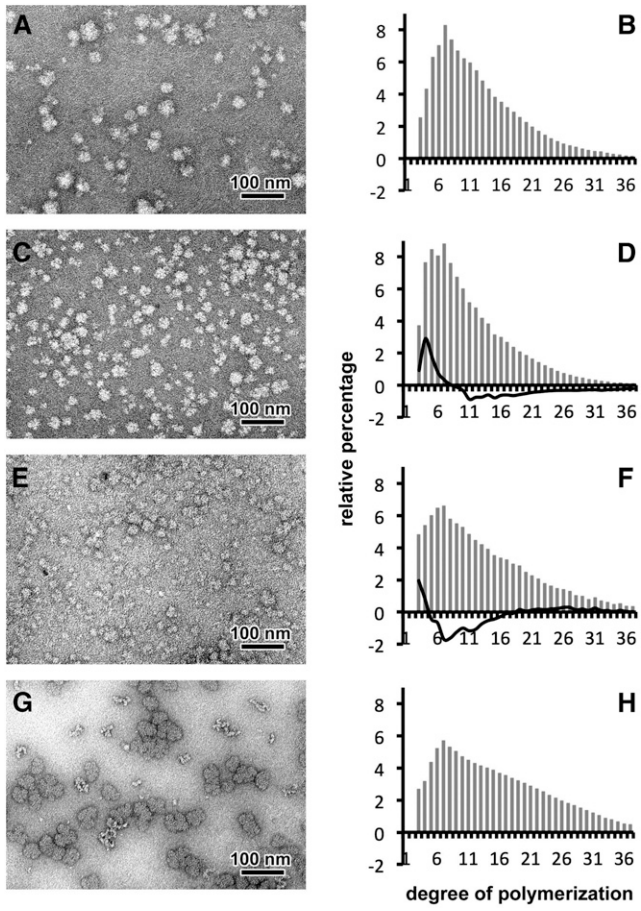


Figure 2. Characterization of WSPs Accumulated in the Wild-Type Strain and Class A Mutants of *Cyanobacterium* sp CLg1.

Negative staining following TEM observations suggest that WSP of the wild type **(A)**, the 99A7 mutant **(C)**, and the 153H12 mutant **(E)** are highly branched polysaccharides with a diameter below 50 nm similar to glycogen particles of rabbit liver **(G)**. After purification and complete digestion with a commercial isoamylase, chains of Glc were separated according to their DP by HPAEC-PAD. From three independent extractions, the relative abundance for each DP (gray bars) was determined for the wild-type strain **(B)**, the severe mutant 99A7 **(D)**, the intermediate 153H12 mutant **(F)**, and glycogen from rabbit liver **(H)**. Subtractive analyses (percentage of each DP of mutant minus percentage of each DP in the wild type), depicted as black lines in **(D)** and **(F)**, reveal an increase of short chains (DP 3 to 7) and a defect in long-chain content (DP 12 to 35) in the WSPs accumulated in the mutant 99A7 (in **[D]**) and in the intermediate mutant 153H12 (in **[F]**). WSP appears to contain fewer chains with a DP ranging between 4 and 16.

distinct α-1,4–linked and α-1,6–branched polysaccharides fractions in the form of both a major (85%) insoluble and a minor (15%) yet significant water-soluble fraction (Table 1). We confirm that the structure of the insoluble polysaccharide fraction (Figures 1A and 1B) was composed of starch-like granules containing both semicrystalline amylopectin and amylose with chain-length distribution consistent with the presence of starch as previously described (Deschamps et al., 2008). Interestingly, in addition to this major starch-like fraction, a small hydrosoluble fraction representing 15% of the total glucan amount under nitrogen starvation conditions was also observed. The chain-length distribution (Figure 2B) and the negative staining transmission electronic microscopy (TEM) observations (Figure 2A) suggest that this material is a highly branched soluble polysaccharide capable of excluding the uranyl acetate molecule in a fashion similar to rabbit liver glycogen particles (Figure 2G). In addition, TEM images of wild-type CLg1 strain ultrathin sections, in which polysaccharides were stained by the periodic acid thio-carbohydrazide-silver proteinate method, confirm the presence not only of the very obvious large-size starch granules but also of the smaller size glycogen-like particles (Figure 3A). To get a better idea of the nature of the biochemical pathways explaining the presence of these two fractions and to better characterize the nature of the CLg1 strain, we examined the recently reported sequence of CLg1 strain genome (Suzuki et al., 2013). The 16 rRNA and NifH phylogenetic analysis revealed a close relationship to the genus *Cyanobacterium* (Falcon et al., 2004; Suzuki et al., 2013). This strain, although containing the genes required for nitrogen fixation, remains unable to grow without reduced nitrogen under laboratory conditions.

Surprisingly, the network evidenced is a rather complex set of bacterial and archean distributed enzymes of glycogen metabolism containing, among others, three glycogen/starch synthases, four glycogen/starch branching enzymes, two GlgX-like DBE sequences, two amylopullulanases, and one indirect DBE–like α-1,6 glucosidase domain (Colleoni and Suzuki, 2012). With two exceptions, no traces of enzymes of eukaryotic affiliation involved in glycogen or starch metabolism were found in this genome. The important exceptions consist of GBSSI and ADP-Glc pyrophosphorylase, which were probably donated to the Archaeplastida through endosymbiotic gene transfer after plastid endosymbiosis. The affiliation of one of the soluble glycogen/starch synthases to the SSIII-IV group of enzymes of Archaeplastida is indirect, and this enzyme groups with a number of other proteobacteria and chlamydia sequences. In this case, either the proteobacteria, or more probably the chlamydia, can be considered as the direct source of the archaeplastidal enzymes (Ball et al., 2013).

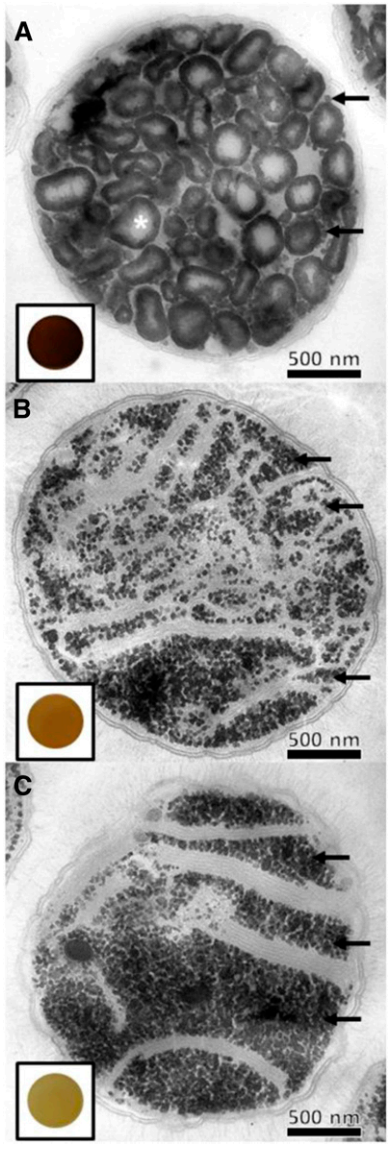


Figure 3. TEM Observations of ultrathin sections (60 nm) of Wild-Type and Class A Mutant Strains.

Polysaccharide contents of the wild-type CLg1 strain **(A)**, and the 99A7 **(B)** and 175E12 **(C)** mutant strains were observed after periodic acid thio-carbohydrazide-silver proteinate staining. Both starch-like granules (white star) and glycogen particles (dots indicated with black arrows) are witnessed in the wild-type strain. Starch granules are missing and substituted by a large amount of glycogen particles in class A mutants. The dark-blue iodine stain from a cell patch of the wild-type strain (framed in **[A]**) is displayed. By contrast, the absence of starch granules in class A mutants is correlated by a yellow-orange stain of cell patches after spraying iodine vapors (framed in **[B]** and **[C]**).

Selection and Characterization of the Polysaccharide Structure from Glycogen-Accumulating Mutants of *Cyanobacterium* sp CLg1

Because of the very slow growth rate of the CLg1 strain and the absence of established transgenesis procedures for this organism, we chose a forward genetic approach (UV mutagenesis) to dissect the genetic determinants of starch metabolism in cyanobacteria (for details, see Methods). Five phenotypic classes of mutants were found after 3 years of screening, segregation, and purification (see Supplemental Figure 1 online). We benefited from the presence of amylose, which yielded a very strong and sensitive iodine stain for the screening of 2.10^4 cell patches. Of relevance to this study was the finding of a class of

mutants (class A), which lacked interaction with iodine and yielded low starch and abundant water-soluble polysaccharides (WSPs). Figure 3 shows that 14 mutants of this class accumulated abundant small size granules resembling the minor glycogen fraction of the wild-type reference. Table 1 shows that 11 out of 14 mutants of this class witness a 94 to 97% decrease of insoluble granules and a replacement by a significant amount of hydrosoluble polysaccharide. However, three additional mutant strains displayed a phenotype intermediate between standard class A mutants and the wild type. The latter accumulated up to 9 to 12% of the wild-type amount of starch and over-accumulated glycogen (Table 1). This will be referred to as the incompletely defective class A mutants. These mutants also gave a clear mutant iodine-stain phenotype of cell patches (see Supplemental Figure 1 online). The substitution of starch by glycogen restricted the total amount of stored glucans in all mutants from two- to threefold. Purification of the water-soluble fraction followed by enzymatic debranching and analysis of the chain-length distribution of the resulting glucans yielded a chain length that mimicked those of glycogen and clearly differed from that of amylopectin-like polymers purified from the wild-type CLg1 reference (Figure 2). Although similar to the chain-length distribution of the minor glycogen fraction from the wild-type reference, the mutant WSP was selectively enriched in very small chains (degree of polymerization [DP] 3 to 6) when compared with the wild type. A very minor amount (2 to 6% of the wild-type mount) of high molecular mass material could still be purified from the glycogen accumulating mutants. This material resembled the polysaccharide accumulating in the single incompletely defective class A mutant 153H12 (12%). Interestingly, the amylopectin-like polysaccharide was equally enriched in very small glucans when compared with the wild type.

The Glycogen-Accumulating Cyanobacteria Are Defective for a Cation Requiring DBE Activity

To get a better understanding of the underlying biochemical cause of the glycogen-accumulating phenotype, we undertook a series of zymogram and enzyme assays for all possible enzymes of bacterial glycogen/starch metabolism. We were unable to find any significant difference in our crude extract assays. A notable increase of one starch phosphorylase isoform was noted through zymograms of the mutants (see Supplemental Figure 2 online). However, this increase was noted in all mutant types that displayed a decrease in total polysaccharide amount and not only in Class A mutants. It thus seems to define

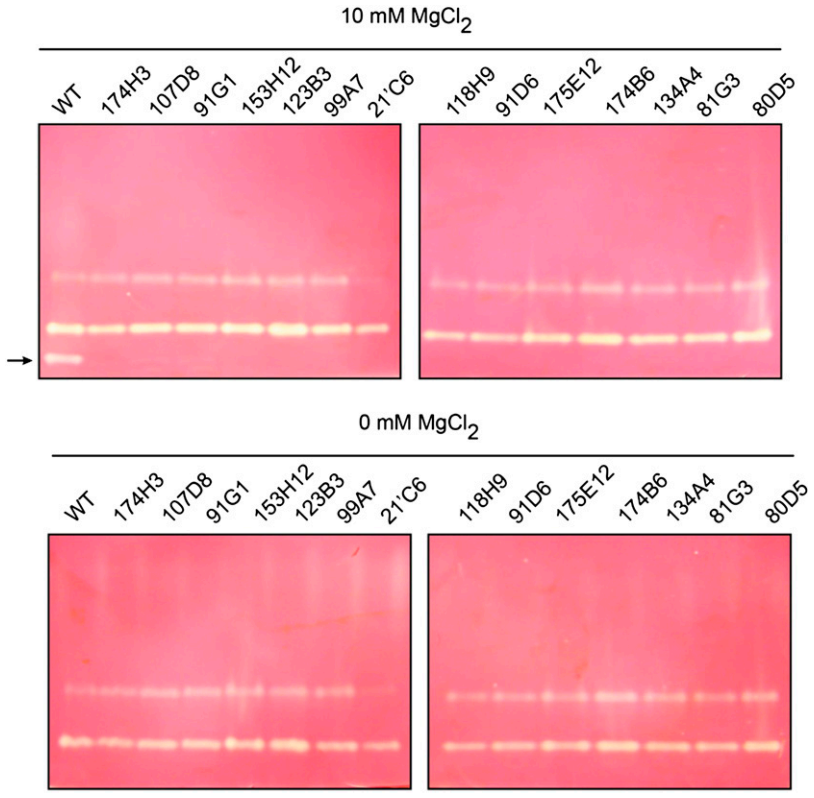


Figure 4. Zymogram Analysis of Starch-Metabolizing Enzymes from Wild-Type and Class A Mutants.

Total protein of wild-type (WT) and class A mutant strains (21'C6, 80D5, 81G3, 91D6, 91G1, 99A7, 107D8, 118H9, 123B3, 134A4, 153H12, 174B6, 174H3, and 175E12) were separated by native PAGE, followed by transfer of proteins to a native PAGE containing 0.6% (p/v) amylopectin. The native gels were then incubated with **(A)** or without **(B)** 10 mM $MgCl_2$. A cation-dependent activity enzyme is witnessed after iodine staining in the wild type's crude extract and disappears in all class A mutants (black arrow).

a universal secondary effect of altered starch metabolism not only in the *Cyanobacterium* sp CLg1 but also in other cyanobacteria species (Fu and Xu, 2006). Interestingly, the presence of white-staining glycogen-degrading bands that selectively disappeared from the class A mutants was evidenced from the glycogen/starch synthase gels but was not mirrored by the starch amylopectin- or glycogen-containing gels designed to emphasize starch modifying enzymes (e.g., amylases, branching enzymes, and α-1,4 glucanotransferase) (Figure 4). Because the procedures differed between the hydrolase and synthase zymograms with respect to pH and incubation buffers, we were able to test and narrow down the differences responsible for the identification of this activity selectively in the glycogen/starch synthase activity gels. We found that the missing glucan hydrolase required high levels of cation (see Supplemental Figure 3 online) that were supplied in the glycogen/starch synthase zymograms (i.e., Mg^{2+}) but not in our other procedures. Figure 3 shows that when starch- or amylopectin-containing gels were supplied with 10 mM $MgSO_4$, a blue staining glucan hydrolase band that disappeared in all class A mutants was evidenced. Interestingly, the incompletely defective class A mutant displays a slower migrating fainter activity. Because such a stain was indicative of the presence of a starch DBE, we purified the activity from wild-type cyanobacteria to near homogeneity. We then subjected amylopectin to the action of the purified enzyme and examined the proton nuclear magnetic resonance (NMR) spectra of the substrate before and after incubation with the glucan hydrolase. Results displayed in Figure 5 demonstrate that amylopectin was completely debranched and that the amount of branch hydrolysis corresponded quantitatively to the appearance of reducing-end resonance signals. These results allow us to identify the purified enzyme as a direct DBE. In addition, we were able to size the activity during our purification procedure as a polypeptide of 82 kD under denaturing conditions and with an apparent molecular mass of 247 kD under nondenaturing conditions (see Supplemental Figure 4 online), which suggests that the native protein eluted as a homotrimeric or as a homodimeric enzyme.

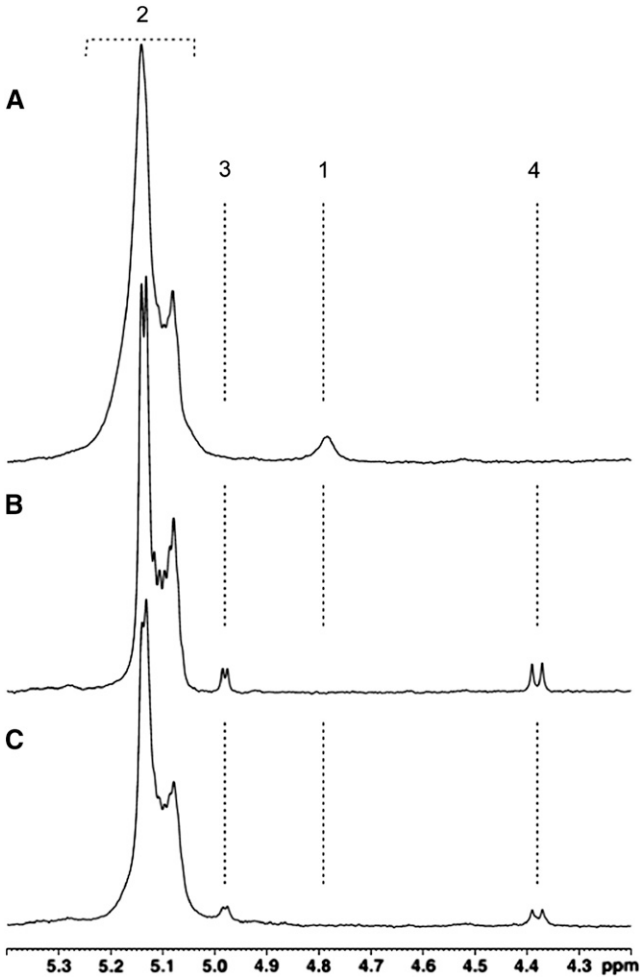


Figure 5. Part of the ^{1}H-NMR Spectra of Amylopectin in Dimethyl-Sulfoxide.

NMR analysis was performed on amylopectin **(A)** and amylopectin samples incubated overnight with a commercial isoamylase **(B)** or with active cation-dependent enzyme purified from wild-type strain **(C)**. Peak #2 (5.2 to 5.08 ppm) and peak #1 (4.79 ppm) represent the signals of ^{1}H engaged in α-1,4 and α-1,6 linkages, respectively. Both incubation experiments with amylopectin (**[B]** and **[C]**) result in the release of reducing ends and the apparition of α- (peak #3) and β-anomeric protons (peak #4) at 4.98 and 4.38 ppm, respectively. The presence of peaks #3 and #4 is correlated with a disappearance of the ^{1}H signal engaged in α-1,6 linkages (peak #1). The NMR spectrometry analysis suggests that cation-dependent activity specifically hydrolyzes α-1,6 linkages.

Glycogen-Accumulating Mutants Carry Mutations in a GlgX-Like DBE Activity

The presence of pure enzyme preparations enabled us to identify a dozen of oligopeptide sequences that matched the protein sequences deduced from the genome (see Supplemental Figure 5 online). We were thus able to identify peptides that corresponded to *GlgX2*, one of the two GlgX-like genes present on the *Cyanobacterium* sp Clg1 genome sequence. Highly pure enzyme preparations enabled us to size the protein and ascertain the possible presence of distinct enzyme subunits. No bands corresponding to the expected size of GlgX1 and no trypsic fragments matching this enzyme sequence could be retrieved. We therefore conclude that this enzyme is a homomultimer (dimer or trimer). We then sequenced all starch metabolism genes from the 11 class A and three incompletely defective class A mutants and systematically found mutations only in the GlgX2 sequence, altering the amino acid sequence of the protein, while no mutations could be found in any of the other starch metabolism genes. These mutations are summarized in Figure 6. Among these mutations, we found four non-sense mutations resulting in three premature stop codons, one of which was preceded by a frame shift; the other 10 mutants defined different missense alleles. Three of these missense alleles were responsible for the incompletely defective phenotype of the 153H12, 123B3, and 176B6 strains. Because we hereby define an allelic series of 11 (plus three incompletely defective) independent mutants, none of which carried additional mutations in starch metabolism genes (see Supplemental Table 1 online), we can safely conclude that all

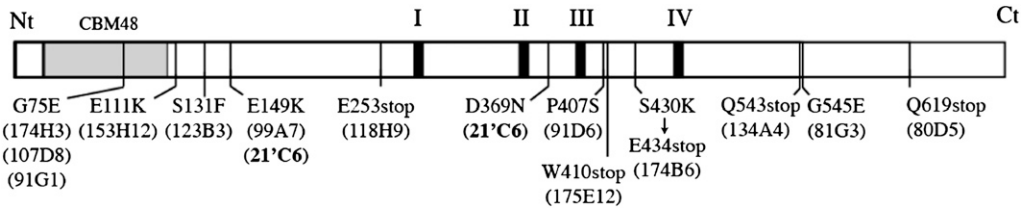


Figure 6. Various Allelic Mutations in the *Cyanobacterium* sp CLg1 glgX2 Gene.

Forward and reverse primers were designed in the untranslated region of the glgX2 gene. PCR reactions were performed on genomic DNA for each class A mutants. PCR products were cloned and DNA was sequenced on both strands. Punctual mutations (vertical black lines) found in each mutant (name in parentheses) are shown on the glgX2 segment. A mutant strain, 21'C6, harbors two point mutations (bold name). Regions I, II, III, and IV represent highly conserved sequences in the α-amylase family that includes acid residues involved in the catalytic site (Suzuki et al., 2007). The Carbohydrate Binding Module 48 (CBM 48; gray box) is observed at the N terminus (Nt) (Janeček et al., 2011).

defective phenotypes recorded in relation to polysaccharide accumulation result directly from the GlgX2 defect and not from other causes, obviating the need for formal complementation of the mutant gene through transgenesis with the wild-type sequence. Successful transgenesis has indeed not been reported in this subgroup of cyanobacteria, and all our attempts were unsuccessful.

The glgX-Like DBE Defines an Isoamylase Type of Activity

In bacteria and cyanobacteria, GlgX-type enzymes display a marked substrate preference for glycogen whose outer chains have been recessed with starch phosphorylase and little activity with either amylopectin or glycogen. The pure GlgX2 enzyme displayed little activity with such substrates or even with glycogen (Figure 7). However, unlike GlgX and in a fashion reminiscent of the plant isoamylase (Hussain et al., 2003), it displayed a marked preference for amylopectin, which it debranched to completion (Figure 7). Hence, the GlgX2 enzyme of *Cyanobacterium* more closely resembles the potato (*Solanum tuberosum*) isoamylase substrate specificity than that of the reference *Pseudomonas* sp enzyme. It certainly deviated from the bacterial GlgX type of activity classically involved in glycogen catabolism.

GlgX2 Defines the Enzyme Responsible for Starch Accumulation in *Cyanobacterium* Only

The mutants defined in this work establish that the *Cyanobacterium* sp CLg1 aggregates starch from a hydrosoluble precursor in a fashion very similar to starch in green algae and land plants. Since in both cases an enzyme from the same CAZy GH 13 subfamily 11 has been recruited to splice out those branches that prevent polysaccharide aggregation during synthesis, it is of interest to examine the phylogeny of these enzyme sequences and correlate the presence of selective enzyme forms to those of starch-like polymers. The phylogeny displayed in Figure 8 (see alignments in Supplemental Data Set 1 online) confirms that the source of the archaeplastidal DBEs cannot be traced to cyanobacteria. The strong monophyletic grouping of the archaeplastida enzymes with those of chlamydiales pathogens establishes the latter as the source of direct DBE, a bacteria-specific activity, in photosynthetic eukaryotes (Ball et al., 2013). In addition, the presence of the particular GlgX2-type enzyme does not correlate with the presence of starch in cyanobacteria. In Figure 8, all other documented starch-accumulating cyanobacteria lack this particular form of the GlgX enzyme, while the closest relatives to GlgX2 are reported to accumulate glycogen.

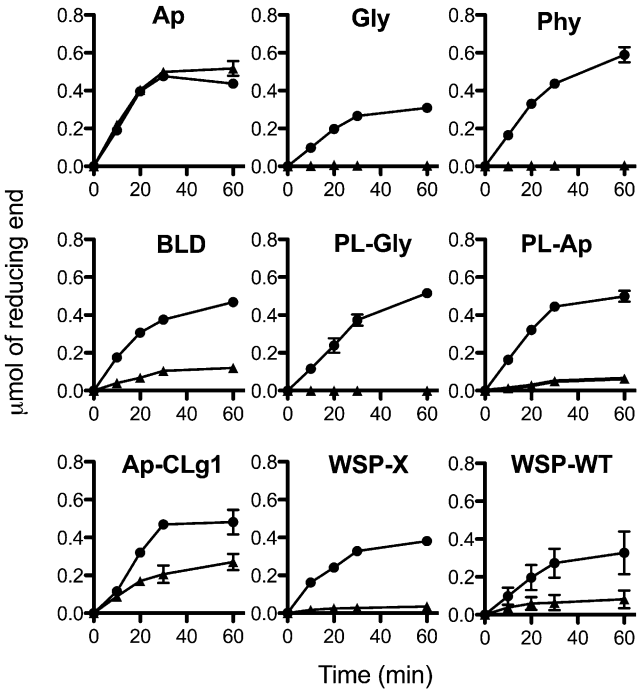


Figure 7. Specificity of the Cation-Dependent Debranching Activity toward Different Polysaccharides.

Both purified DBE activity (black triangle) and commercial isoamylase produced by *Pseudomonas* sp (megazyme; black circle) were incubated with 0.5% of amylopectin (Ap), glycogen (Gly), phytoglycogen (Phy), β-limit dextrin of amylopectin (BLD), phosphorylase-limit dextrin of glycogen (PL-gly), phosphorylase-limit dextrin of amylopectin (PL-Ap), semiamylopectin of the wild-type CLg1 strain (Ap-CLg1), water-soluble polysaccharide of class A mutant (WSP-X), and water polysaccharide of wild-type CLg1 strain (WSP-WT). The release of reducing ends was determined at 10, 20, 30, and 60 min of incubation using the 3,5-dinitro-2-hydroxybenzoic acid method. In order to compare both DBE activities, the specific activity of commercial isoamylase was adjusted at 0.5 μmol of reducing ends min^{-1} mL^{-1} using amylopectin as substrate. Data are the means ± SD of triplicates of two independent purification of cation-dependent debranching activity.

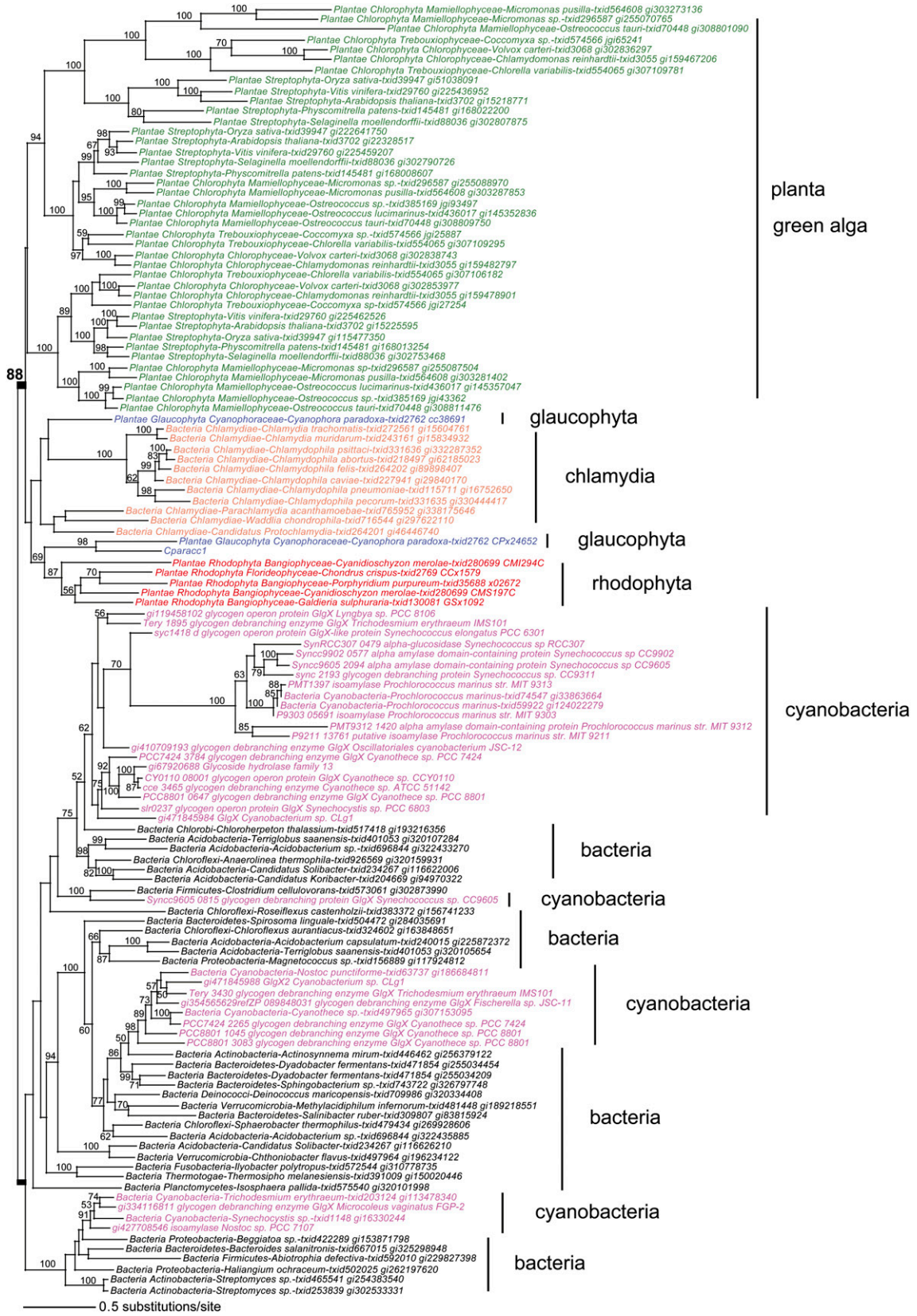


Figure 8. Phylogenetic Analysis of Cyanobacterial, Bacterial, and Archaeplastidal Glycogen DBEs.

Starch DBEs in Chloroplastida play an important role in polysaccharide synthesis and in starch degradation. The maximum likelihood phylogeny of these enzymes shows that the three isoforms of isoamylase in Chloroplastida (green text) and the DBE found in Glaucophyta (blue text) and Rhodophyta

We can therefore conclude that while acquisition of starch in Chloroplastida can still be suggested to be monophyletic (Cenci et al., 2013), transition of glycogen to starch metabolism in cyanobacteria cannot and defines distinctive polyphyletic events.

Substrate Preferences of Recombinant GlgX and Isoamylases from Cyanobacteria and Plants

As detailed above, the GlgX2 enzyme responsible for amylopectin synthesis of *Cyanobacterium* sp CLg1 has converged toward a substrate preference similar to that of the potato ISOAMYLASE1 (ISA1) enzyme. This result would suggest that this enzyme preference may be one of the important features required for the enzyme involved in amylopectin biosynthesis. In order to strengthen this conclusion, we correlated these results to the characterization of recombinant DBEs (rDBEs) produced from organisms well characterized with respect to the structure of the storage polysaccharide they produce. We thus chose to study five distinct rDBEs from four different species: two starch (rice [*Oryza sativa*] and *Cyanothece* ATCC51142) and two glycogen accumulators (*E. coli* and *Synechococcus elongatus* PCC7942). In rice, it is well established that the ISA1 enzyme is involved in the aggregation of starch granules, while ISA3 is devoted to hydrolysis of external branches present in soluble oligosaccharides produced by either phosphorylases or amylases during starch degradation. In this respect, the function performed by ISA3 is very similar to that of the *E. coli* GlgX enzyme. Figure 9 summarizes the substrate preferences of the purified recombinant enzyme, while the purity of the preparations is displayed in Supplemental Figure 6 online. Quite evidently, all enzymes suspected (*Cyanothece* ATCC51142) or proven (rice ISA1) to be involved in amylopectin crystallization display a clear preference for amylopectin debranching, while those involved in glycogen (*E. coli* and *S. elongatus* PCC7942) or starch (rice ISA3) breakdown show dramatically reversed preferences. These results underline the importance of the substrate preferences of direct DBEs for either glycogen or starch accumulation in a given organism.

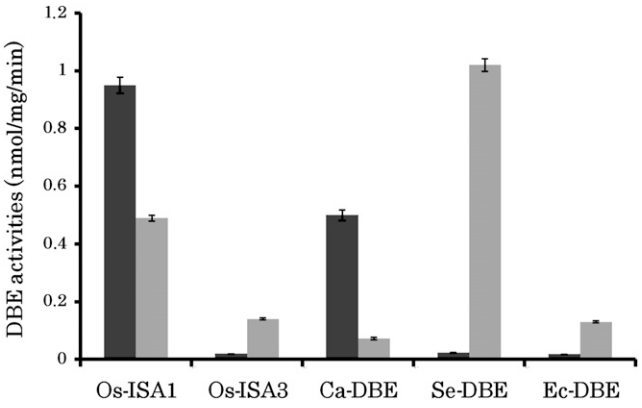


Figure 9. Activities of Isoamylase-Type DBEs from Various Sources toward Amylopectin and Phytoglycogen.

Enzymatic activities were determined by measuring the amounts of malto-oligosaccharides liberated after the enzymatic reactions with amylopectin (blue bars) or phytoglycogen (brown bars) as glucan substrate, as described in the Methods. The enzymes measured were Os-ISA1, rice ISA1; Os-ISA3, rice ISA3; Ce-ISA, *Cyanothece* ATCC51142 DBE; Se-DBE, *S. elongatus* PCC7942 DBE; and Ec-GlgX, *E. coli* GlgX. Vertical axis presents the debranching activity as expressed by nmol equivalents liberated/μg protein/min, except that the activities of Os-ISA1 were reduced to 0.025 as a result of to their much higher activities compared with other enzymes. Each value represents the mean ± SD of three separate measurements.

DISCUSSION

Cyanobacterium sp CLg1 Accumulates Both Glycogen and Cyanophycean Starch

The significant accumulation of both glycogen and starch-like material in wild-type axenized *Cyanobacterium* sp CLg1 is a distinctive feature of these organisms when compared with Archaeplastida. In the latter, with the exception of the red alga *Porphyridium sordidum* (Shimonaga et al., 2008), only starch is known to accumulate with just trace amounts of WSPs and oligosaccharides being sometimes detectable. In this study, however, under conditions of maximal polysaccharide synthesis, 10% of the total storage polysaccharide pool always accumulates in the form of glycogen. We deem this amount to be physiologically significant. The presence of both types of storage polysaccharides may in this case be required to optimize cyanobacterial physiology by making a significant pool of Glc available through a more dynamic and readily accessible form of storage. Indeed, in glycogen, the Glc of the outer chains may be easier to mobilize following fluctuations of physiology both in light and darkness. The crystalline starch-like granules, in turn, will offer a larger pool of osmotically inert carbohydrate stores for delayed use by nitrogenase or cell division at night (Schneegurt et al., 1994). It will be of interest to check for the presence of an analogous fraction in other starch-storing cyanobacteria. We confirm the presence of starch-like granules in *Cyanobacterium* sp CLg1 composed of both an amylopectin-like high mass fraction and a smaller amylose fraction. The chain-length distribution of the high mass polysaccharide complies to the definition given for semiamylopectin, as it contains fewer of those chains exceeding a DP of 40 (Nakamura

Figure 8. (continued).

(red text) do not share the same phylogenetic origin as the cyanobacterial GlgX (magenta text), particularly for *Cyanobacterium* sp CLg1. Rather, GlgX enzymes in Archaeplastida are sister to Chlamydiae (orange text; bootstrap support value is 90%), strongly suggesting a chlamydial origin of the gene in this eukaryotic supergroup (Ball et al., 2013). Alignments are given in Supplemental Data Set 1 online.

et al., 2005). Because the granules also contain a significant amount of amylose (5%), we propose to call this material cyanobacterial starch.

Evolution of Direct DBEs Involved in Amylopectin Maturation May Suffice to Explain the Transition from Glycogen to Starch Metabolism in All Living Cells

Convergent evolution in cyanobacteria and Archaeplastida underlines that polysaccharide debranching may define a common requirement for evolution of aggregated semicrystalline polysaccharides from glycogen metabolism (Cenci et al., 2013). The experiments reported here suggest that *Cyanobacterium* sp CLg1 use a multimeric DBE to aggregate polysaccharides into insoluble granules. In both cyanobacteria and plants, a glycohydrolase of a similar CAZy GH13 subfamily (subfamily 11 also known as TreX-GlgX) has been recruited. GlgX has been documented in proteobacteria and cyanobacteria to be selectively involved in glycogen degradation through both genetic and biochemical characterizations. In *E. coli*, GlgX was proven to be restrictive with respect to its substrate preference by debranching (only or preferably) those chains that have been recessed by glycogen phosphorylase during glycogen degradation (Dauvillée et al., 2005). This restrictive specificity was shown through the determination of the three-dimensional structure of the *E. coli* enzyme to be due to the shorter size of the inferred substrate binding groove (Song et al., 2010). The potato isoamylase was demonstrated to display a preference that accommodates amylopectin with comparatively little activity on glycogen or on either glycogen or amylopectin predigested by glycogen phosphorylase (Hussain et al., 2003). Here, the cyanobacterial enzyme GlgX2 is demonstrated to display the very same substrate preference as that reported for the potato enzyme (Hussain et al., 2003).

To strengthen these conclusions, we characterized rDBEs from two additional cyanobacterial species accumulating either glycogen or semiamylopectin. Both species, unlike *Cyanobacterium* sp CLg1, contain only one candidate CAZy GH13 subfamily 11 type of enzyme. Because the preference toward amylopectin debranching seems to distinguish GlgX2 from its GlgX progenitor relatives, we focused on this property by comparison to the enzyme activity in the presence of glycogen (in this case phytoglycogen). We further compared these preferences to those of the plant ISA1 enzyme and of the GlgX-like ISA3 plant enzyme. The results obtained so far build a convincing case for the evolution of isoamylase from a GlgX limit dextrinase type (GlgX or ISA3) of DBE as a major step to achieve aggregation of branched glucans into insoluble semicrystalline granules.

Indeed, the switch to an isoamylase specificity, together with a particular type of multimeric organization suggested by a plant study (Kubo et al., 2010; Utsumi et al., 2011), might define the two major determinants determining starch versus glycogen accumulation.

Although the convergent evolution reported in this work seems to imply that debranching of a soluble amylopectin precursor is mandatory to obtain aggregation of starch, we believe that this statement should nevertheless be taken with some caution. First, Streb et al. (2008) reported that in a particular mutant background lacking starch DBE, the presence of an additional mutation in a particular plastidial α-amylase restored starch synthesis in *Arabidopsis* leaves. This may imply that starch granule aggregation could dispense with the postulated debranching mechanism, although the crystallinity of this residual material still needs to be ascertained. Second, some of the secondary endosymbiosis lineages derived from Archaeplastida, as well as some little studied semiamylopectin accumulating cyanobacteria, were reported to lack candidate-direct DBE genes (Deschamps et al., 2006; Ball et al., 2011; Colleoni and Suzuki, 2012; Curtis et al., 2012). This observation would imply a nonessential function of polysaccharide debranching in amylopectin synthesis. Nevertheless, at least in one case, bona fide DBE activities were reported in biochemical assays, hinting that other glucan hydrolases may have evolved to ensure this function (Deschamps et al., 2006).

In any case, the work detailed in this article proves that at least in *Cyanobacterium* sp CLg1, a direct DBE has been recruited and evolved an isoamylase specificity and also possibly a multimeric organization that enabled these cyanobacteria to switch from glycogen synthesis to starch metabolism. Because both the chlamydial GlgX recruited by the Archaeplastida ancestors and the GlgX enzyme from cyanobacteria went through a similar convergent path to generate enzymes with similar properties, we propose that polysaccharide debranching could define an essential mechanism that enables conversion of glycogen metabolism into starch in all living cells. Such an evolution might have been favored each time a particular clade would have required a larger sized osmotically-inert storage polysaccharide pool.

METHODS

Media and Culture Conditions

Cyanobacterium sp CLg1 was grown in Artificial Sea Water medium in the absence (AS0 medium) or in the presence of 0.88 mM $NaNO_3$ (ASNIII medium) (Rippka et al., 1979). The axenic strain was grown at 22°C and subjected to light/dark cycles (12 h/12 h) under 0.12 μmol m^{-2} s^{-1}. Wild-type and mutant strains were maintained on solid ASNIII medium (0.8% noble agar) and transferred onto new plates every 3 months.

Cyanobacterium sp CLg1 Mutagenesis and Iodine Screening

The mutagenesis was performed on log phase *Cyanobacterium* sp CLg1 cultures plated on solid ASNIII medium (0.9% noble agar) as described by Deschamps et al. (2008). Cells (>10^5 cells/plate) were subjected to UV irradiation at 0.5, 3.0, or 6.0 cm from the source (Trans-illuminator TS-15 with a peak intensity of 7 mW cm^{-2} at 254 nm; Ultra-violet Products) for various exposure times (15, 30, 60, and 120 s) and immediately transferred to darkness during 24 h to avoid activation of the photodamage reparation system (Golden, 1988). After 1 month of growth in day-night cycles (12 h/12 h), all surviving colonies belonging to one mutagenesis condition (e.g., 30 s, 1 cm) were gently scraped after adding ASNIII medium over the agar slants and transfer in a 1-liter flask containing 300 mL of ASNIII. Mutagenized cells were diluted and plated on solid ASNIII medium. After 1 month, single colonies were transferred to 96-well plates containing 300 μL of liquid ASNIII medium and incubated for another month. A total of 20,000 cell patches were inoculated in duplicate by loading 40 μL of cell suspension on two solid ASNIII medium plates.

After 1 month of growth, mutants impaired in starch metabolism were screened by spraying with iodine crystal vapors. In contrast with the dark-blue staining of wild-type cell patches, cell mutants harboring a distinct iodine staining were picked on sister Petri dishes and transferred to liquid ASNIII medium. The segregation of genome copies was verified for each individual mutant by iodine vapor screening of 48 colonies. Of a total of 20,000 cell patches screened, 88 mutants impaired in starch metabolism were isolated. Among them, 14 mutants harbored a yellow iodine phenotype, indicating that starch was replaced by glycogen synthesis.

Transmission Electron Microscopy Observation

Wild-type and mutant strains were cultivated in 50 mL of liquid AS0 medium and harvested by centrifugation (5 min at 4000*g* at 4°C) after 2 weeks of growth. The pellets were suspended and incubated during 1 h in fixing buffer containing 2.5% glutaraldehyde and 0.1 M cacodylate buffer at pH 7.2. The fixed cells were centrifuged at 16,000*g* during 15 min. Cell pellets were washed three times with 0.1 M cacodylate buffer, pH 7.2. The samples were dehydrated by incubating 20 min in increasing ethanol percentages (25, 50, 70, 95, and 100% three times). The impregnation step was conducted in EPON (epoxy resins) resin/ethanol mixtures at different ratios (1:2, 1:1, and 2:1) for 60 min, in EPON 100% for 30 min twice, and finally in EPON 100% overnight at 4°C. Pellets were then transferred to fresh EPON 100% and incubated for 48 h at 60°C for polymerization. Ultrathin sections were cut using a Leica UC6 ultramicrotome and then treated with 1% periodic acid for 30 min and washed six times in water. The sections were immersed in 20% acetic acid containing 1% thiosemicarbazide for 1 h, and a washing series was performed as follows: 20, 10, 5, and 2% acetic acid for 5 min each. The sections were washed six times in water, stained with 1% silver proteinate for 30 min in the dark, and finally washed six times in water.

As described in the following section, WSPs were purified from wild-type and mutant strains subjected to nitrogen starvation. Rabbit liver glycogen was used as a reference. Droplets of diluted WSP suspension were laid onto glow-discharged carbon-coated copper grids. The excess liquid was blotted with filter paper and a droplet of 2% uranyl acetate was added prior to drying.

Both sections and negatively stained preparations were observed using a Philips CM200 microscope operating at 80 kV. The images were recorded on Kodak SO163 film.

Quantification of Water-Soluble and -Insoluble (Starch-Like) Material Produced in Wild-Type and Mutant Strains

Wild-type and mutant strains were grown for 12 d in liquid AS0 medium and harvested at the middle of the light period by centrifugation at 3600*g* for 15 min at 4°C. Cells were washed three times with extraction buffer (50 mM HEPES, pH 8, 5 mM DTT, 1 mM EDTA, and 0.025% Triton), and the cell suspension (10 mL) was disrupted through a French press. WSP and insoluble polysaccharide (starch-like granules) were separated by spinning the lysate at 16,000*g* for 15 min at 4°C. The pellet was washed three times with sterile water at 4°C and then solubilized in 100% DMSO at 90°C for 10 min. The WSP (supernatant) and the starch-like granules (in the pellet) were quantified with an amyloglucosidase assay (R Biopharm Starch/amidon). The total protein concentration in the supernatant was determined using the Bradford method (Bio-Rad).

Structural Analysis of Soluble Polysaccharide

WSPs were purified using anion-exchange chromatography (Roth DOWEX 50:8) followed by size exclusion chromatography preequilibrated in 10% DMSO (Toyopearl TSK HW 50) as described (Colleoni et al., 1999). Polysaccharide was quantified in each 1-mL fraction by the phenol-sulfuric acid method (Fox and Robyt, 1991). Polysaccharide fractions were pooled and then subjected to second size exclusion chromatography (Toyopearl TSK HW 55S). Polysaccharide detected in the exclusion volume was totally debranched by *Pseudomonas* sp isoamylase (Megazyme) in 55 mM sodium acetate buffer, pH 3.5. The chain length distribution of glucan chains was then analyzed by high performance anion exchange chromatography-pulse amperometric detection (HPAEC-PAD) as described (Colleoni et al., 1999).

Structural Analysis of the Insoluble Polysaccharide Fraction

Insoluble polysaccharides were precipitated in 70% ethanol, centrifuged at 6000*g* for 10 min, and dried at room temperature. The pellet was suspended in 10 mM NaOH and subjected to CL-2B gel permeation chromatography. Each 300-μL fraction was checked for the presence of glucans through their interaction with iodine. Eighty microliters of fraction was incubated with 20 μL of iodine solution (1% KI and 0.1% I_2), and the absorbance spectra were monitored to record the wavelength at maximum OD (λmax). The amylopectin, found in the exclusion volume, was then collected and totally debranched as described above for soluble polysaccharides and subjected to HPAEC-PAD.

Purification and Identification of the Cation-Dependent DBE

After 12 d of growth in ASNIII medium, cells were harvested at the middle of the light period by centrifugation at 3000*g* for 15 min at 4°C and washed three times with Tris-acetate buffer at 4°C (25 mM Tris-acetate, pH 7.5, and 10 mM DTT). Cells were disrupted by sonication or by a French press at 1250 p.s.i. The lysate was centrifuged at 16,000*g* for 15 min at 4°C. The supernatant (20 mL) was loaded on preparative anion-exchange chromatography column (HitrapQ sepharose Fast Flow, 5-mL column volume; GE Healthcare) preequilibrated in buffer A (25 mM Tris-acetate, pH 7.5, and 10 mM DTT). The proteins were eluted at 4 mL min^{-1} using a stepwise gradient of 40, 75, and 100% 1 M NaCl. Cation-dependent DBE activity was monitored for each fraction (2 mL) by zymogram analysis. The proteins were separated by nondenaturing PAGE containing 0.15% of β-limit dextrin (Megazyme). After electrophoresis, gels were incubated overnight at room temperature in buffer A with 10 mM $MgCl_2$. The cation dependent activity was then visualized as a white band in the presence of β-limit dextrin of amylopectin or a blue band in the presence of starch polysaccharides after iodine staining. Fractions containing cation-dependent activity were pooled, and ammonium sulfate was added in order to reach a final concentration of 1 M. After filtration, the pooled fraction was loaded on a hydrophobic exchange chromatography column (Hitrap Butyl sepharose Fast Flow, 5-mL column volume; GE Healthcare) preequilibrated in buffer C [25 mM Tris-acetate, pH 7.5, 10 mM DTT, and 1 M $(NH_4)_2SO_4$]. The proteins were eluted at 4 mL min^{-1} with a linear gradient of buffer A (0 to 100% in 25 mL). DBE activity of each fraction (2 mL) was monitored by zymogram analysis as described above. Fractions containing DBE were further purified using an anion-exchange chromatography column (MonoQ 5/50 GL, 1-mL column volume; GE Healthcare) preequilibrated in buffer A (25 mM Tris-acetate, pH 7.5, and 10 mM DTT). The proteins were eluted at 1 mL min^{-1} with a gradient of buffer A containing 1 M NaCl: 0 to 40% in 5 mL, 40 to 60% in 20 mL, 60 to 80% in 10 mL, 80 to 100% in 5 mL. The enzyme activities of each fraction (1 mL) were monitored by zymogram analysis. Fractions containing debranching activity were further concentrated with Amicon Ultra 0.5-mL centrifugal filters (Merck Millipore). Concentrated samples with activity were stored at −80°C in 20% glycerol.

Biochemical Characterization of Cation-Dependent Activity

Purified samples displaying cation-dependent activity (0.526 μmol of reducing end min^{-1} mL^{-1}) were incubated in 120 mM Tris, 30 mM imidazole, and 30 mM acetic acid, pH 7.5, in presence of 0.5% (w/v) of

polysaccharide (amylopectin, glycogen, β-limit dextrin, and phosphorylase-limit dextrin) and 10 mM $MgCl_2$. The release of glucan chains was followed by measuring the increase in reducing ends using the dinitrosalicylic acid method. After 10 min at 99°C, the OD was measured at 540 nm and compared with a standard curve (0 to 100 μg) measured with 1 mg mL^{-1} of Glc. Glucan chains released during the incubation were analyzed by HPAEC-PAD as described below. The cation-dependent activity was compared with that of isoamylase of *Pseudomonas* sp diluted 1/100 and incubated with 0.5% polysaccharide and 55 mM sodium acetate, pH 3.5.

Identification of the 80-kD Polypeptide by Nano–Liquid Chromatography–Tandem Mass Spectrometry

Samples with cation-dependent activity were further purified by an additional size-exclusion chromatography preequilibrated with 10 mM Tris/acetate buffer, pH 7.5, 10 mM DTT, and 150 mM NaCl (GE Healthcare Superdex 200; D = 0.8 centimorgans; H = 30 centimorgans). Each fraction (300 μL) was analyzed using zymogram procedures. Fractions containing the cation-dependent activity were then precipitated with 20% TCA (60 μL 20% TCA per 1 mL of sample with 25 μL of 2% sodium deoxycholate) and incubated on ice for 10 to 15 min. After centrifugation for 5 min at 16,100*g*, the pellets were suspended in 10 μL of buffer A (0.1 M Na_2CO_3 and 0.1 M DTT). The samples were boiled at 95°C for 5 min in the presence of 5 μL of buffer B (5% SDS, 30% saccharose, and 0.1% blue of bromophenol). Proteins were separated on 7.5% polyacrylamide gels with 0.1% SDS. SDS-PAGE gels were then stained for one night using Roti Blue (Roth) and washed with 25% methanol. The polypeptide was cut off and placed in 60% acetonitrile in bicarbonate ammonium before nano–liquid chromatography–tandem mass spectrometry analysis as described by Gurcel et al. (2008).

Gene Cloning and Sequencing

Starch metabolism genes (see Supplemental Table 2 online), DBE genes (glgX2 and glgX1), branching enzyme genes (be3), and starch/glycogen synthase genes (gbss) were amplified from genomic DNA of wild-type and mutant strains using the following oligonucleotide primers: glgX2F (forward), 5′-ATGTTAATGGGAGATGAATCTATGA-3′; glgX2R (reverse), 5′-TAATTAGTGGTTTTTAGTACTACTAACG-3′; glgX1F, 5′-ATGAACCATAAAACGTTACCTG-3′; glgX1R, 5′-CTATTTTGCCATTAATAAAATGCAAC-3′; gbss, 5′-TCCTCATGAATTGGTGACATAGTATGTT-3′; gbssR, 5′-CAGATACAGGTGAAAATCGTAACGC-3′; be3F, 5′-AGTGAATAGCCAAAAATCAACGAT-3′; and be3R, 5′-TGACCATCCCATTTGGCTCCTA-3′.

PCR (Analytik Jena Flexcycler) was conducted as follows: 95°C for 5 min, 30 cycles of denaturation at 98°C for 30 s, annealing 30 s at 59.6°C for GlgX2, GBSS, and BE3 and at 59.1°C for GlgX1, and extension for 2 min and 30 s at 72°C, and a final elongation step at 72°C for 5 min. The PCR products were cloned into pCR-BluntII-TOPO vector (Invitrogen), transferred into the chemically competent *Escherichia coli* TOP10 Mach1-T^R, and grown on Luria-Bertani agar with kanamycin and X-Gal. Purified plasmids were sequenced by the GATC Biotech Company according to Sanger methods. Each gene was sequenced on both strands using additional primers when required. The presence of a mutation was identified by alignment with the wild-type gene (BLASTn).

Phylogenetic Analysis

Homologs of GlgX were identified in GenBank or other sources using BLASTp and aligned with multiple sequence comparison by log expectation (MUSCLE) (http://www.ebi.ac.uk/Tools/msa/muscle/). The alignment was manually refined using SeAl (http://tree.bio.ed.ac.uk/software/seal/), and blocks of missing data in some taxa or regions of low identity were manually removed (final alignment of 595 amino acids available from Supplemental Data Set 1 online). This reduced alignment was analyzed under maximum likelihood. The best-fitting amino-acid substitution model was selected according to the Akaike informational criterion with ProtTest using the default values (Abascal et al., 2005). The Le Gascuel (Le and Gascuel, 2008) model with heterogeneous gamma rate distribution across sites (+Γ) was selected by ProtTest for this protein data set (see Supplemental Methods 1 online). The Le Gascuel-model parameter values were used under randomized accelerated maximum likelihood (RAxML) v.7.2.8 (Stamatakis 2006) for the maximum likelihood tree searches. The stability of monophyletic groups was assessed using RAxML with 1000 bootstrap replicates. The phylogenetic tree is rooted on the branch uniting the Archaeplastida sequences. This was done to reflect current understanding that supports Archaeplastida monophyly.

Preparation of α-Glucans for rDBE Characterization

The rice (*Oryza sativa*) amylopectin was prepared from starch in rice endosperm from a GBSSI-deficient *waxy* mutant line, EM21, derived from the *japonica*-type rice cultivar 'Kinmaze'. The starch granules from EM21 were extracted from polished rice seeds using the cold-alkali method as described (Yamamoto et al., 1981). Ten grams of the starch granules was dispersed in 500 mL of dimethylsulfoxide at 80°C under N_2 gas. The dispersed starch was precipitated by adding 1500 mL of ethanol and centrifuged at 10,000*g* for 20 min. The resulting precipitate was washed with acetone and diethyl ether, dried, and stored at −30°C until use as amylopectin.

The rice phytoglycogen was prepared from an isa1-deficient *sugary1* mutant line, EM914, derived from the *japonica*-type rice cultivar 'Taichung-65' (Nakamura et al., 1997). The phytoglycogen from polished seeds was extracted with 0.1% (w/v) NaOH solution at 4°C. The extract was filtrated through the 100-μm sieve. The filtrate was collected and phytoglycogen was precipitated with 3 volumes of ethanol. The precipitate was collected by centrifugation at 10,000*g* for 20 min. To remove the low molecular mass sugars and dextrins, the precipitate was rinsed 10 times with 50% (v/v) ethanol. The resulting precipitate was washed with acetone and diethyl ether, dried, and stored at −30°C until use as phytoglycogen.

Cloning of the Regions Encoding Glucan DBEs

Total RNA was isolated from the developing endosperm of the *japonica*-type rice cultivar 'Nipponbare' according to the manufacturer's protocol using an RNeasy plant mini kit (Qiagen). The cDNA was synthesized by Superscript III reverse transcriptase (Invitrogen).

The expression vectors for genes encoding the predicted rice mature ISA1 and Pullulanase proteins in *E. coli* were prepared as described previously (Fujita et al., 2009; Utsumi et al., 2011).

The region encoding the predicted rice mature ISA3 in Nipponbare was amplified by PCR with pGEM-T easy containing the Os-*ISA3* gene using the following primer pairs: forward primer (5′-CACGTGTAGCACCACGGCGAGAG-3′) containing *Bbr*PI, and reverse primer (5′-GTCGACCTAAGGCTTTGCCTTGAGC-3′) containing *Sal*I. The DNA fragment was subcloned into pGEM-T easy vector. The sequence encoding the mature Os-ISA3 protein in the pGEM-T easy vector was excised by restriction enzymes *Bbr*PI and *Sal*I. The DNA fragment was ligated with the pET32b (Novagen) expression vector treated with *Eco*RV and *Sal*I.

The region including the *Synechococcus DBE* gene was amplified by PCR with genomic DNA from *Synechococcus* PCC7942 with the following primer pairs: forward primer (5′-CATATGACTGTTTCATCCCGTCGC-3′) containing *Nde*I and reverse primer (5′-GTCGACCTG CAGGCGGCCGCGAATTC-3′) containing *Sal*I. The DNA fragment was subcloned into pGEM-T easy vector. The *Synechococcus DBE* in pGEM-T

easy was digested with the restriction enzymes *Nde*I and *Sal*I. The DNA fragment was ligated with the pCold TF (Takara) expression vector treated with *Nde*I and *Sal*I.

The DNA fragment encoding the *Cyanothece* ATCC51142 DBE was isolated from genomic DNA of *Cyanothece ATCC51142 DBE* by PCR using the following primer pairs: forward primer (5′-CATATGAGCTCAAAACCCTTC-3′) containing *Nde*I and reverse primer (5′-GTCGACTTATGAATTAGACTTTGCC-3′) containing *Sal*I. The DNA fragment was subcloned into pGEM-T easy vector. The *Cyanothece DBE* in pGEM-T easy was isolated using restriction enzymes *Nde*I and *Sal*I. The DNA fragment was ligated with the pCold TF expression vector treated with *Nde*I and *Sal*I.

The DNA fragment encoding the *E. coli GlgX* was amplified by PCR using the genomic DNA from *E. coli* K12 strain as template and the following primer pairs: forward primer (5′-CATATGACACAACTCGCCATTG-3′), containing *Nde*I, and reverse primer (5′-GTCGACTCATCTCTGGAACACACAC-3′) containing *Sal*I. The DNA fragment was subcloned into pGEM-T easy vector. The *E. coli* GlgX in pGEM-T easy was prepared by hydrolysis with restriction enzymes *Nde*I and *Sal*I. The treated DNA fragment was ligated in the pCold TF expression vector treated by *Nde*I and *Sal*I.

E. coli Overexpression of DBE

All of the constructed vectors including *DBE* genes were introduced into the expression host *E. coli* BL21 (DE3) star strain (Invitrogen) containing pTf16 chaperone plasmid (Takara). The transformed cells were grown at 37°C in 10 mL of Terrific Broth medium containing 100 μg/mL of carbenicillin and 34 μg/mL of chloramphenicol overnight. The precultured cells were subsequently grown at 37°C with 1.0 liters of Terrific Broth medium containing 100 μg/mL of carbenicillin, 34 μg/mL of chloramphenicol, and 1 mg/mL of L-arabinose. The L-arabinose was used for induction of chaperon protein Tf16. When the OD at 600 nm of the culture medium reached 0.4 to 0.5, the culture was preincubated at 15°C for 30 min. Then, the culture medium was supplemented with isopropylthio-β-D-galactoside to a final concentration of 1 mM, and the rDBE was induced by further incubation at 15°C for 20 h. The cells were collected by centrifugation and stored at –80°C until use.

Purification of Recombinant DBE

The frozen transformed *E. coli* cells were thawed on ice and suspended in extraction buffer (50 mM imidazol-HCl, pH 7.4, 8 mM $MgCl_2$, 12.5% glycerol, and 50 mM 2-mercaptoethanol). The cells were disrupted by sonication on ice and centrifuged at 10,000*g* for 30 min at 0°C. The resulting supernatant was applied to the HiTrap HP anion-exchange column (5 mL; GE Healthcare) equilibrated with a buffer A (50 mM imidazol-HCl, pH 7.4, 8 mM $MgCl_2$, and 50 mM 2-mercaptoethanol). After washing the column with the buffer A, the proteins were eluted with a linear gradient of 0.0 to 0.5 M NaCl at a flow rate of 2 mL/min. The peak fractions containing the abundant rDBE protein were collected and were added to one-third volume of 50 mM Na-phosphate, pH7.0, containing 3 M ammonium sulfate. The rDBE protein was applied to the TSKgel Ether-5PW hydrophobic column (7.5 mm in diameter × 75 mm in length; Tosoh Corporation) equilibrated with buffer B (50 mM Na-phosphate, pH 7.0, 50 mM 2-mercaptoethanol, and 1.0 M ammonium sulfate). After washing the column with buffer B, the protein was eluted with a linear gradient of 1.0 to 0 M ammonium sulfate in a buffer solution consisting of 50 mM Na-phosphate, pH 7.0, and 50 mM 2-mercaptoethanol at a flow rate of 1 mL/min. The peak fraction containing the major rDBE protein was collected and concentrated using centrifugal filter units (Ultracel-3K; Millipore). Each DBE preparation was purified to be a near homogeneity using these procedures. The purified DBE preparation was suspended in 50 mM imidazol-HCl, pH 7.4, 8 mM $MgCl_2$, 1 mM DTT, 0.5 M NaCl, and 12.5% glycerol and kept at −80°C until use.

Recombinant ISA Enzymatic Reaction

The reaction mixture routinely contained 10 mM HEPES-NaOH buffer, pH 7.4, 2 mg of glucan, and DBE in a total volume of 400 μL. The reaction was run at 30°C. The amounts of DBE used were 0.13 or 0.13 μg for rOsISA1, 32 or 9.54 μg for rOsISA3, 32 or 32 μg for rCaDBE, 30.3 or 30.3 μg for rSeDBE, 32 or 32 μg for rEcDBE, and 0.11 or 0.11 μg for PaISA when the enzyme was incubated with phytoglycogen or amylopectin, respectively. At appropriate time intervals, the enzymatic reaction was terminated by heating the reaction mixture in a boiling water bath for 5 min.

Assay of DBE Activity

The DBE activity was determined by measuring the amount of reducing ends in the reaction products according to the method of Utsumi et al. (2009). Solution A consisted of 97.1 mg of disodium 2,2'-bicinchoninate, 3.2 g of sodium carbonate monohydrate, and 1.2 g of sodium bicarbonate in a total volume of 50 mL. Solution B consisted of 62 mg of copper sulfate pentahydrate and 63 mg of L-Ser in a total volume of 50 mL. The working reagent was freshly prepared by mixing the equal volume of Solutions A and B (Fox and Robyt, 1991). An amount of 225 μL of the sample was added to 225 μL of the working reagent in a tube, and the mixture was incubated at 80°C for 40 min in a water bath. After incubation, the assay mixture was then cooled to room temperature and incubated for 10 min. An aliquot (150 μL) of the disodium 2,2'-bicinchoninate–treated sample was taken, and its absorbance at 560 nm was measured using a microplate spectrophotometer (Bio-Rad). The absorbance at 560 nm was found to be proportional to the concentration of maltose or Glc in the range of 0 to at least 25 μM of the assay mixture.

Accession Numbers

Sequence data from this article can be found in the GenBank/EMBL libraries under the following accession numbers: GlgX1 (KC422239.1), GlgX2 (KC422240.1), GlgB1 (JX074061.1), GlgB2 (JX074062.1), GlgB3 (JX074063.1), Apu57 (KF717069), BE57 (KF717075), GlgA1 (KF717072), GlgA2 (KF717073), Apu13 (KF717068), GBSS (KF717071), GlgP (KF717074), MalQ (KF717076), and indirect DBE (KF717070).

Supplemental Data

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

Supplemental Figure 1. Cell Patches of the Five Classes of Mutants of *Cyanobacterium* sp CLg1 Stained with Iodine Vapors.

Supplemental Figure 2. Zymogram of Phosphorylase Activity.

Supplemental Figure 3. Effects of Different Cations on Debranching Enzyme Activity.

Supplemental Figure 4. Purification of Cation-Dependent Debranching Enzyme Activity.

Supplemental Figure 5. Identification of Debranching Enzyme Activity.

Supplemental Figure 6. SDS-PAGE of Purified DBE Preparations from Rice, Cyanobacteria, and *E. coli*.

Supplemental Table 1. Summary Table of Starch Metabolizing Genes Checked for Mutations.

Supplemental Table 2. List of Primers Used for Gene Cloning and Sequencing Involved in the Storage Polysaccharide Metabolism of *Cyanobacterium* sp CLg1.

Supplemental Methods 1. Parameter Values for Phylogenetic Analysis.

Supplemental Data Set 1. GlgX and Isoamylase Alignments.

ACKNOWLEDGMENTS

S.G.B. and C.C. were supported by the Centre National de la Recherche Scientifique, the Université des Sciences et Technologies de Lille, the Région Nord Pas de Calais, and the Agence Nationale pour la Recherche (ANR) grants from Starchevol.

AUTHOR CONTRIBUTIONS

C.C. and S.G.B. designed the research and wrote the article. U.C., M.C., M.D., J.N.-R., and C.T. carried out the mutagenesis campaign and the screening process. Y.U., D.K., S.S., E.S., Y.N. expressed recombinant debranching enzymes and sequenced the CLg1 genome. J.-L.P. and A.D.-T. performed TEM observations. X.R. and E.M. were responsible for HPAEC-PAD and proton NMR analysis, respectively. D.B. and M.-C.A. performed phylogenetic analysis. A.-S.V.-E. carried out mass spectrometry analysis. M.P. and L.S. performed molecular modeling of CLg1-GlgX2.

Received September 3, 2013; revised September 3, 2013; accepted October 2, 2013; published October 25, 2013.

REFERENCES

Abascal, F., Zardoya, R., and Posada, D. (2005). ProtTest: Selection of best-fit models of protein evolution. Bioinformatics **21:** 2104–2105.

Ball, S., Colleoni, C., Cenci, U., Raj, J.N., and Tirtiaux, C. (2011). The evolution of glycogen and starch metabolism in eukaryotes gives molecular clues to understand the establishment of plastid endosymbiosis. J. Exp. Bot. **62:** 1775–1801.

Ball, S., Guan, H.P., James, M., Myers, A., Keeling, P., Mouille, G., Buléon, A., Colonna, P., and Preiss, J. (1996). From glycogen to amylopectin: A model for the biogenesis of the plant starch granule. Cell **86:** 349–352.

Ball, S.G., Subtil, A., Bhattacharya, D., Moustafa, A., Weber, A.P. M., Gehre, L., Colleoni, C., Arias, M.C., Cenci, U., and Dauvillée, D. (2013). Metabolic effectors secreted by bacterial pathogens: Essential facilitators of plastid endosymbiosis? Plant Cell **25:** 7–21.

Buléon, A., Colonna, P., Planchot, V., and Ball, S. (1998). Starch granules: Structure and biosynthesis. Int. J. Biol. Macromol. **23:** 85–112.

Cenci, U., Nitschke, F., Steup, M., Minassian, B.A., and Colleoni, C. Ball, S. (2013). Transition from glycogen to starch metabolism in Archaeplastida. Trends Plant Sci. (In Press).

Colleoni, C., et al. (1999). Genetic and biochemical evidence for the involvement of alpha-1,4 glucanotransferases in amylopectin synthesis. Plant Physiol. **120:** 993–1004.

Colleoni, C., and Suzuki, E. (2012). Storage polysaccharide metabolism in Cyanobacteria. In Essential Reviews in Experimental Biology: Starch: Origins, Structure and Metabolism, Vol. 5, I. Tetlow, ed (London: The Society for Experimental Biology), pp. 213–253.

Coppin, A., Varré, J.S., Lienard, L., Dauvillée, D., Guérardel, Y., Soyer-Gobillard, M.O., Buléon, A., Ball, S., and Tomavo, S. (2005). Evolution of plant-like crystalline storage polysaccharide in the protozoan parasite *Toxoplasma gondii* argues for a red alga ancestry. J. Mol. Evol. **60:** 257–267.

Curtis, B.A., et al. (2012). Algal genomes reveal evolutionary mosaicism and the fate of nucleomorphs. Nature **492:** 59–65.

Dauvillée, D., Deschamps, P., Ral, J.P., Plancke, C., Putaux, J.L., Devassine, J., Durand-Terrasson, A., Devin, A., and Ball, S.G. (2009). Genetic dissection of floridean starch synthesis in the cytosol of the model dinoflagellate *Crypthecodinium cohnii*. Proc. Natl. Acad. Sci. USA **106:** 21126–21130.

Dauvillée, D., Kinderf, I.S., Li, Z., Kosar-Hashemi, B., Samuel, M.S., Rampling, L., Ball, S., and Morell, M.K. (2005). Role of the *Escherichia coli* glgX gene in glycogen metabolism. J. Bacteriol. **187:** 1465–1473.

Delrue, B., Fontaine, T., Routier, F., Decq, A., Wieruszeski, J.M., Van Den Koornhuyse, N., Maddelein, M.L., Fournet, B., and Ball, S. (1992). Waxy *Chlamydomonas reinhardtii*: Monocellular algal mutants defective in amylose biosynthesis and granule-bound starch synthase activity accumulate a structurally modified amylopectin. J. Bacteriol. **174:** 3612–3620.

Deschamps, P., et al. (2008). Metabolic symbiosis and the birth of the plant kingdom. Mol. Biol. Evol. **25:** 536–548.

Deschamps, P., et al. (2006). Nature of the periplastidial pathway of starch synthesis in the cryptophyte *Guillardia theta*. Eukaryot. Cell **5:** 954–963.

Falcon, L.I., Lindvall, S., Bauer, K., Bergman, B., and Carpenter, E.J. (2004). Ultrastructure of unicellular N2 fixing cyanobacteria from the tropical north Atlantic and subtropical north Pacific oceans. J. Phycol. **40:** 1074–1078.

Fox, J.D., and Robyt, J.F. (1991). Miniaturization of three carbohydrate analyses using a microsample plate reader. Anal. Biochem. **195:** 93–96.

Fu, J., and Xu, X. (2006). The functional divergence of two glgP homologues in Synechocystis sp. PCC 6803. FEMS Microbiol. Lett. **260:** 201–209.

Fujita, N., et al. (2009). Characterization of pullulanase (PUL)-deficient mutants of rice (*Oryza sativa* L.) and the function of PUL on starch biosynthesis in the developing rice endosperm. J. Exp. Bot. **60:** 1009–1023.

Golden, S.S. (1988). Mutagenesis of cyanobacteria by classical and gene-transfer-based methods. Methods Enzymol. **167:** 714–727.

Gurcel, C., Vercoutter-Edouart, A.S., Fonbonne, C., Mortuaire, M., Salvador, A., Michalski, J.C., and Lemoine, J. (2008). Identification of new O-GlcNAc modified proteins using a click-chemistry-based tagging. Anal. Bioanal. Chem. **390:** 2089–2097.

Hussain, H., Mant, A., Seale, R., Zeeman, S., Hinchliffe, E., Edwards, A., Hylton, C., Bornemann, S., Smith, A.M., Martin, C., and Bustos, R. (2003). Three isoforms of isoamylase contribute different catalytic properties for the debranching of potato glucans. Plant Cell **15:** 133–149.

James, M.G., Robertson, D.S., and Myers, A.M. (1995). Characterization of the maize gene sugary1, a determinant of starch composition in kernels. Plant Cell **7:** 417–429.

Janeček, S., Svensson, B., and MacGregor, E.A. (2011). Structural and evolutionary aspects of two families of non-catalytic domains present in starch and glycogen binding proteins from microbes, plants and animals. Enzyme Microb. Technol. **49:** 429–440.

Jeanningros, R., Creuzet-Sigal, N., Frixon, C., and Cattaneo, J. (1976). Purification and properties of a debranching enzyme from *Escherichia coli*. Biochim. Biophys. Acta **438:** 186–199.

Kubo, A., Colleoni, C., Dinges, J.R., Lin, Q., Lappe, R.R., Rivenbark, J.G., Meyer, A.J., Ball, S.G., James, M.G., Hennen-Bierwagen, T.A., and Myers, A.M. (2010). Functions of

heteromeric and homomeric isoamylase-type starch-debranching enzymes in developing maize endosperm. Plant Physiol. **153:** 956–969.

Kubo, A., Fujita, N., Harada, K., Matsuda, T., Satoh, H., and Nakamura, Y. (1999). The starch-debranching enzymes isoamylase and pullulanase are both involved in amylopectin biosynthesis in rice endosperm. Plant Physiol. **121:** 399–410.

Le, S.Q., and Gascuel, O.L.G. (2008). An improved general amino acid replacement matrix. Mol. Biol. Evol. **25:** 1307–1320.

Meléndez, R., Meléndez-Hevia, E., and Canela, E.I. (1999). The fractal structure of glycogen: A clever solution to optimize cell metabolism. Biophys. J. **77:** 1327–1332.

Mouille, G., Maddelein, M.L., Libessart, N., Talaga, P., Decq, A., Delrue, B., and Ball, S. (1996). Preamylopectin processing: A mandatory step for starch biosynthesis in plants. Plant Cell **8:** 1353–1366.

Nakamura, Y., Kubo, A., Shimamune, T., Matsuda, T., Harada, K., and Satoh, H. (1997). Correlation between activities of starch debranching enzyme and α-polyglucan structure in endosperms of *sugary-1* mutants of rice. Plant J. **12:** 143–153.

Nakamura, Y., et al. (2005). Some Cyanobacteria synthesize semi-amylopectin type alpha-polyglucans instead of glycogen. Plant Cell Physiol. **46:** 539–545.

Rippka, R., Deruelles, J., Waterbury, J.B., Herdman, M., and Stanier, R.Y. (1979). Generic assignments, strain histories and properties of pure cultures of cyanobacteria. J. Gen. Microbiol. **111:** 1–61.

Schneegurt, M.A., Sherman, D.M., Nayar, S., and Sherman, L.A. (1994). Oscillating behavior of carbohydrate granule formation and dinitrogen fixation in the cyanobacterium *Cyanothece sp*. strain ATCC 51142. J. Bacteriol. **176:** 1586–1597.

Shearer, J., and Graham, T.E. (2002). New perspectives on the storage and organization of muscle glycogen. Can. J. Appl. Physiol. **27:** 179–203.

Shimonaga, T., et al. (2008). Variation in storage alpha-glucans of the Porphyridiales (Rhodophyta). Plant Cell Physiol. **49:** 103–116.

Song, H.N., Jung, T.Y., Park, J.T., Park, B.C., Myung, P.K., Boos, W., Woo, E.J., and Park, K.H. (2010). Structural rationale for the short branched substrate specificity of the glycogen debranching enzyme GlgX. Proteins **78:** 1847–1855.

Stamatakis, A. (2006). RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics **22:** 2688–2690.

Streb, S., Delatte, T., Umhang, M., Eicke, S., Schorderet, M., Reinhardt, D., and Zeeman, S.C. (2008). Starch granule biosynthesis in *Arabidopsis* is abolished by removal of all debranching enzymes but restored by the subsequent removal of an endoamylase. Plant Cell **20:** 3448–3466.

Suzuki, E., Onoda, M., Colleoni, C., Ball, S., Fujita, N., and Nakamura, Y. (2013). Physicochemical variation of cyanobacterial starch, the insoluble α-glucans in cyanobacteria. Plant Cell Physiol. **54:** 465–473.

Suzuki, E., Umeda, K., Nihei, S., Moriya, K., Ohkawa, H., Fujiwara, S., Tsuzuki, M., and Nakamura, Y. (2007). Role of the GlgX protein in glycogen metabolism of the cyanobacterium, *Synechococcus elongatus* PCC 7942. Biochim. Biophys. Acta **1770:** 763–773.

Utsumi, Y., Utsumi, C., Sawada, T., Fujita, N., and Nakamura, Y. (2011). Functional diversity of isoamylase oligomers: The ISA1 homo-oligomer is essential for amylopectin biosynthesis in rice endosperm. Plant Physiol. **156:** 61–77.

Utsumi, Y., Yoshida, M., Francisco, P.B., Jr., Sawada, T., Kitamura, S., and Nakamura, Y. (2009). Quantitative assay method for starch branching enzyme with bicinchoninic acid by measuring the reducing terminals of glucans. J. Appl. Glycosci. **56:** 215–222.

Wattebled, F., Dong, Y., Dumez, S., Delvallé, D., Planchot, V., Berbezy, P., Vyas, D., Colonna, P., Chatterjee, M., Ball, S., and D'Hulst, C. (2005). Mutants of Arabidopsis lacking a chloroplastic isoamylase accumulate phytoglycogen and an abnormal form of amylopectin. Plant Physiol. **138:** 184–195.

Yamamoto, K., Sawada, S., and Onogaki, I. (1981). Effects of quality and quantity of alkali solution on the properties of rice starch. Denpun Kagaku **28:** 241–244.

Convergent Evolution of Polysaccharide Debranching Defines a Common Mechanism for Starch Accumulation in Cyanobacteria and Plants

Ugo Cenci, Malika Chabi, Mathieu Ducatez, Catherine Tirtiaux, Jennifer Nirmal-Raj, Yoshinori Utsumi, Daiki Kobayashi, Satoshi Sasaki, Eiji Suzuki, Yasunori Nakamura, Jean-Luc Putaux, Xavier Roussel, Amandine Durand-Terrasson, Debashish Bhattacharya, Anne-Sophie Vercoutter-Edouart, Emmanuel Maes, Maria Cecilia Arias, Monica Palcic, Lyann Sim, Steven G. Ball and Christophe Colleoni

Plant Cell 2013;25;3961-3975; originally published online October 25, 2013;
DOI 10.1105/tpc.113.118174

The Plant Cell, Vol. 26: 4462–4482, November 2014, www.plantcell.org

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]

Maxime Chantreau,[a,b] Antoine Portelette,[c,d] Rebecca Dauwe,[e] Shingo Kiyoto,[c,d,f] David Crônier,[c,d] Kris Morreel,[g,h] Sandrine Arribat,[a,b] Godfrey Neutelings,[a,b] Malika Chabi,[a,b] Wout Boerjan,[g,h] Arata Yoshinaga,[f] François Mesnard,[e] Sebastien Grec,[a,b] Brigitte Chabbert,[c,d] and Simon Hawkins[a,b,1]

[a] Université Lille Nord de France, Lille 1, UMR1281, F-59650 Villeneuve d'Ascq Cedex, France
[b] INRA, UMR1281, Stress Abiotiques et Différenciation des Végétaux Cultivés, F-59650 Villeneuve d'Ascq, France
[c] INRA, UMR614, Fractionnement des AgroRessources et Environnement, F-51100 Reims, France
[d] Université de Reims Champagne-Ardenne, UMR614, Fractionnement des AgroRessources et Environnement, F-51100 Reims, France
[e] Université de Picardie Jules Verne, EA 3900, BIOPI, Laboratoire de Phytotechnologie, F-80037 Amiens Cedex 1, France
[f] Laboratory of Tree Cell Biology, Division of Forest and Biomaterials Science, Graduate School of Agriculture, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan
[g] Department of Plant Systems Biology, VIB, 9052 Gent, Belgium
[h] Department of Plant Biotechnology and Bioinformatics, UGent, 9052 Gent, Belgium

Histochemical screening of a flax ethyl methanesulfonate population led to the identification of 93 independent M2 mutant families showing ectopic lignification in the secondary cell wall of stem bast fibers. We named this core collection the *Linum usitatissimum* (flax) *lbf* mutants for *lignified bast fibers* and believe that this population represents a novel biological resource for investigating how bast fiber plants regulate lignin biosynthesis. As a proof of concept, we characterized the *lbf1* mutant and showed that the lignin content increased by 350% in outer stem tissues containing bast fibers but was unchanged in inner stem tissues containing xylem. Chemical and NMR analyses indicated that bast fiber ectopic lignin was highly condensed and rich in G-units. Liquid chromatography-mass spectrometry profiling showed large modifications in the oligolignol pool of *lbf1* inner- and outer-stem tissues that could be related to ectopic lignification. Immunological and chemical analyses revealed that *lbf1* mutants also showed changes to other cell wall polymers. Whole-genome transcriptomics suggested that ectopic lignification of flax bast fibers could be caused by increased transcript accumulation of (1) the *cinnamoyl-CoA reductase*, *cinnamyl alcohol dehydrogenase*, and *caffeic acid O-methyltransferase* monolignol biosynthesis genes, (2) several lignin-associated peroxidase genes, and (3) genes coding for respiratory burst oxidase homolog NADPH-oxidases necessary to increase H_2O_2 supply.

INTRODUCTION

Lignin is a major component of many plant cell walls and is essential for water transport in vascular tissue, mechanical support, and resistance to pathogens in higher land plants (Baucher et al., 1998; Boerjan et al., 2003; Weng and Chapple, 2010). This phenolic polymer also contributes to the recalcitrance of lignocellulosic biomass for biofuel production and the regulation of lignin biosynthesis has therefore been intensely studied (Whetten and Sederoff, 1995; Fu et al., 2011; Vanholme et al., 2012a). Much information about this process has been obtained by biochemical and genetics studies on mutants showing modified lignification profiles (Anterola and Lewis, 2002; Bonawitz and Chapple, 2010; Vanholme et al., 2012b).

Generally, lignin mutants can be divided into two main groups: (1) those showing reduced cell wall lignin levels and (2) ectopic lignification mutants where the secondary cell wall developmental program is activated. In the first group, lignin is often reduced and/or modified via the downregulation of genes involved in lignin monomer biosynthesis and/or the oxidation of monomers for subsequent polymerization (laccases and peroxidases) (Vanholme et al., 2010; Weng and Chapple, 2010; Zhao et al., 2013). Reduced lignin content is often accompanied by modifications to other cell wall polymers, suggesting the existence of a dynamic relationship between the cell wall matrix and the lignification process (Hu et al., 1999; Van Acker et al., 2013). In the second group, upregulation/downregulation of different transcription factors leads to the activation of the secondary cell wall developmental program and the biosynthesis and deposition of cellulose, hemicellulose, and lignin in parenchyma-type cells that normally only produce nonlignified primary cell walls (Mitsuda et al., 2007; Zhong et al., 2007; Zhao and Dixon, 2011). Alternatively, ectopic lignification can also result

[1] Address correspondence to simon.hawkins@univ-lille1.fr.
The authors responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) are: Brigitte Chabbert (brigitte.chabbert@reims.inra.fr) and Simon Hawkins (simon.hawkins@univ-lille1.fr).
[C] Some figures in this article are displayed in color online but in black and white in the print edition.
[W] Online version contains Web-only data.
www.plantcell.org/cgi/doi/10.1105/tpc.114.130443

Table 1. Number of M2 Families Assigned to Different Bast Fiber Lignification Classes

	Bast Fiber Lignification Class			
Screen	Class 3	Class 2	Class 1	Total
UV	252	156	132	540
P-HCl	150	176	93	319

from perturbations in the biosynthesis of other cell wall polymers (Zhong et al., 2002; Caño-Delgado et al., 2003).

Interestingly, the stems of certain fiber plants (e.g., flax, hemp, ramie, etc.) naturally contain two populations of cells showing highly contrasted secondary cell wall compositions. In outer-stem tissues, specialized cells (bast fibers) possess hypolignified and cellulose-rich thick secondary cell walls, whereas the xylem cells from inner-stem tissues have a more typical lignified secondary cell wall structure. Analyses in flax (*Linum usitatissimum*), for example, show that bast fiber secondary cell walls are almost completely filled with cellulose (~70%) and contain 5 to 15% of noncellulosic polysaccharide (NCP) mainly composed of β-1-4 galactan and arabinogalactan, but only 2 to 4% lignin (Davis et al., 1990; Girault et al., 1997; Day et al., 2005; Gorshkova and Morvan, 2006). By contrast, xylem cell walls contain lower amounts of cellulose (~40%) and much higher amounts (30%) of lignin (Day et al., 2005). Bast fibers are elongated cells that provide mechanical support and allow relatively tall plants with small stem diameters to maintain an erect state (Neutelings, 2011; Guerriero et al., 2013). In flax, the outer stem tissues enriched in hypolignified primary bast fibers can be easily peeled away from the central xylem core-containing lignified cells, and this plant therefore appears to be an excellent model to study secondary cell wall formation and lignification.

We have previously shown that lignification in flax appears to be partially modulated through transcriptional regulation of genes encoding lignin monomer biosynthesis and polymerization (Fenart et al., 2010; Huis et al., 2012). More recently, we generated and characterized a flax ethyl methanesulfonate (EMS) mutant population that has allowed us to obtain mutants for the *CAD* and *C3H*

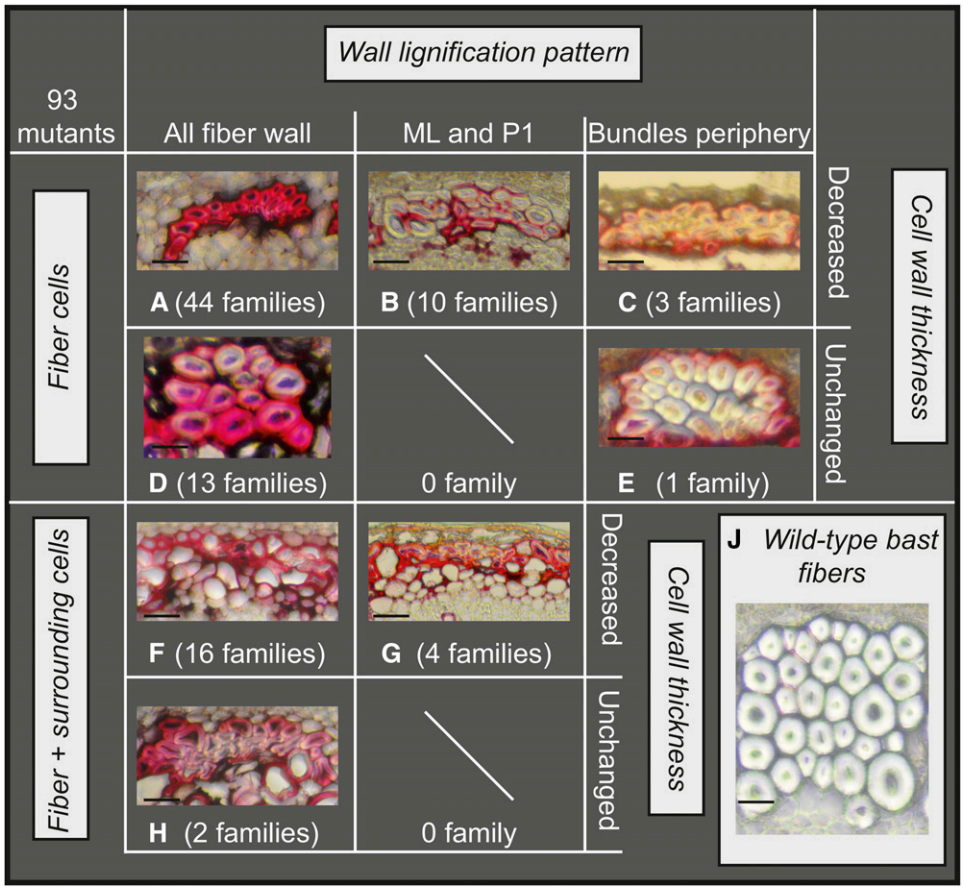


Figure 1. Classification of Flax *lbf* Mutants into Eight Different Groups According to Modified Lignification Pattern.

Groups A to C: Only fiber cells lignified, cell wall thickness decreased. Groups D and E: Only fiber cells lignified, cell wall thickness unchanged. Groups F and G: fibers and surrounding cells lignified, cell wall thickness decreased. Group H: fibers and surrounding cells lignified, cell wall thickness unchanged. Group J: wild-type bast fibers. Number of families in each group is given in brackets. Bar = 10 μm, phloroglucinol-HCl staining of stem cross sections (lignified walls are colored red).

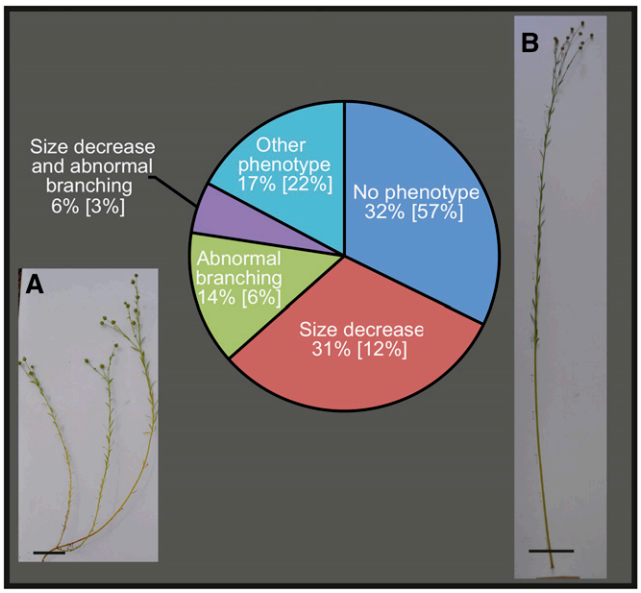


Figure 2. Proportion of Flax *lbf* Mutants in Different Categories Based on Visual Phenotype.

Percentage values refer to the proportion of *lbf* mutants showing a given phenotype. Values in brackets correspond to the proportion of all flax mutants (PT-flax collection) showing the phenotype.
(A) Photo of typical *lbf* mutant showing reduced size and increased branching. Bar = 5 cm.
(B) Photo of wild-type flax plant. Bar = 10 cm.
[See online article for color version of this figure.]

lignin genes through a TILLinG reverse genetics approach (Chantreau et al., 2013). In this article, we report the screening of this population and the identification of 319 independent mutants showing altered lignification profiles in bast fibers. We believe that this collection of flax lignin mutants represents a valuable biological resource for plant cell wall biologists. The detailed characterization of individual mutants should provide information on the different regulatory mechanisms and signaling pathways used by plants to regulate lignin biosynthesis. In addition, the identification of novel key genes involved in this process could provide targets for engineering improved lignocellulosic quality in other plant species. Cell wall analyses of mutants containing bast fibers with variable lignin content will also lead to a better understanding of the dynamic relationship between lignin and other cell wall polymers. As a proof of concept, we report the detailed characterization of a highly lignified flax bast fiber mutant.

RESULTS

Identification and Visual Phenotyping of the Flax Lignified Bast Fiber Mutant Core Collection

To identify mutants showing increased lignification in bast fibers, we first screened 8999 plants from 3391 M2 families (Chantreau et al., 2013). Examination of transversal hand-sections of stem from individual plants by UV microscopy allowed us to identify 540 families showing increased autofluorescence in bast fibers. Families were assigned to three different classes based on a visual estimation of modified autofluorescence: strong, class 1; moderate, class 2; and weak, class 3 (Table 1). In a second round of screening, thin freehand stem sections were prepared from the previously identified 540 families and stained with phloroglucinol-HCl. Families showing a red coloration of bast fibers (indicating potential lignin deposition) were once again assigned to three different classes (strong, moderate, and weak) (Table 1). Two hundred and twenty-one families previously identified in the first round of screening showed only little/no differences when compared with wild-type plants and were therefore not retained for further analyses. Altogether, 319 families showed increased coloration of bast fibers and 93 families showed strong coloration (class 1). We named this core collection of class 1 mutant families

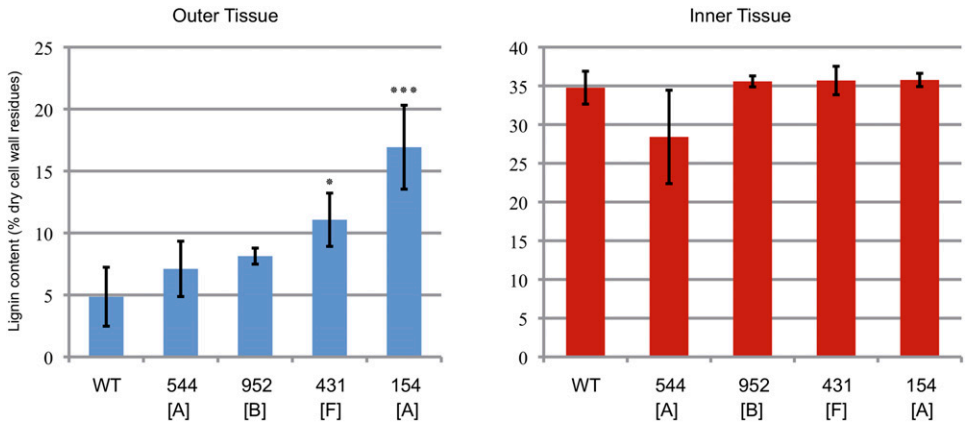


Figure 3. Acetyl Bromide Lignin Content of Inner- and Outer-Stem Tissues of Four *lbf* Mutants and Wild-Type Plants.

Letters in brackets refer to the subclassification presented in Figure 1. Significant differences (Student's *t* test) between wild-type and mutant tissues were observed at $P < 0.001$ (***) and $P < 0.05$ (*); error bars = SD.
[See online article for color version of this figure.]

lignified bast fiber (*lbf*) mutants. The *lbf* core collection was then subclassified into eight different groups according to the type of modified lignification pattern (Figure 1). For example, the groups A to E show increased lignification uniquely in fiber cells, whereas the groups F to H also show increased lignification in surrounding cells.

We then classified the flax *lbf* mutants into different categories based on previously established visual phenotypes of the flax EMS mutant population (Supplemental Table 1; Chantreau et al., 2013). Thirty-two percent of the core collection families showed no obvious morphological phenotype, 31% were smaller to wild-type plants, and 14% showed increased stem branching (Figure 2). Other observed phenotypes included early death and non-erect stems. Comparison of these values with the corresponding values for the overall mutant population (Figure 2) would suggest that *lbf* mutants are generally smaller, show increased branching, and have thinner stems when compared with either wild-type plants or other mutants.

The Flax *lbf1* Mutant Has a Modified Lignin Content

To confirm that the red coloration of bast fibers observed in phloroglucinol-stained sections of *lbf* mutants corresponded to lignin, we determined the acetyl bromide lignin content (Iiyama and Wallis, 1990) in stem tissues from four independent *lbf* mutants belonging to three different groups (A, B, and F). Our results (Figure 3) show that the lignin content of outer stem tissues from two of these mutants was significantly greater than in wild-type plants. In contrast, the lignin content of inner stem tissues (xylem) from all mutants was not significantly increased. We then focused on a single family (154) that showed the largest increase out of the four mutant lines evaluated for further detailed characterization. We named this mutant *lbf1* for *lignified bast fiber1*. This mutant line shows a typical core collection phenotype (reduced plant size and stem diameter and increased basal stem ramification) (Figure 2).

Thioacidolysis of M4 *lbf1* outer tissues revealed significant increases in all three lignin units (H, G, and S) (Table 2). As in wild-type flax lignin, the G unit is the major monomer present in the *lbf1* ectopic lignin and constitutes ~70% of total released lignin units (Table 2, Figure 4A). No significant changes in the S/G ratio were observed, indicating that the ectopic lignin is similar to that previously analyzed in flax (Day et al., 2005; del Río et al., 2011). Thioacidolysis only releases lignin units from noncondensed intermonomeric bonds (β-O-4 and α-O-4); therefore, calculation of the amount of S and G units released, divided by the amount of acetyl bromide lignin values (Table 2; S+G/lignin values), provides an estimate of the relative condensation of the lignin polymer (i.e., the proportion of condensed bonds, such as β-5:phenylcoumaran and β-β:resinol). No significant changes in these values were seen (Table 2, Figure 4B), suggesting that increased lignification in the *lbf1* mutant is not associated with changes in the degree of lignin condensation.

Our previous results (Figure 3) had shown no increase in the lignin content of *lbf1* stem inner tissues. Thioacidolysis showed that lignin composition was much less affected than in outer tissues (Table 2). A slight but significant decrease in S units was observed resulting in a lower S/G ratio. As for *lbf1* outer tissues, no modification in lignin condensation was observed (Figure 4C).

Flax lignin is generally highly condensed, and we therefore used 2D NMR analyses to provide complementary information on the condensed fraction of the lignin polymer. The heteronuclear single quantum coherence (HSQC) spectra (δ_C/δ_H 45 – 125/2.5 – 7.2) of the acetylated cell wall from flax fibers are shown in Figure 5. Our results show the presence of G and S signals in the outer stem tissues of the *lbf1* mutant, whereas barely any lignin signals were detected in the spectra of the wild-type plants. Various signals from cellulose and polysaccharides were also detected in both spectra. For quantification of the different interlinkages, we used the guaiacyl (G_2) C_2-H signal as internal standard. Then all signals assigned to the various interunit linkage types were integrated: β-O-4' alkyl-aryl ether linkages (Aα and Aβ), phenyl coumaran β-5′/α-O-4' linkages (Bα and Bβ), and resinol β-β'/α-O-γ'/γ-O-α' linkages (Cα and Cβ). The main lignin substructures observed in *lbf1* outer tissues correspond to the β-O-4' alkyl-aryl ether, with 52.3% (Aα) and 47.9% (Aβ) relative abundance signals, respectively, followed by phenylcoumaran and resinol with 14.5% (Bα) and 18% (Bβ), and 13.6% (Cα) and 14.7% (Cβ), respectively. These values are in good agreement with previous HSQC analyses of flax fiber milled wood lignin (del Río et al., 2011). Signals were also assigned to other minor substructures (Aγ, Bγ, and Cγ, dibenzodioxocin). The S/G ratio (0.12) estimated by NMR was lower than the molar ratio (0.28) determined by thioacidolysis. H units (data not shown) were difficult to quantify in *lbf1* lignin

Table 2. Lignin Monomeric Composition Determined by Thioacidolysis in Inner- and Outer-Stem Tissues of Flax *lbf1* Mutants and Wild-Type Plants

	Outer Tissues		Inner Tissues	
	Wild Type	*lbf1*	Wild Type	*lbf1*
H	1.14 ± 0.45	**3.33 ± 2.07**	3.11 ± 0.82	2.96 ± 0.70
G	10.10 ± 5.80	**26.60 ± 10.31**	243.37 ± 41.09	243.05 ± 28.09
S	3.25 ± 1.34	**6.94 ± 2.64**	49.02 ± 3.73	**32.85 ± 7.98**
S/G	0.34 ± 0.05	0.26 ± 0.11	0.20 ± 0.03	**0.13 ± 0.02**
S+G	13.35 ± 7.14	**35.03 ± 11.53**	292.39 ± 42.14	275.90 ± 33.73
S+G/lignin	329.90 ± 203.75	301.49 ± 81.87	872.00 ± 79.94	968.70 ± 147.75

H, G, S = yields of the thioethylated products of *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) lignin units expressed as μmoles per gram of CWR. S/G = ratio of G to S lignin units. S+G = total S plus G lignin units expressed as μmoles per gram of CWR. S+G/lignin = yields of total lignin (S+G) expressed as μmoles per gram of lignin. For *lbf1*, values represent the mean values obtained with 16 and 33 individual plants for inner and outer tissues, respectively. For the wild type, inner and outer values are the mean ± SD from three individual plants. Values significantly different from the wild type at $P < 0.05$ are indicated in bold.

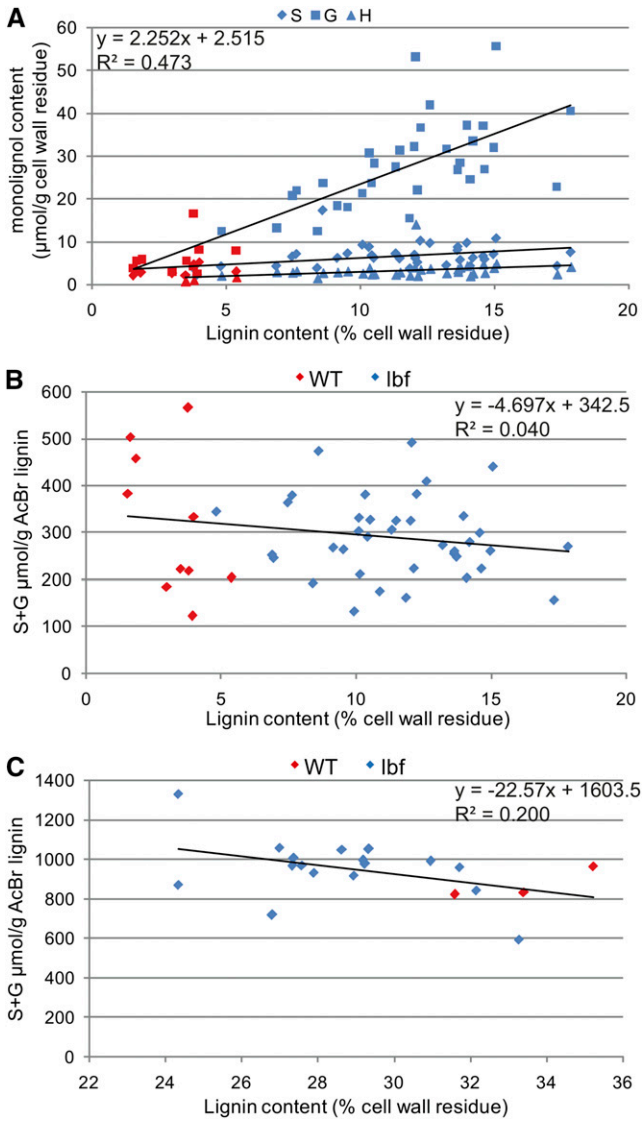


Figure 4. Lignin Analyses in *lbf1* Inner- and Outer-Stem Tissues.

(A) Relationship between lignin content and the composition in H (triangle), G (square), and S (diamond) lignin units in outer tissues of wild-type and *lbf1* mutants. Correlation coefficient given for G units only.
(B) Relationship between lignin condensation and lignin content in outer tissues.
(C) Relationship between lignin condensation and lignin content in inner tissues. Lignin content was determined by acetyl bromide analyses; S/G ratios and lignin composition were determined by thioacidolysis. Blue dots correspond to individual *lbf1* mutants and red dots correspond to wild-type plants.
[See online article for color version of this figure.]

because of masking by polysaccharide signals in the HSQC spectra. For *lbf1* inner-stem tissues, the HSQC spectra confirmed chemical analyses, suggesting that the inner tissue lignin structure of *lbf1* mutants was very similar to that of wild-type plants. No changes in the proportion of lateral chains were observed with aryl ether, phenyl coumaran, and resinol bonds, representing 58.6, 16.0, and 14.2%, respectively. In contrast to outer-stem tissues, the S/G ratio (0.12) estimated from *lbf1*2D NMR spectra (Supplemental Figure 1) was very similar to that estimated by thioacidolysis (0.13).

Further information on *lbf1* lignin was obtained using the KM1 antibody targeted against lignin phenylcoumaran linkages (β-5) (Kiyoto et al., 2013). TEM observation showed that labeling was present in the whole cell wall (primary cell wall, secondary cell wall) of *lbf1* fibers, but was conspicuously absent in wild-type fibers (Figure 6). By contrast, a similar comparison of *lbf1* and wild-type xylem tissues did not show any differences with labeling being detected in the middle lamella and primary and secondary cell walls of both mutant and wild-type plants (Figure 6).

Liquid Chromatography-Mass Spectrometry Profiling Reveals Modifications of the Oligolignol Pool in the *lbf1* Mutant

Oligolignol liquid chromatography-mass spectrometry (LC-MS) profiling provides information about the availability and nature (e.g., glycosylated versus nonglycosylated) of lignin monolignols and di/trilignols present in lignifying tissues. The composition of the oligolignol pool is closely related to polymeric lignin structure and gives complementary insight into the lignification process and metabolic flow (Morreel et al., 2010a). Changes in lignin content are often accompanied by modifications in the soluble oligolignol pool, and we therefore performed an LC-MS-based metabolite profiling of ethanol extracts of inner- and outer-stem tissues of *lbf1* mutants and wild-type plants (Table 3; Supplemental Figure 2). Because of the strong variation in intensities of many of the phenolic compounds observed, only the compounds that were consistently above the detection limit in all samples, or consistently below the detection limit in all samples, within each of the four categories (wild-type inner, mutant inner, wild-type outer, and mutant outer), were taken into account. In the outer tissues, four nonhexosylated dilignols and eight nonhexosylated trilignols that were consistently detected in the wild-type plants were not detected in the mutants. In addition, one nonhexosylated dilignol was significantly less abundant in *lbf1* outer tissues compared with corresponding wild-type samples. In stem inner tissues, six compounds (coniferin, syringin, and four hexosylated dilignols) were only detected in the mutants, and two compounds (two hexosylated dilignols) were significantly more abundant in the mutants than in the wild type. In these eight compounds, coniferyl alcohol moieties were overrepresented with respect to sinapyl alcohol moieties: The intensity of coniferin was 27 times higher than that of syringin, and sinapyl alcohol moieties were present in only two out of the eight upregulated compounds. The MS data and structure of five oligolignols identified for the first time in flax are shown in Supplemental Figure 2.

lbf1 Ectopic Lignification Is Associated with Modifications to Other Cell Wall Polymers

To investigate whether the modified lignin content in flax outer tissues was also associated with changes to other cell wall polymers, we first used antibodies targeted against the main classes of cell wall polymers: (1) pectin (galactan)-LM5 (Jones et al., 1997); (2) hemicellulose-LM10, LM11 (McCartney et al., 2005), and LM21 (Marcus et al., 2010); and (3) arabinogalactan proteins

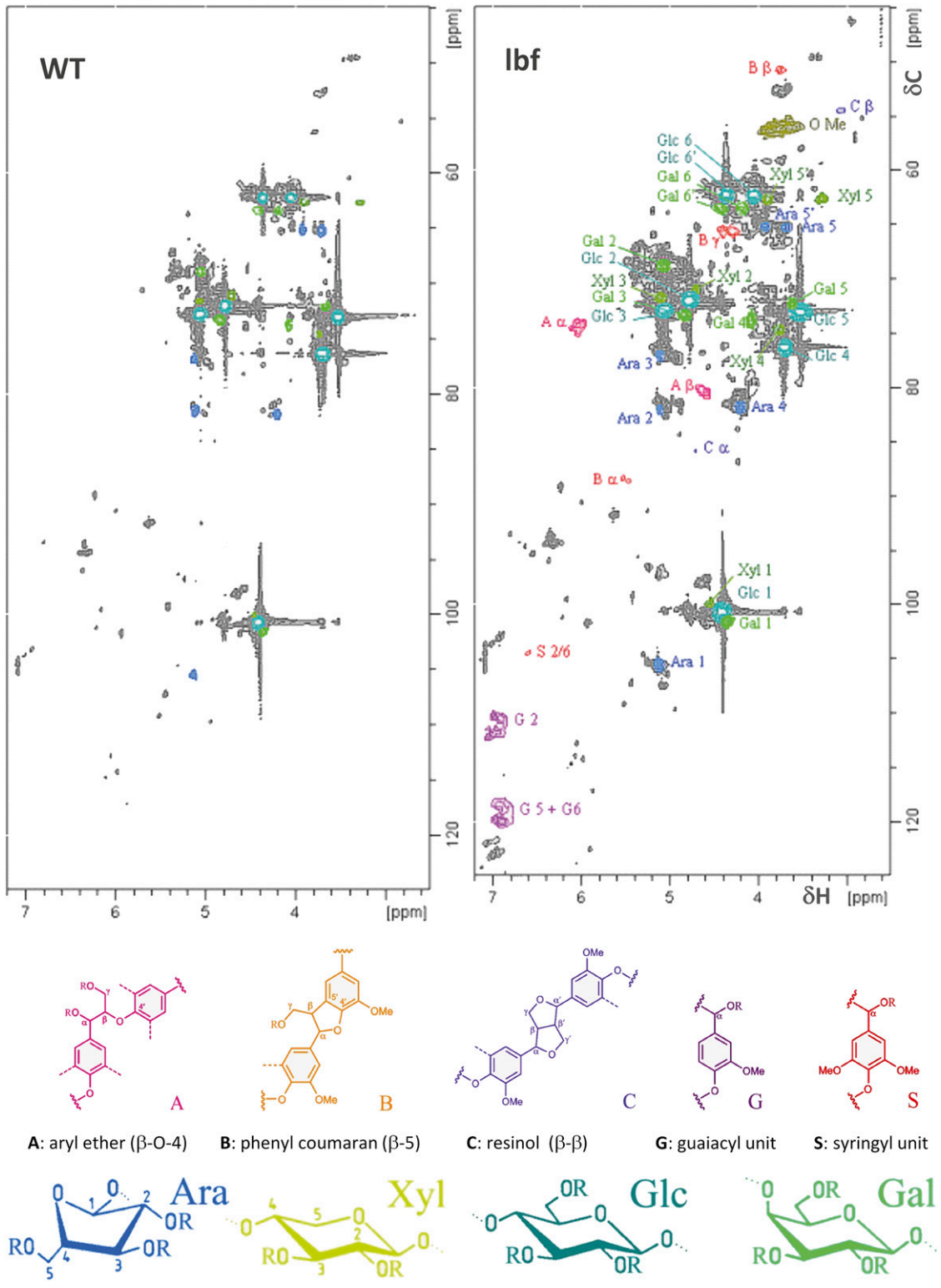


Figure 5. 2D NMR Spectra Revealing Lignin Monomers, Interunit Distribution, and Sugar Signals.

Partial short-range (HSQC) spectra (d_C/d_H 45 – 125/2.5 – 7.2) of the acetylated cell wall from wild-type (left) and *lbf1* (right) outer tissues.

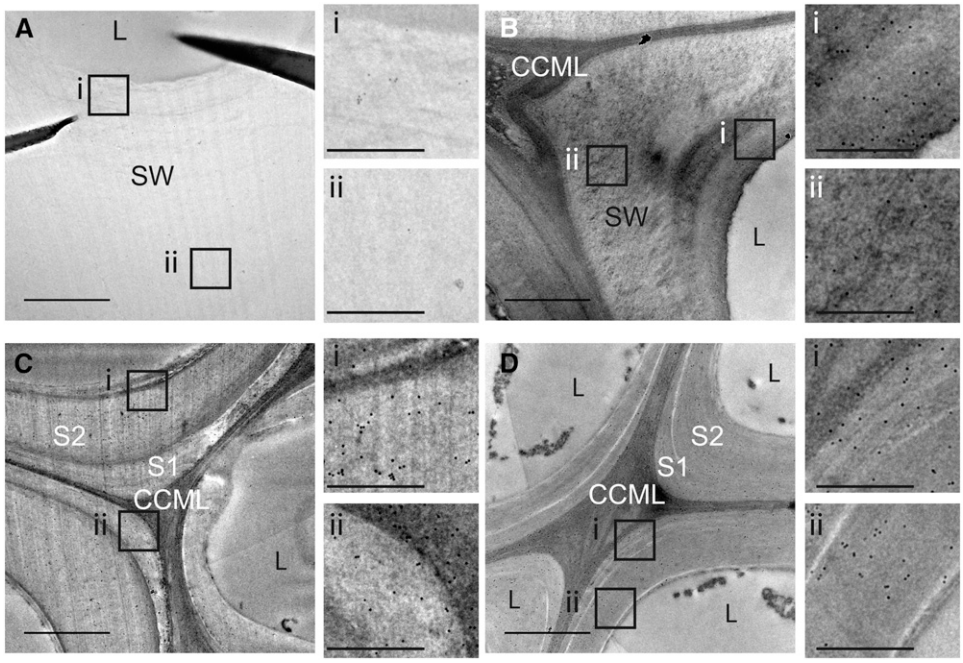


Figure 6. Immunogold Silver Staining of a Transverse Section of the Flax Stem Median Region with KM1 Antibody.

Bast fibers (**[A]** and **[B]**), xylem fibers (**[C]** and **[D]**), wild-type (**[A]** and **[C]**), and *lbf1* (**[B]** and **[D]**). CCML, cell corner middle lamella; SW, secondary wall; L, lumen. Smaller photos (i and ii) on the right side of each main photo (**[A]** to **[D]**) show a zoom of the corresponding regions indicated on main photo. Bars = 1 μm (main photos) and 0.15 μm (small photos).

(AGPs)-LM2 (Smallwood et al., 1996; Yates et al., 1996) and JIM14 (Knox et al., 1991; Yates and Knox, 1994; Yates et al., 1996). LM5, JIM14, LM10, and LM11 showed stronger fluorescence in *lbf1* outer tissues than in wild-type outer tissues (Figure 7), suggesting that cell walls in this mutant are enriched in pectin, hemicellulose, and glycoprotein compared with the wild type. LM5 labeling appeared strongly on the inner part of the *lbf1* secondary wall, whereas labeling appeared weakly on the whole secondary cell wall of the wild type. JIM14 labeling is restricted to the inner part of the secondary wall in the wild type, whereas epitope distribution in the mutant seems more diffuse in the secondary wall. Increased LM10 and LM11 labeling was apparent on the whole secondary cell wall. LM21 gave fluorescence labeling in the thick secondary cell walls and in the inner secondary wall layer in the wild type and *lbf1*, respectively, suggesting modest changes in mannan hemicelluloses content. By contrast, almost no labeling was observed with LM2 antibodies in the bast fibers of both *lbf1* and the wild type. Measurements of bast fiber cell wall thickness also indicated that the cell walls of *lbf1* mutants are generally thinner (3.8 ± 1.1 μm) than corresponding cell walls in wild-type plants (9.9 ± 2.2 μm). When *lbf1* and wild-type inner stem tissues were compared, no differences in antibody labeling were seen (Supplemental Figure 3).

Further information on cell wall polymer modifications in the *lbf1* mutant was obtained by analyzing total sugar content in stem outer tissues. Our results (Figure 8A) showed that increased lignification was correlated with a reduction in total sugar (mainly glucose) content when the latter was expressed as a percentage of the dry cell wall residue content. When the quantities of individual sugars were expressed as a percentage of the total sugar content (Figure 8B), glucose decreased from 87% (wild-type) to 70% (*lbf1*) total sugars, suggesting that cellulose content was reduced in *lbf1* mutants. Further analyses with trifluoroacetic acid that does not degrade crystalline cellulose (Crônier et al., 2005) showed that the amounts of trifluoroacetic acid-released glucose accounted for 10.0% ± 1.5% and 9.8% ± 0.3% of glucose released by total hydrolysis of *lbf1* and wild-type cell walls, respectively, thereby suggesting that the proportion of crystalline to noncrystalline cellulose is unchanged in mutant outer tissues. By contrast, the relative proportion of sugar monomers (Fuc, Ara, Rha, Gal, Xyl, Man, GalA, and GlcA) from other NCPs increased in the outer tissues of *lbf1* mutants (Figure 8B) in agreement with outer tissue NMR data (Figure 5). For *lbf1* inner tissues, there was no significant decrease in total sugar content and only a slight decrease in the relative proportion of glucose (Figure 8C). The relative proportions of other sugars increased, but less significantly than in the outer tissues. Our JIM14 results (Figure 7) indicated increased AGP content in *lbf1* outer stem tissues, and we therefore quantified nitrogen levels in order to estimate relative protein content. Our results (Figure 8D) revealed that increased lignification was correlated with increased protein content.

Transcriptomics Suggests a Role for Lignin-Related Peroxidases in the *lbf1* Phenotype

To obtain information about modifications in gene expression associated with the *lbf1* phenotype, we performed whole-genome

Table 3. Identified Differentially Accumulating Phenolics in Inner- and Outer-Stem Tissues of the *lbf1* Mutant as Revealed by LC-MS

		Outer Tissues		Inner Tissues	
t_R (min)	Compound	Wild Type	*lbf1*	Wild Type	*lbf1*
7.75	Coniferin	n.d.	n.d.	n.d.	**131,603 ± 28,678**
10.48	Syringin	n.d.	n.d.	n.d.	**4,788 ± 695**
12.88	*__G__(8-O-4)__G'__ hex	n.d.	n.d.	n.d.	**4,294 ± 2,419**
13.30	*__G__(8-O-4)__G'__ hex	n.d.	n.d.	n.d.	**6,187 ± 3,826**
14.90	*__G__(8-O-4)__G'__ hex	n.d.	n.d.	401 ± 201	**11,431 ± 3,315**
15.12	__G__(*e*8-O-4)__S__ hex	n.d.	n.d.	n.d.	**25,94 ± 866**
15.79	__G__(8-5)__G__ hex	n.d.	n.d.	9,062 ± 4,375	**266,534 ± 46,161**
15.86	__G__(*t*8-O-4)__G__	1,466 ± 1,078	*n.d.*	13,168 ± 2,265	19,750 ± 10,882
15.87	Lariciresinol hex	n.d.	n.d.	n.d.	**5,803 ± 1,963**
16.36	__G__(*e*8-O-4)__G__	1,738 ± 1,060	*n.d.*	14,266 ± 2,556	23,642 ± 9,543
16.70	*__G__(*e*8-O-4)__FA__	50,087 ± 22,282	*7,725 ± 3,722*	4,327 ± 1,503	11,461 ± 5,925
21.48	Lariciresinol	3,533 ± 980	*n.d.*	2,808 ± 1,866	6,582 ± 2,210
21.57	__G__(*t*8-O-4)secoisolariciresinol	2,153 ± 441	*n.d.*	342 ± 73	985 ± 369
22.17	__G__(*e*8-O-4)lariciresinol	3,208 ± 32	*n.d.*	8,056 ± 2,475	5,673 ± 1,576
22.79	*__G__(*t*8-O-4)lariciresinol	764 ± 83	*n.d.*	1,429 ± 465	925 ± 263
23.65	*__G__(8-5)__FA__	2,474 ± 327	*n.d.*	129 ± 53	1,089 ± 362
24.54	__G__(*t*8-O-4)__S__(8-5)__G__	873 ± 237	*n.d.*	29,468 ± 5,586	28,725 ± 8,515
24.93	__G__(*t*8-O-4)__S__^{red}/__S__(8-8/5)__G__^{red}/__G__	1,377 ± 614	*n.d.*	2,184 ± 522	1,322 ± 373
25.80	__G__(*e*?8-O-4)__G__(8-5)__G'__	338 ± 3	*n.d.*	9,551 ± 1,938	4,253 ± 1,477
27.55	*__G__(*t*8-O-4)__S__(8-8)__S__	1,554 ± 457	*n.d.*	2,457 ± 802	4,462 ± 1,308
28.86	*__G__(*e*8-O-4)__S__(8-8)__S__	232 ± 52	*n.d.*	593 ± 144	963 ± 272

Values represent the relative abundance based on the extracted ion chromatogram, expressed as per mg dry weight tissue; values significantly different (Student's *t* test) from the wild type at $P < 0.05$ are indicated in bold or italics, respectively, when they are higher or lower in abundance. n.d., not detected. Nomenclature is based on Morreel et al. (2004). Guaiacyl units, syringyl units, and units derived from ferulic acid and coniferaldehyde are referred to as **G**, **S**, **FA**, and **G'**, respectively. The linkage type is indicated in parentheses. "red," reduced unit or adjacent linkage (Morreel et al., 2010a). A forward slash indicates that two units or two linkage positions are equally possible at this position in the shorthand name. hex, hexose or hexoside; t_R, retention time. Asterisks indicate compounds that have not been described previously in flax stem tissue. Spectral data and structures of these compounds are given in Supplemental Figure 2. The spectral data of the other compounds are described by Huis et al. (2012). Three wild-type and six *lbf1* plants were analyzed. Values are means ± SE. Only the differentially accumulating metabolites with known identities are shown.

transcriptomics using flax-specific Agilent microarrays. Gene expression patterns in inner- and outer-stem tissues from six individual lignified *lbf1* mutants were compared with corresponding tissues from wild-type plants. Our results (Figure 9A; Supplemental Data Set 1) show that transcripts of 1487 genes were significantly more abundant (P value < 0.05) in the *lbf1* mutants as compared with wild-type plants. Of these 1487 transcripts, 959 were specifically more abundant in stem outer tissues and 277 were specifically more abundant in stem inner tissues; transcripts for 250 genes were more abundant in both tissues (Figure 9A). A total of 1197 transcripts were less abundant in the *lbf1* mutants (Figure 9A; Supplemental Data Set 6), of which 806 were specifically less abundant in stem outer tissues, 294 were specifically less abundant in stem inner tissues, and 96 were less abundant in both inner- and outer-stem tissues of *lbf1* mutants when compared with wild-type plants.

Functional classification using Gene Ontology (GO) (Figure 9B) showed that the differentially accumulated transcripts are implicated in diverse biological processes, molecular functions, and transport. For example, 5.6% (149 genes) of the differential transcript abundance is related to genes involved in biosynthesis and maintenance of the plant cell wall. Examination of the 20 most abundant transcripts in outer tissues showed that the most represented class corresponded to defense genes (5) and oxidation/reduction genes (5) (Table 4). For inner tissues, the two most represented classes were transport and cellular process with four genes in each class. Among the 20 least abundant transcripts in outer-stem tissues (Table 5), the classes biological process and unknown were the two most represented. In mutant inner tissues, the least abundant transcript (Lus10038721) corresponds to a homolog of *CCD8* belonging to the carotenoid cleavage dioxygenase family (Leyser, 2008). In *Arabidopsis thaliana*, a mutation in this gene is associated with a decrease in strigolactone content and increased axillary bud production (Sorefan et al., 2003). It is possible that the reduced transcript accumulation of the flax putative *CCD8* ortholog is related to the branched phenotype of the *lbf1* mutants.

Increased lignification is the major observed cell wall phenotype in *lbf1* mutants, and we therefore focused our attention on transcripts corresponding to two major control points in the lignification process: (1) monolignol biosynthesis and (2) monolignol polymerization. Following interrogation of the *Arabidopsis* database, sequence alignment, and phylogenetic analyses, we identified a total of 48 putative genes involved in monolignol biosynthesis in the flax genome. Transcripts corresponding to 22 of these genes were differentially accumulated between *lbf1* mutants and the wild type. In outer tissues, transcripts corresponding to a *CCR*, a *COMT*, and a *CAD* gene were significantly more abundant. For inner tissues, transcripts corresponding to 19 lignin

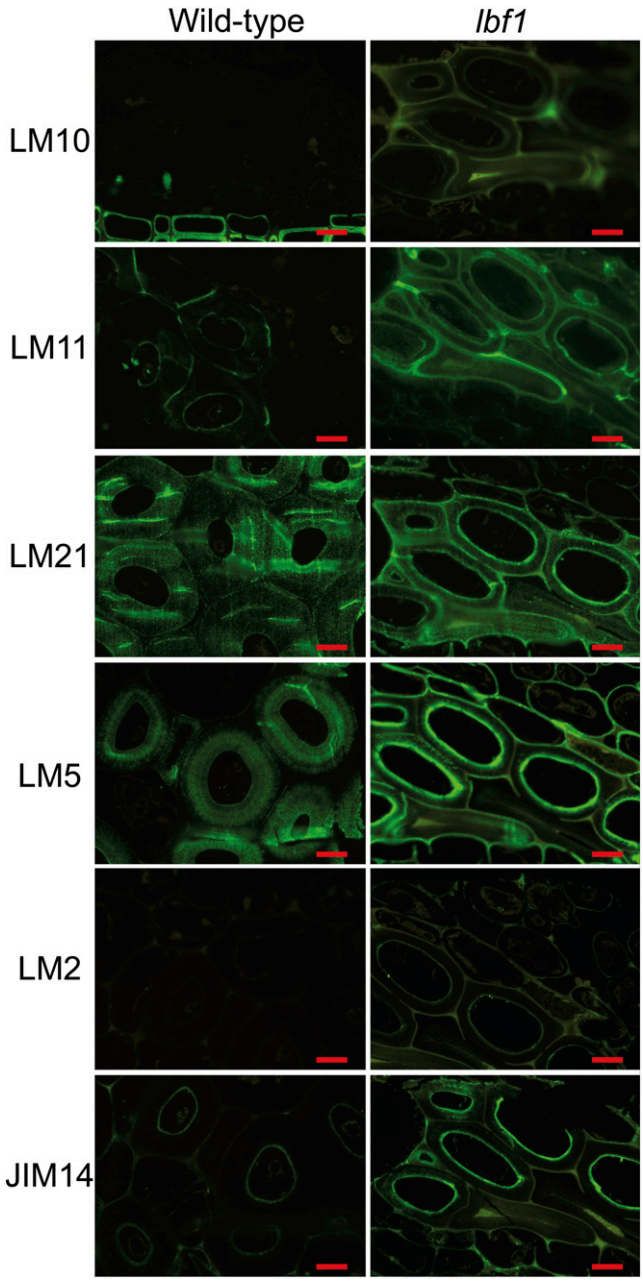


Figure 7. Fluorescent Microscopy Immunolocalization of Cell Wall NCPs with LM10, LM11, LM21, LM5, LM2, and JIM14 Antibodies.

Bars = 10 μm.
[See online article for color version of this figure.]

genes (2 × *PAL*, 4 × *4CL*, 3 × *C3H*, 1 × *F5H*, 2 × *CCoAOMT*, 1 × *CCR*, and 6 × *CAD*) genes were significantly less abundant (Figure 10).

During lignification, the synthesized monolignols are exported to the cell wall where they are oxidized by laccases and/or peroxidases prior to polymerization into the lignin polymer. Analyses of transcriptomics data showed that no laccase transcripts were differentially accumulated between *lbf1* and wild-type outer tissues. By contrast, transcripts corresponding to the *laccase11* (*LAC11*) gene were significantly less abundant in *lbf1* inner tissues. Transcripts corresponding to 16 peroxidase genes showed significant differential accumulation between *lbf1* mutants and wild-type plants (Figure 11A; Supplemental Data Set 2). Transcripts for 11 of these genes were more abundant uniquely in outer stem tissues, transcripts for one gene were more abundant uniquely in inner stem tissues, and transcripts for three genes were more abundant in both tissues. Transcripts for one peroxidase gene were significantly less abundant in the *lbf1* mutant. A phylogenetic tree (Figure 11B; Supplemental Data Set 2) based on an alignment of protein sequences of both flax and *Arabidopsis* peroxidases shows that 9 of the 11 flax peroxidase transcripts specifically more abundant in *lbf1* outer tissues are phylogenetically close to three distinct At-PRXs (At-PRX52, At-PRX53, and At-PRX71) known to oxidize monolignols and therefore are potentially involved in lignin polymerization (Østergaard et al., 2000; Nielsen et al., 2001; Herrero et al., 2013; Shigeto et al., 2013).

Peroxidases require H_2O_2 to oxidize monolignols in order to make lignin. H_2O_2 is produced through the action of two types of enzyme: (1) NADPH-oxidase enzymes, and more specifically RBOH enzymes, that produce superoxide ions; and (2) superoxide dismutase, which converts superoxide ions into H_2O_2 (Karpinska et al., 2001; Karlsson et al., 2005). We identified 14 flax orthologs of the 10 *RBOH* genes identified in *Arabidopsis* (Torres, 2010). Transcripts corresponding to five of these genes were specifically more abundant in *lbf1* outer-stem tissues. Phylogenetic analyses indicated that two of these genes are closely related to At-RBOH-*F* (Figure 12), recently shown to be involved in Casparian strip lignification (Lee et al., 2013). The other three flax *RBOH* genes are orthologs of *AtRBOH-A* and *C* genes involved in defense related apoplastic H_2O_2 production (Schweizer, 2008). Our analyses also showed that transcripts corresponding to another At-*RBOH-F* ortholog were significantly more abundant in both inner- and outer-stem tissues of the mutant compared with wild-type plants (Figure 12; Supplemental Data Set 3). Finally, our data (Supplemental Data Set 1) indicated that a transcript corresponding to a superoxide dismutase gene was specifically more abundant in *lbf1* outer tissues.

DISCUSSION

Lignification plays an important role in plant biology and has a major impact on the quality of a wide range of different products derived from plants. In timber, the presence of lignin is positive as it provides rigidity and mechanical support to fiber cell walls. In contrast, the presence of lignin inhibits saccharification during biofuel production and therefore has a negative effect on the quality of lignocellulosic biomass. The lignin polymer is initially deposited in the preexisting middle lamella and primary wall of cells during the formation of the secondary cell wall. Lignin deposition then continues in the secondary wall with the result that most secondary plant cell walls contain relatively high amounts of lignin. This type of lignification is typical of the cell walls of xylem fibers, vessels, and tracheids. By contrast, bast fiber plants, such as flax, ramie, and jute, have been exploited by man for many thousands of years precisely because their stems also contain

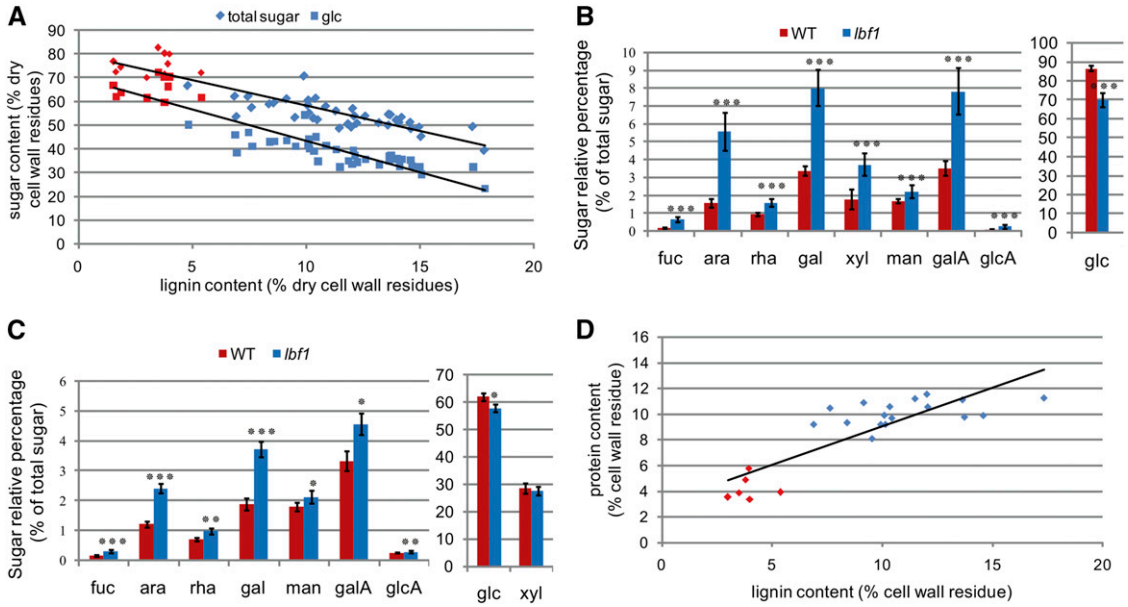


Figure 8. Sugar and Protein Analyses in *lbf1* Inner- and Outer-Stem Tissues.

(A) Relationship between sugar content and lignin content in outer tissues of wild-type and mutant plants. Diamond-shaped dots correspond to total sugar content and square dots correspond to glucose content (red, wild-type plants; blue, *lbf1* mutants).
(B) Relative content of different sugars in outer tissues of wild-type and *lbf1* mutants. Glucose content is separated from other sugars due to the scale difference.
(C) Relative amounts of different sugars in inner tissues of wild-type and lignified mutants. Glucose and xylose are separated from others sugars due to the scale difference.
(D) Relationship between lignin and protein content. Lignin content was determined by acetyl bromide and protein content by nitrogen dosage (red, wild-type plants; blue, *lbf1* mutant plants).
For **(B)** and **(C)**, significant differences (Student's *t* test) between the wild type and mutant were observed at $P < 0.001$ (***), $P < 0.01$ (**), and $P < 0.05$ (*). Error bars = SD.

elongated fiber cells with thick cellulose-rich secondary cell walls but only low amounts of lignin (Day et al., 2005; del Río et al., 2011). It therefore appears that certain plant species possess particular regulatory mechanisms that allow them to construct thick nonlignified secondary cell walls. A better understanding of these mechanisms could provide novel targets for engineering of plant biomass. In flax stems, the outer tissues containing the cellulose-rich bast fibers can be easily separated from the inner tissues containing the lignified secondary xylem cells, thereby allowing comparative studies of cell wall formation in these two tissues (Fenart et al., 2010; Huis et al., 2012). To learn more about the mechanisms regulating cell wall biosynthesis in flax, we used a combination of UV autofluorescence and phloroglucinol-HCl staining to screen a flax EMS mutant population for mutants showing altered bast fiber lignification patterns (Chantreau et al., 2013). This approach allowed us to identify 93 families showing increased lignification in bast fibers, and we then went on to characterize one of these mutants (*lbf1*) in detail.

Characterization of the Flax *lbf1* Mutant

(1) Lignin and Oligolignols

Chemical analyses of bast fiber ectopic lignin monomeric composition in *lbf1* mutants showed significant increases in the amounts of all three lignin monomers with no significant modification in the S/G ratio, indicating that lignin structure was unchanged. Flax lignin is particularly condensed and therefore only ~10% of outer tissue lignin and 20% of inner tissue lignin are probably accessible via thioacidolysis disruption of noncondensed alkyl-aryl ether linkages (Day et al., 2005). We therefore used NMR analysis of solubilized cell wall samples to complete the chemical data. These results confirmed that the chemical composition of bast fiber ectopic lignin was rich in G units. NMR data provided a lower S/G ratio than that obtained with thioacidolysis, suggesting a preferential involvement of S units in alkyl-aryl ether in agreement with previous NMR analysis of milled wood lignin (del Río et al., 2011). Immunolabeling of phenylcoumaran in *lbf1* mutant bast fiber walls was in good agreement with lignin analysis showing a noticeable amount of side chains involved in this structure.

Oligolignol profiling indicated that ectopic lignification in the outer stem tissues of the *lbf1* mutant was accompanied by a strong decrease in the accumulation of nonhexosylated oligolignols in that tissue. We have previously shown that a wide range of (mono)oligolignols normally accumulates in this tissue in wild-type flax, and it is possible that their levels decrease in the mutant because they are incorporated into the lignin polymer (Huis et al., 2012). This hypothesis was supported by the observation that the depleted lignin oligolignols in *lbf1* bast fibers were mainly composed of G units and several contained phenylcoumaran linkages

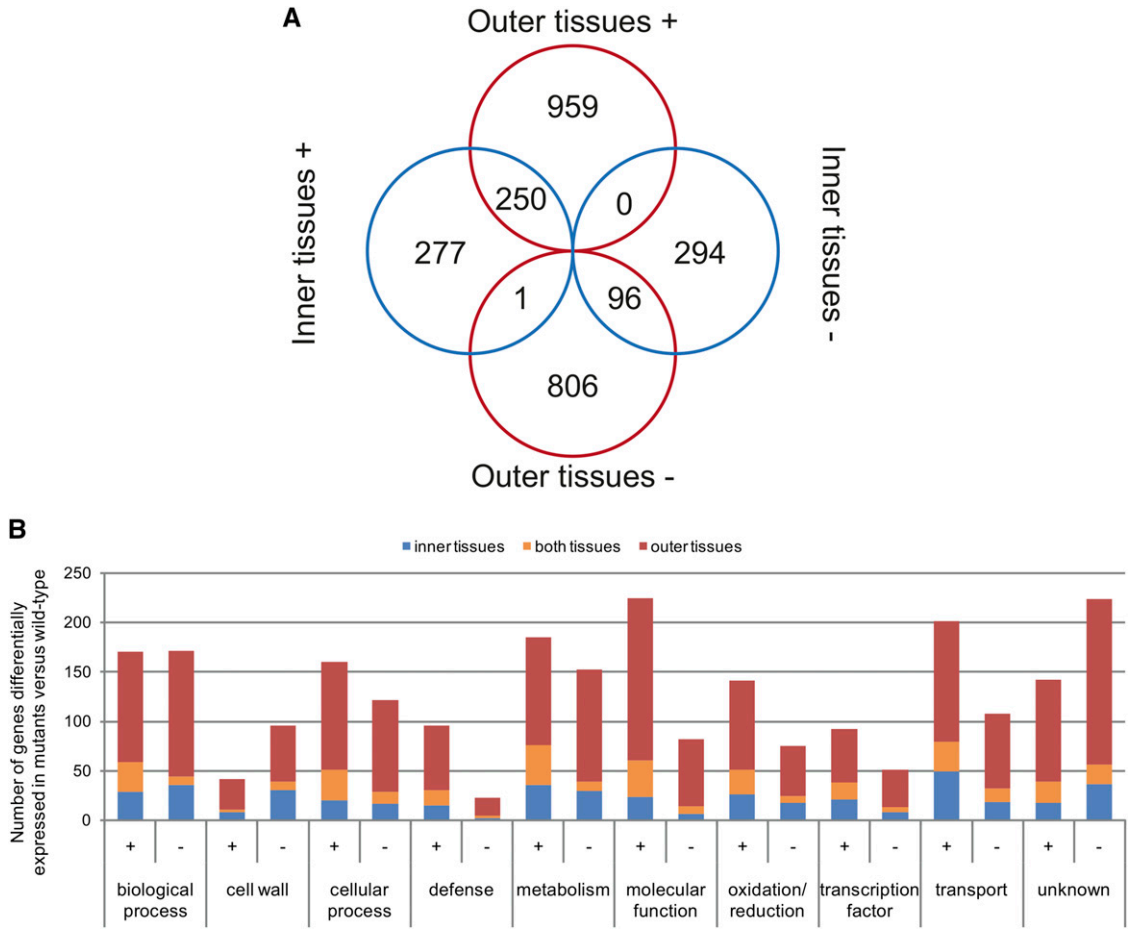


Figure 9. Differentially Accumulated Transcripts in *lbf1* Mutants versus the Wild Type in Separated Tissues.

(A) Venn diagram; ±, over/underaccumulated (P value < 0.05, Bonferroni method) transcripts in *lbf1* stem tissues versus corresponding tissues of the wild type.

(B) GO classification of differentially accumulated transcripts in outer (red), inner (blue), and both (orange) tissues. GO classification was determined using blast2go on protein sequences (NCBI) and verified by expert curation.

in agreement with the chemical, NMR, and immunological analyses. These results would suggest that hypolignification in wild-type flax bast fibers is not so much caused by a lack of lignin precursors but is rather due to insufficient polymerization.

The polymerization of lignin occurs via radical coupling of monolignol and oligolignol radicals, which are formed by peroxidase and/or laccase activity (Zhao et al., 2013). In support of peroxidase involvement in ectopic bast fiber lignification, we observed increased transcript levels of nine lignin-related peroxidase genes specifically in the outer tissues of *lbf1* mutants compared with wild-type plants. Peroxidase activity was previously reported to be associated with the onset of lignification in flax fibers (McDougall, 1991, 1992) and peroxidase ESTs/genes are highly represented/expressed in flax outer stem cDNA libraries (Day et al., 2005; Roach and Deyholos, 2007) and tissues (Fenart et al., 2010; Huis et al., 2012). Based on microarray data, laccase genes are probably more closely associated with lignification of flax xylem tissues, but not bast fibers (Huis et al., 2012). The transcriptomics data from the *lbf1* mutant suggest that flax outer-stem peroxidases and not laccases are responsible for the increased lignification. It would obviously be interesting to characterize other flax *lbf* mutants and/or create laccase overexpressors to investigate whether lignified bast fibers could be induced by upregulating laccase gene expression.

Peroxidases, but not laccases, require H_2O_2 for radical production, and we also observed increased transcript levels in *lbf1* outer-stem tissues of five NADPH oxidase genes. Interestingly, two of these flax NADPH-oxidase genes are homologs of *Arabidopsis* type RBOH-F NADPH-oxidases recently shown to be involved in the polymerization of lignin within the Casparian strip of the endodermis (Lee et al., 2013). The highly localized Casparian strip lignification in *Arabidopsis* occurs through docking proteins, called Casparian strip domain proteins, which are targeted to the area of the Casparian strip and recruit both an NADPH oxidase and a peroxidase. Such enzyme assemblies then direct localized oligolignol polymerization to form the Casparian strip. The coordinated overexpression of NADPH oxidases and

Table 4. List of the 20 Most Highly Abundant Transcripts in *lbf1* Mutants versus the Wild Type

Reference	Name	Delta	P Value	GO Annotation	*Arabidopsis* Correspondence
20 Most Abundant Transcripts in Outer Tissues					
Lus10020493	Pathogenesis-related gene 1	7.41	0.00E+00	Defense	AT2G14610.1
Lus10003264	Pathogenesis-related 4	6.98	0.00E+00	Defense	AT3G04720.1
Lus10006925	Terpenoid cyclases/protein prenyltransferases superfamily protein	6.70	0.00E+00	Metabolism	AT4G02780.1
Lus10028898	Cytochrome P450, family 76, subfamily C, polypeptide 4	6.64	2.22E-16	Oxidation/reduction	AT2G45550.1
Lus10022642	LYS/HIS transporter 7	6.36	4.44E-16	Transport	AT4G35180.1
Lus10003339	Transmembrane amino acid transporter family protein	6.35	2.22E-16	Transport	AT1G47670.1
Lus10004958	Somatic embryogenesis receptor-like kinase 2	6.35	4.44E-16	Molecular function	AT1G34210.1
Lus10012684	Peroxidase superfamily protein	6.29	0.00E+00	Oxidation/reduction	AT2G41480.1
Lus10020826	Peroxidase superfamily protein	6.15	1.78E-15	Oxidation/reduction	AT2G41480.1
Lus10032178	Unknown	6.14	2.22E-16	Unknown	
Lus10014508	Unknown	6.08	3.77E-15	Unknown	
Lus10030945	Nitrate transporter 1.5	6.03	4.46E-14	Transport	AT1G32450.1
Lus10015339	Unknown	5.96	1.11E-15	Defense	
Lus10035241	Glutathione *S*-transferase tau 7	5.95	5.33E-15	Molecular function	AT2G29420.1
Lus10039454	MLP-like protein 423	5.91	3.49E-14	Defense	AT1G24020.1
Lus10022415	2-Oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	5.90	4.44E-16	Oxidation/reduction	AT1G06620.1
Lus10035221	Matrixin family protein	5.54	2.75E-14	Cellular process	AT1G24140.1
Lus10004410	Pathogenesis-related thaumatin superfamily protein	5.51	3.88E-11	Defense	AT1G20030.2
Lus10008173	Peroxidase superfamily protein	5.47	4.88E-15	Oxidation/reduction	AT5G06730.1
Lus10025253	Protein of unknown function (DUF567)	5.45	1.39E-11	Transport	AT5G01750.2
20 Most Abundant Transcripts in Inner Tissues					
Lus10000453	Homolog of carrot EP3-3 chitinase	6.38	4.44E-16	Cell wall	AT3G54420.1
Lus10016323	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	5.87	2.22E-16	Transport	AT5G48490.1
Lus10002741	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	5.79	0.00E+00	Transport	AT5G48490.1
Lus10031759	Plant natriuretic peptide A	5.72	0.00E+00	Defense	AT2G18660.1
Lus10007270	P-loop-containing nucleoside triphosphate hydrolases superfamily protein	5.67	8.44E-15	Molecular function	AT3G28540.1
Lus10005395	Unknown	5.14	1.11E-11	Unknown	
Lus10021102	Glutathione *S*-transferase TAU 8	5.12	1.78E-15	Cellular process	AT3G09270.1
Lus10032930	2-Oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	5.11	6.66E-16	Oxidation/reduction	AT5G24530.1
Lus10010702	MLP-like protein 423	4.87	1.22E-14	Defense	AT1G24020.1
Lus10034484	Chaperone DnaJ-domain superfamily protein	4.82	2.27E-13	Cellular process	AT4G36040.1
Lus10015350	Disease resistance protein (TIR-NBS-LRR class) family	4.82	5.46E-14	Defense	AT4G12010.1
Lus10005523	Unknown	4.75	6.20E-12	Unknown	
Lus10025060	Chaperone DnaJ-domain superfamily protein	4.71	2.11E-12	Cellular process	AT4G36040.1
Lus10023142	Cytochrome P450, family 79, subfamily B, polypeptide 2	4.51	4.67E-09	Oxidation/reduction	AT4G39950.1
Lus10033041	Leucine-rich repeat protein kinase family protein	4.40	8.88E-16	Molecular function	AT2G31880.1
Lus10022547	Phosphate transporter 1;5	4.38	1.09E-13	Transport	AT2G32830.1
Lus10016635	Phosphate transporter 1;7	4.37	3.82E-13	Transport	AT3G54700.1
Lus10015933	Unknown	4.36	6.88E-11	Metabolism	AT5G61820.1
Lus10040328	α/β-Hydrolases superfamily protein	4.34	1.01E-11	Cellular process	AT2G39420.1
Lus10034312	NIM1-interacting 2	4.34	2.22E-12	Molecular function	AT3G25882.1

peroxidases specifically in *lbf1* outer tissues would suggest a similar concerted action of these two enzymes. However, no evidence exists as yet, based on the comparative microarray data set of *lbf1* and wild-type flax, of the involvement of a scaffolding protein homologous to the *Arabidopsis* endodermis Casparian strip domain proteins in lignification in flax stem tissues. Further evidence for a potential role of peroxidases and NADPH oxidases in *lbf1* lignification was provided by the observation that the other three flax NADPH oxidase genes are all orthologs of the *Arabidopsis RBOH-A* and *RBOH-C* genes involved in the generation of apoplastic H_2O_2 during the defensive oxidative burst (Schweizer, 2008). Increased accumulation of RBOH-A and -C transcripts could therefore also contribute to apoplastic H_2O_2 content and stimulate lignification. Somewhat intriguingly, we also observed a significant accumulation of transcripts corresponding to another flax RBOH-F ortholog in both outer- and inner-stem tissues

Table 5. List of the 20 Least Abundant Transcripts in *lbf1* Mutants versus the Wild Type

Reference	Name	Delta	P Value	GO Annotation	*Arabidopsis* Correspondence
20 Least Abundant Transcripts in Outer Tissues					
Lus10011872	Tetratricopeptide repeat (TPR)-like superfamily protein	−5.08	1.58E-10	Biological process	AT5G48850.1
Lus10022806	Ethylene-dependent gravitropism-deficient and yellow-green-like 2	−5.08	1.47E-09	Cellular process	AT5G05740.2
Lus10012353	Unknown	−5.02	1.69E-09	Unknown	
Lus10041133	Purine permease 3	−4.90	2.75E-10	Transport	AT1G28220.1
Lus10009917	Expansin A8	−4.48	1.29E-07	Cell wall	AT2G40610.1
Lus10006996	Unknown	−4.33	4.43E-11	Unknown	AT2G27830.1
Lus10006759	Gibberellin 2-oxidase 8	−4.30	1.97E-12	Oxidation/reduction	AT4G21200.1
Lus10038566	Dynein light chain type 1 family protein	−4.29	2.04E-11	Biological process	AT4G27360.1
Lus10000385	Unknown	−4.16	4.72E-10	Unknown	AT2G27830.1
Lus10003913	Urophorphyrin methylase 1	−4.12	3.55E-09	Molecular function	AT5G40850.1
Lus10023289	Unknown	−3.90	6.35E-11	Unknown	AT1G30260.1
Lus10025278	Aluminum sensitive 3	−3.86	2.41E-11	Transport	AT2G37330.1
Lus10009069	Aluminum sensitive 3	−3.84	2.41E-13	Transport	AT2G37330.1
Lus10038821	Nodulin MtN3 family protein	−3.78	7.64E-09	Biological process	AT5G53190.1
Lus10010529	Unknown	−3.72	5.15E-11	Unknown	AT5G19340.1
Lus10008485	Protein of unknown function (DUF567)	−3.72	6.79E-12	Unknown	AT3G14260.1
Lus10028947	Xyloglucan endotransglucosylase/hydrolase 15	−3.65	1.03E-11	Cell wall	AT4G14130.1
Lus10038517	Unknown	−3.64	5.59E-10	Unknown	AT1G30260.1
Lus10037476	Urophorphyrin methylase 1	−3.64	8.61E-08	Molecular function	AT5G40850.1
Lus10023377	Sec14p-like phosphatidylinositol transfer family protein	−3.56	4.58E-11	Transport	AT1G30690.1
20 Least Abundant Transcripts in Inner Tissues					
Lus10038721	Carotenoid cleavage dioxygenase 8	−5.34	3.59E-12	Metabolism	AT4G32810.1
Lus10002073	Protein of unknown function, DUF584	−4.08	3.00E-08	Unknown	AT1G61930.1
Lus10023311	Gibberellin 2-oxidase	−3.98	7.66E-09	Oxidation/reduction	AT1G30040.1
Lus10017253	RING/U-box superfamily protein	−3.77	3.03E-13	Molecular function	AT5G42200.1
Lus10034238	NAD-dependent glycerol-3-phosphate dehydrogenase family protein	−3.77	6.99E-07	Metabolism	AT2G40690.1
Lus10025771	Peptide-N4-(*N*-acetyl-β-glucosaminyl)asparagine amidase A protein	−3.74	1.31E-10	Cell wall	AT3G14920.1
Lus10034206	Major facilitator superfamily protein	−3.72	9.36E-11	Transport	AT2G40460.1
Lus10005617	RING/U-box superfamily protein	−3.65	2.94E-12	Molecular function	AT5G42200.1
Lus10008304	Pathogenesis-related thaumatin superfamily protein	−3.60	3.75E-09	Defense	AT5G40020.1
Lus10035519	HXXXD-type acyl-transferase family protein	−3.56	6.39E-08	Molecular function	AT5G01210.1
Lus10043404	Unknown	−3.55	3.63E-10	Unknown	AT3G11600.1
Lus10017817	Major facilitator superfamily protein	−3.50	6.18E-11	Transport	AT1G68570.1
Lus10013489	Late embryogenesis abundant protein (LEA) family protein	−3.50	2.54E-10	Biological process	AT1G52690.1
Lus10025278	Aluminum sensitive 3	−3.43	1.45E-09	Transport	AT2G37330.1
Lus10030457	Glucose-methanol-choline (GMC) oxidoreductase family protein	−3.43	1.47E-08	Metabolism	AT1G14185.1
Lus10013401	Branched-chain α-keto acid decarboxylase E1 β-subunit	−3.42	7.33E-09	Oxidation/reduction	AT1G55510.1
Lus10029063	Major facilitator superfamily protein	−3.37	5.23E-10	Transport	AT2G40460.1
Lus10023189	Laccase 11	−3.33	3.46E-08	Oxidation/reduction	AT5G03260.1
Lus10037164	Expansin A1	−3.32	6.67E-08	Cell wall	AT1G69530.1
Lus10041338	Serine carboxypeptidase-like 48	−3.31	9.73E-08	Cellular process	AT3G45010.1

of the *lbf1* mutant despite the fact that increased lignification was only observed in outer tissues. Further work is necessary to understand the significance of increased NADPH oxidase accumulation in *lbf1* inner stem tissues.

Although our results suggest that ectopic lignification in *lbf1* bast fibers is related to modified polymerization, the increased transcript abundance of the monolignol biosynthesis genes *COMT*, *CCR*, and *CAD* suggests that the supply of monolignols to these fibers is also increased. The Lu-*CCR* gene is phylogenetically close to At-*CCR* involved in developmental lignification in *Arabidopsis* and *CCR* downregulation drastically reduces lignin biosynthesis (Lacombe et al., 1997; Dauwe et al., 2007; Leplé et al., 2007). By contrast, the Lu-*COMT* and Lu-*CAD* genes are not part of the bona fide lignin group responsible for developmental lignification (Supplemental Figure 3 and Supplemental Data Sets 4 to 6) but rather belong to gene groups involved in the response to stress or pathogen attack (Barakat et al., 2010, 2011). This observation is interesting since among the 20 most abundant transcripts in the

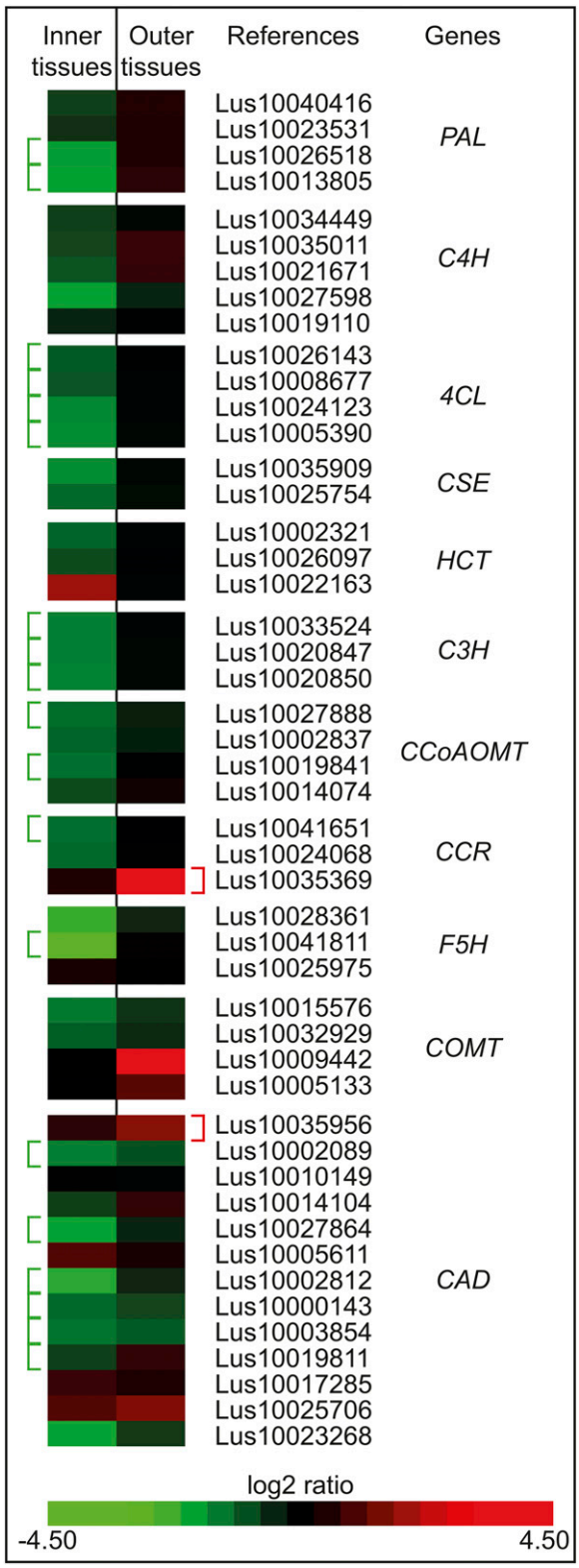


Figure 10. Heat Map Representing Comparative Accumulation of Monolignol Biosynthetic Transcripts in Outer- and Inner-Stem Tissues of *lbf1*.

Transcripts annotated by a bracket are significantly overaccumulated (red) or underaccumulated (green) (P value < 0.05, Bonferroni method) in the mutant compared with the wild type.

inner- and outer-stem tissues, several are potentially involved in defense, raising the possibility that increased lignification in the *lbf1* mutant could be caused by a mutation affecting the defense signaling and/or response pathway.

The idea that bast fiber ectopic lignification in *lbf1* outer tissues could be associated with increased monolignol production and/or availability is also supported by the decreased accumulation of transcripts corresponding to a *UGT* (*UDP-glucosyltransferase*) gene. This gene encodes a putative ortholog of the *Arabidopsis* UGT72E1 protein that glucosylates monolignols (Lanot et al., 2006, 2008). Monolignol glucosylation is believed to play a role in detoxifying monolignols and is also involved in addressing monolignols to the vacuole for storage (Miao and Liu, 2010; Tsuyama et al., 2013). This process therefore represents a potential control point in lignification and the observed reduction in *UGT* transcript abundance in outer tissues of the flax *lbf1* mutant could be expected to increase monolignol availability for subsequent lignification. Alternatively, increased incorporation of monolignols into the lignin polymer could reduce the necessity for detoxification and/or vacuolar storage leading to *UGT* downregulation. Interestingly, decreased lignification in *Arabidopsis* triple laccase mutants is associated with increased expression of genes encoding the 72E2 and 72E3 UGT proteins (Zhao et al., 2013), suggesting the existence of a relationship between modified lignification and regulation of monolignol supply via glycosylation.

In contrast to outer-stem tissues, levels of lignin and non-glycosylated oligolignols remained unchanged in *lbf1* inner-stem tissues compared with wild-type plants. Nevertheless, quantities of the hexosylated monolignols coniferin and syringin as well as of several hexosylated dilignols were significantly higher in mutant xylem tissues. Transcriptomics data indicated that transcripts of four peroxidase genes were more abundant in *lbf1* inner tissues. Whereas none of these peroxidase genes belong to the same clades as the lignin-related peroxidase genes, they exhibit increased transcript abundance in *lbf1* outer tissues, suggesting that they are not involved in lignification. In the absence of a significant increase in the capacity to oxidize monolignols for polymerization, it is possible that monolignols cannot be incorporated into the lignin polymer and must be detoxified by other mechanisms such as glycosylation. The observed decrease in the abundance of transcripts corresponding to a laccase gene (*LAC11*) implicated in lignification also suggests a reduced monolignol oxidizing capacity (Zhao et al., 2013). Although no significant change in *UGT* expression was observed, transcript abundance was reduced for 19 genes in the lignin biosynthetic pathway. This massive decrease in transcripts corresponding to 7 of the 11 lignin gene families could be interpreted as an attempt to regulate monolignol production and cellular toxicity. Further work is necessary to clarify this point.

Altogether our observations indicate that the *lbf1* mutation results in contrasted tissue-specific effects on transcript abundance of a range of lignin-related genes (i.e., genes encoding enzymes involved in monolignol biosynthesis, a UGT, peroxidases,

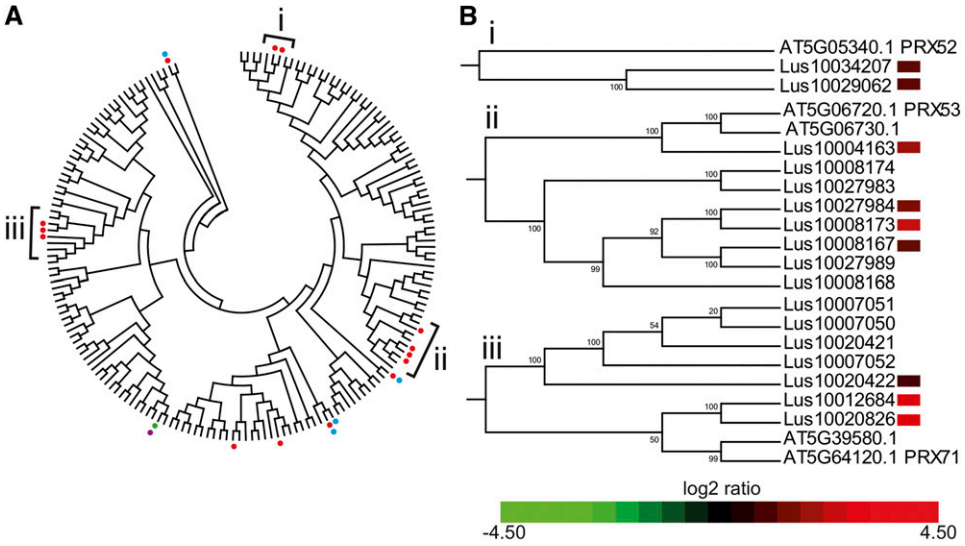


Figure 11. Phylogenetic and Expression Analyses of Peroxidases in *lbf1* Mutants.

(A) Phylogenetic unrooted tree of *Arabidopsis* and flax peroxidase proteins. Branches marked by a dot correspond to individual peroxidase transcripts significantly more abundant (P value < 0.05, Bonferroni method) in inner (blue dot) or outer tissues (red dot) or less abundant in inner (pink dot) or outer tissues (green) of *lbf1* mutant compared with the wild type.

(B) Heat map of transcript accumulation corresponding to flax peroxidases in clades A, B, and C containing known *Arabidopsis* lignin-related peroxidases.

NADPH-oxidases, and a superoxide dismutase), oligolignol content, and lignin quantity in flax stems. These observations not only suggest the existence of complex tissue-specific regulation mechanisms, but also underline the importance of taking into account organ and tissue specificity when interpreting expression data.

(2) Other Cell Wall Polymers

The thick secondary walls of mature flax bast fibers largely consist of cellulose and pectic galactan as the main incrusting NCP together with AGPs (His et al., 2001; Morvan et al., 2003). Both chemical analyses and immunolabeling suggested that *lbf1* bast fibers contain less cellulose and significantly higher amounts of NCPs and AGPs compared with the wild type and provide strong evidence that increased lignification is accompanied by changes in polysaccharide architecture as previously observed in different *Arabidopsis* lignin mutants (Van Acker et al., 2013).

Although at first view such changes could be due to the higher lignin content in the mutant, another intriguing possibility is that crosstalk between cell wall polymers during biosynthesis may favor the formation of a cell wall matrix more favorable to lignification. Higher hemicellulose deposition concomitant with lower cellulose content, for example, would lead to a looser cell wall structure and/or less crystalline cellulose, both of which could facilitate monolignol transport and subsequent lignin polymerization within a xylan hemicellulose matrix. In agreement with this idea is the observation that disruption of cellulose biosynthesis, either by chemical inhibition with isoxaben, or by mutations in the *CESA3* gene leads to ectopic lignification in *Arabidopsis* wild type and *eli1* mutants (Caño-Delgado et al., 2003). Similarly, ectopic lignification in the *elp1 Arabidopsis* mutant is due to a mutation in a chitinase-like (*CTL*) gene (Zhong et al., 2002). While only a small number of transcripts corresponding to genes directly involved in cell wall biosynthesis were differentially accumulated between flax *lbf1* mutant and wild-type plants, these included several glucosyltransferases, possibly accounting for the observed changes in cell wall matrix polysaccharides. In addition, transcripts for a COBRA4-like extracellular glycosyl-phosphatidyl inositol-anchored protein were less abundant in the outer tissues of the flax mutant. This protein has been proposed to modulate cellulose assembly through interaction with cellulose microfibrils (Liu et al., 2013), and reduced expression of this gene is associated with lower cellulose content and higher lignification in mature stem tissues of the maize (*Zea mays*) *bk2* mutant (Sindhu et al., 2007).

In addition to modifications in cell wall chemical composition, *lbf1* fiber cell walls were also significantly thinner than wild-type ones and could be related to the differential transcript abundance of putative flax orthologs corresponding to *Arabidopsis* expansin (AtEXP8) and xyloglucan endotransglycosylase/hydrolase (AtXTR7) genes (Cosgrove, 2005; Sasidharan et al., 2008).

Further analyses, not only of the spatial distribution of different cell wall components, but also fiber morphology and cell wall thickness at different stages of fiber development, are needed to obtain better insight into the relationship between polysaccharides and lignin deposition in the growing cell wall.

Conclusions and Perspectives

In conclusion, we generated a core collection of flax *lbf* mutants that represent an interesting biological resource for investigating the regulatory mechanisms used by fiber plants to produce poorly lignified, thick, secondary cell walls. As a proof of concept, we

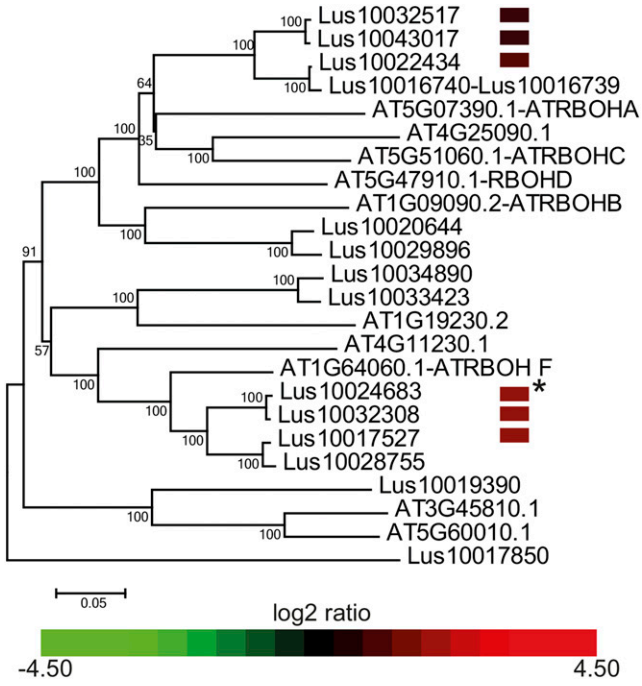


Figure 12. Phylogenetic and Expression Analyses of RBOH Proteins in *lbf1* Mutants.

Phylogenetic unrooted tree of *Arabidopsis* and flax RBOH proteins. Heat map expression data of the six flax genes overexpressed in *lbf1* mutants compared with the wild type are given in front of their corresponding Phytozome references. Transcripts of all genes are specifically more abundant in *lbf1* outer tissues except where transcripts are more abundant in outer and inner *lbf1* tissues (asterisk).

undertook a detailed characterization of the *lbf1* mutant. Our results suggest that the main regulatory point occurs at the oxidative polymerization step and that the typical low lignification observed in wild-type bast fibers is related to the absence of different actors necessary for monolignol oxidation. Recent analyses of peroxidase gene promoters have suggested that these genes are regulated by a number of different transcription factors (NAC, MYB, AP2, and class I and III HD-ZIP) previously associated with vascular tissue formation and/or secondary cell wall formation (Herrero et al., 2014), and it is possible that increased bast fiber lignification is associated with a mutation in such a gene(s). Alternatively, peturbations in the biosynthesis of other cell wall polymers affecting cell wall integrity and/or activation of defense signaling could also be responsible for the ectopic lignification in the *lbf1* mutant. The flax genome at ~390 Mb is relatively small, and recent advances in NGS technology should allow the development of a "mapping by sequencing" approach (Wang et al., 2012; Allen et al., 2013; Wijnen and Keurentjes, 2014) in this species and the subsequent identification of the gene(s) associated with increased lignification and other interesting phenotypes. We observed that for *lbf1* the lignified phenotype is heritable over several generations, and we are currently generating F2 backcrossed material for such an approach. Heritability of the lignified phenotype has also been confirmed for 6 out of 10 other *lbf* families that we are multiplying.

The systematic exploitation of the flax *lbf* collection will allow us to improve our understanding of the functional relationship between lignin and other cell wall polymers, thereby leading to a better understanding of cell wall dynamics. Finally, our collection can be used to gain a better knowledge of how increased lignification modifies different fiber mechanical and physical properties. For example, preliminary saccharification analyses using a commercial cellulase cocktail (Novozymes) indicate that 30% less glucose is released from flax *lbf* mutant outer-stem tissues when compared with wild-type tissues.

METHODS

Plant Material

Flax (*Linum usitatissimum*) EMS mutants used in this study come from the PT-Flax Collection (Chantreau et al., 2013). M2 to M5 plants were grown in greenhouses or outside at the University of Lille, France. For chemical, metabolomic, and transcriptomic analyses, stem outer tissues were separated from inner tissues by peeling as previously described (Day et al., 2005). For transcriptomics, tissues were harvested before flowering and were immediately frozen in liquid nitrogen. For chemistry and metabolomics, tissues were harvested at grain maturity and lyophilized before analyses.

Microscopy

UV-based lignin screening was made on thick freehand cross sections from the median part of M2 mutant stems. Outer-tissue fluorescence was determined using an inverted microscope (Nikon Eclipse TS100) coupled with an UV irradiation system (λ excitation, 365 nm; λ emission, 420 nm). Phloroglucinol-HCl staining was made on semi-thin freehand cross sections and examined with a Nikon Eclipse TS100 and/or a LEICA DM2000 microscope. Photographs were taken with a Nikon D5000 camera.

Immunohistochemical Analyses

Ethanol-fixed specimens of the median region of flax stems were dehydrated using an ethanol series and acetone prior to epoxy resin impregnation and embedding (epoxy embedding medium, EEM hardener DDSA, and EEM hardener NMA; Fluka). Immunolabeling was done on semithin (0.5 μm) and ultrathin (200 nm) transverse sections of resin-embedded block prior to observations by fluorescence microscopy (Nikon Eclipse TE300) and transmission electron microscopy at 200 kV (JEM2100F; JEOL), respectively.

Immunogold labeling of 8-5′ Linked Lignin Structure for Transmission Electron Microscopy

Transverse ultrathin sections were cut from the Epoxy resin-embedded block and mounted on nickel grids (200 mesh). Sections were floated on a drop of blocking buffer (1% BSA, and 0.1% NaN_3 in TBS) for 30 min at room temperature and then floated on a drop of KM1 ascites fluid diluted 1:100 in blocking buffer for 2 d at 4°C. Following washing thrice for 15 min on drops of blocking buffer, sections were incubated with immunogold conjugate EM goat anti-mouse IgG, 10 nm (EM.GAM10; BB International), diluted 1:100 in blocking buffer for 4 h at room temperature. Finally, the sections were washed six times for 15 min on drops of blocking buffer and then washed with ultra pure water. Sections were observed under a JEM2100F transmission electron microscope (JEOL) without poststaining.

Immunolabeling for Fluorescence Microscopy

Sections were mounted on silanized slides and incubated with 3% protein milk in PBS (0.1 M phosphate containing 0.9% NaCl, pH 7.6) for 30 min at room temperature to avoid nonspecific binding of antibody. Sections were then washed with PBS and incubated with LM10, LM11, or LM21 diluted 1:20 in blocking buffer (PBS containing 1% BSA and 0.01% sodium azide) or LM5, LM2, or JIM14 diluted 1:10 in blocking buffer for 3 h at room temperature and 1 d at 4°C. After washing twice for 5 min with PBS, the sections were incubated at room temperature for 4 h with Alexa Fluor 488 goat anti-rat IgG (H+L) (Life Technologies) diluted 1:100 in TBS. They were again washed three times for 5 min with PBS and washed with ultrapure water. Sections were mounted in Eukit (Sigma-Aldrich).

Chemical Analyses and NMR

Cell Wall Residue Preparation

All chemical analyses were performed on extractive-free cell wall residue (CWR) obtained from manually separated outer- and inner-stem tissues. CWR was obtained by extracting tissues (7-fold) with 80% ethanol (6 mL/100 mg CWR) prior to grinding.

Lignin, Sugar, and Protein Determination

Acetyl bromide lignin was determined by measuring absorbance at 280 nm as previously described (Iiyama and Wallis, 1990). Thioacidolysis and subsequent gas chromatography-mass spectrometry analyses of β-O-4 ether-linked lignin monomers (analyzed as their trimethylsilylated derivatives) were performed as previously described using a Hewlett-Packard HP6890 Series gas chromatograph-flame ionization detector and a Thermo Focus gas chromatograph coupled with a Polaris Q gas chromatograph-mass spectrometer (Day et al., 2005).

Sugar analysis was performed by high-performance anion-exchange chromatography (Dionex DX 500; Thermo Scientific) after a two-step sulfuric acid hydrolysis of CWR using 2 deoxyribose as internal standard (Belmokhtar et al., 2013).

Protein content was determined in triplicate by measuring the total N contents (N*6.25) of 3 mg of ball-milled samples using an elemental analyzer (NA 1500; Carlo Erba) coupled to a mass spectrometer (Euro EA elemental analyzer).

NMR Analysis

Approximately 200 mg of CWR was ball-milled in a 25-mL jar with 20 × 20-mm ZrO_2 ball bearings using a Retsch MM2000 mixer mill, for 1 h and 50 min using 20-min milling intervals with 10-min breaks. DMSO (1.8 mL) and *N*-methylimidazole (0.9 mL) were added to 100 mg of each ball-milled cell wall sample for cell wall dissolution (Hedenström et al., 2009). After acetylation and precipitation into water, samples were centrifuged in a Beckman JLA-10.500 rotor at 18,600*g* for 10 min. The pellets were washed twice with water and then centrifuged as previously. Around 80 mg of acetylated cell wall was dissolved in 0.6 mL of $CDCl_3$ in a 5-mm NMR tube prior to NMR acquisition. NMR spectra were acquired on a Bruker Biospin Avance III 600 MHz spectrometer, using a 5-mm TCI cryoprobe equipped with cold preamplifiers for ^1H, ^{13}C, and ^{15}N. Adiabatic HSQC (hsqcedetgpsisp2.2) spectra widths were 5102 and 24,147 Hz for the ^1H- and ^{13}C-dimensions, respectively. The number of collected complex points was 1024 for the ^1H-dimension using a relaxation delay of 1 s. The number of scans was 64, and 386 time increments were always recorded in the ^{13}C-dimension. The spectra were processed using Topspin 3.1 Bruker Biospin. All spectra were manually phase corrected and calibrated with $CDCl_3$peak (δ_C, 77.2; δ_H, 7.26 ppm) used as internal reference. Signals were assigned by comparison with 2D NMR spectra reported in the literature (del Río et al., 2011; Mansfield et al., 2012; Ralph et al., 2012) and recorded on acetylated standards dehydrogenation polymers (Cathala et al., 1998), galactan (lupin) P-GALLU; 1,5-α-L-arabinan (sugar beet) P-LARB (Megazyme).

Oligolignol Profiling

Sample Preparation

Phenolic profiling was independently performed on inner- and outer-stem tissues of three wild-type plants and six *lbf1* mutants. For wild-type plants, three inner- and three outer-stem tissues were analyzed, and for *lbf1* mutants, five inner- and six outer-stem tissues were analyzed. Ethanolic extracts from CWR preparations were mixed and filtered through a paper filter then evaporated at 40°C to dryness under reduced pressure. Dry extracts were resuspended with ~1.5 mL of a mix of diethyl ether and Milli-Q Water and then transferred into 2.5-mL vials. Vials were kept open and diethyl ether evaporated at room temperature under a stream of ambient air. Vials were stored at 4°C and then evaporated with Centrivap LABONCO at 50°C prior to analysis.

Oligolignol Profiling by HPLC-High-Resolution Mass Spectrometry

Phenolic profiling was performed using 10 μL of the water phase. Extracts were analyzed with a DionexUltiMate 3000 LC module equipped with a LPG-3400 pump, UV-Vis detector (model VWD-3400), and an autosampler (model WPS-3000 SL) and further hyphenated to an LTQ Orbitrap XL hybrid FTMS mass spectrometer (MS) (Thermo Electron) consisting of a linear ion trap MS connected with a Fourier transform Orbitrap MS. The separation was performed on a reversed phase Sunfire C18 column (150 mm × 3 mm, 3.5 μm; Waters) with aqueous 0.1% acetic acid and acetonitrile/water (99/1, v/v, acidified with 0.1% acetic acid) as solvents A and B. A gradient of 0 min 5% B, 40 min 45% B, and 45 min 100% B was applied using a flow rate of 300 μL/min and a column temperature of 40°C. The autosampler temperature was 10°C. Analytes were negatively ionized with an electrospray source using the following parameter values: source voltage 5.00 kV, source current 100.00 μA, capillary temperature 300°C, sheath gas 20 (arb), aux gas 10 (arb), and sweep gas 2 (arb). Full Fourier transform-mass spectrometry spectra between 120 and 1400 *m/z* were recorded at a resolution of 100,000. In parallel, three data-dependent MSn spectra were recorded on the ion trap MS using the preliminary low-resolution data obtained during the first 0.1 s of the previous full Fourier transform-mass spectrometry scan: a MS2 scan of the most abundant *m/z* ion of the full Fourier transform-mass spectrometry scan, followed by two MS3 scans of the most abundant first product ions. MSn scans were obtained with 35% collision energy.

Elucidation of MS^2 Spectra

Elucidation of the MS^2 spectra and the sequencing terminology of the first product ions was based on the lignin oligomer/(neo)lignan sequencing approach mentioned by Morreel et al. (2010a) and on the fragmentation rules of the different linkage types described by Morreel et al. (2010b). Briefly, the three types of linkages, i.e., 8-O-4 (β-aryl ether), 8-5 phenylcoumaran), and 8-8 (resinol), either loose, small, neutral molecules that are indicative of the type of linkage (referred to as pathway I, I) or are cleaved, hence yielding information on the units that are connected by the linkage (referred to as pathway II, II). In the case of a β-aryl ether, pathway II cleavage leads to first product ions corresponding with the phenolic 8-end (A- ion) and aliphatic 4-end (B- ion) moieties. The structures of the described compounds are given in Supplemental Figure 2.

Agilent Microarray Transcriptomics

Total RNA was extracted from separated inner- and outer-stem tissues of six individual *lbf1* mutants and three wild-type plants using the TriReagent method (Molecular Research Center). RNA integrity and concentration were evaluated with RNA StdSens Chips using the Experion automated capillary electrophoresis system (Bio-Rad). RNA processing and hybridization were performed following the manufacturer's instruction for One-Color Microarray-Based Gene Expression Analysis (Agilent Technologies). Samples were hybridized to the Agilent-045382 UGSF flax 45K v1.0 array based upon flax genome coding sequence (Wang et al., 2012) available at Phytozome (http://phytozome.org). The array contains 45,220 60-mer in situ oligonucleotides per block. All nine samples were analyzed independently. Following hybridization, washing was performed following the manufacturer's instruction, and slides were immediately scanned at 5-mm pixel^{-1} resolution using an Axon GenePix 4000B scanner (Molecular Devices) piloted by GenePix Pro 6.0 software (Axon). Grid alignment and expression data analyses were made with the same software. After background noise elimination, median values of overall hybridization were normalized by robust local regression (Yang et al., 2002). Artifact spots were manually eliminated. Differential analysis was performed with the method varmixt (Delmar et al., 2005), available in the package anapuce of the software R. A double-sided, unpaired *t* test was computed for each gene between the two conditions. Variance of the difference in gene expression (transcript abundance) was split between subgroups of genes with homogeneous variance (Delmar et al., 2005). The raw P values were adjusted by the Bonferroni method, which controls the family-wise error rate (Ge et al., 2003). A gene is declared differentially expressed if the Bonferroni-corrected P value is <0.05.

Bioinformatics

Phylogenetic trees were made using a neighbor-joining method implemented in MEGA5. Bootstrap consensus tree were inferred from 1000 replicates. Branches corresponding to partitions reproducing <50% bootstrap replicates are collapsed. The evolutionary distances were computed using the *p*-distance method.

Accession Numbers

All data are available through the Gene Expression Omnibus repository at NCBI (Barrett et al., 2007) under accession numbers GSE61311 and GPL19181.

Supplemental Data

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

Supplemental Figure 1. 2D NMR Spectra of Lignin from Flax Wild-Type and *lbf1* Inner Tissues.

Supplemental Figure 2. Structures of Oligolignols Previously Unidentified in Flax.

Supplemental Figure 3. Immunolocalization of Xylem Cell Wall NCPs.

Supplemental Figure 4. Phylogenetic Trees of Lignin Genes Overexpressed in *lbf1* Outer Tissues.

Supplemental Table 1. Visual Phenotyping Classes for Flax *lbf* Mutants, as Previously Described (Chantreau et al., 2013).

Supplemental Data Set 1. Over- and Underaccumulated Transcripts in Outer Tissues.

Supplemental Data Set 2. Alignments Used to Generate the Peroxidase Phylogenies Presented in Figure 11.

Supplemental Data Set 3. Alignments Used to Generate the RBOH Phylogenies Presented in Figure 12.

Supplemental Data Set 4. Alignments Used to Generate the CCR Phylogenies Presented in Supplemental Figure 4A.

Supplemental Data Set 5. Alignments Used to Generate the COMT Phylogenies Presented in Supplemental Figure 4B.

Supplemental Data Set 6. Alignments Used to Generate the CAD Phylogenies Presented in Supplemental Figure 4C.

ACKNOWLEDGMENTS

M. Chantreau gratefully acknowledges the University Lille1 and the Nord-Pas de Calais Region for a PhD fellowship. S.K. gratefully acknowledges the financial support of the Kyoto University Foundation. This work was carried in the context of and financed by the French national project PT-Flax (ANR-09-GENM-020). We thank Marie-Laure Martin-Magniette (URGV France) for her advice on transcriptomics data analyses. Authors acknowledge the technical support of the PICT IBiSA biological imaging center (transmission electron microscopy) and the PLANET analytical platform (NMR) at the University of Reims Champagne-Ardenne.

AUTHOR CONTRIBUTIONS

S.H. and B.C. conceived the project and decided on the scientific strategy. A.P. performed cell wall chemical analyses, and D.C. realized the NMR analyses. R.D. performed the oligolignol analyses and interpreted all MS data together with K.M. S.K. performed lignin and cell wall light microscopy and transmission electron microscopy immunolocalization. M. Chantreau, S.G., B.C., S.H., and G.N. collected plant material. M. Chantreau produced plant material, screened the mutant population, and undertook all bioinformatic analyses and transcriptomics. S.A. and M. Chabi assisted with transcriptomics and data analyses. G.N. validated microarray data. W.B., A.Y., and F.M. provided important scientific criticism and input during the writing of this article. This article was written by M. Chantreau and S.H. with important contributions from B.C. and R.D. All authors read, reviewed, and approved the final article.

Received July 25, 2014; revised September 12, 2014; accepted October 19, 2014; published November 7, 2014.

REFERENCES

Allen, R.S., Nakasugi, K., Doran, R.L., Millar, A.A., and Waterhouse, P.M. (2013). Facile mutant identification via a single parental backcross method and application of whole genome sequencing based mapping pipelines. Front. Plant Sci. **4:** 362.

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.

Barakat, A., Bagniewska-Zadworna, A., Frost, C.J., and Carlson, J.E. (2010). Phylogeny and expression profiling of CAD and CAD-like genes in hybrid Populus (*P. deltoides* x *P. nigra*): evidence from herbivore damage for subfunctionalization and functional divergence. BMC Plant Biol. **10:** 100.

Barakat, A., Choi, A., Yassin, N.B.M., Park, J.S., Sun, Z., and Carlson, J.E. (2011). Comparative genomics and evolutionary analyses of the *O*-methyltransferase gene family in *Populus.* Gene **479:** 37–46.

Barrett, T., Troup, D.B., Wilhite, S.E., Ledoux, P., Rudnev, D., Evangelista, C., Kim, I.F., Soboleva, A., Tomashevsky, M., and

Edgar, R. (2007). NCBI GEO: mining tens of millions of expression profiles—database and tools update. Nucleic Acids Res. **35:** D760–D765.

Baucher, M., Monties, B., Montagu, M.V., and Boerjan, W. (1998). Biosynthesis and genetic engineering of lignin. Crit. Rev. Plant Sci. **17:** 125–197.

Belmokhtar, N., Habrant, A., Ferreira, N.L., and Chabbert, B. (2013). Changes in phenolics distribution after chemical pretreatment and enzymatic conversion of *Miscanthus x giganteus* internode. BioEnergy Res. **6:** 506–518.

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

Bonawitz, N.D., and Chapple, C. (2010). The genetics of lignin biosynthesis: connecting genotype to phenotype. Annu. Rev. Genet. **44:** 337–363.

Caño-Delgado, A., Penfield, S., Smith, C., Catley, M., and Bevan, M. (2003). Reduced cellulose synthesis invokes lignification and defense responses in Arabidopsis thaliana. Plant J. **34:** 351–362.

Cathala, B., Saake, B., Faix, O., and Monties, B. (1998). Evaluation of the reproducibility of the synthesis of dehydrogenation polymer models of lignin. Polym. Degrad. Stabil. **59:** 65–69.

Chantreau, 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.

Cosgrove, D.J. (2005). Growth of the plant cell wall. Nat. Rev. Mol. Cell Biol. **6:** 850–861.

Crônier, D., Monties, B., and Chabbert, B. (2005). Structure and chemical composition of bast fibers isolated from developing hemp stem. J. Agric. Food Chem. **53:** 8279–8289.

Dauwe, R., et al. (2007). Molecular phenotyping of lignin-modified tobacco reveals associated changes in cell-wall metabolism, primary metabolism, stress metabolism and photorespiration. Plant J. **52:** 263–285.

Davis, E.A., Derouet, C., Herve Du Penhoat, C., and Morvan, C. (1990). Isolation and an NMR study of pectins from flax (*Linum usitatissimum* L.). Carbohydr. Res. **197:** 205–215.

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.

Delmar, P., Robin, S., and Daudin, J.J. (2005). VarMixt: efficient variance modelling for the differential analysis of replicated gene expression data. Bioinformatics **21:** 502–508.

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.

Fenart, S., et al. (2010). Development and validation of a flax (*Linum usitatissimum* L.) gene expression oligo microarray. BMC Genomics **11:** 592.

Fu, C., Mielenz, J.R., Xiao, X., Ge, Y., Hamilton, C.Y., Rodriguez, M., Jr., Chen, F., Foston, M., Ragauskas, A., Bouton, J., Dixon, R.A., and Wang, Z.Y. (2011). Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. Proc. Natl. Acad. Sci. USA **108:** 3803–3808.

Ge, Y., Dudoit, S., and Speed, T.P. (2003). Resampling-based multiple testing for microarray data analysis. Test **12:** 1–77.

Girault, R., Bert, F., Rihouey, C., Jauneau, A., Morvan, C., and Jarvis, M. (1997). Galactans and cellulose in flax fibres: putative contributions to the tensile strength. Int. J. Biol. Macromol. **21:** 179–188.

Gorshkova, T., and Morvan, C. (2006). Secondary cell-wall assembly in flax phloem fibres: role of galactans. Planta **223:** 149–158.

Guerriero, G., Sergeant, K., and Hausman, J.-F. (2013). Integrated -omics: a powerful approach to understanding the heterogeneous lignification of fibre crops. Int. J. Mol. Sci. **14:** 10958–10978.

Hedenström, M., Wiklund-Lindström, S., Öman, T., Lu, F., Gerber, L., Schatz, P., Sundberg, B., and Ralph, J. (2009). Identification of lignin and polysaccharide modifications in Populus wood by chemometric analysis of 2D NMR spectra from dissolved cell walls. Mol. Plant **2:** 933–942.

Herrero, J., Esteban-Carrasco, A., and Zapata, J.M. (2013). Looking for *Arabidopsis thaliana* peroxidases involved in lignin biosynthesis. Plant Physiol. Biochem. **67:** 77–86.

Herrero, J., Esteban Carrasco, A., and Zapata, J.M. (2014). *Arabidopsis thaliana* peroxidases involved in lignin biosynthesis: *in silico* promoter analysis and hormonal regulation. Plant Physiol. Biochem. **80:** 192–202.

His, I., Andème-Onzighi, C., Morvan, C., and Driouich, A. (2001). Microscopic studies on mature flax fibers embedded in LR white: immunogold localization of cell wall matrix polysaccharides. J. Histochem. Cytochem. **49:** 1525–1536.

Huis, R., Morreel, K., Fliniaux, O., Lucau-Danila, A., Fénart, S., Grec, S., Neutelings, G., Chabbert, B., Mesnard, F., Boerjan, W., and Hawkins, S. (2012). Natural hypolignification is associated with extensive oligolignol accumulation in flax stems. Plant Physiol. **158:** 1893–1915.

Hu, W.J., Harding, S.A., Lung, J., Popko, J.L., Ralph, J., Stokke, D.D., Tsai, C.J., and Chiang, V.L. (1999). Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees. Nat. Biotechnol. **17:** 808–812.

Iiyama, K., and Wallis, A.F. (1990). Determination of lignin in herbaceous plants by an improved acetyl bromide procedure. J. Sci. Food Agric. **51:** 145–161.

Jones, L., Seymour, G.B., and Knox, J.P. (1997). Localization of pectic galactan in tomato cell walls using a monoclonal antibody specific to (1 [->] 4)-[beta]-D-galactan. Plant Physiol. **113:** 1405–1412.

Karlsson, M., Melzer, M., Prokhorenko, I., Johansson, T., and Wingsle, G. (2005). Hydrogen peroxide and expression of hipl-superoxide dismutase are associated with the development of secondary cell walls in *Zinnia elegans.* J. Exp. Bot. **56:** 2085–2093.

Karpinska, B., Karlsson, M., Schinkel, H., Streller, S., Süss, K.-H., Melzer, M., and Wingsle, G. (2001). A novel superoxide dismutase with a high isoelectric point in higher plants. expression, regulation, and protein localization. Plant Physiol. **126:** 1668–1677.

Kiyoto, S., Yoshinaga, A., Tanaka, N., Wada, M., Kamitakahara, H., and Takabe, K. (2013). Immunolocalization of 8-5′ and 8-8′ linked structures of lignin in cell walls of *Chamaecyparis obtusa* using monoclonal antibodies. Planta **237:** 705–715.

Knox, J.P., Linstead, P.J., Peart, J., Cooper, C., and Roberts, K. (1991). Developmentally regulated epitopes of cell surface arabinogalactan proteins and their relation to root tissue pattern formation. Plant J. **1:** 317–326.

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

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.

Lee, Y., Rubio, M.C., Alassimone, J., and Geldner, N. (2013). A mechanism for localized lignin deposition in the endodermis. Cell **153:** 402–412.

Leplé, J.-C., et al. (2007). Downregulation of cinnamoyl-coenzyme A reductase in poplar: multiple-level phenotyping reveals effects on cell wall polymer metabolism and structure. Plant Cell **19:** 3669–3691.

Leyser, O. (2008). Strigolactones and shoot branching: a new trick for a young dog. Dev. Cell **15:** 337–338.

Liu, L., Shang-Guan, K., Zhang, B., Liu, X., Yan, M., Zhang, L., Shi, Y., Zhang, M., Qian, Q., Li, J., and Zhou, Y. (2013). Brittle Culm1, a COBRA-like protein, functions in cellulose assembly through binding cellulose microfibrils. PLoS Genet. **9:** e1003704.

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.

Marcus, S.E., et al. (2010). Restricted access of proteins to mannan polysaccharides in intact plant cell walls. Plant J. **64:** 191–203.

McCartney, L., Marcus, S.E., and Knox, J.P. (2005). Monoclonal antibodies to plant cell wall xylans and arabinoxylans. J. Histochem. Cytochem. **53:** 543–546.

McDougall, G.J. (1991). Cell-wall-associated peroxidases and lignification during growth of flax fibres. J. Plant Physiol. **139:** 182–186.

McDougall, G.J. (1992). Changes in cell wall-associated peroxidases during the lignification of flax fibres. Phytochemistry **31:** 3385–3389.

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. USA **107:** 22728–22733.

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.

Morreel, K., Dima, O., Kim, H., Lu, F., Niculaes, C., Vanholme, R., Dauwe, R., Goeminne, G., Inzé, D., Messens, E., Ralph, J., and Boerjan, W. (2010a). Mass spectrometry-based sequencing of lignin oligomers. Plant Physiol. **153:** 1464–1478.

Morreel, K., Kim, H., Lu, F., Dima, O., Akiyama, T., Vanholme, R., Niculaes, C., Goeminne, G., Inzé, D., Messens, E., Ralph, J., and Boerjan, W. (2010b). Mass spectrometry-based fragmentation as an identification tool in lignomics. Anal. Chem. **82:** 8095–8105.

Morreel, K., Ralph, J., Kim, H., Lu, F., Goeminne, G., Ralph, S., Messens, E., and Boerjan, W. (2004). Profiling of oligolignols reveals monolignol coupling conditions in lignifying poplar xylem. Plant Physiol. **136:** 3537–3549.

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.

Neutelings, G. (2011). Lignin variability in plant cell walls: contribution of new models. Plant Sci. **181:** 379–386.

Nielsen, K.L., Indiani, C., Henriksen, A., Feis, A., Becucci, M., Gajhede, M., Smulevich, G., and Welinder, K.G. (2001). Differential activity and structure of highly similar peroxidases. Spectroscopic, crystallographic, and enzymatic analyses of lignifying *Arabidopsis thaliana* peroxidase A2 and horseradish peroxidase A2. Biochemistry **40:** 11013–11021.

Østergaard, 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.

Ralph, J., Akiyama, T., Coleman, H.D., and Mansfield, S.D. (2012). Effects on lignin structure of coumarate 3-hydroxylase downregulation in poplar. BioEnergy Res. **5:** 1009–1019.

Roach, M.J., and Deyholos, M.K. (2007). Microarray analysis of flax (*Linum usitatissimum* L.) stems identifies transcripts enriched in fibre-bearing phloem tissues. Mol. Genet. Genomics **278:** 149–165.

Sasidharan, R., Chinnappa, C.C., Voesenek, L.A., and Pierik, R. (2008). The regulation of cell wall extensibility during shade avoidance: a study using two contrasting ecotypes of *Stellaria longipes*. Plant Physiol. **148:** 1557–1569.

Schweizer, P. (2008). Tissue-specific expression of a defence-related peroxidase in transgenic wheat potentiates cell death in pathogen-attacked leaf epidermis. Mol. Plant Pathol. **9:** 45–57.

Shigeto, J., Kiyonaga, Y., Fujita, K., Kondo, R., and Tsutsumi, Y. (2013). Putative cationic cell-wall-bound peroxidase homologues in Arabidopsis, AtPrx2, AtPrx25, and AtPrx71, are involved in lignification. J. Agric. Food Chem. **61:** 3781–3788.

Sindhu, A., Langewisch, T., Olek, A., Multani, D.S., McCann, M.C., Vermerris, W., Carpita, N.C., and Johal, G. (2007). Maize Brittle stalk2 encodes a COBRA-like protein expressed in early organ development but required for tissue flexibility at maturity. Plant Physiol. **145:** 1444–1459.

Smallwood, M., Yates, E.A., Willats, W.G., Martin, H., and Knox, J.P. (1996). Immunochemical comparison of membrane-associated and secreted arabinogalactan-proteins in rice and carrot. Planta **198:** 452–459.

Sorefan, K., Booker, J., Haurogné, K., Goussot, M., Bainbridge, K., Foo, E., Chatfield, S., Ward, S., Beveridge, C., Rameau, C., and Leyser, O. (2003). MAX4 and RMS1 are orthologous dioxygenase-like genes that regulate shoot branching in Arabidopsis and pea. Genes Dev. **17:** 1469–1474.

Torres, M.A. (2010). ROS in biotic interactions. Physiol. Plant. **138:** 414–429.

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 plants. Plant Physiol. **162:** 918–926.

Van Acker, R., Vanholme, R., Storme, V., Mortimer, J.C., Dupree, P., and Boerjan, W. (2013). Lignin biosynthesis perturbations affect secondary cell wall composition and saccharification yield in *Arabidopsis thaliana*. Biotechnol. Biofuels **6:** 46.

Vanholme, R., Demedts, B., Morreel, K., Ralph, J., and Boerjan, W. (2010). Lignin biosynthesis and structure. Plant Physiol. **153:** 895–905.

Vanholme, R., Morreel, K., Darrah, C., Oyarce, P., Grabber, J.H., Ralph, J., and Boerjan, W. (2012a). Metabolic engineering of novel lignin in biomass crops. New Phytol. **196:** 978–1000.

Vanholme, R., Storme, V., Vanholme, B., Sundin, L., Christensen, J.H., Goeminne, G., Halpin, C., Rohde, A., Morreel, K., and Boerjan, W. (2012b). A systems biology view of responses to lignin biosynthesis perturbations in Arabidopsis. Plant Cell **24:** 3506–3529.

Wang, Z., et al. (2012). The genome of flax (Linum usitatissimum) assembled de novo from short shotgun sequence reads. Plant J. **72:** 461–473.

Weng, J.-K., and Chapple, C. (2010). The origin and evolution of lignin biosynthesis. New Phytol. **187:** 273–285.

Whetten, R., and Sederoff, R. (1995). Lignin biosynthesis. Plant Cell **7:** 1001–1013.

Wijnen, C.L., and Keurentjes, J.J. (2014). Genetic resources for quantitative trait analysis: novelty and efficiency in design from an Arabidopsis perspective. Curr. Opin. Plant Biol. **18:** 103–109.

Yang, Y.H., Dudoit, S., Luu, P., Lin, D.M., Peng, V., Ngai, J., and Speed, T.P. (2002). Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. Nucleic Acids Res. **30:** e15.

Yates, E.A., and Knox, J.P. (1994). Investigations into the occurrence of plant cell surface epitopes in exudate gums. Carbohydr. Polym. **24:** 281–286.

Yates, E.A., Valdor, J.-F., Haslam, S.M., Morris, H.R., Dell, A., Mackie, W., and Knox, J.P. (1996). Characterization of carbohydrate structural features recognized by anti-arabinogalactan-protein monoclonal antibodies. Glycobiology **6:** 131–139.

Zhao, Q., and Dixon, R.A. (2011). Transcriptional networks for lignin biosynthesis: more complex than we thought? Trends Plant Sci. **16:** 227–233.

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.

Zhong, R., Kays, S.J., Schroeder, B.P., and Ye, Z.-H. (2002). Mutation of a chitinase-like gene causes ectopic deposition of lignin, aberrant cell shapes, and overproduction of ethylene. Plant Cell **14:** 165–179.

Zhong, R., Richardson, E.A., and Ye, Z.-H. (2007). The MYB46 transcription factor is a direct target of SND1 and regulates secondary wall biosynthesis in Arabidopsis. Plant Cell **19:** 2776–2792.

Ectopic Lignification in the Flax *lignified bast fiber1* Mutant Stem Is Associated with Tissue-Specific Modifications in Gene Expression and Cell Wall Composition

Maxime Chantreau, Antoine Portelette, Rebecca Dauwe, Shingo Kiyoto, David Crônier, Kris Morreel, Sandrine Arribat, Godfrey Neutelings, Malika Chabi, Wout Boerjan, Arata Yoshinaga, François Mesnard, Sebastien Grec, Brigitte Chabbert and Simon Hawkins

Plant Cell 2014;26;4462-4482; originally published online November 7, 2014;
DOI 10.1105/tpc.114.130443