Université catholique de Louvain

Faculté d'ingénierie biologique, agronomique et environnementale

Molecular investigation of cell wall formation in hemp stem tissues

---

Marc BEHR

Doctoral thesis in Agronomical sciences and biological engineering

PhD committee:

Prof. Hans van Dyck, UcL (chairperson)

Prof. Stanley Lutts, UcL (supervisor)

Dr. Jean-François Hausman, LIST (supervisor)

Prof. Pierre Bertin, UcL

Prof. Patrick Gerin, UcL

Dr. Gea Guerriero, LIST

Prof. Simon Hawkins (Lille1)

Abstract

The interest around hemp is currently renewed for the production of bio-based and renewable materials in the context of global warming. This fibre crop is able to produce large amounts of fibres with different features and industrial application. The most valuable fibres, found in the phloem tissue, are rich in cellulose while relatively poor in lignin, by contrast with the xylem tissue. This thesis aims at studying the events related to the biogenesis of these two types of fibre. Using two different systems, we will investigate the molecular mechanisms related to elongation, secondary growth, deposition and maturation of the cell wall and bast fibre development.

Our first objective is to provide a comprehensive overview of the transcriptional factors and phytohormones involved in primary and secondary growth. To this end, the development of the hemp hypocotyl is investigated by a high throughput transcriptomic approach, in addition to proteomics, phytohormone and lignin analyses and microscopy. We show that elongation and secondary growth are characterised by specific patterns of gene expression. The consequences on the biogenesis and modification of the cell walls are widely discussed.

The second objective is to decipher the molecular actors associated with the development of the bast fibres. We first show that the contrasting composition of xylem and bast fibres is regulated at the transcriptional level. Next, we highlight the evolution of the transcriptome during the development of the bast fibres, from intrusive growth to thickening. We put a special emphasis on the study of cell wall-related genes (cellulose synthase, fasciclin-like arabinogalactan,

transcription factors) and phytohormones. We also formulate several hypotheses to explain the hypolignification of the bast fibres.

Finally, this thesis ends with the perspectives raised by our results, notably concerning the deposition of cellulose, non-cellulosic polysaccharides and lignin in the bast fibres.

Remerciements

Une thèse de doctorat est toujours une aventure. En premier lieu pour le principal intéressé, mais aussi pour celles et ceux qui y contribuent, y participent et apportent leur soutien à qui entreprend cette folle expérience. Les mots peuvent dans ces cas-là paraître dérisoires, mais je tiens tout de même à ne pas en faire l'économie…

Si j'en suis aujourd'hui à écrire ces lignes (ainsi que les quelques pages à suivre), Gea Guerriero y a contribué de manière décisive. Merci Gea pour m'avoir apporté, en toutes circonstances, bien des paroles et des actes qui m'ont conduit à ce niveau de développement scientifique et humain. Au-delà des aspects nécessairement académiques de notre relation, j'ai en toutes circonstances trouvé une oreille attentive (et très réactive !) pour écouter mes turpitudes et me faire rebondir lors des épreuves qui ont émaillé ces quatre années. Tu m'as grandement fait bénéficier de tes multiples connaissances scientifiques et compétences méthodologiques, tout particulièrement en biologie moléculaire, sur les parois, les plantes à fibres et en histologie. Le RNA-Seq est si facile avec toi ! Et que dire du bond réalisé en rédaction grâce à toi ! Tout le chocolat du monde ne suffirait pas à te dire ma gratitude pour tout ce que tu m'as apporté. *Grazie tante !*

Il me revient de remercier égalitairement Stanley Lutts et Jean-François Hausman, pour leur confiance et leur idée un peu saugrenue d'avoir confié à un simple ingénieur viticole, œnologue de surcroît, ce poste de doctorant dans un projet si complexe. Mes promoteurs ont souvent été de bon conseil, au niveau scientifique comme au niveau organisationnel. Après huit ans de présence cumulée au LIST, quatre ans à l'Université catholique de Louvain, je suis fier du parcours que

vous m'y avez fait emprunter. Et il n'est pas donné à tout le monde d'avoir un promoteur passionné de Révolution française. Quelle chance pour un doctorant français d'être reçu sabre au clair dans un bureau à la gloire de ces grandes heures historiques ! Tu m'as toujours chaleureusement accueilli, te rendant disponible quand cela était nécessaire et œuvrant à mon intégration au sein du GRPV.

J'ai eu la chance d'avoir un fil rouge tout au long de ma présence au LIST depuis 2009, en la personne de Sylvain Legay. De très proche voisin de bureau lors de l'épopée Nutox, à ta présence et tes conseils rassurants ces derniers mois, en passant par mon initiation aux lois incertaines de la biologie moléculaire et à nos discussions musicales, politiques et scientifiques, c'est aussi grâce à toi que je me suis pleinement intégré dans cet environnement. Cela n'est certainement pas étranger à mon parcours, et je t'en remercie !

L'ambition scientifique du projet Cancan m'a amené à travailler avec un grand nombre de personnes, avec des outils à la complexité bien trop grande pour le béotien que je suis dans un certain nombre de domaines. Je remercie particulièrement Kjell Sergeant pour ses lumières et son travail en matière de protéomique, ainsi que ses conseils dans bien des domaines de ma thèse ; merci également à Céline Leclercq, Sébastien Planchon et Jenny Renaut pour leur aide technique sur ces expériences. Un grand merci à l'équipe de chimie analytique, Audrey Lenouvel, Cédric Guignard et Emmanuelle Cocco, à qui mes recherches ont donné bien du fil à retordre depuis les analyses de vin jusqu'à la caractérisation de la lignine. La chimie fait partie de mes fidèles compagnons de route, et ce fut un plaisir de travailler en votre compagnie. Je remercie également Eva Žižková, Václav Motyka et Petre Dobrev (Institute of Experimental Botany from the Czech Academy of Science) pour les analyses de phytohormones,

qui ont apporté une réelle plus-value à mes résultats. Merci également à Christelle André et Sophie Charton pour leur implication dans les analyses métabolomiques, et leurs explications à ce sujet.

Le travail au laboratoire ne serait pas ce qu'il est sans le concours d'Aude Corvisy, de Laurent Solinhac et de Camille Renault, dont j'ai bien apprécié l'aide et les conseils en plusieurs occasions. Merci à vous trois ! Ainsi qu'aux autres membres de l'équipe technique, dont Boris, Servane, Marine, François et Denis.

J'adresse aussi mes plus sincères remerciements aux techniciens du GRPV à Louvain, Baudouin, Brigitte et Jacques, pour leur aide lors des campagnes de récolte de chanvre, qui n'étaient pas mince affaire en termes de logistique et d'épluchage de tiges !

Une pensée affectueuse pour les doctorants-assistants du GRPV, Marie, Anne-Sophie, Delphine, Emna, Willy et les autres, pour leur gentillesse et leur amitié. Je vous souhaite pleine réussite dans vos travaux de recherche et pour votre avenir.

Je remercie respectueusement André Lejeune, l'autre pilier du GRPV, pour les discussions que nous avons eu, ses bons conseils et son humeur joyeuse. C'est toujours un plaisir de partager quelques moments avec vous André.

A titre plus officiel, que les membres de mon jury de thèse, son président Prof. Hans van Dyck, son secrétaire Prof. Pierre Bertin, ses examinateurs les Prof. Simon Hawkins et Patrick Gerin, soient aussi remerciés pour le travail accompli sur ce document. Je remercie le Fonds National de la Recherche du Luxembourg pour le financement du projet Cancan.

Une thèse étant encore une période étudiante, cela va sans dire que je garderai toujours le souvenir des personnes qui ont partagé les quelques moments de fête et de convivialité inhérents à ce statut. Un grand merci à mes chers collègues grecs, Bella et Christos, la crème de la crème, d'avoir illuminé les folles nuits luxembourgeoises et animé de si nombreuses discussions au bureau ! Merci à Xuan, Kahina, Annelie, Daniel, Sébastien, Gilles, Valentin et tous les autres. Yes, this is to exercise your French ;)

Enfin, je remercie sans retenue aucune toute ma famille pour leur soutien, leurs encouragements et leur confiance en moi, en toutes circonstances. Mes chers parents, vous m'avez souvent répété que « personne ne peut travailler à ta place », j'espère en avoir fait la démonstration ! Annaeg et Artus, mes petits neveux, Christelle et Amaury, si souvent intéressés par mes avancées scientifiques, merci ! Chers cousins et oncles, chères cousines et tantes, à toute ma famille, je vous adresse mes plus chaleureux remerciements.

Preface

This PhD thesis was carried out in the context of the research project "CANCAN" which ran between 2014 and 2018 at Luxembourg Institute of Science and Technology (LIST). The PhD study has been co-supervised by:

- Prof. Stanley Lutts (Université catholique de Louvain)
- Dr. Jean-François Hausman (LIST)

CANCAN (A multidisciplinary approach to understand the molecular basis of the heterogeneous cell wall composition in hemp stems) was financially supported as a CORE project by the Fonds National de la Recherche, Luxembourg (C13/SR/5774202). Researchers from LIST and several institutions (Université de Lille, University of Vienna and UcL) were involved in this project. It aims to study the factors regulating the heterogeneous lignification pattern in hemp stem by putting a special emphasis on the phenylpropanoid pathway and the cell wall formation by a core transcriptomic approach (headed by Dr. Gea Guerriero), together with proteomics (Dr. Kjell Sergeant), metabolomics (Dr. Christelle André), wet chemistry and imaging techniques.

As part of my PhD, I have conducted a significant portion of the research in CANCAN, including laboratory experiments, data interpretation, article writing and cooperation with other researchers partaking in the project.

Thesis achievements

a. Articles published in the thesis frame

Behr M, Legay S, Žižková E, Motyka V, Dobrev PI, Hausman JF, *et al.* (2016). Studying secondary growth and bast fiber development: The hemp hypocotyl peeks behind the wall. *Frontiers in Plant Science* 7. (Chapter 1).

Guerriero G, Behr M, Legay S, Mangeot-Peter L, Zorzan S, Ghoniem M, *et al.* (2017). Transcriptomic profiling of hemp bast fibres at different developmental stages. *Scientific Reports* 7. (Chapter 4).

Guerriero G, Mangeot-Peter L, Legay S, Behr M, Lutts S, Siddiqui KS, *et al.* (2017). Identification of fasciclin-like arabinogalactan proteins in textile hemp (*Cannabis sativa* L.): In silico analyses and gene expression patterns in different tissues. *BMC Genomics* 18. (Chapter 3).

Behr M, Legay S, Hausman JF, Lutts S and Guerriero G. (2017). Molecular investigation of the stem snap point in textile hemp. *Genes* 8. (Chapter 5).

Behr M, Sergeant K, Leclercq C, Planchon S, Guignard C, Lenouvel A, *et al.* (2018). Insights into the molecular regulation of monolignol-derived product biosynthesis in the growing hemp hypocotyl. *BMC Plant Biology* 18. (Chapter 2).

b. Articles submitted in the thesis frame

Behr M, Lutts S, Hausman JF and Guerriero G. (2018). Jasmonic acid to boost secondary growth in hemp hypocotyl. *Planta*. (Appendix).

Behr M, Lutts S, Hausman JF and Guerriero G. (2018). The “ins and outs” of hemp stem tissue development. *Fibers*. (Chapter 3).

c. Articles about cell walls

Behr M, Legay S, Hausman JF and Guerriero G. (2015). Analysis of cell wall-related genes in organs of *Medicago sativa* L. under different abiotic stresses. *International Journal of Molecular Sciences* 16**,** 16104-16124.

Guerriero G, Behr M, Hausman JF and Legay, S. (2017). Textile hemp *vs*. salinity: Insights from a targeted gene expression analysis. *Genes* 8.

Sergeant K, Printz B, Gutsch A, Behr M, Renaut J and Hausman JF. (2017). Didehydrophenylalanine, an abundant modification in the beta subunit of plant polygalacturonases. *PLoS ONE* 12.

Guerriero G, Behr M, Backes A, Faleri C, Hausman JF, Lutts S and Cai G. (2017). Bast fibre formation: insights from Next-Generation Sequencing. *Procedia Engineering* 200, 229-235.

d. Articles published out of the thesis scope

Molitor D, Behr M, Hoffmann L and Evers D. (2012). Impact of grape cluster division on cluster morphology and bunch rot epidemic. *American Journal of Enology and Viticulture* 63**,** 508-514.

Behr M, Cocco E, Lenouvel A, Guignard C and Evers D. (2013). Earthy and fresh mushroom off-flavors in wine: Optimized remedial treatments. *American Journal of Enology and Viticulture* 64**,** 545-549.

Behr M, Legay S and Evers D. (2013). Molecular identification of *Botrytis cinerea*, *Penicillium* spp. and *Cladosporium* spp. in Luxembourg. *Journal International des Sciences de la Vigne et du Vin* 47**,** 239-247.

Behr M, Serchi T, Cocco E, Legay S, Molitor D, Guignard C, *et al.* (2013). Genomic and proteomic studies on the pathogenic grape flora in Luxembourg. *Journal of Plant Pathology* 1.

Camara M, Gharbi N, Lenouvel A, Behr M, Guignard C, Orlewski P, *et al.* (2013). Detection and quantification of natural contaminants of wine by gas chromatography-differential ion mobility spectrometry (GC-DMS). *Journal of Agricultural and Food Chemistry* 61**,** 1036-1043.

Behr M, Serchi T, Cocco E, Guignard C, Sergeant K, Renaut J, *et al.* (2014). Description of the mechanisms underlying geosmin production in *Penicillium expansum* using proteomics. *Journal of Proteomics* 96**,** 13-28.

Content

Acronyms

$(GlcNAc)_n$	*N*-acetylglucosamine oligosaccharide
2D-DiGE	Two-dimensional difference gel electrophoresis
4CL	4-coumaric acid-coa ligase
ABA	Abscisic acid
ABP1	Auxin-binding protein 1
ACC	1-aminocyclopropane-1-carboxylic acid
ACN	Acetonitrile
AFB	Auxin signaling f-box
AGP	Arabinogalactan protein
AIF	Atbs1 interacting factor
AIL1	Ant-like 1
ANT	Aintegumenta
ARF	Auxin response factor
ARR	Arabidopsis response regulator
ASD	α-L-Arabinofuranosidase
ASD1/ARAF1	α-L-Arabinofuranosidase/β-D-Xylosidase
ASP	Internode above the snap point
ATMS1	Methionine synthase
BBSP	Internode below-below the snap point
BGAL	β-Galactosidase
BGLU	β-Glucosidase
bHLH	Basic helix loop helix
BLH	BEL1-like homeodomain 1
BR	Brassinosteroids
BSP	Internode below the snap point
BSTFA	Bis(trimethylsilyl)trifluoroacetamide
BXL	β-Xylosidase
bZIP	Basic leucine zipper
BZR	Brassinazole resistant
C3'H	*P*-Coumaroyl Shikimate 3’ Hydroxylase
C4H	Cinnamate-4-hydroxlase
CAB	Chlorophyll a/b binding protein
CAD	(Hydroxy)cinnamyl alcohol dehydrogenase
CBM3a	Carbohydrate binding module 3a

CCoAOMT	Caffeoyl-CoA 3-O-Methyltransferase
CCR	(Hydroxy)cinnamoyl-CoA reductase
CESA	Cellulose synthase
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
CHS	Chalcone synthase
CK	Cytokinins
CNRQ	Calibrated normalized relative quantities
COBL	Cobra-like protein
COI	Coronatine insensitive
COMT	Caffeic acid/5-hydroxyferulic acid o-methyltransferase
COP	Constitutive photomorphogenesis
CRF	Cytokinin response factor
CSC	Cellulose synthase complex
CSL	Cellulose synthase like
CTAB	Cetyltrimethyl ammonium bromide
CTL	Chitinase-like
CWI	Cell wall invertase
CWR	Cell wall residue
CYP	Cytochrome P450
DAB	3,3'-diaminobenzidine
DCG	Dehydrodiconiferyl alcohol glucoside
DDCG	Dihydro-dehydrodiconiferyl alcohol glucoside
DIR	Dirigent protein
DLP	Dirigent-like protein
DRT	Dna-damage-repair/toleration protein
DTT	Dithiothreitol
DUF	Domain of Unknown Function
DZ	Dihydrozeatin
EDTA	Ethylenediaminetetraacetic acid
EIN	Ethylene insensitive
ER	Endoplasmic reticulum
ERF	Ethylene response factor
EXPA	Class α Expansin
F3H	Flavanone 3-hydroxylase
F5H	Ferulate 5-hydroxylase
FAD	Fatty acid desaturase

FAS	Fasciclin
FLA	Fasciclin-like arabinogalactan protein
FRK	Fructokinase
G6PI	Glucose-6-phosphate isomerase
GA	Gibberellins
GA1/CPS	Ent-copalyl diphosphate synthase
GALS	Galactan synthase
GA-OX	2-oxoglutarate-dependent dioxygenase
GEO	Gene expression omnibus (NCBI)
GH	Glycosyl hydrolase
GH3	Gretchen hagen 3
GID1	GA insensitive dwarf 1
G-Layer	Gelatinous layer
Gn-Layer	Galactan layer
GOE	Gene ontology term enrichment
GPI	Glycosylphosphatidylinositol
GT	Glycosyl transferase
GUX	Glucuronic acid substition of xylan
GXM	Glucuronoxylan methyltransferase
H6 to H20	Hypocotyls from 6 to 20 days after sowing
HA	H+-ATPase
HB	Homeobox domain
HCT	Hydroxycinnamoyl-CoA Shikimate/Quinate Hydroxycinnamoyl-Transferase
HK	Hexokinase
HPT	Homogentisate phytyltransferase
HY	Long hypocotyl
IAA	Indole-3-acetic acid
IBH1	ILI1 Binding bHLH 1 Protein
ICA	Independent component analysis
INV	Invertase
iP	N6-(δ2-isopentenyl)adenine
IRX	Irregular xylem phenotype
JA	Jasmonic acid
JA-Ile	Jasmonoyl-isoleucine
JAZ	Jasmonate-ZIM-domain
KAO2	*Ent*-kaurenoic acid oxidase 2
KCR	β-Ketoacyl Reductase

KCS	3-Ketoacyl-CoA Synthase
KDR	Kidari
KFB	Kelch-repeat F-box protein
KOR1	Korrigan
LAC	Laccase
LBD	Lob-domain containing protein
LCC	Lignin-carbohydrate complex
LHCB	Light harvesting complex
LHW	Lonesome highway
Lignin G	Guaiacyl
Lignin H	*P*-Coumaryl
Lignin S	Syringyl
LTP	Non-specific lipid transfer protein
LTPG	LTP-GPI
MAP	Microtubule-associated protein
MCM2	Minichromosome maintenance protein
MTHFR	Methylenetetrahydrofolate reductase
MYB	Myeloblastosis transcription factor
NCED	Nine-cis-epoxycarotenoid dioxygenase
NINJA	Novel interactor of JAZ
NodC	*N*-acetylglucosaminyltransferase
NSAF	Normalized spectral abundance factor
NST	NAC secondary wall thickening promoting factor
OPDA	12-oxo-phytodienoic acid
ORC	Origin recognition complex subunit
PAA	Phenylacetic acid
PAL	Phenylalanine ammonia lyase
PBS	Phosphate buffer saline
PCA	Principal component analysis
PCBER	Phenylcoumaran benzylic ether reductase
PCNA	Proliferating cell nuclear antigen
PCW	Primary cell wall
PDF	Protodermal factor gene
PGA	Polygalacturonic acid
PGIP	Polygalacturonase inhibiting protein
PGM	Phosphoglucomutase
Phe	Phenylalanine

PIF	Phytochrome-interacting factor
PIN1	Pin-formed 1
PLATZ	Plant AT-rich sequence and zinc-binding protein
PLP	Patatin-related phospholipase
PLR	Pinoresinol lariciresinol reductase
PME/I	Pectin methylesterase / inhibitor
PR	Pathogenesis-related gene
PRR	Pinoresinol reductase
PRX	Peroxidase
PUX5	Pentosylated side chain
PVP	Polyvinylpyrrolidone
PXC	PXY/TDR-correlated genes
PXY	Phloem intercalated with xylem
RG	Rhamnogalacturonan
RIN	RNA integrity number
RNA-Seq	RNA sequencing
RPKM	Reads per kilobase transcript per million reads
RWA	Reduced wall acetylation
SAHH	S-adenosyl-L-homocysteine hydrolase
SAM	S-adenosylmethionine
SAMS	S-adenosylmethionine synthase
SANS	Small-angle neutron scattering
SAUR	Small auxin-up RNA
SCW	Secondary cell wall
SDG	(+)-Secoisolariciresinol diglucoside
SHYG	Speedy hyponastic growth
SND	Secondary wall associated NAC domain protein
SOC1	Suppressor of overexpression of CO1
SP	Internode containing the snap point
ssNMR	Solid-state nuclear magnetic resonance
SUS	Sucrose synthase
TAIR	The Arabidopsis information resource
TALE	Three amino acid loop extension
TBL	Trichome birefringence-like
TBS	Tris buffer saline
TCA	Trichloroacetic acid
TDIF	Tracheary element differentiation inhibitory factor

TDR	TDIF receptor
TF	Transcription factor
TFA	Trifluoroacetic acid
THFS	10-formyltetrahydrofolate synthetase
TIR1	Transport inhibitor response 1
TT	Transparent testa
TW	Tension wood
tZ	Trans-zeatin
UDG	UDP-glucose dehydrogenase
UGT	UDP glucosyltransferase
UVI	UV-b-insensitive
UXS	UDP-xylose synthase
VDAC	Voltage dependent anion channel
VLCFA	Very long chain fatty acid
VND	Vascular-related nac-domain
WAT	Walls are thin
WAXS	Wide-angle x-ray scattering
WOX	Wuschel-related homeobox
WRI	Wrinkled
XEH	Xyloglucan endohydrolase
XET	Xyloglucan endotransglucosylase
XND	Xylem NAC domain
XRES	Xylan reducing-end sequence
XT	Xylosyltransferase
XTH	Xyloglucan endotransglucosylase/hydrolase
XYL	α-xylosidase
XYP	Xylogen protein
ZFP	Zinc finger protein
β-ME	β-Mercaptoethanol

List of Figures

List of Tables

List of Supplementary data

Supplementary data 1-1: Validation of RNA-Seq data by RT-qPCR analysis of 10 selected transcripts

Supplementary data 1-2: Genes differentially expressed by RNA-Seq

Supplementary data 1-3: Transcription factors differentially expressed

Supplementary data 1-4: Genes differentially expressed in H6 vs H20 involved in cell wall biogenesis

Supplementary data 2-1: Protein sequences used for the phylogenetic analysis

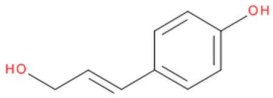
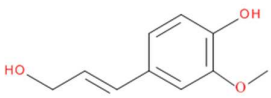
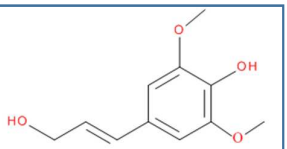
Supplementary data 2-2: RT-qPCR and gene amplification efficiency data

Supplementary data 2-3: Gel-based and gel-free proteomics data

Supplementary data 4-1: Assembly and expression analysis of TOP, MID and BOT samples assessed by RNA-Seq and RT-qPCR

Supplementary data 5-1: Gene expression analysis of SP1, SP2 and SP3 assessed by RT-qPCR

List of recurrent molecules

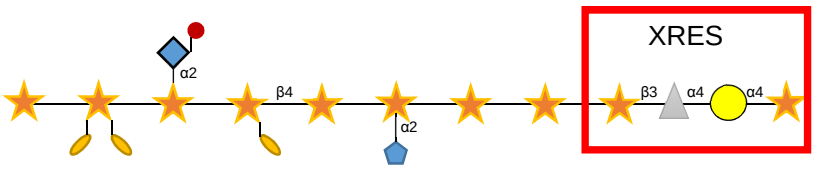
		
p-coumaryl alcohol (lignin H-unit)	Coniferyl alcohol (lignin G-unit)	Sinapyl alcohol (lignin S-unit)

Typical structure of Dicot polysaccharides

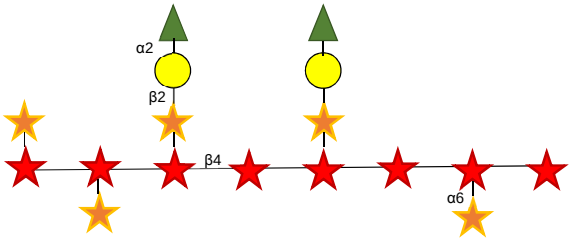
a) xylan (glucuronoxylan with xylan reducing end sequence XRES),

b) xyloglucan (fucogalactoxyloglucan)

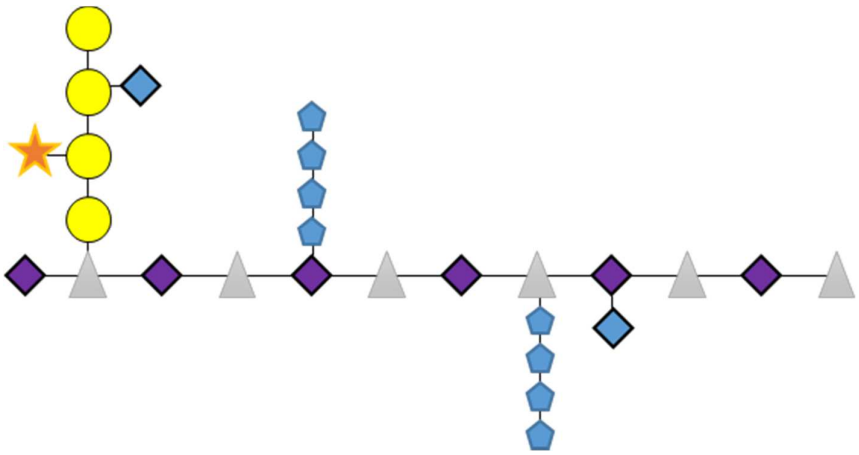
c) Rhamnogalacturonan-I



a) Glucuronoxylan

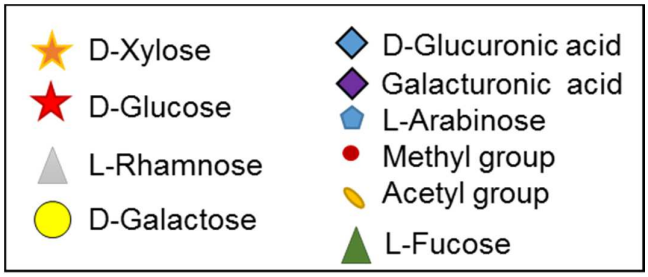


b) Fucogalactoxyloglucan



c) Rhamnogalacturonan-I

Key for the composition of hemicelluloses and pectin



Glossary

Diffuse growth: Biological process occurring during the very early stage of bast fibre elongation. All the tissues grow symplastically, at the same rate. Hemp primary bast fibres grow symplastically in the region located between 1 and 2 mm from the shoot apical meristem.

Fibre: Histologically, a fibre is a sclerenchyma cell showing the following characteristics: i) an elongated shape, ii) a thick secondary cell wall; iii) tapered ends, and iv) an important role in plant mechanical properties. The term Fibre also corresponds to the manufactured product originating from the processing of natural fibres (such as hemp or flax bast fibres) or artificial fibres (such as carbon or glass). This thesis mainly focus on the first acceptance of this word.

Gelatinous-type secondary cell wall (abbreviated as G-type SCW): Gelatinous cell walls are built in the bast fibres of flax, hemp and ramie, and in tension wood, and consist in a S1+S2+G ultrastructure. This name came from the swelling of this layer during its preparation with alkali for microscopic analysis, but also refers to its mechanical performance and related mesoporosity. It is rich in crystalline cellulose (>75%) while poor in lignin (<5%).

Hempcrete: Bio-composite material obtained by mixing hemp hurds (xylan-type secondary cell walls) with lime. The high silica content of hurds favours the binding with lime. It is used as a material in the construction industry. It is very light in weight.

Intrusive growth: Biological process accounting for the extreme elongation of bast fibres and other tissues (in hemp and other species). The tip of the fibre invades the middle lamella of neighbouring cells while elongation takes place at the whole cell surface (by contrast with

tip growth). Hemp primary bast fibres undergoes diffuse growth from 2 mm to 7-8 cm from the shoot apical meristem. Deposition of the secondary cell wall starts at the end of intrusive growth.

Secondary growth: biological process leading to girth increase through cambial activity. Secondary growth results in the formation of secondary phloem and xylem and accounts for the major part of biomass production on Earth.

Snap point: Spot of the stem corresponding to the transition from elongation (above it) to primary bast fibre thickening (below it). This results in changes in fibre mechanical properties. The snap point is widely used to study the cell wall-related changes in fibre crops such as hemp and flax.

Xylan-type secondary cell wall: this type of SCW is found in the xylem of hemp and flax stem, and in woody species. By contrast with G-type SCW, it is rich in xylan and lignin (>15%). It consists in a S1+S2+S3 ultrastructure. The presence of both xylan and lignin provides physical resistance through various covalent linkages (xylan-cellulose, xylan-lignin).

Introduction

1. Hemp

a. Hemp and humankind

Originating from central Asia, plants of the genus *Cannabis* were first harvested 8500 years ago, a finding suggesting that they were among the first actively grown crops, 4000 to 6000 years ago (Fike, 2016). In Europe, hemp found its way around 1500 BC and was grown in Great Britain for grain and fibre. Later on, hemp fibres have been used in whole Europe to make cloth, sails, ropes and bags. Hemp fibres were also of special importance for the maritime powers such as England and France, to produce canvas and ropes for sailing boats (Fike, 2016). Hemp was at this moment a very important crop which was widely cultivated, with a crop growing surface higher than 176,000 hectares in France in the beginning of the 19th century (Razous, 1944). Declining demand for fibres, lowered cotton processing costs and importation of jute and abaca fibres from overseas in the mid-19th started to decrease its interest, as well as the bad publicity resulting from its use for hallucinogenic purposes (Fike, 2016). In the first half of the 20th, several multilateral treaties started to prohibit the use of drugs (e.g., cocaine, heroin and *Cannabis*). These treaties resulted in the adoption of the Single Convention on Narcotic Drugs in 1961 under the depositary of the Secretory-General of the United Nations. This Convention, together with the beginning of the petrol era and the synthetic fibre production, has marked the fall from grace of hemp, with a minimum cropping area of 700 hectares in France in 1960. Nowadays we can note a renewed interest for this crop. Environmental and geopolitical concerns about abusive use of fossil energies favours

the cultivation of fibre plants. Global warming and climate change are now widely considered as a massive threat to humankind. The anthropogenic contribution to this dramatic phenomenon is documented by several high-scale studies (Fischer and Knutti, 2015; Stips *et al.*, 2016). This change is massively fuelled by the combustion of fossil energies such as oil and coal which emits large quantities of CO_2. Because, globally, a third of petroleum reserves, half of gas reserves and over 80 % of coal reserves should remain unused from 2010 to 2050 to maintain the global warming below 2°C (McGlade and Ekins, 2015), there is an urgent need to decarbonise our way of life. Investing in renewable resources is therefore of strategic importance to maintain the global warming below 2°C (Dincer, 2000; Arent *et al.*, 2011; Dhillon and von Wuehlisch, 2013; Ellabban *et al.*, 2014). Bio-based materials are part of this answer (Hermann *et al.*, 2007).

b. Valorisation of hemp

Petroleum, in addition to be an important source of energy (electricity, heating systems), is also used as raw resource for the production of several materials (tar, asphalt, nylon, tires, plastic wool, etc…). Some of these materials, such as plastic wool and nylon, can be advantageously substituted with plant fibres (Fike, 2016). Fibre composite partly made with hemp fibres are already used in the automotive industry and in isolative concrete materials for the construction sector (Fike, 2016). Hemp phloem fibres match with other uses because of their length; nonetheless the associated processing methods are costly, which limit their application in specific area, such as fine papers (cigarette, banknote) and specialty cloth manufacturing (Fike, 2016). Hemp primary bast fibres are preferred to secondary bast fibres for composition reinforcement because of their length and crystalline cellulose content; the thermoplasticity of lignin (softening at

temperatures > 90 °C) modifies the physical properties of secondary bast fibres and hence their applications (Meshram and Palit, 2013).

Hemp is also industrially grown for seed grains and oil, to be used as food additives or feed for cattle, birds or fisheries (Fike, 2016). In human nutrition and health, hemp seeds are considered as a functional food because they are rich in essential fatty acids (α- and γ-linolenic acid, n-3:n-6 ratio about 3:1) and in proteins. Hemp flowers could also have interesting properties for the production of essential oil, with antimicrobial and insecticidal activities (Fike, 2016). The draft genome of a hemp cultivar grown for its seed production, Finola, has been released recently (van Bakel *et al.*, 2011), however it has not been annotated. Besides these aspects, the hundred or so of cannabinoids synthesised by several species of *Cannabis* may be useful pharmaceutical drugs, but their medical applications is hindered by their psychoactive properties (André *et al.*, 2016). For instance, cannabinoids have analgesic properties (Russo, 2011).

c. Agronomy of hemp

Hemp yield and quality rely on several agronomical and physiological features. The first aspect that farmers take into consideration is the choice of the variety. The European Union states that, in order for a *Cannabis* plant to be eligible for financial support as industrial hemp, the tetrahydrocannabinol (THC) content must be lower than 0.2 % (w/w) *(règlement CEE n° 619/71)*, a maximal value applied in France to authorise the culture of hemp. The majority of legal varieties are monoecious. The flowering of male (XY chromosomes) and female (XX chromosomes) plants in dioecious cultivars is not synchronised, causing problems during the mechanical harvest (Faux, 2014). Monoecious plants have both male and female flowers on the same plant. These monoecious varieties (e.g. Santhica, Fédora) have higher

productivity and growth homogeneity, as compared to the dioecious ones. They behave like female plants, *i.e.*, they have an extended growth period and grow more homogeneously as compared to male plants but are smaller and differentiate less fibres (Schluttenhofer and Yuan, 2017 and references therein). However, the establishment of a population with a dominant female character is not straight forward for the breeders and seeds producers, because it requires the manual elimination of male plants before flowering. Male specific markers in both monoecious and dioecious varieties have been identified, allowing the identification of the male character with a simple PCR (Salentijn *et al.*, 2014). Sex determination also depends on the earliness of the cultivar, with the latest being the most feminised (Faux, 2014), and can be altered by chemicals, such as phytohormones or silver thiosulfate (Salentijn *et al.*, 2014).

The most critical point in the agronomy of hemp is the establishment of the seedlings. An excess in humidity during the early days after sowing is highly detrimental to the establishment of the seedlings. It has been observed that increasing sowing density results in higher bast fibre content, as well as a higher proportion of primary bast fibres (unpublished data). Importantly, hemp is a short day plant, implying that the flowering induction occurs below a critical photoperiod threshold (between 14 and 15.5 hours, depending on the variety) (Struik *et al.*, 2000). The precocity of the variety has therefore to be adapted to the eco-region where it is grown. If the growing conditions are good, choosing a late variety will increase the straw yield because of the later date of flowering. By contrast, the highest seed yields are observed for early or mid-early varieties, highlighting the antagonism between the productions of straw / fibres and seeds. Hemp is a very-well adapted crop under our latitudes, requiring a considerable lower

quantity of water to fulfil its development, as compared to cotton or corn for example. Growing hemp is relatively easy because it does not require pesticides, effectively suppresses weeds, has a robust root system and grows very rapidly. Grown under favourable conditions, hemp may yield up to 25 tons of dry matter per hectare with as much as 12 tons of cellulose per hectare (Struik *et al.*, 2000).

Fibre quality strongly relies on the proportion between bark and core tissues from one side, and the proportion between primary and secondary bast fibres on the other side. The bast fibres are separated from the woody core using a microbial process occurring in the field after the stem is cut, the retting (Fike, 2016). Retting time should be long enough to disaggregate lignin, pectin and waxes while preventing the degradation of the cellulose. The xylem tissue constitutes the core (a.k.a hurds or shivs). The fibres in the core are more lignified than the bast fibres (15 % of the cell wall versus 4 %) (van den Broeck *et al.*, 2008; Neutelings, 2011). The highest quality is found in the primary bast fibres, because they are long, poor in lignin (3 %) and rich in crystalline cellulose (75 %) (Crônier *et al.*, 2005; Fernandez-Tendero *et al.*, 2017). An overview of the chemical and physical properties of the main natural and synthetic fibres, as well as some of their potential applications, is provided in Table 1.

Table 1 (next page): Main properties and industrial applications of plant (hemp, flax, ramie, jute) and synthetic (glass and carbon) fibres. Data from Day *et al.*, 2005, Chabi *et al.*, 2017, Crônier *et al.*, 2005, Goudenhooft *et al.*, 2018, Sarkar *et al.*, 2010, Barta *et al.*, 2010, Moxley *et al.*, 2008, Al-Maadeed and Labidi, 2014. na not applicable, nd not determined.

	Bast fibres				Core fibres		Synthetic fibres	
	Hemp	*Flax*	*Ramie*	*Jute*	*Hemp*	*Flax*	*Glass*	*Carbon*
Cellulose (%)	67-75	62-71	68-76	59-71	40-45	50-60	na	na
Lignin (%)	2.7-4.4	1.5-4.2	0.6-0.7	11.8-12.9	15-20	15-30	na	na
Xylan (%)	4	3-5	3.6-4.4	8-12	18-19	10-15	na	na
Young's modulus (GPa)	70	18-30	62-120	26.5	nd	nd	70-86	230-240
Tensile strength (MPa)	690	500-1500	400-938	393-773	nd	nd	2000-4500	4000
Density (g/cm3)	1.47	1.5	nd	1.3	nd	nd	2.5	1.4
Applications	Clothes, biocomposite, yarn	Epoxy matrix composite, car parts	Forage crop, textile, composite	Textile and industrial uses	Hempcrete	Particleboard, valorisation of antioxidant molecules, concrete	Insulation, car parts, tanks	Composite, textile

2. Bast fibres: biogenesis and composition

Hemp produces raw material which can be used in the construction and automotive sectors among others. The hemp stem is of particular interest as a source of industrial fibres in this context. Histologically, a fibre is a sclerenchyma cell showing the following characteristics: i) an elongated shape, ii) a thick secondary cell wall; iii) tapered ends, and iv) an important role in plant mechanical properties (Gorshkova *et al.*, 2012). In the stem, fibres are either associated with the xylem tracheary elements and xylem parenchyma (wood-like core fibres) or with phloem sieve elements and phloem parenchyma, both of primary (procambium) and secondary (cambium) origins. The fibres supporting the phloem are referred to as bast fibres. The bast fibres are separated from the core and cleaned. After setting the desired core/cortex ratio, the fibres are carded to the suitable fineness. Additional steps depends on the final utilisation (paper, composites…) and lead to the so-called technical fibres.

An important difference between hemp fibres located in the xylem or in the phloem is the nature of their secondary cell wall (SCW). In the xylem, this is a xylan-type SCW, while in phloem it is gelatinous (G-layer) (Gorshkova *et al.*, 2012). The biologist can take advantage of this difference to study the mechanism leading to the two types of SCW formation. Xylan-type SCW of the xylem and G-type SCW of the bast fibres contain 15 % and 4 % of lignin, respectively (van den Broeck *et al.*, 2008; Neutelings, 2011). The thickness of the xylan-type SCW is between 1 and 4 µm (Mikshina *et al.*, 2013). In fibre crops, the cortex rich in bast fibres can easily be separated manually, allowing differential analysis of the two regions. This experimental setup has been used to study the hypolignification pattern in flax bast fibres (Day *et al.*, 2005; Chantreau *et al.*, 2014; Mokshina *et al.*, 2017) and the

development of hemp bast fibres (De Pauw *et al.*, 2007; Guerriero *et al.*, 2017).

a. From fibre initiation to fibre maturation

The primary vascular meristem is called procambium and produces primary xylem and phloem (Nieminen *et al.*, 2015). The procambium is located between primary xylem and phloem. Secondary fibres are produced via cambial activity (Figure 1).

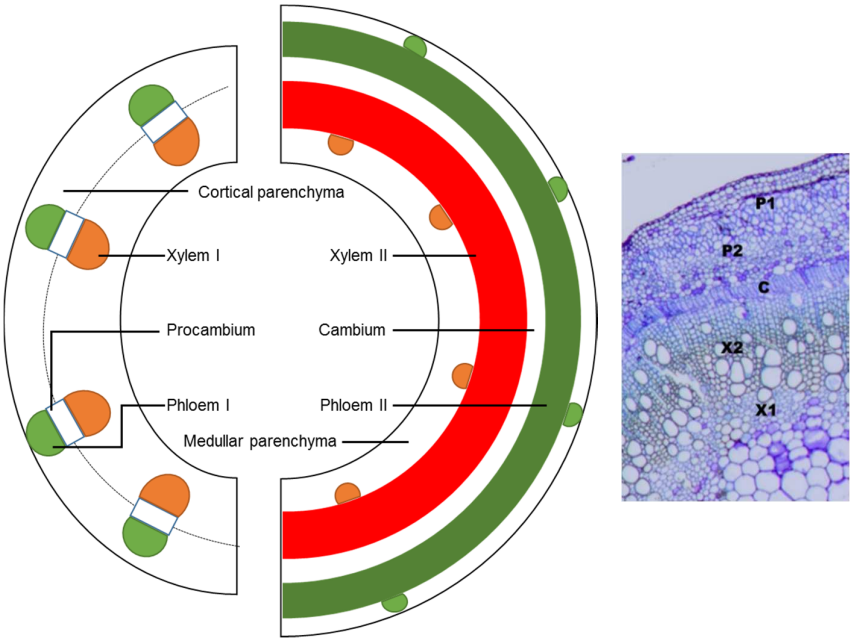


Figure 1: Primary structure (left), secondary structure (middle) of Eudicot stem and cross section of hemp stem (right). C cambium, P1 phloem I, P2 phloem II, X1 xylem I, X2 xylem II.

Bast fibre initiation is regulated by complex interactions between auxins, brassinosteroids, cytokinins, gibberellins, jasmonic acid and ethylene and several transcription factors (reviewed in Gorshkova *et al.*, 2012). Auxins play an important role in vascular

development. During secondary growth, auxins preferentially accumulate in the xylem side of the cambium, in the xylem mother cells. This localisation is mediated by the transcription factor MONOPTEROS and the auxin transporter PIN1. This auxin acts upstream of primary and secondary xylem proliferation. Several transcription factors involved in the biosynthesis of cytokinins are expressed in the cambium and in the phloem, linking cytokinins to secondary phloem development.

Primary fibres first grow symplastically (i.e., at the same rate of longitudinal growth than surrounding cells) and continue to elongate intrusively via diffuse growth (i.e., at a higher rate of longitudinal growth). The primary fibres reach their final length mainly by intrusive growth: at the end of symplastic growth, hemp bast fibres do not exceed 200 µm in length, for an average final size of 17.5 mm (Snegireva *et al.*, 2015). During intrusive growth, the number of fibres observed in a time-course series of stem cross-section results in an increase in the diffusion decay (spin echo NMR), indicating higher density of plasma membranes and cell walls (Figure 2, Snegireva *et al.*, 2006). Secondary fibres originating from the cambium do not elongate symplastically and are shorter than primary ones (7.6 mm in average).

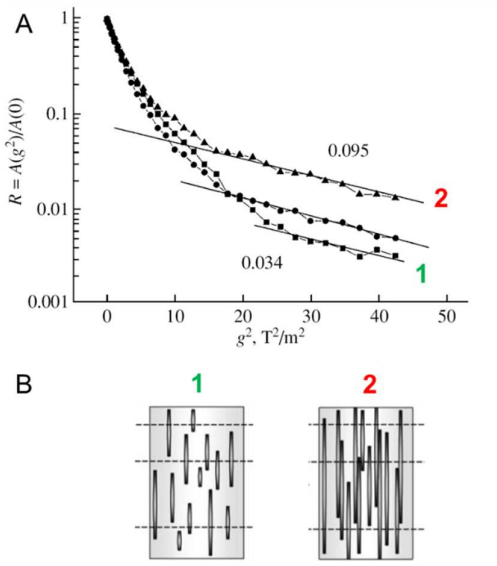


Figure 2: Elongation of flax bast fibre characterised by NMR echo decay.

A: NMR diffusion decay in the stem region above the snap point, 1: immediately after stem cutting; 2: 24 hours after stem cutting. R: echo relative amplitude; g: magnetic field. B: longitudinal view of the stem in time 1 and time 2. Adapted from Snegireva *et al.*, 2006.

After primary and secondary fibres reached their final size, cell wall thickening is initiated. In hemp, primary fibres stop intrusive growth approximatively 6 centimetres below the shoot apical meristem (SAM) and start cell wall thickening, from the periphery to the core (Snegireva *et al.*, 2015). This mostly indicates that bast fibres thickening is regulated at the tissue level, as e.g. xylem cells have already deposited their SCW. Eight centimetres below the SAM, the cambium starts to form secondary xylem. This region may correspond to the empirically determined “snap-point”, which marks the transition between elongation and cell wall thickening (Gorshkova *et al.*, 2003). Cambium starts to form secondary phloem fibres only 60 to 70 centimetres below the SAM (Snegireva *et al.*, 2015). One metre down from the SAM, all

the primary fibres have started their thickening. In bast fibres, cell wall thickening is mainly due to the deposition of the G-layer, consisting in a high amount of cellulose embedded in a rhamnogalacturonan-I (RG-I) matrix and almost completely devoid of lignin.

Bast fibre maturation consists in the crystallisation of the cellulose microfibrils allowed by the enzymatic digestion of the pectic matrix. The two stages of maturation are illustrated in the Figure 3. RG-I composition is highly variable between species, tissues and time and is one of the most complex molecules found in plants. Depending on its molecular structure, several enzymes may partake in its digestion (e.g., β-galactosidase), leading to the compaction of cellulose microfibrils and final organisation of the G-layer, such as hydrolysis of the galactan-rich RG-I in flax bast fibres (Roach *et al.*, 2011).

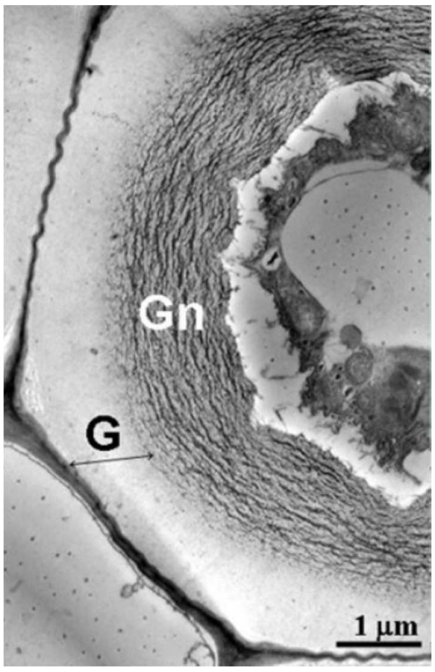


Figure 3: Transition in the maturation of flax bast fibre observed by electron microscopy.

The RG-I matrix of the Gn-layer is progressively transformed in G-layer, resulting in lateral interaction between cellulose microfibrils. Adapted from Mikshina *et al.*, 2013.

b. Cell wall organisation of elongating bast fibre

The primary cell wall (PCW) is the common envelope surrounding and providing their shapes to all plant cells and is therefore a central element for the extreme elongation of hemp bast fibres. Its composition, mainly cellulose (linear (1,4)-linked β-D-glucan), hemicelluloses and pectins (chains of galacturonic acid, glucuronic acid with sugar residues) allows anisotropic cell growth while still maintaining intercellular and external interactions (Keegstra, 2010). The primary wall extensibility relies on the contacts between cellulose, xyloglucan (a cellulose-similar backbone decorated with xylose branches on 3 out of 4 glucose residues with galactose and fucose terminal residues), pectins (RGs, homogalacturonan, xylogalacturonans, arabinan and arabinogalactans) and cell wall proteins. The cell wall can be considered as a natural composite consisting of three networks (Figure 4) : a first one composed of cellulose and hemicellulose; a second one embedding the first one in pectic polysaccharides; and finally a third one consisting in *O*-glycosylated cell wall proteins (arabinogalactan proteins and extensins). These three networks are strongly interconnected (Nguema-Ona *et al.*, 2014).

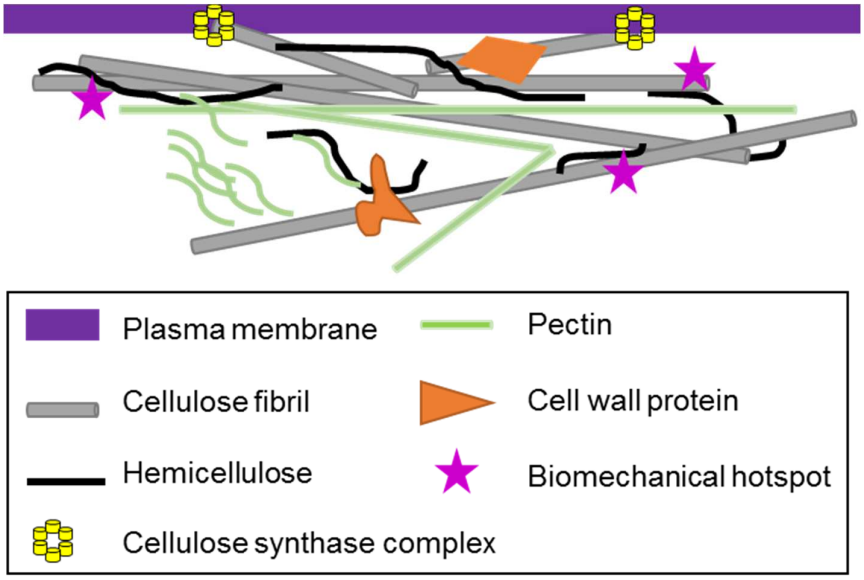


Figure 4: Schematic view of the three networks of the primary cell wall.

The updated model of primary wall growth states that its extensibility is controlled at the precise site where xyloglucan closely interacts with the cellulose microfibril, defined as “biomechanical hotspot” (Park and Cosgrove, 2015). Expansins target a spot consisting of cellulose in close proximity with xyloglucan, a structure remarkably similar to the biomechanical hotspot. Therefore, loosening of this hotspot allows cell extension, at a rate influenced by the bulk viscoelasticity of the matrix network conferred by pectins (Figure 4).

Cellulosic chains are organised in microfibril of various length, area and shape. PCW cellulose displays low crystallinity and high dispersion as a result of the pectin and xyloglucan crosslinks. The cellulose microfibril size’s itself is a source of questioning. Recent experiments using high resolution techniques, such as wide-angle X-ray scattering (WAXS), small-angle neutron scattering (SANS) or solid-state nuclear magnetic resonance (ssNMR), suggest that the microfibril

are most likely made of 18-24 β-(1,4)-glucan chains (Cosgrove, 2014). Interestingly, it has been observed that some microfibril regions are twinned, providing an experimental element in favour of the biomechanical hotspot model.

Xyloglucan is the most abundant hemicellulose in dicot PCWs (ca. 20 % of the dry mass of growing cell walls, Park and Cosgrove, 2015). To follow the cellular growth, the xyloglucan matrix is constantly produced, rearranged and degraded by a large set of enzymes, explaining the variable length and side chain patterns.

The second network of the cell wall is made of pectins of various composition. RG-I are characteristic pectins of PCWs. The plant metabolism is able to adapt the RG-I composition to the development stage of the cell. RG-I in those cells functions as a sort of adhesive between neighbouring cells (Mikshina *et al.*, 2015). The nature and the proportion of RG-I side chains are related to the stage of development: arabinans are typically found in dividing cells while galactans are mainly found during cell elongation. The structure of PCW RG-I is rather stable between species. Pointing to an important role in cell wall mechanics, pectic sugars are close to the cellulose surface (≈ 1 nm), while ca. 50 % of the cellulose surface is in contact with the pectin (Cosgrove, 2014). In flax, the presence of RG-I in developing fibre cell has been detected using monoclonal antibodies specific for several RG-I epitopes: CCRCM2 and anti-polygalacturonic acids / RG-I signals are present in the cell junctions of elongating hypocotyls, suggesting an important role in fibre-cell junctions (Andème-Onzighi *et al.*, 2005, Figure 5).

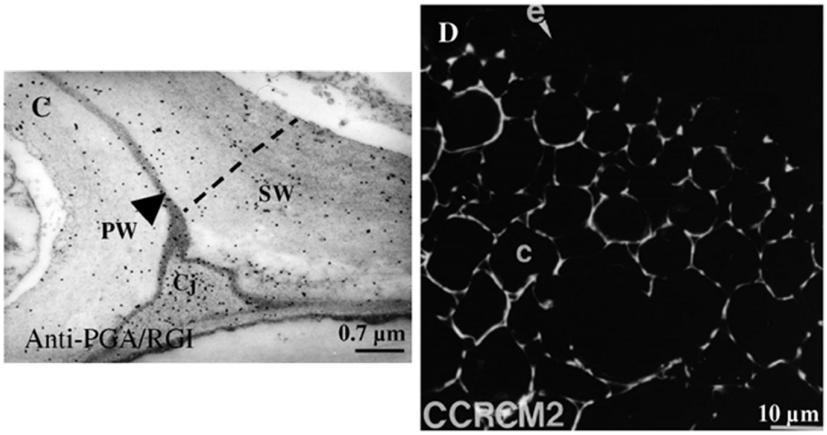


Figure 5: Detection of pectic epitope in flax hypocotyl. Left picture: anti-polygalacturonic acids (PGA) / RG-I is detected in the cell junction (Cj) and PCW of bast fibre. Right picture: RG-I is detected by CCRM2 in the cell junction bast fibre / cortical cells (c). Adapted from Andème-Onzighi *et al.*, 2005.

Homogalacturonans from middle lamella and PCWs of primary phloem fibres of flax have a low degree of methylesterification, resulting from the pectin methylesterase (PME) activity. Galacturonic backbones are secreted into the cell wall in a highly methylated form (Snegireva *et al.*, 2010). Two biochemical options may follow PME activity. First, pectin demethylation could enhance polygalacturonase and pectin-lyase activities, supporting fibre intrusive growth. Alternatively, the demethylated sites may favour the formation of the Ca^{2+} - pectate gel within the “shifted egg boxes” (Braccini and Perez, 2001), therefore decreasing cell extensibility. In hemp, immunohistochemical analysis with antibodies JIM5, JIM7 (partially methylesterified homogalacturonan) and PAM1 (unesterified homogalacturonan) has revealed that homogalacturonans of the stem were largely methylesterified, although only weak JIM5 and JIM7 signals were recovered in developing bast fibres (Blake *et al.*, 2008).

Besides polysaccharides, cell wall (glyco)proteins also have a relevant role in the morphology and the functions of the PCW. More precisely, arabinogalactan proteins (AGPs) and extensins (*O*-hydroxyproline cell wall proteins) were found to have critical roles in cell elongation, root and hypocotyl length or cell plate formation (Nguema-Ona *et al.*, 2014). Both AGP and extensins have been detected in the hemp phloem fibre cells. Anti-AGP probes (LM2, JIM14) display a signal in young stem, inside the developing fibres (Blake *et al.*, 2008), pointing to an important role in cell wall and G-layer deposition. Because of the huge diversity of AGP and extensins, one may presume that they are involved in different process at various stages of development.

c. Cell wall organisation of thickening bast fibre

The transition from bast fibre elongation via intrusive growth to SCW (S1 and S2) and G-layer deposition occurs at the snap point (Gorshkova *et al.*, 2003). This change has important consequences not only for the mechanical properties of the bast fibres, but also for the plant metabolism, because these three new compartments are powerful carbon sink and constitute the bulk of plant biomass. Xylan rich-SCW is typically arranged in three layers (S1, S2 and S3; Figure 6) and is mainly made of cellulose, glucuronoxylan and lignin, with minor amount of pectins. Each layer has a specific angle between cellulose microfibrils and the longitudinal cell axis (Gorshkova *et al.*, 2012). S1 is composed of cellulose microfibrils oriented approximately perpendicular to the cell axis. In S3, these microfibrils are, in average, perpendicular to the cell axis but with higher angular variations than in S1. S2 is the thickest layer and is deposited almost parallel to the cell elongation axis (Zhong and Ye, 2015). The xylan-type SCW is present in bast fibres of jute and kenaf (Gorshkova *et al.*, 2012).

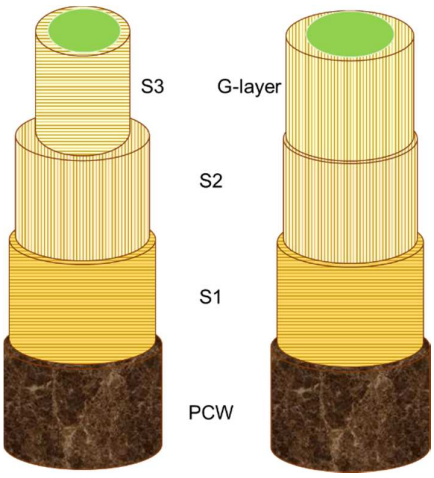


Figure 6: Schematised structure of the cell wall with a xylan-type SCW (S1, S2 and S3, left) and a G-type SCW (S1, S2 and G-layer).

The SCW of the G-layer type is deposited in three layers (Figure 6). S1 and S2 compositions are similar to xylan-type wall. In contrast, the G-layer has a very different organisation (Mikshina *et al.*, 2013): it displays a high content of cellulose (> 75 % and up to 90 %), the cellulose fibrils are axially oriented, it does not contain xylan nor lignin and is of considerable thickness (> 10 µm). The relative number of glucan chains in cellulose fibrils is considered to be four times larger in the G-layer than in the S-layer (Mellerowicz and Gorshkova, 2012). Gelatinous cell walls are built in the bast fibres of flax, hemp and ramie (Gorshkova *et al.*, 2012). This name came from the swelling of this layer during its preparation with alkali for microscopic analysis, but also refers to its mechanical performance and related mesoporosity (Mikshina *et al.*, 2013). The different polysaccharide composition of xylan-type SCW and G-layer SCW can be visualised in Figure 7.

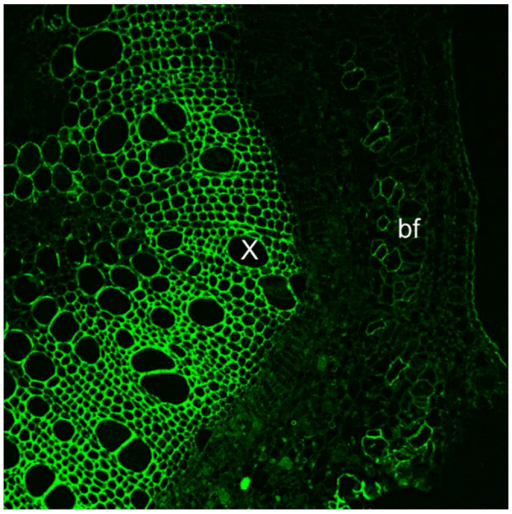


Figure 7: Cross-section of hemp stem stained with LM10 (unsubstituted xylan).

Xylan-type SCW is found in the xylem (X) while G-type SCW is found in the bast fibre (bf).

In hemp mature bast fibres, S1 and S2 have a modest proportion as compared to the G-layer. However, they play an important role in providing strength to the stem. Given that an adult hemp plant may reach 2-3 meters, sclerenchymatic tissues from both xylem and phloem are critical to confer rigidity. Xylan and lignin, the two chief macromolecules of the SCW, can clearly be observed in this tissue (Blake *et al.*, 2008; Crônier *et al.*, 2005).

At the ultrastructural level, xylan and lignin cross-link cellulose microfibrils. Xylan has a much more essential role than xyloglucan in wall assembly, as a reduction in its content severely impairs secondary wall thickening (irregular xylem phenotype, IRX) and plant growth (Persson *et al.*, 2007), while xyloglucan double mutant *xylosyltransferase* 1 (*xt1) xt2* lacked a significant morphological phenotype (Cavalier *et al.*, 2008). However, the molecular mechanisms explaining this important role are still unresolved. The presence of decorations [(4-*O*-methyl) glucuronic acid, acetylation, and methylation] is supposed to play an important role in this respect. Xylan decoration may regulate interactions between xylan, lignin and

cellulose (Busse-Wicher *et al.*, 2016; Martínez-Abad *et al.*, 2017). Low or unsubstituted forms of xylan favours the formation of the cellulose microfibril matrix (Donaldson and Knox, 2012; Ruel *et al.*, 2001). Another level of xylan reactivity is provided by modifications occurring in the cell wall under the action of endo-1,4-β-xylanases, xylan endotransglycosylases, 1,4-β-xylosidases and bifunctional α-arabinofuranosidases/β-xylosidases. Such *in muro* enzymatic activities may be considered as a way for the plant to adapt the organisation of the cell wall to growth requirements. Hemp primary and secondary bast fibres show contrasted lignin and xylan amounts which may be explained by the histological origin of those two tissues (i.e. procambium and cambium).

Lignin provides additional mechanical strength by impregnating the cellulose-xylan network. This microfibrillar template is characterised by a restricted volume (ca. 30Å) which can be filled by other molecules (Ruel *et al.*, 2001). Such an organisation impacts the molecular structure of the lignin impregnating this network. Lignin is a polymer mainly made of three different monomeric blocks: *p*-coumaryl (H), guaiacyl (G) and syringyl (S) units. H units are capable of forming condensed units at the 3 and 5 positions (Terashima and Fukushima, 1988), resulting in a bulk and unordered structure that fits in the middle lamella / PCW. By contrast, S and G units display a non-condensed structure with an extended and ordered conformation and are deposited after H units. Subsequently, the lignin with a high proportion of G and S units is able to impregnate the xylan-cellulose matrix of the SCW, leading to the formation of the lignin-carbohydrate complex (LCC, Barakat *et al.*, 2007). This process is described in the Figure 8.

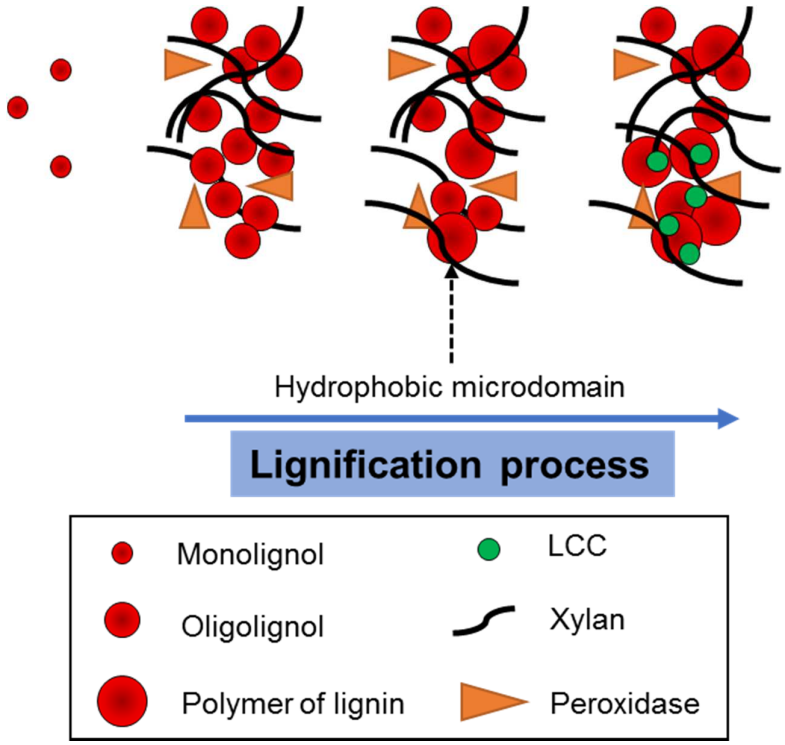


Figure 8: Formation of the lignin carbohydrate complex in SCW. Monolignols diffuse into the cell wall and are polymerised. Lignin then aggregates with xylan by covalent interaction, forming hydrophobic domains. Finally, xylan reacts with the transient quinone methide intermediate to form the LCC.

Oinonen and colleagues (2015) suggest that hemicelluloses with phenolic moieties are first incorporated into the lignification initiation site. They are then cross-linked first by excreted laccases and, later on, monolignols are oxidatively coupled to this matrix by class III peroxidases, forming the LCC. Xylan is therefore a key polymer regulating both the quantity and the structure of the lignin present in the SCW.

The interest of hemp as plant fibre relies on the presence of a thick G-layer whose high crystalline cellulose content and axial

orientation of cellulose microfibrils are the main features. It has to be noted that this specific cell wall shares some similarities with the tension wood of specific woody species like poplar. In trees, tension wood supports the bending of stems (branches and trunk) when exposed to gravitational force (Mellerowicz and Gorshkova, 2012). The cells of this tension wood exhibit a thick G-layer almost depleted in hemicellulose and lignin but with a significant proportion of non-cellulosic matrix polysaccharides such as xyloglucan and RG-I (ca. 20 %; Mellerowicz and Gorshkova, 2012). Gelatinous fibres can occur in both xylem and phloem, or solely in phloem, as in hemp or flax. Since the G-layers of woody species are easier to separate than extraxylary gelatinous fibres of fibre crop plants, they have been extensively studied to understand the basics of their formation (reviewed by Groover, 2016). The proportion of the non-cellulosic polysaccharides found in the G-layer differs between species: xyloglucan and pectic galactan (RG-I) are, respectively, the major forms in poplar and flax. In trees, xyloglucan is deposited in the developing G-layer and is trapped inside the cellulose crystalline domain to ensure mechanical strength, and at the interface between the G- and S2-layers to transfer the tensile stress to the G-fibres and ultimately to the wood. In aspen tension wood, β-(1→4)-galactan and RG-I backbone have been isolated in all polysaccharide fractions (ammonium oxalate, 4M KOH and 2M trifluoroacetic acid), indicating a localisation even within the cellulose microfibrils of the G-layer (Gorshkova *et al.*, 2015). In fibre crops, the exact structure of the G-layer pectic galactan *in muro* is still unknown because of the complexity of the RG-I molecule. By analysing the soluble sugars trafficking by the Golgi vesicles at the time corresponding to G-layer deposition, it is possible to infer the composition of this pectic galactan. The Golgi apparatus is indeed the chief cellular compartment determining the non-cellulosic composition

of the cell wall (Driouich *et al.*, 2012; van de Meene *et al.,* 2017). In flax, as assessed in the Golgi vesicles, the major sugar of the G-layer pectin is galactose (55-85 %), followed by rhamnose and galacturonic acid (Gorshkova and Morvan, 2006). This polymer is characterised by a high degree of rhamnose branching and side chains of β-(1 → 4)-linked galactose of variable length and Gal:Rha ratio. This side chain is detected by the LM5 antibody (specific for relatively long 1 → 4 – β-galactosyl residues, DP ≥ 4), which displays a strong signal in the flax galactan-layer (Gn-layer) (Andème-Onzighi *et al.*, 2000; Gorshkova and Morvan, 2006). The LM5 signal can be observed in Figure 9.

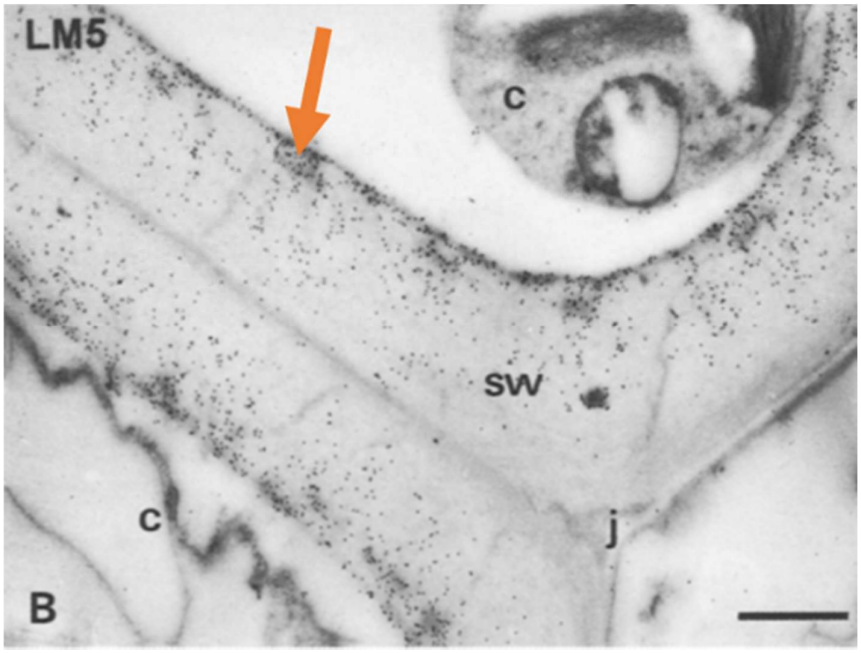


Figure 9: Immunogold labelling of flax fibre cell walls with LM5. Note the intense signal in the inner part of the G-layer (orange arrow). c cytosol, j cell junction, sw secondary cell wall. Image adapted from Andème-Onzighi *et al.* (2000).

As the Gn-layer matures, the LM5 signal decreases in intensity, pointing to a remodelling of the pectic galactan, resulting in the transition to the G-layer (His *et al.*, 2001; Gorshkova and Morvan, 2006; Roach *et al.*, 2011; Hobson and Deyholos, 2013). The cell wall monosaccharide analysis of isolated bast fibres from flax hypocotyls suggests that pectins enrobing cellulose microfibrils are half PGA and half RG-I, with an overall PGA to rhamnose ratio around 6 (Andème-Onzighi *et al.*, 2005). To date, the nature of the complex polysaccharide matrix surrounding crystalline cellulose microfibrils (RG-I, RG-II, PGA) has not been elucidated in hemp, although similar soluble galactans have been detected (Chernova and Gorshkova, unpublished results). The LM5 signal found in flax is not observed in hemp (Blake *et al.*, 2008; Behr *et al.*, 2016), thus the structure of the main non-cellulosic polymer from the Gn-layer is likely not the same. The LM6 antibody, recognising 1,5-arabinan epitope, is weakly present in the inner region of the hemp mature G-layer (Blake *et al.*, 2008).

3. Molecular players involved in SCW formation and G-layer thickening

a. Transcription factors

The biogenesis of the SCW and G-layer is a highly compartmented process. Because of this compartmentalisation, and because SCW and G-layer have to be deposited at a precise time, i.e. at the end of cell elongation, the transcriptional regulation of the multiple genes involved in these processes needs to be finely tuned. This coordination is ensured by a specific class of proteins, the transcription factors (TFs). The timing of SCW deposition is tissue-dependent, thus implying

that each tissue has its own gene regulation. In the xylem conducting cells, important transcription factors regulating SCW biosynthesis are members of the VASCULAR-RELATED NAC-DOMAINS (VNDs) family (Zhou *et al.*, 2014). VND1 to VND5 are transcriptional regulators of secondary wall biosynthesis in the vessels of the stem and roots (Zhou *et al.*, 2014); VND6 and VND7 are transcriptional switches for root and shoot metaxylem and protoxylem vessels, respectively (Kubo *et al.*, 2005; Yamaguchi *et al.*, 2011). In xylem fibres, SCW deposition is redundantly activated by SECONDARY WALL ASSOCIATED NAC DOMAIN PROTEIN 1 (SND1) and NAC SECONDARY WALL THICKENING PROMOTING FACTOR 1 (NST1) (Hussey *et al.*, 2013). VND6, VND7, NST1 and SND1 are referred to as Tier 3 transcription factors because they are able to induce the expression of two supplemental levels of transcription factors (Tiers 1 and 2, illustrated in Figure 10). Overall, this transcriptional cascade regulates the expression of structural genes involved in cellulose, hemicellulose and lignin biosyntheses. Notably, lignification of the SCW is regulated by three Tier 1 MYB proteins, MYB58, MYB63 and MYB85 (Nakano *et al.*, 2015).

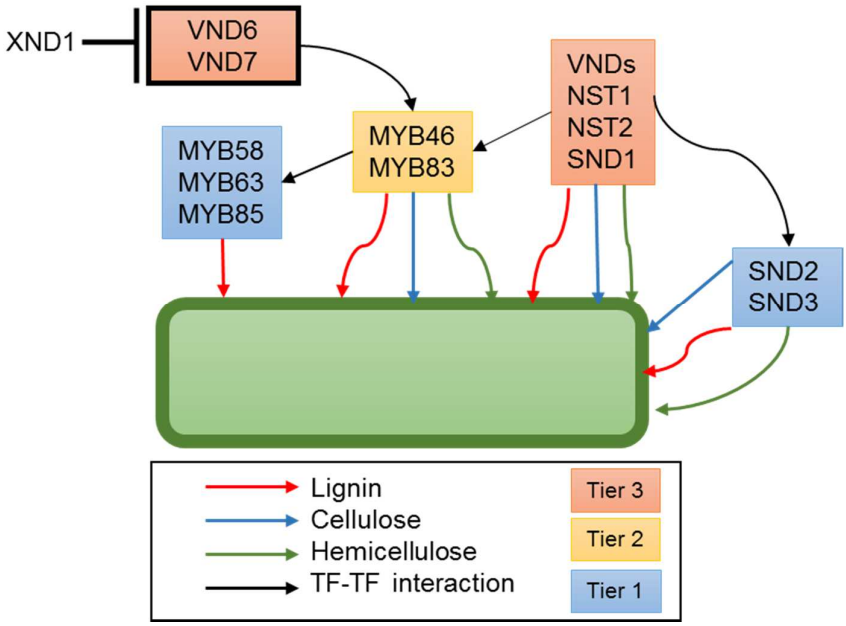


Figure 10: Simplified overview of the transcriptional regulation of SCW. The TF interactions are shown with black arrows. For simplicity, many Tier 2 and Tier 1 level TFs are not shown.

Given the different composition of SCW and G-layer, the transcriptional biosynthetic regulation of these two cell walls is largely different. RNA-Seq analysis of developing flax stem points to this contrast (Gorshkov *et al.*, 2017). Importantly, several transcription factors involved in SCW deposition are strongly down-regulated in the isolated phloem fibres. Among them are the orthologs of *NST1*, *SND1*, *SND2*, *MYB46* and *MYB63*. By contrast, transcription factors from *G2-like*, *bZIP*, *bHLH*, and other *MYB*-related families were up-regulated in the phloem fibres. The function of those transcription factors remains to be determined (Gorshkov *et al.*, 2017).

b. Cellulose deposition

Cellulose deposition in SCW is governed by a whole set of proteins, assembled within the cellulose synthase complex (CSC). The CSC is thought to be assembled in the Golgi, secreted through the *trans* Golgi network and located at the plasma membrane by the actin network (Bashline *et al.*, 2014). The polymerisation of the (1,4)-linked β-D-glucan residues is performed by a set of three cellulose synthase, namely CESA4, CESA7 and CESA8 (Hill *et al.*, 2014). The SCW CSC is likely made from a hexamer of CESA trimers, with an equimolar stoichiometry for each CESA (Hill *et al.*, 2014; Figure 11).

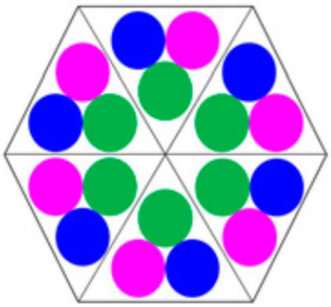


Figure 11: Model of the rosette with hexamer of trimers, with a 1:1:1 stoichiometry between CESA4, CESA7 and CESA8 (one colour for each CESA). Hill *et al.*, 2014.

Other genes have been implicated in cellulose deposition and crystallization. Sucrose synthase provides the UDP-glucose to CESA and is therefore a central element in cellulose synthesis (Fujii *et al.*, 2010; Brill *et al.*, 2011). COBRA-like proteins (COBL) are implicated in SCW thickening and cellulose deposition (Zhong and Ye, 2015) while CHITINASE-LIKE 2 (CTL2) likely play a key role in cellulose structure and assembly of cellulose microfibrils in the SCW, based on its co-expression pattern with secondary wall *CesA*s (Sánchez-Rodríguez *et al.*, 2012). KORRIGAN (KOR1) interacts with CESA4 and CESA8 and is directly involved in cellulose biosynthesis (Mansoori *et al.*, 2014). Several proteins, including those containing a fasciclin-like domain, have also been implicated in this process: in eucalypt trees, fasciclin-

like arabinogalactan protein (FLA) EgrFLA2 is involved in cellulose deposition, while an insertion mutant of *atfla11* displayed an irregular xylem phenotype and an alteration in cellulose content (Persson *et al.*, 2005; Ito *et al.*, 2005). The expression of *LusFLA11* is enriched during phloem fibre development. A role for FLA11 may thus be invoked for the deposition of the G-layer (Gorshkova *et al.*, 2012).

c. Xylan biosynthesis

Xylan is the main hemicellulose of the SCW. It is synthesized in the Golgi and secreted via vesicles into the wall. It has a linear backbone composed of (1 → 4)-linked β-D-xylosyl residues which may be substituted with glucuronic acid and acetylated (Hao and Mohnen, 2014). In Dicots, glucuronoxylan is the most abundant form of xylan. A specific oligosaccharide, named xylan reducing-end sequence (XRES), is supposed to function as an initiator or terminator of xylan chain synthesis (Hao and Mohnen, 2014). Two glucuronoxylan biosynthesis models have been proposed based on this XRES assumption (Hao and Mohnen, 2014). In the first model, XRES acts as a primer and chain elongation occurs via the transfer of D-xylose from UDP-xylose onto the non-reducing end of XRES. In the second model, XRES is *de novo* synthesized. The xylan backbone is elongated by the reducing-end and XRES is transferred at the end of the elongation process when the backbone has reached a certain length. These two mechanisms are not exclusive in different species, tissues or cell type (Hao and Mohnen, 2014). To date, all the enzymes from the xylan biosynthetic pathway have not yet been functionally characterised, nor have their substrates precisely been defined. Notably, the enzymes involved in the XRES synthesis are partially characterised (e.g., FRAGILE FIBRE 8 FRA8, PARVUS, IRX7 and IRX8; Hao and Mohnen, 2014). The GLUCURONIC ACID SUBSTITION OF XYLAN (GUX1 to

GUX5) enzymes (CAZy family GT8) are involved in adding the glucuronic acid side chains. The *REDUCED WALL ACETYLATION* genes (*RWA1* to *RWA4*) redundantly participate in the acetylation of xylan during secondary wall biosynthesis and are regulated by the transcription factor SND1 (Lee *et al.*, 2011). Acetylated polysaccharides have reduced solubility and thus are more easily deposited and cross-linked with cellulose (Lee *et al.*, 2011). An overview of some enzymes hypothetically involved in glucuronoxylan biosynthesis is shown in Figure 12.

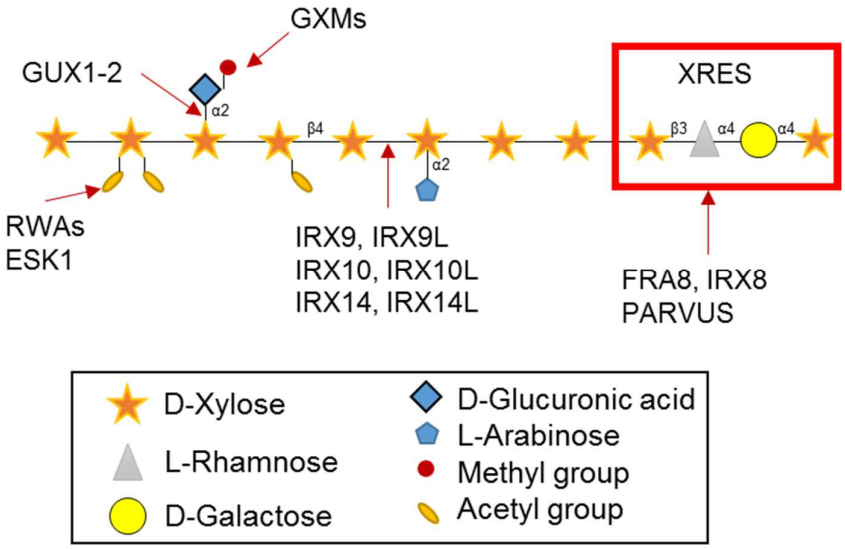


Figure 12: Typical structure of glucuronoxylan and enzymes involved in its biosynthesis (adapted from Kumar *et al.*, 2015). Abbreviations are as in the text. ESK1 ESKIMO 1 (Xiong *et al.*, 2013), GXM GLUCURONOXYLAN METHYLTRANSFERASE (Hao and Mohnen, 2014).

d. Monolignol and lignin polymerisation

Lignin biosynthesis is performed in two main steps taking place in the cytoplasm and in the apoplast. The production of the monolignols is

located both in the cytosol and in endoplasmic reticulum (ER)-based structures (Bassard *et al.*, 2012). After excretion to the apoplast, monolignols or oligolignols are oxidatively polymerised by laccases and class III peroxidases. The entry molecule for monolignol production is phenylalanine, which is a key aromatic aminoacid for secondary metabolism (e.g., flavonoids biosynthesis). Phenylalanine (Phe) is the first molecule of the phenylpropanoid pathway which leads to the production of *p*-coumaroyl-CoA by a suite of 3 reactions, catalysed by Phenylalanine ammonia lyase (PAL), Cinnamate-4-hydroxlase (C4H, ER-based) and 4-coumaric acid-CoA ligase (4CL). In Arabidopsis, PAL1 and PAL2 are the primary isoforms (Fraser and Chapple, 2011). 4CL1 is likely the best candidate to drive the flux of *p*-coumaric acid to the monolignol pathway because of its high expression level in the stem and root (Li *et al.*, 2015). *p*-coumaroyl CoA is either directed to the biosynthesis of *p*-coumaryl alcohol (H unit) via (hydroxy)cinnamoyl-CoA reductase (CCR) and (hydroxy)cinnamyl alcohol dehydrogenase (CAD) activities, or to the methylated monolignol coniferyl alcohol (G unit) and sinapyl alcohol (S unit). These methylated monolignols require the activities of the following enzymes: hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl-transferase (HCT, partially ER-based), *p*-coumaroyl shikimate 3' hydroxylase (C3'H, ER-based), Caffeoyl-CoA 3-*O*-methyltransferase (CCoAOMT), Ferulate 5-hydroxylase (F5H), Caffeic acid/5-hydroxyferulic acid *O*-methyltransferase (COMT) and CAD (Bonawitz and Chapple, 2010; Fraser and Chapple, 2011; Bassard *et al.*, 2012). The monolignol ability to diffuse from cell to cell is a crucial aspect of lignification. Indeed, unlike sclerenchyma cells, xylem vessels rapidly undergo programmed cell death and rely on the production of monolignols from the surrounding parenchyma or cambial cells to complete their lignification (Barros *et al.*, 2015). This non-cell

autonomous lignification is known as the “good neighbour” hypothesis (Smith *et al.*, 2013; Voxeur *et al.*, 2015). The good neighbours produce and export the monolignol to lignifying cells. This specificity may be regulated at the transcriptional level by different signatures of transcription factors: in the good neighbour cells, only the genes involved in monolignol biosynthesis have to be activated, while enzymes polymerising the monolignols in the apoplast have to be specifically present in the lignifying cells. Several MYB transcription factors (e.g., MYB61 in protoxylem and MYB63 in the fibres) are likely regulating this cell-dependent program (Voxeur *et al.*, 2015).

After monolignols (RH) are excreted to the apoplast, they are oxidatively polymerised by phenol-oxidoreductase enzymes laccases and class III peroxidases. Laccases and peroxidases use O_2 and H_2O_2 as oxidant, respectively, to produce the phenoxy radical forms ($R^{\bullet}$) of *p*-coumaryl, coniferyl and sinapyl alcohols. These radicals react with the free-phenolic end of the growing lignin polymer (Ralph *et al.*, 2008). H_2O_2 synthesis may be regulated by cell wall-associated peroxidase or by plasma membrane-localised NADPH oxidase. Both enzymes yield $O_2^{\bullet-}$, a strong oxidant, which is stabilised by superoxide dismutase under the form of H_2O_2 for subsequent monolignol phenoxy radical synthesis (Novo-Uzal *et al.*, 2013).

In *Arabidopsis*, 17 laccases and 73 class III peroxidases have been found (Cosio and Dunand, 2009; Turlapati *et al.*, 2011). The great number of class III peroxidases may indicate a functional specialization of these enzymes (e.g., different substrates). However, the key amino acid residues and protein size are rather constant within one species and in distant plant families (Cosio and Dunand, 2009). The isoelectric point (pI) may be another variable to characterise this family. Most acidic peroxidases dimerize *p*-coumaryl alcohol and coniferyl alcohol

while most basic peroxidases may oxidise the three monolignol units (Herrero *et al.*, 2013). The *Arabidopsis* putative homologues of the poplar cationic peroxidase AtPRX2, AtPRX25 and AtPRX71 are involved in lignin polymerisation. Mutants deficient in one of those three peroxidases have either reduced lignin content or altered S/G ratio (Shigeto *et al.*, 2013). Knock-out mutants of cationic *atprx4*, *atprx52* and *atprx72* show the same altered lignification pattern, with reduced lignin content and lower S/G ratio (Herrero *et al.*, 2013; Fernández-Pérez *et al.*, 2014; Fernández-Pérez *et al.*, 2015). As AtPRX52 and AtPRX72 mutants have lower S content in the late stage of lignification characteristic of the mature stem, they are clearly involved in the biosynthesis of syringyl monomers. The anionic peroxidases AtPRX37 and AtPRX53 might be involved in cell wall phenolic cross-linking and lignin deposition (Østergaard *et al.*, 2000; Pedreira *et al.*, 2011). By contrast with AtPRX52 and AtPRX72, AtPRX53 prefers *p*-coumaryl alcohol and coniferyl alcohol as substrates and is unable to oxidise sinapyl alcohol (Østergaard *et al.*, 2000). Similarly to peroxidases, laccases have a high degree of conservation between species. The assumption that laccases are involved in lignification is based on their co-expression with genes from the monolignol pathway and their activities in cells undergoing lignin deposition (Barros *et al.*, 2015). To date, 5 *Arabidopsis* laccases have been identified to participate in lignification: AtLAC4, AtLAC11, AtLAC12, AtLAC15 and AtLAC17 (Liang *et al.*, 2006; Berthet *et al.*, 2012; Zhao *et al.*, 2013; Schuetz *et al.*, 2014). AtLAC4, AtLAC11 and AtLAC17 are responsible for the lignification of xylem cells, as knock-out mutants of two of those genes results in decreased G content (in both vessels and fibres; Berthet *et al.*, 2011; Zhao *et al.*, 2013). The triple mutant stops its development after the appearance of the two first pairs of true leaves (Zhao *et al.*, 2013). Laccase-driven lignification may also be regulated post-

transcriptionally by specific microRNAs. In flax and poplar, *miR397* is able to cleave several laccase transcripts, such as *LusLAC1*, *LusLAC5*, *LusLAC9*, *LusLAC38*, *LusLAC39* and *LusLAC44* (Le Roy *et al., 2017)* and 29 poplar laccases, 5 being homologs of *AtLAC4* and 6 being homologs of *AtLAC17 (*Barros *et al.*, 2015). More generally, miRNA may be an important element regulating the lignification both temporally and spatially, in the light of the model of laccase and peroxidase sequential actions described above.

e. Gelatinous layer

In hemp and flax, the G-layer is almost completely depleted in lignin and xylan. Instead, a RG-I matrix is surrounding the cellulose microfibrils. Xylan-type SCW and G-layer are thus present in the stem. Flax has been used as a model plant to study this special feature. Few molecular factors have emerged to explain the different properties of xylan-type and G-layer cell walls. Both primary and SCW-related *CesA*s are expressed in the flax fibres depositing the G-layer (Chantreau *et al.*, 2015; Mokshina *et al.*, 2017). *LusCesA3B*, *LusCesA3C* and *LusCesA8B* are the most highly coexpressed genes within a network of mixed primary and secondary CesAs (Mokshina *et al.*, 2017). As CSC with mixed primary and secondary CESAs does not occur naturally (Carroll *et al.*, 2012), the presence of two functional CSCs in fibres deposing the G-layer may reasonably be foreseen (Song *et al.*, 2010). Chitinases and CTL proteins seem to be involved in bast fibre biogenesis: a set of three CTL genes (*LusCTL19*, *LusCTL20* and *LusCTL21*) is specifically expressed in phloem fibres (Mokshina *et al.*, 2014). CTLs have been associated with SCW deposition and cellulose assembly in *Arabidopsis* (Sánchez-Rodríguez *et al.*, 2012). A *FLA* gene is expressed in flax maturating bast fibre, *LusFLA11*, but its role in particular, and all other FLA in general,

remains to be determined (Hobson and Deyholos, 2013). Considering the pectic matrix, RG-I is secreted during gelatinous cell wall formation (Gorshkova and Morvan, 2006). Several genes involved in RG-I formation and degradation have been described in flax (Gorshkov *et al.*, 2017). Glycosyl transferases (GT2, GT47, GT92) are the best candidates for RG-I synthesis because of their activities and high expression in bast fibre tissues. In *Arabidopsis*, the GT92 GALACTAN SYNTHASE 1, (GALS1), GALS2 and GALS3 synthesise β-1,4-galactan and GALS1 specifically forms β-1,4-galactosyl linkages (Liwanag *et al.*, 2012). RG lyase cleaves the α-(1→4)-glycoside bonds between L-rhamnose and D-galacturonic acids in the RG-I backbone and its coding gene is up-regulated in the mature fibres and at the snap point, as compared to the apical part of the stem (Gorshkov *et al.*, 2017). RG-I may also be degraded by beta-galactosidase activity, a family with 43 different genes in flax (Hobson and Deyholos, 2013). Specific beta-galactosidases are expressed in the flax bast fibre. *LusBGAL3* is highly expressed in the phloem rich outer tissue of the upper stem, while *LusBGAL22* is more expressed in the same tissue in the lower stem (Hobson and Deyholos, 2013). A microarray analysis of flax developing stem highlighted the up-regulation of three *BGAL* genes at the snap point (Roach and Deyholos, 2007). Notably, one of those three genes (probe set 4738, renamed *LusBGAL1*) has been functionally characterised by RNA interference. This study came to the conclusion that *LusBGAL1* structurally and chemically modified the Gn-layer into the G-layer by degrading the galactan bound to RG-I (Roach *et al.*, 2011). In addition, *LusBGAL1* is significantly more expressed in the mature fibres and at the snap point, as compared to the apical part of the stem (Gorshkov *et al.*, 2017). An illustration of the action of the beta-galactosidase is provided in Figure 13.

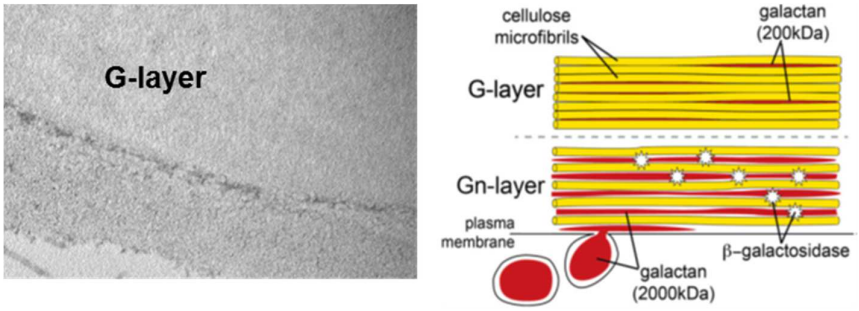


Figure 13: Maturation of the G-layer in bast fibre. Left: TEM image of hemp bast fibre during the G-layer maturation. Right: Model proposed by Roach *et al.* (2011). A galactan moiety of the RG-I is digested by LusBGAL1, leading to the association of cellulose microfibrils in a highly crystalline matrix.

PME and PMEI are also differentially expressed according to the developmental stage of the flax bast fibres. *LusPME1* and *LusPME61* have a fold change > 20 between thickening bast fibres located below the snap point and elongating bast fibres. In this context, PME activity may strengthen the cell wall after the end of cell elongation (Pinzon-Latorre and Deyholos, 2014). Considering that flax RG-I has rhamnose and galacturonic acid in its backbone and side chains with (1→4)-β-D-galactan (Makshakova *et al.*, 2017), a mechanism associating both PME/PMEI and beta-galactosidase may be supposed for the maturation of the bast fibre. Overall, the deposition and maturation of the Gn/G-layer is regulated by a plethora of genes, often belonging to large multi-genes families such as XET / XTH, AGP and FLA. Apart from the canonical cell wall-related genes, other families seem to be involved in G-layer deposition. Metallothioneins and cell wall-located lipid transfer proteins are highly expressed in phloem tissue of flax (Gorshkov *et al.*, 2017) and hemp (De Pauw *et al.*, 2007). Metallothioneins may be involved in the regulation of the CSC formation by controlling the rate of SCW rosette / CESA turnover (Jacob-Wilk *et al.*, 2006), while lipid transfer proteins may be involved

in cellulose deposition in elongating fibres by mediating the transport of phosphatidylinositol monophosphates. (Deng *et al.*, 2016).

4. Roles of the phytohormones in cell wall biogenesis

Plant growth and development are regulated by different classes of phytohormones. Phytohormones are signal molecules which act as messengers between distant organs. The biosynthesis of a specific phytohormone is regulated by both environmental and developmental signals, such as abiotic stress (e.g. drought, osmotic stress), growth condition (light / shade), embryogenesis, flowering or sink / source strengths (Crozier *et al.*, 2000). Phytohormones are active through transcriptional regulation of many biological processes, from tissue elongation to secondary growth and fruit development. For instance, stem elongation implies, at the cellular level, the biosynthesis of the plasma membrane, the deposition and remodelling of the PCW and the establishment of a turgor pressure driving cellular elongation. We will hereafter describe the roles of selected phytohormones within this elaborate framework.

a. Auxins

Auxins have been largely acknowledged to play a significant role in cell division, cell elongation, vascular differentiation, lignification and tropic response to light (Hayashi, 2012). Several types of auxin response are described (Hayashi, 2012). In the first one, auxins activate the auxin-responsive genes by promoting the degradation of the Aux / IAA repressor and removing the inhibition of auxin response factors (ARF) activity. In this pathway, auxin binds to the nuclear proteins TRANSPORT INHIBITOR RESPONSE 1 (TIR1) and AUXIN

SIGNALING F-BOX (AFB). TIR1 / AFB form an SCF E3 ubiquitin ligase complex between SKP and CULLIN1. This activated complex, so called $SCF^{TIR1/AFB}$, mediates the degradation of the Aux / IAA repressors by uquitination and induce the auxin response. In the second response, auxins induce the *SMALL AUXIN-UP RNA* (*SAUR*) genes. The third auxin response invokes the *GRETCHEN HAGEN 3 (GH3*) family, responsible for the conjugation of IAA with an amino acid to deactivate this form of auxin. Another transcriptional regulation is also present, based on AUXIN-BINDING PROTEIN 1 (ABP1). ABP1 is present at the cell and ER surfaces and is involved in the fast regulation of membrane potential and ion fluxes at the plasma membrane in response to auxin. This protein is involved in the auxin-induced hypocotyl elongation (Hayashi, 2012). Importantly, ABP1 negatively regulates the $SCF^{TIR1/AFB}$ pathway by controlling the Aux / IAA stability (Tromas *et al.*, 2013). As such, the positive and negative regulations of the SCF – ubiquitin pathway constitute a tight regulation of the auxin-mediated response. Functional inactivation of ABP1 strongly impacts the expression of genes involved in cell wall remodelling (XTHs, expansins, glycosyl hydrolases, peroxidases) but affects to a lower extent the genes involved in cell wall biosynthesis (Paque *et al.*, 2014). ABP1 inactivation acts on the xyloglucan structure, resulting in higher rate of *O*-acetylation and fucosylation and ultimately in reduced cell expansion in the hypocotyl. The authors therefore suggest that auxin mediates cell expansion (partially) via fucosidase activity mediated by ABP1 activity.

Auxin has also been the central element of the acid-growth theory, where cell wall acidification is mediated by the proton extrusion via an H^+-ATPase transporter. Lower apoplastic pH increases the activity of expansins and causes plasma membrane hyperpolarization,

leading to solutes and water uptake and finally increased turgor pressure. Acid growth is also regulated by the $SCF^{TIR1/AFB}$ pathway, but a knock-out *abp1* mutant shows no difference in auxin-induced hypocotyl elongation (Fendrych *et al.*, 2016).

SAUR19 is another factor regulating the H^+-ATPase activity – the stabilisation of this protein induces a higher plasma membrane H^+-ATPase activity and increased cell elongation (Fendrych *et al.*, 2016). Therefore, the auxin-mediated apoplastic acidification seems to be regulated by the phosphorylation of the H^+-ATPase protein, the abundance of SAUR19 and the $SCF^{TIR1/AFB}$ – Aux / IAA signalling pathway.

Auxins are also involved in the control of xylogenesis and lignin biosynthesis and deposition (Schuetz *et al.*, 2013). Several lines of evidence from phenotypical to molecular observations strongly suggest an important role of auxin in the regulation of the genes involved in lignification, as well as in the resulting lignin content and monomeric composition. By applying IAA alone or in combination with gibberellic acid, Aloni and colleagues (1990) highlighted the dual impact of these two plant growth regulators on the pattern of lignification of xylem and phloem fibres. Indeed, auxin sprayed alone or with a low proportion of GA resulted in an increase in the S/G ratio in the phloem fibres, while this ratio decreased in xylem. The lignin content of the phloem fibres was also significantly lower, because auxin promotes the differentiation of short primary phloem fibres with thick secondary walls (Aloni, 1979; Aloni *et al.*, 1990). The hormonal regulation of the genes involved in lignification is so far little understood (Didi *et al.*, 2015). The gain of function mutant *IAA3/SHY2-2*, causing auxin insensitivity by inactivating multiple *ARF* genes, showed contrasted expression of genes involved in lignification (Oh *et al.*,

2014). Assigning a precise regulatory function to a single phytohormone is therefore very complex and does not reflect the complexity of the multi-phytohormone response that occurs *in planta*. For instance, cell expansion is controlled through multiple interactions between auxin, brassinosteroids, gibberellin and external stimuli (light and temperature). In this circuit, ARF6, the light / temperature-regulated transcription factor PHYTOCHROME-INTERACTING FACTOR4 (PIF4) and the brassinosteroid-regulated transcription factor BRASSINAZOLE RESISTANT1 (BZR1) cooperatively regulate common target genes, with the gibberellin-inactivated repressor DELLA acting as an antagonist (Oh *et al.*, 2014), as shown in Figure 14.

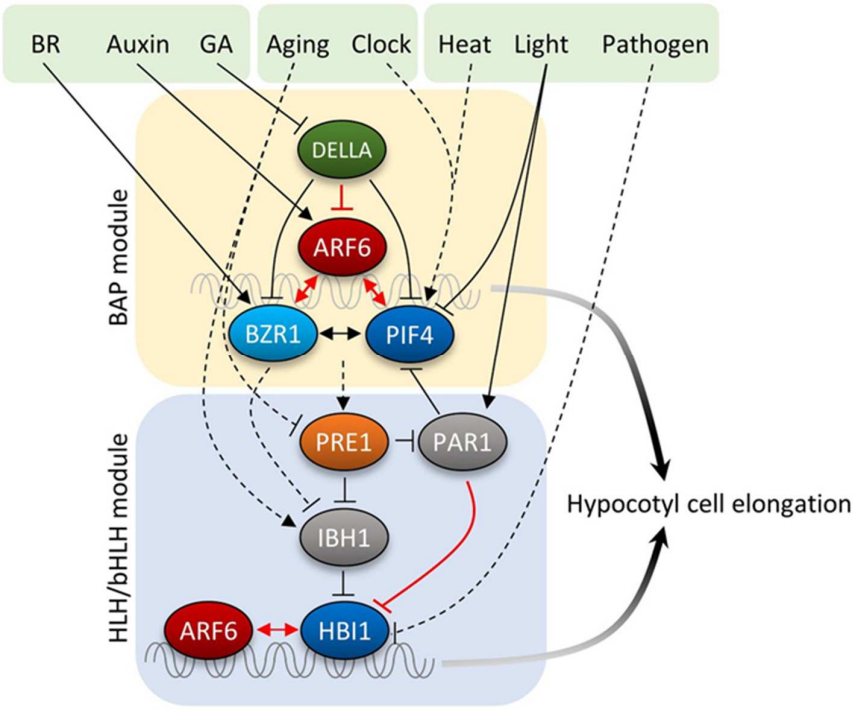


Figure 14: Diagram of the regulation of hypocotyl cell elongation. Solid lines indicate protein–protein interaction or post-translational modification, and

dashed lines indicate transcriptional regulation. BR, auxin and GA cooperatively regulate gene expression through the BZR1-ARF6-PIF4 (BAP) DELLA module. The BAP module positively regulates the expression of *PRE1* in a feedback loop. HBI1 Homolog of BEE1 interacting with IBH1, PAR1 Phytochrome rapidly regulated 1, PRE1 Paclobutrazol resistance 1. From Oh *et al.*, 2014.

b. Brassinosteroids

Brassinosteroids (BR) interact synergistically with auxins to regulate cell expansion and xylogenesis. BR activate BZR1 and BZR2, which bind to the promoter region of the BR-regulated genes in the nucleus (Zhu *et al.*, 2013). BZR1 and BZR2 regulate approximatively 1000 genes and notably several involved in cell wall biogenesis and modification. Among them are several expansins, XTHs, PMEs and cellulose synthases (Didi *et al.*, 2015). A significant overlap with auxin-induced genes is observed. Rice mutants in either signalling or biosynthetic BR pathways have a reduced pool of tonoplast H^+-ATPase, therefore causing perturbation in the vacuole- and Golgi-dependent secretory pathways and reduced cell elongation as a result of the decreased polysaccharide transport to the cell wall (Kutschera and Wang, 2012). A scheme depicting the interactions between auxin, BR and gibberellic acid (GA) and impact on hypocotyl cell elongation is shown Figure 14.

c. Gibberellins

Gibberellins (GAs) are another important growth-promoting hormone and has consistent overlapping activity with BR (Bai *et al.*, 2012). GAs bind to the GA INSENSITIVE DWARF1 receptor (GID1). GID1 interacts with GA-repressor DELLA proteins to create a complex which will be ubiquitylated, degrading the DELLA proteins. When GAs content is low, DELLA proteins inactivate several transcription factors,

among them are the bHLH PIFs and BZR1 (Davière and Achard, 2013 and Figure 14). The presence of BR and the activation of BZR1 / BZR2 is required for GAs-driven hypocotyl elongation (Bai *et al.*, 2012). In line with this result, GAs- and BR-regulated genes significantly overlap (67 %), the remaining part being set by ethylene and jasmonate regulators. Gene ontology analysis of the GAs-BR regulated genes revealed a strong enrichment in cell wall- and cell elongation-related genes (Bai *et al.*, 2012). The role of GAs in stem elongation is acknowledged since the beginning of phytohormone research. Aloni (1979) has observed that the stem with a combination of high GA and low IAA produces long and thin phloem fibres and undergoes maximal elongation. GA impacts the cell wall composition at several levels. It is required for pectin biosynthesis and homogalacturonan esterification and is involved in the modification of the cellulose-xyloglucan network (Sánchez-Rodriguez *et al.*, 2010). GA stimulates the production of G lignin in the phloem fibres, resulting in a higher lignin content and a lower S / G ratio (Aloni *et al.*, 1990). Finally, GA up-regulates the expression of the three secondary wall *CesA*s, thereby significantly increasing the cellulose content (Huang *et al.*, 2015), and promotes the secondary growth of the stem (Hedden and Sponsel, 2015).

d. Cytokinins

Cytokinins (CKs) regulate numerous developmental events such as the activity of shoot meristem, vascular development, nutrient metabolism and primary metabolism (Hwang *et al.*, 2012; Brenner *et al.*, 2012). The *Arabidopsis* response regulators (ARR) are the final acceptor of the CK signalling cascade. Type A ARRs are generally considered as negative regulators of cytokinin signalling. The type B ARRs (11 genes in *Arabidopsis*) bind to the cytokinin-responsive genes and induce their transcription (Hwang *et al.*, 2012). They belong to the MYB family of

transcription factors and show functional redundancy (Brenner *et al.*, 2012). CKs signalling is also triggered by the Cytokinin response factors (CRFs). CRFs mediate the CK signal in a type B ARR—dependent mechanism. Target genes of these 2 families overlap. Overall, the meta-analysis of transcriptomic studies indicates that more than 3000 *Arabidopsis* genes are regulated by cytokinin. Those genes are involved in transcription regulation, phytohormone processes (auxin, ethylene and GA action, CK homeostasis), biotic and abiotic stresses, chloroplast development, maintenance of the activity of vascular cambium and cell wall biogenesis/modification (Brenner *et al.*, 2012; Hwang *et al.*, 2012). In this last class, 43 genes including expansins, laccases, PMEs, cell wall invertase and XETs are differentially regulated. CKs are negative regulator of cell wall thickness and hypocotyl elongation but enhance the expression of *EXPA1* and *CELL WALL INVERTASE 1* (*CWI1*). CWI1 is involved in the regulation of the source / sink carbohydrate transport. In addition, CKs play key roles during vascular development (Figure 15).

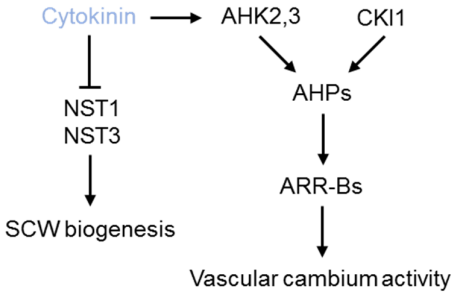


Figure 15: Roles of CKs in secondary growth and SCW biogenesis. CKs signalling from CK independent 1 (CKI1) and *Arabidopsis* His kinases (AHK2-AHK3) positively regulate vascular cambium activity (Hwang *et al.*, 2012) but inhibits *NST1* and *NST3* expression (Didi *et al.*, unpublished results) and finally SCW deposition.

e. Abscisic acid

Abscisic acid (ABA) has major physiological functions. One of the first described properties of ABA is to partially antagonise GA effects (stem elongation). ABA signalling pathway leads to the phosphorylation of basic leucine zipper (bZIP) transcription factors which finally activate the ABA-responsive genes (Cutler *et al.*, 2010). In *Arabidopsis*, over 10 % of the genome is ABA-regulated, corresponding to two to six times the number of genes regulated by other phytohormones. Among the processes up-regulated by ABA are genes conferring tolerance to stress, involved in cellular communication / organisation, energy and metabolism or transcription factors (e.g., MYB52) (Hoth *et al.*, 2002; Nemhauser *et al.*, 2006; Cutler *et al.*, 2010). ABA-repressed genes are associated with plant growth and cell wall assembly. ABA up-regulates several genes involved in lignification, such as *CAD*, *4CL*, *LAC5* and *LAC12* (Nemhauser *et al.*, 2006), as well as transcription factors involved in SCW deposition (*NST3*, *VND2*, *VND4*, *VND5*) (Didi *et al.*, 2015). Contrary to auxin and GA, ABA causes longitudinal orientation of cortical microtubules and subsequent cellulose deposition in the same direction (Shibaoka, 1994). This pattern of cellulose orientation is a feature of the cells which have stopped their elongation and / or have started the deposition of their secondary cell wall. The up-regulation of the transcription factors *NST3*, *VND2*, *VND4* and *VND5* is in line with this developmental stage.

f. Ethylene

Despite its simple structure, ethylene is involved in a variety of processes – seed growth, gravitropism and stress responses, among others (Wang *et al.*, 2013). In *Arabidopsis*, ethylene action is mediated through enhancing the expression of the *Ethylene response factor 1*

(*ERF1*) transcription factor. ERF1 interacts with ethylene-responsive gene promoter region. When exposed to light in the presence of ethylene or its precursor (1-aminocyclopropane-1-carboxylic acid, or ACC), a strong increase in hypocotyl length due to cell elongation is observed (Smalle *et al.*, 1997). Later on, the mechanism regulating *in planta* ethylene-driven cell elongation has been studied (Qin *et al.*, 2007; Yu *et al.*, 2013). Ethylene biosynthesis is activated by a signal produced by saturated very long chain fatty acids (VLCFA, such as C24:0). VLCFA promote the expression of several genes from the ethylene biosynthesis pathway. Cell expansion may be induced by the hydrogen peroxide signalling pathway (Qin *et al.*, 2007). In addition, ethylene enhances the translocation of Constitutive photomorphogenesis 1 (COP1) to the nucleus and subsequent degradation of the transcription factor Long hypocotyl 5 (HY5), contributing to the hypocotyl elongation (Yu *et al.*, 2013). It seems that the ethylene-induced elongation is mediated by auxin, as auxin-resistant mutants show weak response to ACC exposure (Pierik *et al.*, 2009). This elongation is driven by a set of cell wall genes up-regulated by ethylene, such as peroxidases, extensins and xylosidases (Yan Zhong and Burns, 2003).

g. Jasmonates

Jasmonic acid (JA) and related molecules, referred to as jasmonates (JAs), are active in many developmental processes, including secondary metabolism, growth inhibition, mechanotransduction, secondary growth and male fertility (Wasternack and Song, 2017). JAs are also an important element to the response to (a)biotic stresses (insect attack, fungal infection, resistance to salt, drought, chilling and osmotic stress). Jasmonic acid and methyl jasmonate have long been considered as the bioactive forms of JAs. The recent characterisation

of genes responsible for jasmonic acid conjugation with amino acids has demonstrated that these conjugates are the bioactive forms of JAs (Meesters *et al.*, 2014). More precisely, (+)-7-iso-JA-L-Ile, (+)-7-iso-JA-Leu, (+)-7-iso-JA-Val, (+)-7-iso-JA-Met, and (+)-7-iso-JA-Ala are the endogenous bioactive forms of JAs (Yan *et al.*, 2016). The expression of jasmonate-responsive genes is induced by several bHLH, MYB, NAC and WRKY transcription factors (reviewed in De Geyter *et al.*, 2012 and Goossens *et al.*, 2016), the best characterised of which are the bHLH-Leu zipper MYC2 and MYC3. A simplified model of jasmonate signalling is presented in Figure 16.

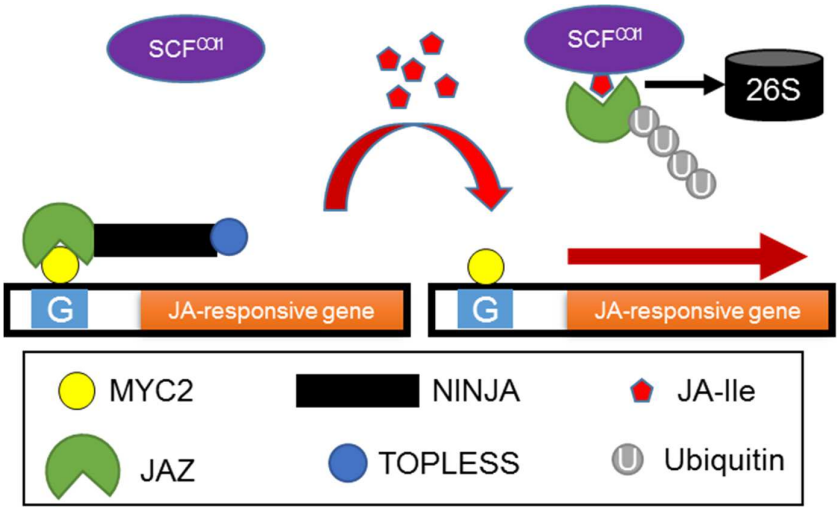


Figure 16: A simplified model of JA-Ile signalling through the SCF^{COI1}-JAZ co-receptor and the transcription factor MYC2. JAZ JASMONATE-ZIM-DOMAIN, NINJA NOVEL INTERACTOR OF JAZ, COI1 CORONATINE INSENSITIVE 1.

JAs activate the auxin biosynthetic genes and auxin up-regulates JA biosynthetic genes (Kazan and Manners, 2008). Since jasmonate signalling directly targets transcription factors, a consequent number of genes is induced or repressed by bioactive JAs.

In *Arabidopsis*, 3611 different genes were differentially regulated by methyl jasmonate (Hickman *et al.*, 2017). The transcriptomic response to JAs impacts many developmental processes. Notably, COI1 inhibits hypocotyl elongation (Huang *et al.*, 2017) and, together with MYC2, JAZ7 and JAZ10, regulates cambial activity (Sehr *et al.*, 2010). The inhibitory effect of JAs on stem and hypocotyl is partially due to a cross-talk with the GA signal. Indeed, JAs derepress the DELLA protein, thereby inhibiting PIFs and suppressing growth (Huang *et al.*, 2017). Several cell expansion deficient mutants have increased level of JA and show impaired cellulose biosynthesis (due to mutations in *CesA1* or *CesA3*) (Sánchez-Rodríguez *et al.*, 2010). It seems that JAs decrease cell wall deposition by inhibiting the synthesis of cell wall polysaccharides through reduced UDP-sugar abundance (Miyamoto *et al.*, 1997). The up-regulation of the phenylpropanoids / monolignols pathways has been extensively studied and may shed light on the lignification pattern of fibre crops. The abundance of coniferyl alcohol and higher-order oligolignols and lignin is significantly higher in methyl jasmonate-treated *Arabidopsis* cells (Pauwels *et al.*, 2008). This over-production of monolignols and lignin may result from a more global response to this hormone. Indeed, JA is a signal of herbivore attack. Accordingly, it has been observed a JA-dependent cell wall thickening in fibre and xylem tissues, as well as additional phloem fibres (Sehr *et al.*, 2010) and higher biosynthesis of oligolignols (Pauwels *et al.*, 2008).

5. Secondary growth and its regulation

Most herbaceous and woody Dicots undergo primary and secondary development (Nieminen *et al.*, 2015). Primary growth is responsible for the elongation at the tip of the organ whereas secondary growth is responsible for the thickening of the stem and roots through the vascular cambium. The primary vascular tissue originates from the shoot apical meristem (SAM) in the stem and from the provascular tissue in the root and hypocotyl. Procambial cells begin to divide at the onset of secondary growth. The division of the (pro)cambial cells gives rise to secondary xylem, secondary phloem and the new vascular cambium. The vascular cambium is composed of meristematic cells, called initials. An initial divides into two cells, another initial and a mother cell. This division perpetually regenerates the cambium and produces the mother cells which will differentiate into secondary xylem cells inwards and secondary phloem cells outwards (Nieminen *et al.*, 2015). More precisely, the cambium initials divide periclinally to produce xylem and phloem mother cells, and anticlinally to create new initials to follow the increased circumference of the stem (Figure 17). Considering that a hemp plant may reach a size of 2 to 3 metres, the mechanical strength conferred by the sclerenchyma from primary growth is not sufficient to provide the necessary stem rigidity. Secondary growth is therefore an indispensable process for the development of hemp, as well as other herbaceous species.

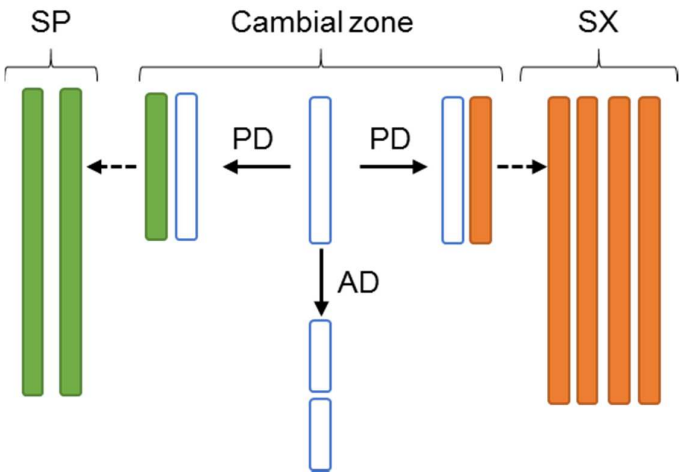


Figure 17: Overview of cambial cell specification and differentiation. Cambial initial (in white) divides, i. periclinally (PD) to form a new cambium initial and a phloem mother cell (small green rectangle) or a xylem mother cell (small orange rectangle), or ii. anticlinally (AD) to form two cambial initials. Phloem mother cells can differentiate in sieve-elements, phloem fibres and phloem parenchyma, within the secondary phloem (SP, large green rectangle). Xylem mother cells can differentiate in vessels, xylem fibres and xylem parenchyma cells (SX, large orange rectangle).

Several model plants have been used to unveil the actors of secondary growth and understand their regulation. Trees are the most evident system for such study, as their girth depends on this process (Liu *et al.*, 2014). Several transcriptional analysis have highlighted the molecular actors involved in this process (Schrader *et al.*, 2004; Sundell *et al.*, 2017). However, the hypocotyl is also a good model for studying secondary growth (Ragni and Hardtke, 2014). Besides the evident gain of time resulting from working with the hypocotyl, another element makes this system useful for molecular studies: elongation and thickening are temporally uncoupled. As a result, the two types of

growth may be accurately followed with a kinetic study, as already done with the hypocotyl of *Arabidopsis* (Baucher *et al.*, 2007).

The regulation of secondary growth is achieved by a suite of physiological and structural events: photosynthesis, xylem and phloem transport, sugar concentration, resistance to gravity. Vascular development relies on the tight regulation of the cambial cell proliferation and differentiation. The dodecapeptide TRACHEARY ELEMENT DIFFERENTIATION INHIBITORY FACTOR (TDIF) plays a key role in this respect (Figure 18). TDIF is secreted from the phloem and suppresses the differentiation of secondary xylem while promoting procambium proliferation and cambium identity (Nieminen *et al.*, 2015). TDIF is perceived by TDIF RECEPTOR (TDR) and PHLOEM INTERCALATED WITH XYLEM (PXY). This peptide induces the expression of the transcription factor *WUSCHEL-RELATED HOMEOBOX4* (*WOX4*) and negatively regulates the activity of the transcription factor *BES1*, resulting in cambial cell proliferation and inhibition of secondary xylem differentiation (Nieminen *et al.*, 2015). Secondary xylem fibre formation requires the expression of the *PXY/TDR-CORRELATED GENES1* (*PXC1*), which is coexpressed with *PXY*.

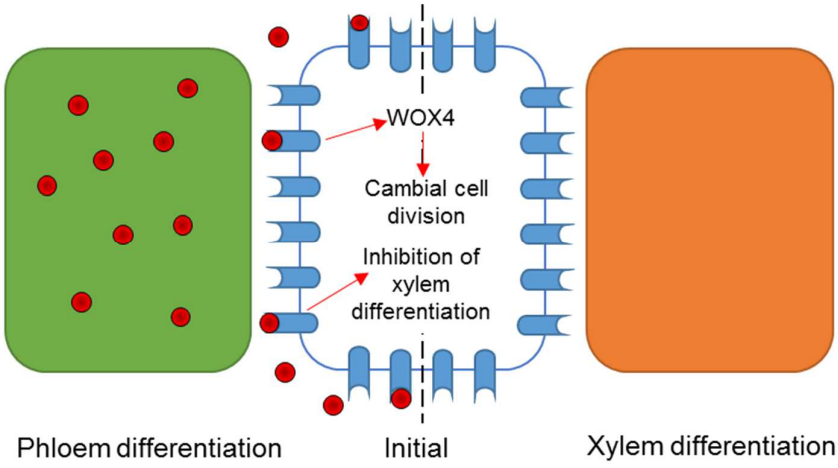


Figure 18: Simplified model of cambial differentiation. TDIF peptides (red sphere) from phloem cells activate the PXY / TDR receptor (in blue) that promotes WOX4-mediated cambial cell proliferation and inhibits xylem differentiation. Abbreviations as in the text.

Phytohormones are key molecules regulating secondary growth and affecting the tissue transcriptomes during plant development (Nieminen *et al.*, 2015 and Figure 19). Auxin has been shown to promote the cambial cell division by TDIF-WOX4 and other signalling pathways (Nieminen *et al.*, 2015). Mutants with severely decreased CK levels have a strong phenotype, with stem dwarfism and greatly reduced diameter. CK has a concentration maximum in the phloem and activates the transcription factor AINTEGUMENTA (ANT) in the cambium. The poplar ortholog of ANT, ANT-LIKE 1 (AIL1) up-regulates the expression of D-type cyclins in the cambium, inducing cell proliferation (Bhalerao and Fischer, 2017). GA stimulate cell division in the cambium and induce xylem fibre formation and elongation (Ursache *et al.*, 2013). Synergisms and crosstalks between auxin, GA and CK are also important regulator of cambial development (Ursache *et al.*, 2013). Ethylene also has a positive effect on cambial

activity. Increased ethylene biosynthesis leads to increased cambial activity (Ursache *et al.*, 2013), while this activity is reduced in a loss-of-function mutant of a gene positively involved in ethylene signalling. However, no phenotype was observed in two mutants with impaired ethylene signalling, suggesting the possibility of another molecular regulation (Nieminen *et al.*, 2015). Finally, jasmonic acid and jasmonates (JA) positively regulate secondary growth through the SCF^{COI1} pathway. Silencing the negative regulator of jasmonate signalling JAZ10 results in an earlier and more pronounced interfascicular cambial activity in *Arabidopsis* (Sehr *et al.*, 2010).

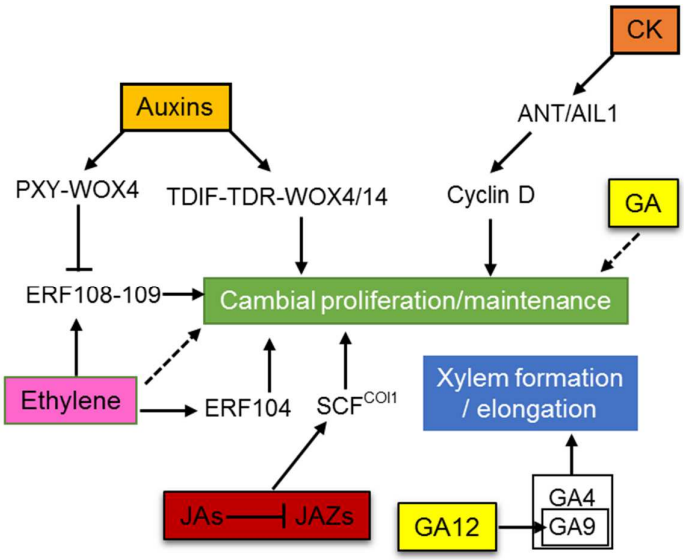


Figure 19: Hormonal regulation of cambial activity and xylem formation / elongation. The roles of auxin, CK, GA, JAs and ethylene are represented. The modulation of ERF108 and ERF109 is an example of the crosstalk between auxin and ethylene. Full arrow indicates a positive regulation and dotted arrow shows positive interaction for which molecular actors are not yet elucidated. ERF ETHYLENE RECEPTION FACTOR. All other abbreviations are as in the text. Data from Ursache *et al.*, 2013; Nieminen *et al.*, 2015; Bhalerao and Fischer, 2017.

6. Objectives

The interesting properties of plant fibres rely on the abundance and type of their SCW. The most striking difference between PCW and SCW (xylan-type) is the presence of lignin in the latter. The pattern of cellulose deposition also greatly differs between PCW and SCW, which is of special interest for the plants which are cultivated for their fibres. During PCW synthesis, the CSC synthesizes cellulose microfibrils with low aggregation; cellulose microfibrils are instead highly aggregated and parallel oriented in the SCW (Li *et al.*, 2016). In addition, the assembly of SCW cellulose relies on specific proteins, such as COBL, CTL or KOR (Zhong and Ye, 2015), the roles of which are not really understood. Importantly, the composition of xylem and phloem SCWs are rather different in hemp. Some models already tentatively explain the interaction between xylan and cellulose (Busse-Wicher *et al.*, 2014) and G-layer-polysaccharides – cellulose (notably xyloglucan – cellulose and RG-I – cellulose in tension wood, Mellerowicz and Gorshkova, 2012 and Gorshkova *et al.*, 2015), but the interactions lignin - cellulose is so far not documented. In G-layer, the establishment of such a model is further hampered by the complexity of the matrix polysaccharides. The biosynthesis pathway of these polysaccharides (e.g., xylan and RG-I) displays an equivalent complexity and is far from being elucidated (Hao and Mohnen, 2014; Gorshkov *et al.*, 2017). It is therefore of the utmost importance to decipher first the composition of the macromolecules composing the SCW and G-layer and their biosynthetic pathways, before being able to explain the interaction between those molecules and finally the overall cell wall organisation determining the properties of the fibres. Based on the above-mentioned observations, I would like to answer to the following questions:

1) What are the transcriptional and hormonal signatures of primary and secondary growth in hemp?
2) Which molecular factors are associated with the development of hemp bast fibres?

These two questions will be addressed using two biological systems. In the first part of the thesis (Chapters 1 and 2), elongation and secondary growths are studied with the hypocotyl model, taking advantage of the temporal uncoupling of elongation and secondary thickening of this system. Four time-points spanning the elongation-thickening transition will be investigated: 6, 9, 15 and 20 days after sowing. In the second part of the thesis (Chapters 3 to 5), I will investigate the mechanisms underlying stem development using bast fibres and core tissues of developing internodes. This system complements the results relative to primary and secondary growth obtained with the hypocotyl model with data documenting the different developmental stages of successive internodes and the contrasted composition of core and cortical tissues of the adult hemp stem. This second part aims at providing the necessary biological insights required to be able to modulate, in the future, the characteristics needed for specific industrial uses of hemp. In both parts, we will take advantage of the power of next-generation sequencing (RNA-Sequencing) to study the development of vascular tissues in our biological systems.

Chapter 1 will describe the contrasted cell wall composition of hypocotyl cross section by optical microscopy and immunohistochemistry, transcriptomics and phytohormone analysis of the four time points, with an emphasis on cell wall-related elements. In this respect, the genes involved in primary and secondary growth, as well as PCW, SCW and G-layer formation will be investigated by RNA-

Seq. These transcriptomic results will be put in perspective with phytohormone contents and fingerprinting of the cell wall molecules by immunohistochemistry.

Chapter 2 will more deeply investigate the lignification of the growing hypocotyl by chemical analysis, laccase and peroxidase activities and a proteomic approach. The quantity and monomeric composition of lignin are shown to be age-dependent. The holistic analysis of the soluble proteome, together with the *in situ* activities of enzymes involved in lignification, aim at explaining this differential pattern of lignification and its functional impact on the hypocotyl growth.

In Chapter 3, the major differences between the lignified core tissue and the fibre-bearing cortical tissues are discussed through a gene expression analysis. Considering the importance of FLA proteins in cell wall deposition and stem biomechanics, a specific expression study of *FLA* genes in the core tissue, cortical tissue and in the growing hypocotyls is also carried out in this Chapter.

The main events accompanying bast fibre elongation and thickening in the adult stem will be investigated with a transcriptomic approach in Chapter 4. A special emphasis is given to the study of genes involved in cell wall biosynthesis and remodelling, phytohormone biosyntheses and signalling and secondary metabolism. The data also show that the major transcriptional changes in the cell wall-related processes occur at the snap point.

Considering the transcriptional changes, occurring at the snap point, of genes involved in secondary cell wall biosynthesis, a molecular “dissection” of this stem region will be carried out by a targeted gene expression analysis (Chapter 5). The snap point is an empirical reference widely used to study the structural aspects of cell

wall development in fibre crops, but to date it has not been analysed at a high spatial resolution with molecular methods. The results show that the gene expression at the snap point marks also a transition between cell elongation and cell wall thickening. Notably, the SCW-*CesA*s are up-regulated below the snap point. In accordance with the results of Chapter 3, we tentatively predict the functions of some *FLA* genes in either stem elongation or thickening.

The thesis structure is presented in a schematic way in Figure 20.

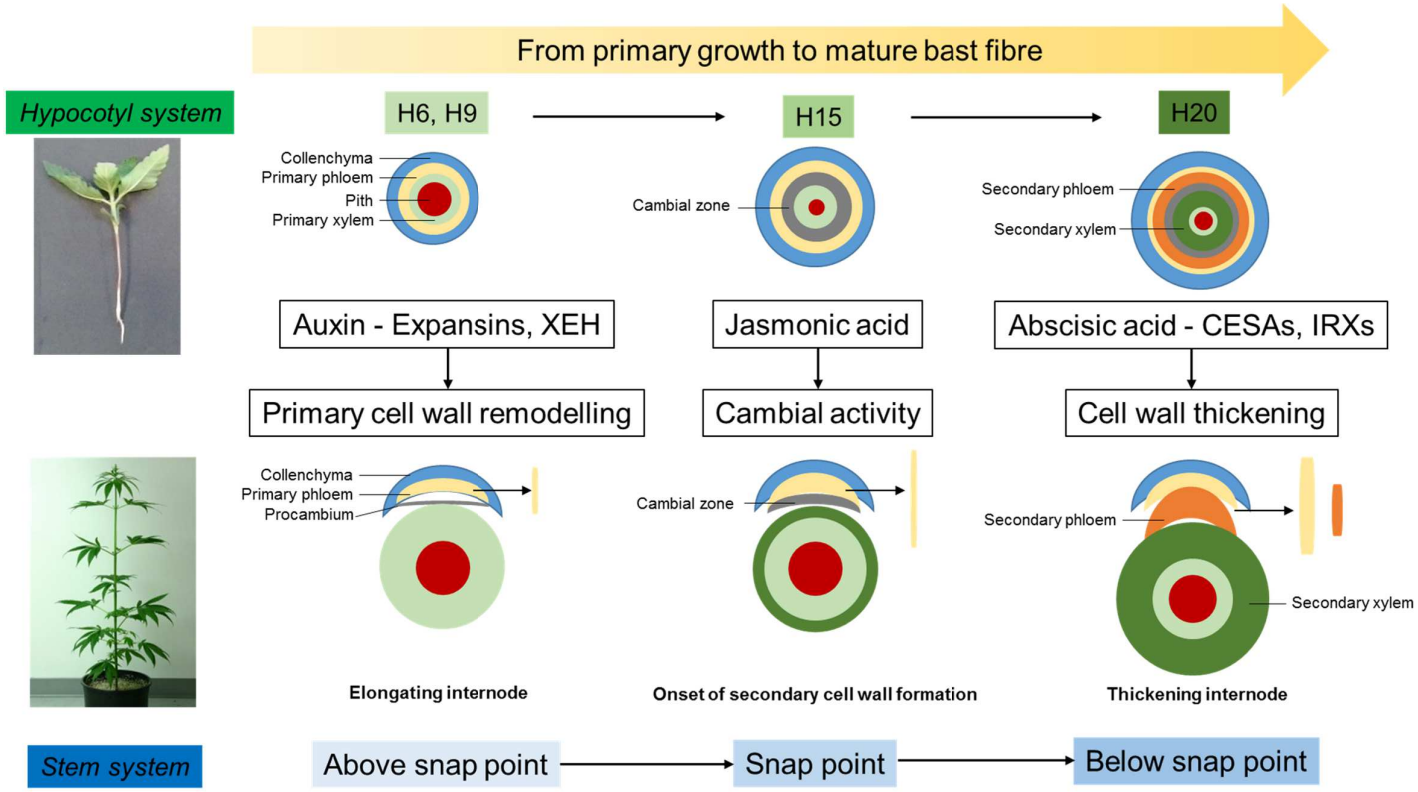

From primary growth to mature bast fibre
Hypocotyl system
H6, H9
H15
H20
Collenchyma
Primary phloem
Pith
Primary xylem
Cambial zone
Secondary phloem
Secondary xylem
Auxin - Expansins, XEH
Jasmonic acid
Abscisic acid - CESAs, IRXs
Primary cell wall remodelling
Cambial activity
Cell wall thickening
Collenchyma
Primary phloem
Procambium
Cambial zone
Secondary phloem
Secondary xylem
Elongating internode
Onset of secondary cell wall formation
Thickening internode
Stem system
Above snap point
Snap point
Below snap point

Figure 20 *(previous page)*: Scheme summarising the main events occurring in the two models used in this thesis. Primary growth is studied with H6 and H9 and the internode above the snap point. Beginning of secondary growth and cell wall thickening is studied with H15 and the internode including the snap point. Finally, G-layer thickening and later stages of secondary growth are described in H20 and in the internode below the snap point. The main hormones and genes involved in specific developmental stages are indicated.

References

Al-Maadeed, M.A., and Labidi, S. (2014). "4 - Recycled polymers in natural fibre-reinforced polymer composites," in *Natural Fibre Composites*. Woodhead Publishing, 103-114.

Aloni, R. (1979). Role of Auxin and Gibberellin in Differentiation of Primary Phloem Fibers. *Plant Physiol* 63**,** 609-614.

Aloni, R. (1982). Role of Cytokinin in Differentiation of Secondary Xylem Fibers. *Plant Physiol* 70**,** 1631-1633.

Aloni, R., Tollier, M.T., and Monties, B. (1990). The Role of Auxin and Gibberellin in Controlling Lignin Formation in Primary Phloem Fibers and in Xylem of *Coleus blumei* Stems. *Plant Physiol.* 94**,** 1743-1747.

Andème-Onzighi, C., Douchiche, O., Driouich, A., and Morvan, C. (2005). Composition of Flax Hypocotyl Fibres. *J. Nat. Fib.* 2**,** 1-15.

Andème-Onzighi, C., Girault, R., His, I., Morvan, C., and Driouich, A. (2000). Immunocytochemical characterization of early-developing flax fiber cell walls. *Protoplasma* 213**,** 235-245.

Arent, D.J., Wise, A., and Gelman, R. (2011). The status and prospects of renewable energy for combating global warming. *Energ. Econ.* 33**,** 584-593.

Babst, B.A., Ferrieri, R.A., Gray, D.W., Lerdau, M., Schlyer, D.J., Schueller, M., *et al.* (2005). Jasmonic acid induces rapid changes in carbon transport and partitioning in *Populus*. *New Phytol.* 167**,** 63-72.

Bai, M.Y., Shang, J.X., Oh, E., Fan, M., Bai, Y., Zentella, R., *et al.* (2012). Brassinosteroid, gibberellin and phytochrome impinge on a common transcription module in Arabidopsis. *Nat. Cell Biol.* 14**,** 810-817.

Barakat, A., Winter, H., Rondeau-Mouro, C., Saake, B., Chabbert, B., and Cathala, B. (2007). Studies of xylan interactions and cross-linking to synthetic lignins formed by bulk and end-wise polymerization: a model study of lignin carbohydrate complex formation. *Planta* 226**,** 267.

Barros, J., Serk, H., Granlund, I., and Pesquet, E. (2015). The cell biology of lignification in higher plants. *Ann. Bot*. 115, 1053-1074.

Barta, Z., Oliva, J.M., Ballesteros, I., Dienes, D., Ballesteros, M., and Réczey, K. (2010). Refining hemp hurds into fermentable sugars or ethanol. *Chem. Bioch. Eng. Quarter.* 24**,** 331-339.

Bashline, L., Li, S., and Gu, Y. (2014). The trafficking of the cellulose synthase complex in higher plants. *Ann. Bot.* 114**,** 1059-1067.

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

Baucher, M., El Jaziri, M., and Vandeputte, O. (2007). From primary to secondary growth: origin and development of the vascular system. *J. Exp. Bot.* 58**,** 3485-3501.

Behr, M., Legay, S., Zizková, E., Motyka, V., Dobrev, P., I., Hausman, J.F., *et al.* (2016). Studying Secondary Growth and Bast Fiber Development: The Hemp Hypocotyl Peeks behind the Wall. *Front. Plant Sci.* 7**,** 1733.

Berthet, S., Demont-Caulet, N., Pollet, B., Bidzinski, P., Cézard, L., Le Bris, P., *et al.* (2011). Disruption of LACCASE4 and 17 Results in Tissue-Specific Alterations to Lignification of Arabidopsis thaliana Stems. *Plant Cell* 23**,** 1124-1137.

Berthet, S., Thevenin, J., Baratiny, D., Mont-Caulet, N., Debeaujon, I., Bidzinski, P., *et al.* (2012). "Chapter 5 - Role of Plant Laccases in Lignin Polymerization," in *Advances in Botanical Research, Lignins Biosynthesis, Biodegradation and Bioengineering,* ed. L. Jouanin and C. Lapierre. Volume 61 ed: Academic Press, 145-172.

Beveridge, C.A., Symons, G.M., and Turnbull, C.G.N. (2000). Auxin Inhibition of Decapitation-Induced Branching Is Dependent on Graft-Transmissible Signals Regulated by Genes *Rms1* and *Rms2*. *Plant Physiol.* 123**,** 689-698.

Bhalerao, R.P., and Fischer, U. (2017). Environmental and hormonal control of cambial stem cell dynamics. *J. Exp. Bot.* 68**,** 79-87.

Blake, A.W., Marcus, S.E., Copeland, J.E., Blackburn, R.S., and Knox, J.P. (2008). In situ analysis of cell wall polymers associated with phloem fibre cells in stems of hemp, *Cannabis sativa* L. *Planta* 228**,** 1-13.

Bonawitz, N.D., and Chapple, C. (2010). The Genetics of Lignin Biosynthesis: Connecting Genotype to Phenotype. *Annu. Rev. Genet.* 44**,** 337-363.

Bootten, T.J., Harris, P.J., Melton, L.D., and Newman, R.H. (2004). Solid-state 13C-NMR spectroscopy shows that the xyloglucans in the primary cell walls of mung bean (Vigna radiata L.) occur in different domains: a new model for

xyloglucan–cellulose interactions in the cell wall. *J. Exp. Bot.* 55**,** 571-583.

Braccini, I., and Pérez, S. (2001). Molecular Basis of Ca2+-Induced Gelation in Alginates and Pectins: The Egg-Box Model Revisited. *Biomacromol.* 2**,** 1089-1096.

Brenner, W., Ramireddy, E., Heyl, A., and Schmülling, T. (2012). Gene Regulation by Cytokinin in Arabidopsis. *Front. Plant Sci.* 3.

Brill, E., Van Thournout, M., White, R.G., Llewellyn, D., Campbell, P.M., Engelen, S., *et al.* (2011). A Novel Isoform of Sucrose Synthase Is Targeted to the Cell Wall during Secondary Cell Wall Synthesis in Cotton Fiber. *Plant Physiol.*157**,** 40-54.

Busse-Wicher, M., Gomes, T.C.F., Tryfona, T., Nikolovski, N., Stott, K., Grantham, N.J., *et al.* (2014). The pattern of xylan acetylation suggests xylan may interact with cellulose microfibrils as a twofold helical screw in the secondary plant cell wall of Arabidopsis thaliana. *Plant J.* 79**,** 492-506.

Busse-Wicher, M., Grantham, N. J., Lyczakowski, Jan j., Nikolovski, N., and Dupree, P. (2016). Xylan decoration patterns and the plant secondary cell wall molecular architecture. *Biochem. Soc. Trans.* 44**,** 74-78.

Carroll, A., Mansoori, N., Li, S., Lei, L., Vernhettes, S., Visser, R.G.F., *et al.* (2012). Complexes with Mixed Primary and Secondary Cellulose Synthases Are Functional in Arabidopsis Plants. *Plant Physiol.* 160**,** 726-737.

Cavalier, D.M., Lerouxel, O., Neumetzler, L., Yamauchi, K., Reinecke, A., Freshour, G., *et al.* (2008). Disrupting Two Arabidopsis thaliana Xylosyltransferase Genes Results in Plants Deficient in Xyloglucan, a Major Primary Cell Wall Component. *Plant Cell* 20**,** 1519-1537.

Chabi, M., Goulas, E., Leclercq, C.C., De Waele, I., Rihouey, C., Cenci, U., *et*

al. (2017). A Cell Wall Proteome and Targeted Cell Wall Analyses Provide Novel Information on Hemicellulose Metabolism in Flax. *Mol. Cell. Prot.* 16**,** 1634-1651.

Chakraborty, A., Sarkar, D., Satya, P., Karmakar, P.G., and Singh, N.K. (2015). Pathways associated with lignin biosynthesis in lignomaniac jute fibres. *Mol. Genet. Genom.* 290**,** 1523-1542.

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

Chen, J., Pei, Z., Dai, L., Wang, B., Liu, L., An, X., and Peng, D. (2014). Transcriptome profiling using pyrosequencing shows genes associated with bast fiber development in ramie (Boehmeria nivea L.). *BMC Genomics* 15**,** 919.

Cosgrove, D.J. (2014). Re-constructing our models of cellulose and primary cell wall assembly. *Curr. Opin. Plant Biol.* 22**,** 122-131.

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

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.

Crozier, A., Kamiya, Y., Bishop, G., and Yokota, T. (2000). "Biosynthesis of Hormones and Elicitor Molecules," in *Biochemistry and Molecular Biology of Plants,* eds. B. Buchanan, W. Gruissem & R. Jones. American Society of Plant Physiologists, 850-929.

Cutler, S.R., Rodriguez, P.L., Finkelstein, R.R., and Abrams, S.R. (2010). Abscisic Acid: Emergence of a Core Signaling Network. *Annu. Rev. Plant Biol.*61**,** 651-679.

Davière, J.-M., and Achard, P. (2013). Gibberellin signaling in plants. *Development* 140**,** 1147-1151.

Day, A., Ruel, K., Neutelings, G., Cronier, 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.

De Geyter, N., Gholami, A., Goormachtig, S., and Goossens, A. (2012). Transcriptional machineries in jasmonate-elicited plant secondary metabolism. *Trends Plant Sci.* 17**,** 349-359.

De Pauw, M., Vidmar, J., Collins, J., Bennett, R., and Deyholos, M. (2007). Microarray analysis of bast fibre producing tissues of *Cannabis sativa* identifies transcripts associated with conserved and specialised processes of secondary wall development. *Funct. Plant Biol.* 34**,** 737-749.

Deng, T., Yao, H., Wang, J., Wang, J., Xue, H., and Zuo, K. (2016). GhLTPG1, a cotton GPI-anchored lipid transfer protein, regulates the transport of phosphatidylinositol monophosphates and cotton fiber elongation. *Sci. Rep.* 6**,** 26829.

Derba-Maceluch, M., Awano, T., Takahashi, J., Lucenius, J., Ratke, C., Kontro, I., *et al.* (2015). Suppression of xylan endotransglycosylase PtxtXyn10A affects cellulose microfibril angle in secondary wall in aspen wood. *New Phytol.* 205**,** 666-681.

Dhillon, R.S., and Von Wuehlisch, G. (2013). Mitigation of global warming through renewable biomass. *Biomass Bioenerg.* 48**,** 75-89.

Didi, V., Jackson, P., and Hejátko, J. (2015). Hormonal regulation of secondary cell wall formation. *J. Exp. Bot.* 66**,** 5015-5027.

Dincer, I. (2000). Renewable energy and sustainable development: a crucial review. *Renew. Sust. Energ. Rev.* 4**,** 157-175.

Donaldson, L.A., and Knox, J.P. (2012). Localization of Cell Wall Polysaccharides in Normal and Compression Wood of Radiata Pine: Relationships with Lignification and Microfibril Orientation. *Plant Physiol.*158**,** 642-653.

Driouich, A., Follet-Gueye, M.L., Bernard, S., Kousar, S., Chevalier, L., Vicré, M., *et al.* (2012). Golgi-mediated synthesis and secretion of matrix polysaccharides of the primary cell wall of higher plants. *Front. Plant Sci.* 3.

Ellabban, O., Abu-Rub, H., and Blaabjerg, F. (2014). Renewable energy resources: Current status, future prospects and their enabling technology. *Renew. Sust. Energ. Rev.* 39**,** 748-764.

Fendrych, M., Leung, J., and Friml, J. (2016). TIR1/AFB-Aux/IAA auxin perception mediates rapid cell wall acidification and growth of Arabidopsis hypocotyls. *eLife* 5**,** e19048.

Fernández-Pérez, F., Pomar, F., Pedreno, M.A., and Novo-Uzal, E. (2014). The suppression of AtPrx52 affects fibers but not xylem lignification in Arabidopsis by altering the proportion of syringyl units. *Physiol. Plant.*154**,** 395-406.

Fernández-Pérez, F., Pomar, F., Pedreño, M.A., and Novo-Uzal, E. (2015). Suppression of Arabidopsis peroxidase 72 alters cell wall and phenylpropanoid metabolism. *Plant Sci.* 239**,** 192-199.

Fernandez-Tendero, E., Day, A., Legros, S., Habrant, A., Hawkins, S., and

Chabbert, B. (2017). Changes in hemp secondary fiber production related to technical fiber variability revealed by light microscopy and attenuated total reflectance Fourier transform infrared spectroscopy. *PLOS ONE* 12**,** e0179794.

Faux, A.-M. (2014). Quantitative approach of the genetic determinism of sex expression in monoecious hemp (Cannabis sativa l.), and its relationship with flowering phenology and stem and seed yields. PhD thesis in agronomical sciences and biological engineering, Université catholique de Louvain.

Fike, J. (2016). Industrial Hemp: Renewed Opportunities for an Ancient Crop. *Crit. Rev. Plant Sci.* 35**,** 406-424.

Fischer, E.M., and Knutti, R. (2015). Anthropogenic contribution to global occurrence of heavy-precipitation and high-temperature extremes. *Nat. Clim. Chang.* 5**,** 560-564.

Fraser, C.M., and Chapple, C. (2011). The Phenylpropanoid Pathway in Arabidopsis. *The Arabidopsis Book* 9**,** e0152.

Fujii, S., Hayashi, T., and Mizuno, K. (2010). Sucrose Synthase is an Integral Component of the Cellulose Synthesis Machinery. *Plant Cell Physiol.* 51**,** 294-301.

Gabaldón, C., López-Serrano, M., Pomar, F., Merino, F., Cuello, J., Pedreño, M.A., *et al.* (2006). Characterization of the last step of lignin biosynthesis in Zinnia elegans suspension cell cultures. *FEBS Letters* 580**,** 4311-4316.

Goossens, J., Fernández-Calvo, P., Schweizer, F., and Goossens, A. (2016). Jasmonates: signal transduction components and their roles in environmental stress responses. *Plant Mol. Biol.* 91**,** 673-689.

Gorshkov, O., Mokshina, N., Gorshkov, V., Chemikosova, S., Gogolev, Y.,

and Gorshkova, T. (2017). Transcriptome portrait of cellulose-enriched flax fibres at advanced stage of specialization. *Plant Mol. Biol.* 93**,** 431-449.

Gorshkova, T., Brutch, N., Chabbert, B., Deyholos, M., Hayashi, T., Lev-Yadun, S., *et al.* (2012). Plant Fiber Formation: State of the Art, Recent and Expected Progress, and Open Questions. *Crit. Rev. Plant Sci.* 31**,** 201-228.

Gorshkova, T., Mokshina, N., Chernova, T., Ibragimova, N., Salnikov, V., Mikshina, 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.* 169**,** 2048-2063.

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

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

Goudenhooft, C., Siniscalco, D., Arnould, O., Bourmaud, A., Sire, O., Gorshkova, T., *et al.* (2018). Investigation of the Mechanical Properties of Flax Cell Walls during Plant Development: The Relation between Performance and Cell Wall Structure. *Fibers* 6**,** 6.

Groover, A. (2016). Gravitropisms and reaction woods of forest trees – evolution, functions and mechanisms. *New Phytol.* 211**,** 790-802.

Guerriero, G., Behr, M., Legay, S., Mangeot-Peter, L., Zorzan, S., Ghoniem, M., *et al.* (2017). Transcriptomic profiling of hemp bast fibres at different developmental stages. *Sci. Rep.* 7**,** 4961.

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.

Hao, Z., and Mohnen, D. (2014). A review of xylan and lignin biosynthesis: Foundation for studying Arabidopsis irregular xylem mutants with pleiotropic phenotypes. *Crit. Rev. Biochem. Mol. Biol.* 49**,** 212-241.

Hayashi, K.I. (2012). The Interaction and Integration of Auxin Signaling Components. *Plant Cell Physiol.* 53**,** 965-975.

Hedden, P. and Sponsel, V. (2015). A Century of Gibberellin Research. *J. Plant Growth Regul*. 34, 740-760.

Hentrich, M., Sánchez-Parra, B., Pérez Alonso, M.M., Carrasco Loba, V., Carrillo, L., Vicente-Carbajosa, J., *et al.* (2013). YUCCA8 and YUCCA9 overexpression reveals a link between auxin signaling and lignification through the induction of ethylene biosynthesis. *Plant Signal. Behav.* 8**,** e26363.

Hermann, B.G., Blok, K., and Patel, M.K. (2007). Producing Bio-Based Bulk Chemicals Using Industrial Biotechnology Saves Energy and Combats Climate Change. *Environ. Sci. Technol.* 41**,** 7915-7921.

Herrero, J., Fernández-Pérez, F., Yebra, T., Novo-Uzal, E., Pomar, F., Pedreno, M., *et al.* (2013). Bioinformatic and functional characterization of the basic peroxidase 72 from Arabidopsis thaliana involved in lignin biosynthesis. *Planta* 237**,** 1599-1612.

Hill, J.L., Hammudi, M.B., and Tien, M. (2014). The Arabidopsis Cellulose Synthase Complex: A Proposed Hexamer of CESA Trimers in an Equimolar Stoichiometry. *Plant Cell* 26**,** 4834-4842.

His, I., Andème-Onzighi, C., Morvan, C., and Driouich, A. (2001). Microscopic

Studies on Mature Flax Fibers Embedded in LR White. *J. Histochem. Cytochem.* 49**,** 1525-1535.

Hobson, N., and Deyholos, M.K. (2013). LuFLA1PRO and LuBGAL1PRO promote gene expression in the phloem fibres of flax (*Linum usitatissimum*). *Plant Cell Rep.* 32**,** 517-528.

Hoth, S., Morgante, M., Sanchez, J.P., Hanafey, M.K., Tingey, S.V., and Chua, N.H. (2002). Genome-wide gene expression profiling in *Arabidopsis thaliana* reveals new targets of abscisic acid and largely impaired gene regulation in the abi1-1 mutant. *J. Cell Sci.* 115**,** 4891-4900.

Huang, D., Wang, S., Zhang, B., Shang-Guan, K., Shi, Y., Zhang, D., *et al.* (2015). A Gibberellin-Mediated DELLA-NAC Signaling Cascade Regulates Cellulose Synthesis in Rice. *Plant Cell* 27**,** 1681-1696.

Huang, H., Liu, B., Liu, L., and Song, S. (2017). Jasmonate action in plant growth and development. *J. Exp. Bot.* 68**,** 1349-1359.

Hussey, S.G., Nicky M., Creux N. M., and Myburg A., A. (2013). Navigating the transcriptional roadmap regulating plant secondary cell wall deposition. *Front. Plant Sci.* 4**,** 1-21.

Hwang, I., Sheen, J., and Müller, B. (2012). Cytokinin Signaling Networks. *Annu. Rev. Plant Biol.* 63**,** 353-380.

Ikeda, M., Mitsuda, N., and Ohme-Takagi, M. (2013). ATBS1 INTERACTING FACTORs negatively regulate *Arabidopsis* cell elongation in the triantagonistic bHLH system. *Plant Signal. Behav.* 8**,** e23448.

Ito, S., Suzuki, Y., Miyamoto, K., Ueda, J., and Yamaguchi, I. (2005). AtFLA11, a Fasciclin-Like Arabinogalactan-Protein, Specifically Localized in Screlenchyma Cells. *Biosci. Biotechnol. Biochem.* 69**,** 1963-1969.

Jacob-Wilk, D., Kurek, I., Hogan, P., and Delmer, D.P. (2006). The cotton fiber zinc-binding domain of cellulose synthase A1 from Gossypium hirsutum displays rapid turnover in vitro and in vivo. *Proc. Natl. Acad. Sci. U.S.A.* 103**,** 12191-12196.

Kazan, K., and Manners, J.M. (2008). Jasmonate Signaling: Toward an Integrated View. *Plant Physiol.* 146**,** 1459-1468.

Keegstra, K. (2010). Plant Cell Walls. *Plant Physiol.*154**,** 483-486.

Khan, S., and Stone, J. (2007). Arabidopsis thaliana GH3.9 in Auxin and Jasmonate Cross Talk. *Plant Signal. Behav.* 2**,** 483-485.

King, K.E., Moritz, T., and Harberd, N.P. (2001). Gibberellins Are Not Required for Normal Stem Growth in *Arabidopsis thaliana* in the Absence of GAI and RGA. *Genetics* 159**,** 767-776.

Kiyoto, S., Yoshinaga, A., and Takabe, K. (2015). Relative deposition of xylan and 8-5'-linked lignin structure in *Chamaecyparis obtusa*, as revealed by double immunolabeling by using monoclonal antibodies. *Planta* 241**,** 243-256.

Komiyama, H., Kato, A., Aimi, H., Ogihara, J., and Shimizu, K. (2008). Chemical structure of kenaf xylan. *Carbohydr. Polym.* 72**,** 638-645.

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

Kumar, M., Campbell, L., and Turner, S. (2016). Secondary cell walls: biosynthesis and manipulation. *J. Exp. Bot.* 67**,** 515-531.

Kutschera, U., and Wang, Z.Y. (2012). Brassinosteroid action in flowering

plants: a Darwinian perspective. *J. Exp. Bot.* 63**,** 3511-3522.

Le Roy, J., Blervacq, A.-S., Créach, A., Huss, B., Hawkins, S., and Neutelings, G. (2017). Spatial regulation of monolignol biosynthesis and laccase genes control developmental and stress-related lignin in flax. *BMC Plant Biol.* 17**,** 124.

Le Roy, J., Huss, B., Creach, A., Hawkins, S., and Neutelings, G. (2016). Glycosylation is a major regulator of phenylpropanoid availability and biological activity in plants. *Front. Plant Sci.* 7.

Lee, C., Teng, Q., Zhong, R., and 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**,** 1289-1301.

Li, S., Bashline, L., Zheng, Y., Xin, X., Huang, S., Kong, Z., *et al.* (2016). Cellulose synthase complexes act in a concerted fashion to synthesize highly aggregated cellulose in secondary cell walls of plants. *Proc. Natl. Acad. Sci. U.S.A.* 113**,** 11348-11353.

Li, Y., Kim, J.I., Pysh, L., and Chapple, C. (2015). Four Isoforms of Arabidopsis 4-Coumarate:CoA Ligase Have Overlapping yet Distinct Roles in Phenylpropanoid Metabolism. *Plant Physiol.* 169**,** 2409-2421.

Liang, M., Davis, E., Gardner, D., Cai, X., and Wu, Y. (2006). Involvement of AtLAC15 in lignin synthesis in seeds and in root elongation of Arabidopsis. *Planta* 224**,** 1185-1196.

Liu, H., Wu, W., Hou, K., Chen, J., and Zhao, Z. (2016). Deep sequencing reveals transcriptome re-programming of *Polygonum multiflorum* thunb. roots to the elicitation with methyl jasmonate. *Mol. Genet. Genomics.* 291**,** 337-348.

Liu, L., Filkov, V., and Groover, A. (2014). Modeling transcriptional networks regulating secondary growth and wood formation in forest trees. *Physiol. Plant.*151**,** 156-163.

Liu, L., Shang-Guan, K., Zhang, B., Liu, X., Yan, M., Zhang, Li, *et al.* (2013). Brittle Culm1, a COBRA-Like Protein, Functions in Cellulose Assembly through Binding Cellulose Microfibrils. *PLOS Genet.* 9**,** e1003704.

Liwanag, A.J.M., Ebert, B., Verhertbruggen, Y., Rennie, E.A., Rautengarten, C., Oikawa, A., *et al.* (2012). Pectin Biosynthesis: GALS1 in *Arabidopsis thaliana* Is a β-1,4-Galactan β-1,4-Galactosyltransferase. *Plant Cell* 24**,** 5024-5036.

Makshakova, O.N., Gorshkova, T.A., Mikshina, P.V., Zuev, Y.F., and Perez, S. (2017). Metrics of rhamnogalacturonan I with β-(1→4)-linked galactan side chains and structural basis for its self-aggregation. *Carbohydr. Polym.*158**,** 93-101.

Mansoori, N., Timmers, J., Desprez, T., Kamei, C.L.A., Dees, D.C.T., Vincken, J.-P., *et al.* (2014). KORRIGAN1 Interacts Specifically with Integral Components of the Cellulose Synthase Machinery. *PLOS ONE* 9**,** e112387.

Martínez-Abad, A., Berglund, J., Toriz, G., Gatenholm, P., Henriksson, G., Lindström, M., *et al.* (2017). Regular Motifs in Xylan Modulate Molecular Flexibility and Interactions with Cellulose Surfaces. *Plant Physiol*.

Mason, M.G., Ross, J.J., Babst, B.A., Wienclaw, B.N., and Beveridge, C.A. (2014). Sugar demand, not auxin, is the initial regulator of apical dominance. *Proc. Natl. Acad. Sci. U.S.A.* 111**,** 6092-6097.

Mccarthy, R.L., Zhong, R., Fowler, S., Lyskowski, D., Piyasena, H., Carleton, *et al.* (2010). The Poplar MYB Transcription Factors, PtrMYB3 and PtrMYB20, are Involved in the Regulation of Secondary Wall Biosynthesis. *Plant Cell*

Physiol. 51**,** 1084-1090.

Mccartney, L., Marcus, S.E., and Knox, J.P. (2005). Monoclonal Antibodies to Plant Cell Wall Xylans and Arabinoxylans. *J. Mol. Histol.* 53**,** 543-546.

Mcglade, C., and Ekins, P. (2015). The geographical distribution of fossil fuels unused when limiting global warming to 2°C. *Nature* 517**,** 187-190.

Meesters, C., Mönig, T., Oeljeklaus, J., Krahn, D., Westfall, C.S., Hause, B., *et al.* (2014). A chemical inhibitor of jasmonate signaling targets JAR1 in Arabidopsis thaliana. *Nature Chem. Biol.* 10**,** 830-836.

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.

Mellerowicz, E.J., Immerzeel, P., and Hayashi, T. (2008). Xyloglucan: The Molecular Muscle of Trees. *Ann. Bot.* 102**,** 659-665.

Mikshina, P., Chernova, T., Chemikosova, S.B., Ibragimova, N., Mokshina, N., and Gorshkova, T. (2013). "Cellulosic Fibers: Role of Matrix Polysaccharides in Structure and Function," in *Cellulose - Fundamental Aspects,* eds. T. Van De Ven & L. Godbout. InTech), 91-112.

Mikshina, P.V., Idiyatullin, B.Z., Petrova, A.A., Shashkov, A.S., Zuev, Y.F., and Gorshkova, T.A. (2015). Physicochemical properties of complex rhamnogalacturonan I from gelatinous cell walls of flax fibers. *Carbohydr. Polym.* 117**,** 853-861.

Miyamoto, K., Oka, M., and Ueda, J. (1997). Update on the possible mode of action of the jasmonates: Focus on the metabolism of cell wall polysaccharides in relation to growth and development. *Physiol. Plant.* 100**,** 631-638.

Mokshina, N., Gorshkov, O., Ibragimova, N., Chernova, T., and Gorshkova, T. (2017). Cellulosic fibres of flax recruit both primary and secondary cell wall cellulose synthases during deposition of thick tertiary cell walls and in the course of graviresponse. *Funct. Plant Biol.* 44**,** 820-831.

Mokshina, N., Gorshkova, T., and Deyholos, M.K. (2014). Chitinase-Like and Cellulose Synthase Gene Expression in Gelatinous-Type Cellulosic Walls of Flax (*Linum usitatissimum* L.) Bast Fibers. *PLOS ONE* 9**,** e97949.

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.

Moxley, G., Zhu, Z., and Zhang, Y.H.P. (2008). Efficient Sugar Release by the Cellulose Solvent-Based Lignocellulose Fractionation Technology and Enzymatic Cellulose Hydrolysis. *JAFC* 56**,** 7885-7890.

Mutwil, M., Debolt, S., and Persson, S. (2008). Cellulose synthesis: a complex complex. *Curr. Opin. Plant Biol.* 11**,** 252-257.

Nakano, Y., Yamaguchi, M., Endo, H., Rejab, N.A., and Ohtani, M. (2015). NAC-MYB-based transcriptional regulation of secondary cell wall biosynthesis in land plants. *Front. Plant Sci.* 6.

Nemhauser, J.L., Hong, F., and Chory, J. (2006). Different Plant Hormones Regulate Similar Processes through Largely Nonoverlapping Transcriptional Responses. *Cell* 126**,** 467-475.

Neutelings, G. (2011). Lignin variability in plant cell walls: Contribution of new models. *Plant Sci.* 181**,** 379-386.

Nguema-Ona, E., Vicré, M., Gotté, M., Plancot, B., Lerouge, P., Bardor, M.,

et al. (2014). Cell wall O-glycoproteins and N-glycoproteins: biosynthesis and some functional aspects. *Front. Plant Sci.* 5.

Nieminen, K., Blomster, T., Helariutta, Y., and Mähönen, A.P. (2015). Vascular Cambium Development. *The Arabidopsis Book*, e0177.

Novo-Uzal, E., Fernández-Pérez, F., Herrero, J., Gutiérrez, J., Gómez-Ros, L.V., Bernal, M.Á., *et al.* (2013). From Zinnia to Arabidopsis: approaching the involvement of peroxidases in lignification. *J. Exp. Bot.* 64, 3499-3518.

Oh, E., Zhu, J.Y., Bai, M.Y., Arenhart, R.A., Sun, Y., and Wang, Z.Y. (2014). Cell elongation is regulated through a central circuit of interacting transcription factors in the Arabidopsis hypocotyl. *eLife* 3, e03031.

Oinonen, P., Zhang, L., Lawoko, M., and Henriksson, G. (2015). On the formation of lignin polysaccharide networks in Norway spruce. *Phytochemistry* 111, 177-184.

Østergaard, L., Teilum, K., Mirza, O., Mattsson, O., Petersen, M., Welinder, K.G., *et al.* (2000). Arabidopsis ATP A2 peroxidase. Expression and high-resolution structure of a plant peroxidase with implications for lignification. *Plant Mol. Biol.* 44, 231-243.

Paque, S., Mouille, G., Grandont, L., Alabadí, D., Gaertner, C., Goyallon, A., *et al.* (2014). AUXIN BINDING PROTEIN1 Links Cell Wall Remodeling, Auxin Signaling, and Cell Expansion in *Arabidopsis*. *Plant Cell* 26, 280-295.

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.

Pauwels, L., and Goossens, A. (2011). The JAZ Proteins: A Crucial Interface in the Jasmonate Signaling Cascade. *The Plant Cell* 23, 3089-3100.

Pauwels, L., Morreel, K., De Witte, E., Lammertyn, F., Van Montagu, M., Boerjan, W., *et al.* (2008). Mapping methyl jasmonate-mediated transcriptional reprogramming of metabolism and cell cycle progression in cultured Arabidopsis cells. *Proc. Natl. Acad. Sci. U.S.A.* 105**,** 1380-1385.

Pedreira, J., Herrera, M.T., Zarra, I., and Revilla, G. (2011). The overexpression of AtPrx37, an apoplastic peroxidase, reduces growth in Arabidopsis. *Physiol. Plant.* 141**,** 177-187.

Persson, S., Caffall, K.H., Freshour, G., Hilley, M.T., Bauer, S., Poindexter, P., *et al.* (2007). The *Arabidopsis* irregular xylem8 Mutant Is Deficient in Glucuronoxylan and Homogalacturonan, Which Are Essential for Secondary Cell Wall Integrity. *Plant Cell* 19**,** 237-255.

Persson, S., Wei, H., Milne, J., Page, G.P., and Somerville, C.R. (2005). Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proc. Natl. Acad. Sci. U.S.A.* 102**,** 8633-8638.

Pierik, R., Djakovic-Petrovic, T., Keuskamp, D.H., De Wit, M., and Voesenek, L.a.C.J. (2009). Auxin and Ethylene Regulate Elongation Responses to Neighbor Proximity Signals Independent of Gibberellin and DELLA Proteins in Arabidopsis. *Plant Physiol.* 149**,** 1701-1712.

Pinzon-Latorre, D., and Deyholos, M.K. (2014). Pectinmethylesterases (PME) and Pectinmethylesterase Inhibitors (PMEI) Enriched during Phloem Fiber Development in Flax (*Linum usitatissimum*). *PLOS ONE* 9**,** e105386.

Qin, Y.M., Hu, C.Y., Pang, Y., Kastaniotis, A.J., Hiltunen, J.K., and Zhu, Y.X. (2007). Saturated Very-Long-Chain Fatty Acids Promote Cotton Fiber and Arabidopsis Cell Elongation by Activating Ethylene Biosynthesis. *Plant Cell* 19**,** 3692-3704.

Ragni, L., and Hardtke, C.S. (2014). Small but thick enough-the *Arabidopsis*

hypocotyl as a model to study secondary growth. *Physiol. Plant.* 151**,** 164-171.

Ralph, J., Brunow, G., Harris, P.J., Dixon, R.A., Schatz, P.F., and Boerjan, W. (2008). "Lignification: are Lignins Biosynthesized via simple Combinatorial Chemistry or via Proteinaceous Control and Template Replication?," in *Recent Advances in Polyphenol Research*. Wiley-Blackwell, 36-66.

Razous, P. (1944). L'évolution de l'agriculture française métropolitaine à travers l'histoire. *Journal de la société statistique de Paris* 85**,** 56-86.

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.

Roach, M.J., Mokshina, N.Y., Badhan, A., Snegireva, A.V., Hobson, N., Deyholos, M.K., *et al.* (2011). Development of Cellulosic Secondary Walls in Flax Fibers Requires ß-Galactosidase. *Plant Physiol.* 156**,** 1351-1363.

Ruel, K., Chabannes, M., Boudet, A.M., Legrand, M., and Joseleau, J.P. (2001). Reassessment of qualitative changes in lignification of transgenic tobacco plants and their impact on cell wall assembly. *Phytochemistry* 57**,** 875-882.

Saks, Y., Feigenbaum, P., and Aloni, R. (1984). Regulatory Effect of Cytokinin on Secondary Xylem Fiber Formation in an *In Vivo* System. *Plant Physiol.* 76**,** 638-642.

Salentijn, E.M.J., Zhang, Q., Amaducci, S., Yang, M., and Trindade, L.M. (2015). New developments in fiber hemp (*Cannabis sativa* L.) breeding. *Ind. Crop. Prod.* 68**,** 32-41.

Salnikov, V.V., Ageeva, M.V., and Gorshkova, T.A. (2008). Homofusion of

Golgi secretory vesicles in flax phloem fibers during formation of the gelatinous secondary cell wall. *Protoplasma* 233**,** 269-273.

Sánchez-Rodríguez, C., Bauer, S., Hématy, K., Saxe, F., Ibánez, A.B., Vodermaier, V., *et al.* (2012). CHITINASE-LIKE1/POM-POM1 and Its Homolog CTL2 Are Glucan-Interacting Proteins Important for Cellulose Biosynthesis in Arabidopsis. *Plant Cell* 24**,** 589-607.

Sánchez-Rodríguez, C., Rubio-Somoza, I., Sibout, R., and Persson, S. (2010). Phytohormones and the cell wall in *Arabidopsis* during seedling growth. *Trends Plant Sci.* 15**,** 291-301.

Schluttenhofer, C., and Yuan, L. (2017). Challenges towards Revitalizing Hemp: A Multifaceted Crop. *Trends Plant Sci.* 22**,** 917-929.

Schrader, J., Nilsson, J., Mellerowicz, E., Berglund, A., Nilsson, P., Hertzberg, M., *et al.* (2004). A High-Resolution Transcript Profile across the Wood-Forming Meristem of Poplar Identifies Potential Regulators of Cambial Stem Cell Identity. *Plant Cell* 16**,** 2278-2292.

Schuetz, M., Benske, A., Smith, R.A., Watanabe, Y., Tobimatsu, Y., Ralph, J., *et al.* (2014). Laccases Direct Lignification in the Discrete Secondary Cell Wall Domains of Protoxylem. *Plant Physiol.* 166**,** 798-807.

Schuetz, M., Smith, R., and Ellis, B. (2013). Xylem tissue specification, patterning, and differentiation mechanisms. *J. Exp. Bot.* 64**,** 11-31.

Sehr, E.M., Agusti, J., Lehner, R., Farmer, E.E., Schwarz, M., and Greb, T. (2010). Analysis of secondary growth in the *Arabidopsis* shoot reveals a positive role of jasmonate signalling in cambium formation. *Plant J.* 63**,** 811-822.

Shibaoka, H. (1994). Plant Hormone-Induced Changes in the Orientation of

Cortical Microtubules: Alterations in the Cross-linking Between Microtubules and the Plasma Membrane. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 45**,** 527-544.

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.

Smalle, J., Haegman, M., Kurepa, J., Van Montagu, M., and Straeten, D.V.D. (1997). Ethylene can stimulate Arabidopsis hypocotyl elongation in the□light. *Proc. Natl. Acad. Sci. U.S.A.* 94**,** 2756-2761.

Smith, R.A., Schuetz, M., Roach, M., Mansfield, S.D., Ellis, B., and Samuels, L. (2013). Neighboring Parenchyma Cells Contribute to Arabidopsis Xylem Lignification, while Lignification of Interfascicular Fibers Is Cell Autonomous. *Plant Cell* 25**,** 3988-3999.

Snegireva, A., Chernova, T., Ageeva, M., Lev-Yadun, S., and Gorshkova, T. (2015). Intrusive growth of primary and secondary phloem fibres in hemp stem determines fibre-bundle formation and structure. *AoB Plants* 7.

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 Physl+.* 57**,** 342-355.

Snegireva, A.V., Ageeva, M.V., Vorob'ev, V.N., Anisimov, A.V., and Gorshkova, T.A. (2006). Plant fiber intrusive growth characterized by NMR method. *Russian Journal of Plant Physiol.* 53**,** 163-168.

Song, D., Shen, J., and Li, L. (2010). Characterization of cellulose synthase complexes in Populus xylem differentiation. *New Phytol.* 187**,** 777-790.

Spartz, A.K., Ren, H., Park, M.Y., Grandt, K.N., Lee, S.H., Murphy, A.S., *et al.* (2014). SAUR Inhibition of PP2C-D Phosphatases Activates Plasma Membrane H^+-ATPases to Promote Cell Expansion in *Arabidopsis*. *Plant Cell* 26**,** 2129-2142.

Stips, A., Macias, D., Coughlan, C., Garcia-Gorriz, E., and Liang, X.S. (2016). On the causal structure between CO2 and global temperature. *Sci. Rep.* 6**,** 21691.

Struik, P.C., Amaducci, S., Bullard, M.J., Stutterheim, N.C., Venturi, G., and Cromack, H.T.H. (2000). Agronomy of fibre hemp (*Cannabis sativa* L.) in Europe. *Ind Crops Prod.* 11**,** 107-118.

Sundell, D., Street, N.R., Kumar, M., Mellerowicz, E.J., Kucukoglu, M., Johnsson, C., *et al.* (2017). AspWood: High-Spatial-Resolution Transcriptome Profiles Reveal Uncharacterized Modularity of Wood Formation in *Populus tremula*. *Plant Cell* 29**,** 1585-1604.

Takahashi, K., Hayashi, K.-I., and Kinoshita, T. (2012). Auxin Activates the Plasma Membrane H+-ATPase by Phosphorylation during Hypocotyl Elongation in *Arabidopsis*. *Plant Physiol.* 159**,** 632-641.

Terashima, N., and Fukushima, K. (1988). Heterogeneity in formation of lignin-XI: An autoradiographic study of the heterogeneous formation and structure of pine lignin. *Wood Sci. Technol.* 22**,** 259-270.

Timell, T.E. (1967). Recent progress in the chemistry of wood hemicelluloses. *Wood Sci. Technol.* 1**,** 45-70.

Tromas, A., Paque, S., Stierlé, V., Quettier, A.-L., Muller, P., Lechner, E., *et al.* (2013). Auxin-Binding Protein 1 is a negative regulator of the $SCF^{TIR1/AFB}$ pathway. *Nature Commun.* 4**,** 2496.

Turlapati, P., Kim, K.W., Davin, L., and Lewis, N. (2011). The laccase multigene family in Arabidopsis thaliana: towards addressing the mystery of their gene function(s). *Planta* 233**,** 439-470.

Ursache, R., Nieminen, K., and Helariutta, Y. (2013). Genetic and hormonal regulation of cambial development. *Physiol. Plant.*147**,** 36-45.

Van De Meene, A.M.L., Doblin, M.S., and Bacic, A. (2017). The plant secretory pathway seen through the lens of the cell wall. *Protoplasma* 254**,** 75-94.

Van Den Broeck, H.C., Maliepaard, C., Ebskamp, M.J.M., Toonen, M.a.J., and Koops, A.J. (2008). Differential expression of genes involved in C1 metabolism and lignin biosynthesis in wooden core and bast tissues of fibre hemp (*Cannabis sativa* L.). *Plant Sci.* 174**,** 205-220.

Verhertbruggen, Y., Marcus, S.E., Chen, J., and Knox, J.P. (2013). Cell Wall Pectic Arabinans Influence the Mechanical Properties of Arabidopsis thaliana Inflorescence Stems and Their Response to Mechanical Stress. *Plant Cell Physiol.* 54**,** 1278-1288.

Voxeur, A., Wang, Y., and Sibout, R. (2015). Lignification: different mechanisms for a versatile polymer. *Curr. Opin. Plant Biol.* 23**,** 83-90.

Wang, F., Cui, X., Sun, Y., and Dong, C.-H. (2013). Ethylene signaling and regulation in plant growth and stress responses. *Plant Cell Rep.* 32**,** 1099-1109.

Wasternack, C., and Song, S. (2017). Jasmonates: biosynthesis, metabolism, and signaling by proteins activating and repressing transcription. *J. Exp. Bot.* 68**,** 1303-1321.

Xiao, C., Zhang, T., Zheng, Y., Cosgrove, D.J., and Anderson, C.T. (2016).

Xyloglucan Deficiency Disrupts Microtubule Stability and Cellulose Biosynthesis in Arabidopsis, Altering Cell Growth and Morphogenesis. *Plant Physiol.* 170**,** 234-249.

Xiong, G., Cheng, K., and Pauly, M. (2013). Xylan O-Acetylation Impacts Xylem Development and Enzymatic Recalcitrance as Indicated by the Arabidopsis Mutant tbl29. *Mol. Plant* 6**,** 1373-1375.

Yamaguchi, M., Mitsuda, N., Ohtani, M., Ohme-Takagi, M., Kato, K., and Demura, T. (2011). VASCULAR-RELATED NAC-DOMAIN 7 directly regulates the expression of a broad range of genes for xylem vessel formation. *Plant J.* 66**,** 579-590.

Yan Zhong, G., and Burns, J.K. (2003). Profiling ethylene-regulated gene expression in *Arabidopsis thaliana* by microarray analysis. *Plant Mol. Biol.* 53**,** 117-131.

Yu, Y., Wang, J., Zhang, Z., Quan, R., Zhang, H., Deng, X.W., Ma, L., and Huang, R. (2013). Ethylene Promotes Hypocotyl Growth and HY5 Degradation by Enhancing the Movement of COP1 to the Nucleus in the Light. *PLOS Genet.* 9**,** e1004025.

Zhang, F., Yao, J., Ke, J., Zhang, L., Lam, V.Q., Xin, X.-F., *et al.* (2015). Structural basis of JAZ repression of MYC transcription factors in jasmonate signaling. *Nature* 525**,** 269-273.

Zhang, L.Y., Bai, M.Y., Wu, J., Zhu, J.Y., Wang, H., Zhang, Z., *et al.* (2009). Antagonistic HLH/bHLH Transcription Factors Mediate Brassinosteroid Regulation of Cell Elongation and Plant Development in Rice and *Arabidopsis*. *Plant Cell* 21**,** 3767-3780.

Zhao, Q., Gallego-Giraldo, L., Wang, H., Zeng, Y., Ding, S.Y., Chen, F., *et al.* (2010). An NAC transcription factor orchestrates multiple features of cell wall

development in *Medicago truncatula*. *Plant J.* 63**,** 100-114.

Zhao, Q., Nakashima, J., Chen, F., Yin, Y., Fu, C., Yun, J., *et al.* (2013). LACCASE Is Necessary and Nonredundant with PEROXIDASE for Lignin Polymerization during Vascular Development in *Arabidopsis*. *Plant Cell* 25**,** 3976-3987.

Zhong, R., and Ye, Z.H. (2015). Secondary Cell Walls: Biosynthesis, Patterned Deposition and Transcriptional Regulation. *Plant Cell Physiol.* 56**,** 195-214.

Zhou, J., Zhong, R., and Ye, Z.H. (2014). Arabidopsis NAC Domain Proteins, VND1 to VND5, Are Transcriptional Regulators of Secondary Wall Biosynthesis in Vessels. *PLOS ONE* 9**,** e105726.

Zhou, M., and Memelink, J. (2016). Jasmonate-responsive transcription factors regulating plant secondary metabolism. *Biotechnol. Adv.* 34**,** 441-449.
Zhu, J.Y., Sae-Seaw, J., and Wang, Z.Y. (2013). Brassinosteroid signalling. *Development* 140**,** 1615-1620.

Chapter 1

Studying secondary growth and bast fibre development: the hemp hypocotyl peeks behind the wall

Adapted from the published article: Marc Behr, Sylvain Legay, Eva Žižková, Václav Motyka, Petre I Dobrev, Jean-Francois Hausman, Stanley Lutts, Gea Guerriero. Studying secondary growth and bast fibre development: the hemp hypocotyl peeks behind the wall. 2016. *Frontiers in Plant Science* 7:1733.

Author contributions:

Marc Behr analysed the RNA-Seq data and phytohormone profiles, performed RT-qPCR, optical microscopy and immunohistochemistry experiments and wrote the manuscript. Sylvain Legay participated in the RNA-Seq experiment and performed the bioinformatic analysis. Eva Žižková, Václav Motyka and Petre Dobrev performed the phytohormone analysis. Jean-François Hausman and Stanley Lutts conceived the experiment. Gea Guerriero conceived the experiment, performed the RNA-Seq, and wrote the manuscript.

Abstract

C*annabis sativa* L. is an annual herbaceous crop grown for the production of long extraxylary fibres, the bast fibres, rich in cellulose and used both in the textile and biocomposite sectors. This Chapter focuses on the mechanisms marking the transition from primary to secondary growth in the hemp hypocotyl by analysing the suite of events accompanying vascular tissue and bast fibre development (i.e. 6-9-15-20 days after sowing). Transcriptomics provides cell wall-related snapshots of the growing hemp hypocotyl and identifies marker genes associated with the young (expansins, β-galactosidases and transcription factors involved in light-related processes) and the older hypocotyl (transcription factors and genes involved in cellulose, xylan and lignin biosyntheses, as well as in secondary growth). Imaging highlights the different steps of cell wall formation in bast fibres and xylem, with massive deposition of cellulose in the secondary cell wall and differential xylan abundance between these two tissues. Finally, we suggest that the phenotypes and gene expression patterns are regulated by phytohormones. Jasmonic acid may positively regulate cambial activity, while abscisic acid may have a key role in the organisation of the cytoskeleton and secondary cell wall formation.

Introduction

Hemp (*Cannabis sativa* L.) is a herbaceous plant native to Central Asia known for its medicinal and textile applications for over 8000 years. There is currently a renewed interest in hemp, because of its multitude of applications, notably in the construction and biocomposite sectors (Andre *et al.* 2016, Guerriero *et al.* 2016). As an example, the woody fibres of industrial hemp are used to manufacture a concrete-like material known as “hempcrete”, which is very light in weight, while the cellulosic fibres find application in the biocomposite sector for the creation of bioplastics.

The stem of *C. sativa* provides two types of fibres, i.e. the xylem and the phloem fibres. The latter are known as bast fibres. Average straw yield lies between 7 and 10 tons per hectare, for a bast fibre content of 30-40%. The cell walls of xylem fibres are impregnated with lignin to ensure strength and resistance to negative sap pressure (Wang *et al.*, 2013), while those of the bast fibres are mainly made of crystalline cellulose (and can account for up to 75-80% of the bast fibre dry mass; Guerriero *et al.* 2013). The other components of bast fibre cell walls are hemicelluloses (4%, including xyloglucan in the primary cell wall and xylan in the secondary cell wall), pectins (4%), proteins (3%), lignin (2%) and traces of phenolic acids (<0.01%) (Crônier *et al.*, 2005).

The cell wall is a highly dynamic and complex compartment, because its composition undergoes massive modifications as the plant develops (Cosgrove, 2005; Zhang *et al.*, 2014). A typical example is the transition from primary to secondary growth in plants, a process requiring the shift from a phase characterized by active elongation to a stage of stem thickening (van Raemdonck *et al.*, 2005; Zhong and Ye,

2007). Secondary growth is a relevant process, as it generates wood, an important renewable resource for humankind (Guerriero *et al.*, 2014). Following the primary elongation phase, it is responsible for the radial growth of the stem and the root. Secondary growth is a complex phenomenon to study from a molecular perspective, as it is tuned by both exogenous factors (Moura *et al.*, 2010) and phytohormones (via transcriptional master regulators; Didi *et al.*, 2015). The diversity of cell wall composition has inspired the molecular study of various models for secondary growth, from trees (Tocquard *et al.*, 2014) to specific tissues of herbaceous plants (Zhong *et al.*, 2008; Taylor-Teeples *et al.*, 2015).

In trees, the annual girth increase is the result of the activity of a secondary meristem, the vascular cambium. Both the vascular cambium and the genes involved in secondary growth are not restricted to trees, but are present in herbaceous species such as the small weed *Arabidopsis thaliana*. Secondary growth has been extensively studied in both the hypocotyl and root of this model organism (Sibout *et al.*, 2008). In the *A. thaliana* hypocotyl the phases of elongation and girth increase are temporally uncoupled (Ragni *et al.* 2011), which enables experimenters to focus on different cell wall-related aspects of the same structure.

Among herbaceous plants, fibre crops also present secondary growth. In particular, the hypocotyl of flax, *Linum usitatissimum* L., was shown to be a valid system to study the molecular processes occurring during the stages of elongation and subsequent thickening of bast fibres (Roach and Deyholos, 2008). The authors showed indeed that bast fibre development in adult flax stems occurs following the same two stages of elongation and secondary wall deposition arising in the young hypocotyls. These two systems are therefore complementary for

the study of bast fibre development. In the adult stem of fibre crops, a lignification gradient is observed: the basal internodes of the stem are more lignified than the younger elongating ones. The transition between the rapidly elongating and the lignified internodes of the stem occurs at a characteristic spot in flax, the so-called “snap point” (Gorshkova *et al.*, 2003). The snap point therefore physically marks a zone below which the mechanical properties of bast fibres change considerably (Gorshkova *et al.*, 2003). Hemp stems show the same basipetal gradient of lignification, with older regions of the stems characterized by a well-developed xylem and both primary and secondary bast fibres (Crônier *et al.*, 2005). Notably, the developing hemp hypocotyl shows the same tissue organization, i.e. at younger stages the cells are elongating, while at older ones the tissues are more lignified, with the appearance of secondary bast fibres. As these cells originate from the cambium rather than the pro-cambium as in flax (Snegireva *et al.*, 2015), the hemp hypocotyl is suitable to study their biosynthesis from a molecular point of view.

With the aim of proposing an alternative simple system to study secondary growth, the suitability of the hemp hypocotyl is here validated. In line with the study of Roach and Deyholos (2008), four sequential time points (6-9-15-20 days after sowing) are here investigated using a cross-disciplinary approach: transcriptomics (RNA-Seq) and the quantification of phytohormones have been coupled to optical and confocal microscopy observations. The goal is to decipher the molecular events involved in the radial growth and bast fibre development of the hemp hypocotyl. The results here shown contribute to add fundamental knowledge to the study of secondary growth and fibre formation in an economically important plant.

I. Material and Methods

a. Plant material and RNA extraction

C. sativa (cv. Santhica 27) hypocotyls of 6, 9, 15 and 20 days after sowing were grown in controlled conditions in incubators following a cycle of 16h light 25°C / 8h dark 20°C. Three biological replicates per time point were analysed. Each biological replicate consisted of 20 hypocotyls randomly selected among all incubators. The pooling of hypocotyls was needed because of the quantity of material required for transcriptomics. In order to bring to a minimum this pooling bias, a high number of hypocotyls were pooled together. By doing so, the power to detect differentially expressed genes within the four populations increased (Rajkumar *et al.*, 2015). Sampling was performed on a single experimental batch to minimize technical variability. Samples were immediately frozen in liquid nitrogen and conserved at -80°C until RNA extraction. The sampled hypocotyls were crushed to a fine powder using a mortar, a pestle and liquid nitrogen. Total RNA was extracted using the RNeasy Plant Mini Kit (Qiagen), treated with DNase I on column, quantified and quality-checked using a Qubit 2.0 Fluorometer (Invitrogen) with Qubit RNA Assay Kit (Molecular Probes), a NanoDrop 1000 Spectrophotometer (Thermo Scientific) and a 2100 Bioanalyzer (Agilent Life Sciences). All the RNAs displayed a RIN above 7.

b. Library preparation and sequencing

Libraries were prepared from 3µg of total RNA using the SMARTer Stranded RNA-Seq kit (Clontech). The isolation of mRNAs was performed using the Illumina beads and the TruSeq protocol (Illumina). The final elution of the mRNAs was performed using the Illumina elution buffer (19.5 µl). The isolated mRNAs were quantified using a Qubit fluorometer, as described above. Ten ng mRNA were used for

the cDNA synthesis and shearing, following the manufacturer's instruction. Indexing was performed using the Illumina indexes 1-12. The enrichment step was carried out using 12 cycles of PCR. Subsequently, libraries were checked using 2100 Bioanalyzer (DNA High sensitivity Kit) to evaluate the mean fragment size. Quantification was performed using the KAPA library quantification kit (KAPA Biosystems) using a ViiA7 Real-Time PCR System (Life technologies). The pooled libraries (20 pM) were sequenced on an Illumina MiSeq in 6 consecutive runs (MiSeq reagent kit V3, 150 cycles) generating 76 base pairs paired-end reads. Raw sequences have been deposited at the NCBI Gene Expression Omnibus (GEO), http://www.ncbi.nlm.nih.gov/geo, accession number: GSE85144.

c. Assembly, mapping and data analysis

Raw sequences reads were uploaded in CLC Genomics Workbench 8.0.3. Sequences were filtered and trimmed as follows: sequence length > 55bp, sequence quality score < 0.01, no ambiguity in the sequence, trimming using Illumina adaptors, hard trim of 10 base pairs (bp) at the 5' end and 3 bp at the 3' end, resulting in final sequence average length of 61 bp. Duplicated reads were removed from each library using the duplicate read removal plugin. The *de novo* assembly was performed with a wording size ranging from 20 to 54. The reads were mapped back to the assembly with a mismatch, insertion and deletion cost of 3, a coverage > 0.8 and similarity > 0.95. The optimal parameters were obtained with the auto-wording mode (24), bubble sizing in automatic mode. The assembly was then annotated using Blast2GO PRO version 3.0 against the *Arabidopsis thaliana* non-redundant database. For each library, the mapping was performed with CLC Genomics Workbench 8.0.3 according to the following criteria: a maximum hit per reads of 3, similarity fraction > 0.95, a length fraction

> 0.7, a mismatch, insertion and deletion cost of 3. Mapping (using the same above-described parameters) was also performed using the available transcriptome for the variety Finola (van Bakel *et al.* 2011), which was annotated using Blast2GO PRO version 3.0 against the *A. thaliana* non-redundant database. The expression values were then calculated using the RPKM method (Mortazavi *et al.* 2008). Genes with less than 10 mapped reads, with no specifically mapped reads in at least one of the libraries were removed from the dataset. In order to highlight the differentially expressed genes, an ANOVA with 4 groups (H6, H9, H15, H20) composed by three biological replicates was performed. A false discovery rate (FDR) correction was then applied to the dataset and genes with FDR corrected *p*-value below 0.05 were selected. Hard cut-offs were finally performed on the fold change (FC absolute value > 2) and the RPKM minimum difference (RPKM difference > 10).

Independent Component Analysis (ICA) was performed to assess the differences between the biological replicates within the online service MetaGeneAlyse v1.7.1 (Daub *et al.*, 2003). MapMan functional annotations of H6 and H20 were performed with a subset of genes (-2 > fold-change > 2) showing a TAIR ID. Gene ontology term enrichment (GOE) analysis of H6 and H20 was realised using locus with -2 > log2 FC > 2 using ClueGO (v2.1.1) and CluePedia (v1.1.1) (Bindea *et al.*, 2009, 2013) within Cytoscape (v3.2.1) with the following parameters: gene ontology from level 3 to level 9, using All_Experimental_evidence, kappa score set at 0.2, Benjamini-Hochberg correction. The presence of a signal peptide cleavage site in the proteins presumably excreted to the wall was determined with PrediSi (www.predisi.de).

d. RT-qPCR validation

The RNA-Seq data were validated with RT-qPCR using a subset of 10 genes involved in cell wall biogenesis. Reverse transcription was carried out with the ProtoScript II Reverse Transcriptase (NEB) following the manufacturer's instructions. Primers were designed with Primer3 and validated for the absence of dimers and secondary structures (hairpin) using OligoAnalyzer 3.1 (http://eu.idtdna.com/calc/analyzer). qPCR runs were performed in 384 well-plates, on a ViiA7 Real-Time PCR System (Applied Biosystems) with the Takyon SYBR Green low ROX (Eurogentec). A melt curve was realised at the end of each experiment to ensure the specificity of the products. Relative gene expressions were determined with the qBasePLUS software v2.5 (Biogazelle). The geNorm analysis designated *CsaETIF4E* and *CsaGAPDH* as the most suitable genes for normalisation (among *Histone*, *EF2*, *Actin*, *Cyclophilin*, *Ubiquitin*, *GAPDH*, *Tubulin*, *ETIF4E*, *ETIF3H* and *ETIF3E*). Calibrated Normalised Relative Quantities (CNRQ) from RT-qPCR were used for the final comparison with the RPKM values.

e. Phytohormones quantification

The concentrations of phytohormones were determined with the method of Dobrev and Vankova (2012) using technical duplicates on biological triplicates (in H6, H9, H15 and H20). Briefly, 5-10 mg of lyophilised material was extracted with methanol/formic acid/water (15:1:4; v/v) with prior addition of 10 pmol [^{2}H]-labelled internal standards described previously (Djilianov *et al.*, 2013). Extracts were evaporated in vacuum concentrator and purified through mixed mode reversed phase – cation exchange SPE column (Oasis-MCX, Waters). The first fraction containing ABA, auxins and JAs and the second

fraction containing cytokinins were concentrated to dryness prior to LC-MS. An aliquot (10 µl) of purified sample was analyzed on LC-MS consisting of HPLC (Ultimate 3000) coupled to hybrid triple quadrupole/linear ion trap mass spectrometer (3200 Q TRAP) set in selected reaction monitoring mode. Quantification of hormones was done using isotope dilution method with multilevel calibration curves. Data processing was carried out with Analyst 1.5 software (Applied Biosystems).

f. Resin-embedded microscopy and immunohistochemistry

Hemp hypocotyls were embedded in Technovit 7100 resin (Kulzer). Briefly, sections of 5 mm were fixed in glutaraldehyde/paraformaldehyde/caffeine (1% / 2% / 1% v/v in Milli-Q water) under vacuum for 15 minutes and 24h at 4°C, dehydrated in an ethanol series (70%-95%-100%), impregnated in resin containing PEG 400 (2% v/v) and dimethacrylate ethylene glycol (0.4% w/v) and finally included. Cross sections of 10 µm thickness were cut using a microtome (Leica) and stained with toluidine blue or used for immunohistochemistry (IHC). Image acquisition was performed with a Leica DMR for toluidine blue and with a confocal microscope LSM 510 Meta (Zeiss) for IHC.

LM5 (β-1,4-galactan), LM10 (xylan), LM15 (xyloglucan) (Plant Probes) and RU2 (INRA) antibodies were diluted 10-fold in milk protein (MP)/PBS (5% w/v). Sections were then incubated for 1.5h, rinsed three times in PBS and incubated for 1.5h with the anti-rat IgG coupled to FITC (Sigma) diluted 100-fold in MP/PBS. Before observation, three washings with PBS were performed. CBM3a (crystalline cellulose, Plant Probes) was diluted to 10 µg/mL in MP/PBS, incubated in mouse

anti-His monoclonal antibody (1% in MP/PBS, Sigma) and finally incubated in 50-fold diluted anti-mouse IgG coupled to FITC (Sigma). Each incubation lasted for 1.5h. Between each step, three washes with PBS were performed. The slides were mounted in Möwiol 4-88 (Sigma) and observed with the following settings: excitation at 488 nm, filter HFT 488/594 and emission recorded with LP 505. The microscope settings were kept rigorously constant between the different observations for a given epitope. Negative controls where either the primary or secondary antibody was omitted resulted in a very weak and negligible signal.

II. Results

a. Cell wall fingerprinting of the growing hemp hypocotyl

In order to study the development of the secondary tissues in hemp, hypocotyls aged between 6 (H6) and 20 days (H20) were studied. The development of stem tissues during development was observed on hypocotyl cross sections stained with toluidine blue (Figure 1-1). In H6 the cortical parenchyma constitutes the bulk of the section, with collenchyma acting as supporting tissue and medullar parenchyma occupying a limited area (Figure 1-1A). In H9 the cambium has started multiple divisions, with the primary xylem being more developed. Simultaneously, the cells of the medullary parenchyma have grown, leading to an equilibrium between this tissue and cortical parenchyma (Figure 1-1B). At this point, the hypocotyl has stopped its elongation and secondary growth takes place. In H12, vessels of the secondary xylem are visible (Figure 1-1C). In H20, bundles of phloem fibres are well visible (Figure 1-1F). Between H15 and H20 (Figure 1-1D-F), no major changes occur, secondary xylem being the most prominent

tissue and phloem bast fibres with thicker gelatinous walls getting distributed regularly between the cambial zone and the cortical parenchyma.

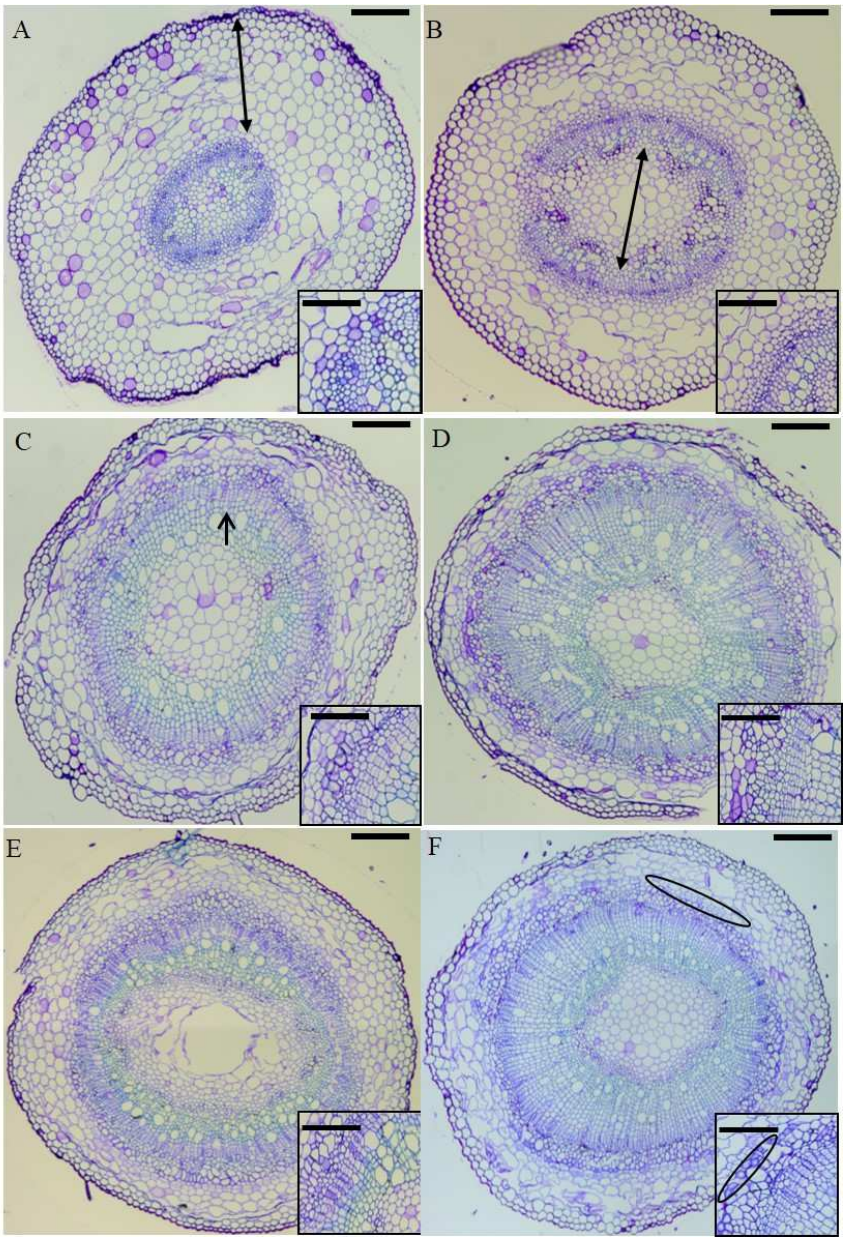


Figure 1-1 *(previous page)*: Cross-sections of H6 (**A**), H9 (**B**), H12 (**C**), H15 (**D**), H17 (**E**) and H20 (**F**) stained with toluidine blue. In **A**, the double-pointed arrow spans the cortical parenchyma. In **B**, the double-pointed arrow spans the medullary parenchyma. In **C**, the arrow points to the developing xylem. In **F**, the phloem bast fibres are circled in black. Scale bar: 200 µm in the main pictures, 50 µm in the insets. The figure insets show zoomed regions of the hypocotyls where bast fibres gradually differentiate.

Confocal microscope observations were carried out on 4 representative time points, i.e. H6, H9, H15 and H20. Monoclonal antibodies recognizing β-1,4-D-galactan (LM5), xyloglucan (LM15) and xylan (LM10) and His-tagged recombinant protein recognizing crystalline cellulose (CBM3a) were used to study the distribution of key cell wall components during hemp hypocotyl development. We decided to use these specific antibodies to: 1) be able to differentiate primary and secondary growth using “marker” cell wall components (in these specific case hemicelluloses, i.e. xyloglucan and xylan) and 2) to provide a visual inspection of the bast fibre development (by following cellulose accumulation).

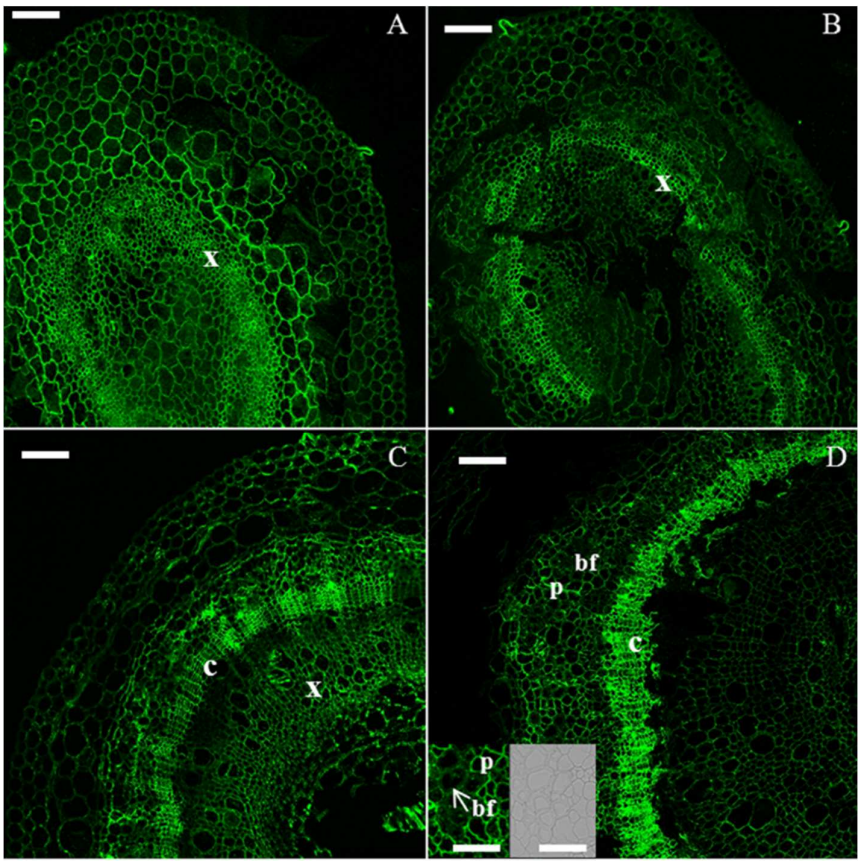


Figure 1-2: Immunodetection of the LM5 epitope specific for β-1,4-D-galactan. H6 (A), H9 (B), H15 (C), H20 (D). bf bast fibre, c cambium, p parenchyma, x xylem. Scale bar: 100 µm in the main pictures and 50 µm in the insets. The inset shows a zoomed detail of the cortex, where bast fibres are visible. For clarity of presentation, the corresponding differential interference contrast image is provided in panel D.

Galactan was detected in cambial cells and young xylem cells in H6 and H9 (Figure 1-2A-B), as well as in the parenchyma surrounding developing bast fibres in H15 and H20 (Figure 1-2C-D). The bast fibres did not show labelling (Figure 1-2D, inset).

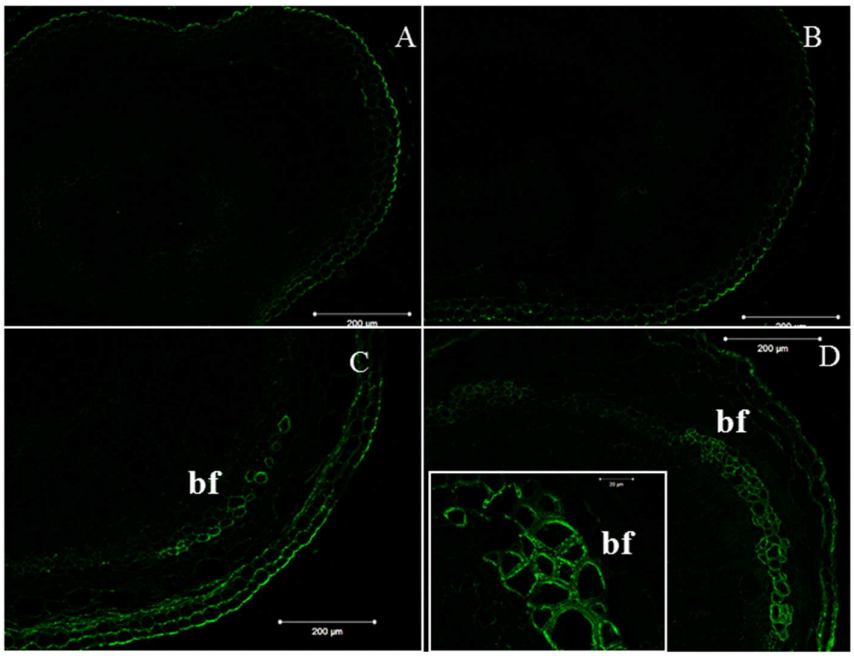


Figure 1-3: Immunodetection of the RU2 epitope specific for rhamnogalacturonan-1. H6 (A), H9 (B), H15 (C), H20 (D). bf bast fibre. The inset shows a zoomed detail of the bast fibres. Scale bar: 200µm in the main pictures, 20µm in the inset.

The rhamnogalacturonan-I epitope detected by RU2 was observed in the gelatinous layer of the bast fibres (Figure 1-3C-D). The signal is highly specific for this tissue, as it is barely visible in other parts of the cross-section, with the exception of the collenchyma.

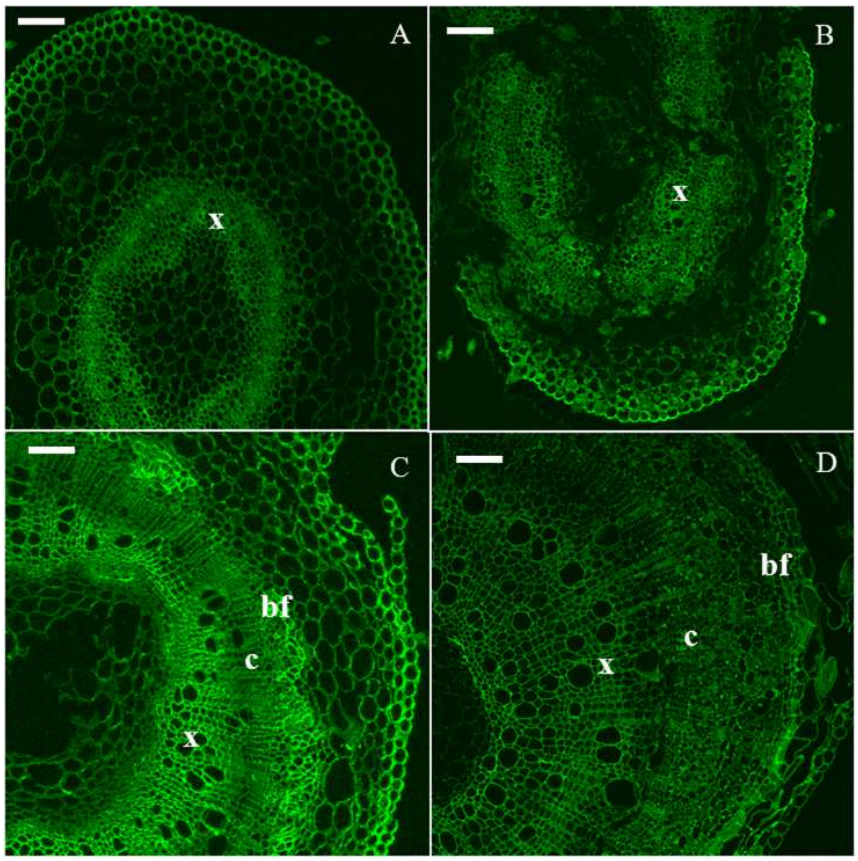


Figure 1-4: Immunodetection of the LM15 epitope specific for xyloglucan. H6 (A), H9 (B), H15 (C), H20 (D). c cambium, bf bast fibre, x xylem. Scale bar: 100 µm.

The xyloglucan epitope was detected in all the time points (Figure 1-4). In H6 and H9 (Figure 1-4A-B), the strongest signal was observed in developing xylem cells. In H15, the young primary bast fibres displayed an intense signal (Figure 1-4C). In older bast fibres, this signal decreased, leading to a comparable intensity in all the tissues of H20 (Figure 1-4D).

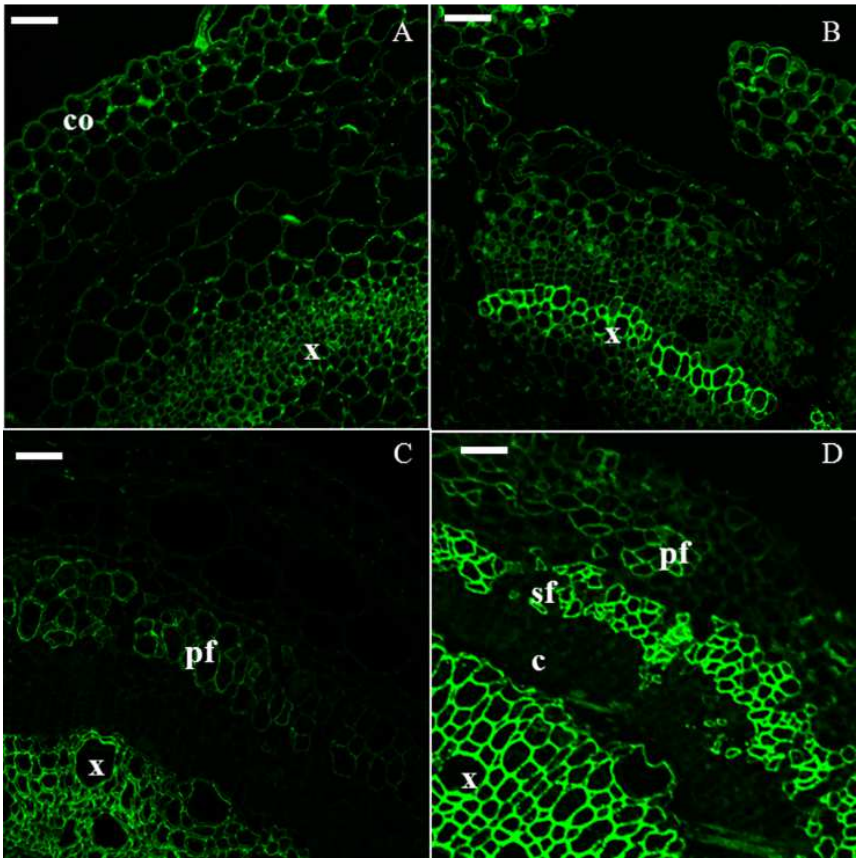


Figure 1-5: Immunodetection of the LM10 epitope specific for xylan. H6 (A), H9 (B), H15 (C), H20 (D). c cambium, co collenchyma, pf primary bast fibre, sf secondary bast fibre, x xylem. Scale bar: 50 µm.

The xylan epitope was associated with the xylem of H6 and H9 (Figure 1-5A-B). In H15, xylem cells were heavily stained and a weaker signal was detected in the cell wall of primary bast fibres (Figure 1-5C). In H20, secondary bast fibres displayed a signal which was more intense than the one detected in the primary bast fibres (Figure 1-5D).

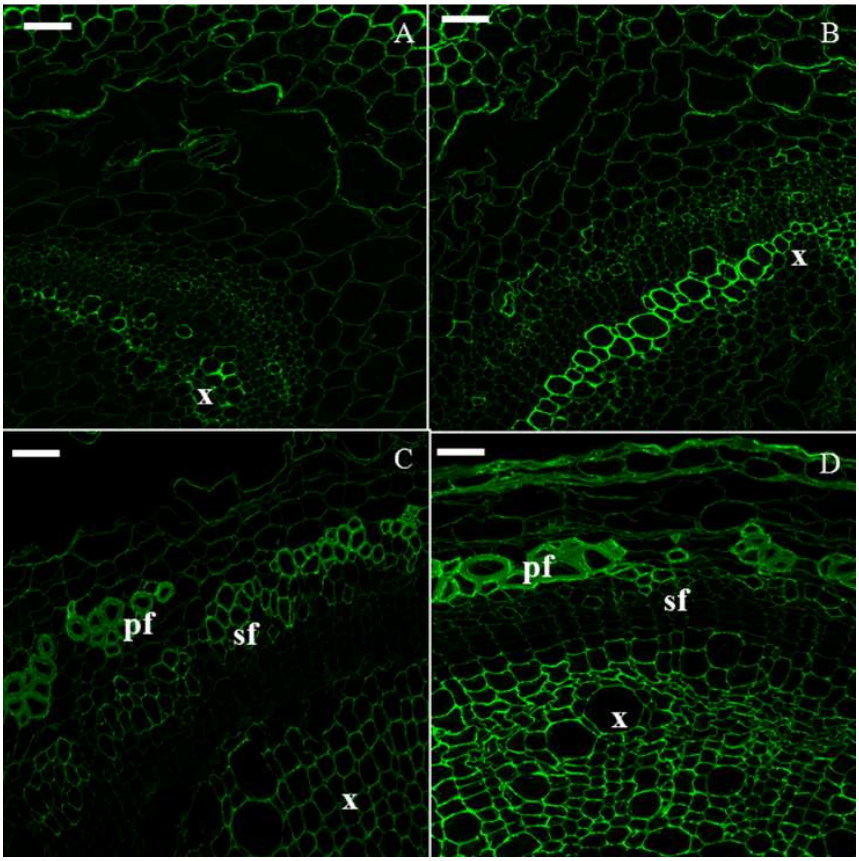


Figure 1-6: Immunodetection of the CBM3a epitope specific for crystalline cellulose. H6 (A), H9 (B), H15 (C), H20 (D). pf primary bast fibre, sf secondary bast fibre, x xylem. Scale bar: 50 µm.

Crystalline cellulose was visualised with CBM3a (Figure 1-6). It was observed in all the tissues at each time point, particularly in xylem cells in H6 and H9 (Figure 1-6A-B) and in bast fibres in H15 (Figure 1-6C) and H20 (Figure 1-6D). Indeed, in the older hypocotyl, the gelatinous walls of both primary and secondary bast fibres were strongly labelled by the antibody.

b. Transcriptome overview of the growing hemp hypocotyl

In order to get a detailed picture of the gene expression changes during the development of the hemp hypocotyl, RNA-Seq was performed on 4 representative time points (H6, H9, H15 and H20). As can be seen in Figure 1-1, these time-points correspond to key developmental stages of the hemp hypocotyl (elongation and secondary growth, development of bast fibres). After filtering and trimming, duplicated reads were removed (4-7% of the filtered reads). The sequences were assembled into 28433 contigs with size ranging from 282 to 15600 bp, an average length of 1136 bp and a N50 of 1660 bp. 84 to 88% of the reads were successfully mapped to the Santhica assembly with more than 99% of reads specifically mapped. A previous study based on microarrays of cDNA library (de Pauw *et al.*, 2007) represented 7053 independent genes, to be compared with the 28433 contigs obtained in our experiment. We may thus draw a more detailed picture of the biological processes occurring in hypocotyls with RNA-Seq. Our study also provides an accurate transcriptome assembly of another industrial hemp variety, Santhica 27, which will be of considerable interest and help to understand the growth and the development of hemp. After an ANOVA statistical analysis, 3622 differentially expressed contigs were selected. When mapping against the Finola assembly (van Bakel *et al.* 2011), 3096 contigs were differentially expressed.

The RNA-Seq results were verified by quantifying the gene expression levels of 10 representative genes by RT-qPCR. Genes involved at different levels of cell wall biogenesis, displaying significant RPKM fold change between the time-points and covering a large range of expression (from 3 to 1600+ RPKM), were chosen. A good

correlation was obtained between the RNA-Seq and RT-qPCR data (R^2 = 0.91) (Supplementary data 1-1).

Independent Component Analysis (ICA; Daub *et al.*, 2003) performed on RPKM values resulted in a good separation between each time-point when using two independent components (Figure 1-7A). The expression profiles of the 3622 differentially regulated genes (Supplementary data 1-2) were assessed by hierarchical clustering using a Euclidean distance matrix in complete linkage. Thirteen clusters were obtained, with the four most important (C1 to C4) accounting for 81% of the genes (Figure 1-6B). Genes up-regulated at 15 and 20 days were mainly distributed in C1 and C2. The transcripts more abundant at 6 and 9 days were distributed in C3 and C4 (Figure 1-7B).

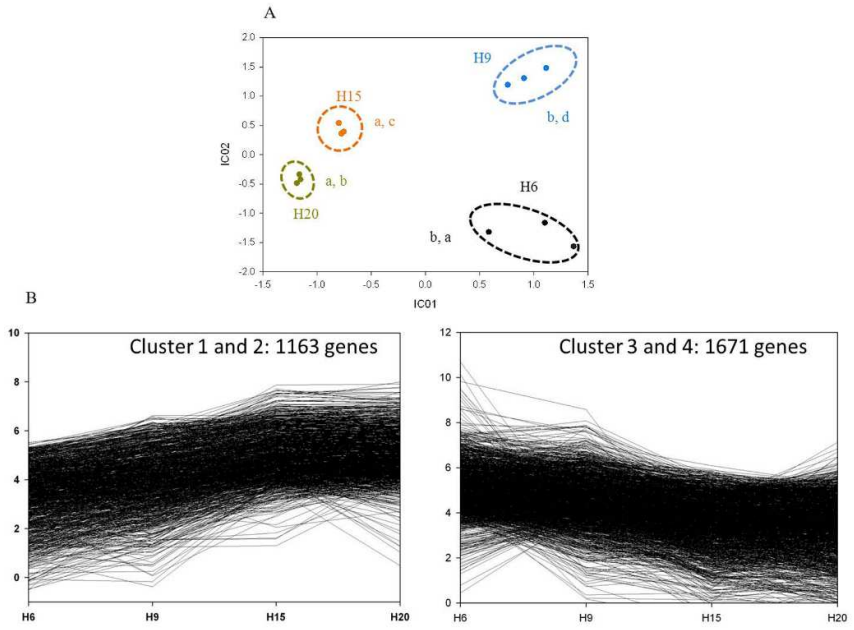


Figure 1-7: General analysis of H6, H9, H15 and H20 transcriptomes. A: ICA of H6 (black dots), H9 (blue dots), H15 (orange dots) and H20 (dark yellow

dots). Variance explained by the two independent components: 89.36%. The three biological replicates of each time point are surrounded by a dotted line. Letters correspond to homogeneous groups determined by the Tukey post hoc test on IC1 and IC2 coordinates respectively. B: clusters with up-regulated and down-regulated genes through the time-points.

c. Transcription factor inventory of the hemp hypocotyl

We looked for differentially expressed genes coding for putative transcription factors (TFs) in hemp by searching for their *Arabidopsis* orthologs (Jin *et al.*, 2014). A total of 131 contigs were obtained (Supplementary file 1-3). As compared to microarrays, the bioinformatic pipeline following RNA-Seq analysis annotates these genes, by homology with the protein sequence of Viridiplantae or the model plant *Arabidopsis*. The presence of an annotation is of considerable importance to understand the transcriptional regulation leading to a biological process, such as secondary growth or lignification. Most of the differentially expressed transcription factors (34%) were more expressed in H6, with proportions decreasing along time (26, 23 and 17% for H9, H15 and H20, respectively). The basic helix-loop-helix (bHLH) family was the most represented across the time-points studied, with 23 members, followed by the homeobox domain (HB) and C2H2 (12 members each).

We could discriminate between TFs expressed at earlier and later time-points and hence get a TF signature of the developing hemp hypocotyl. Following emergence from the soil, several TFs related to circadian rhythm, response to light, photosynthesis, pigment biosynthesis, cell expansion and vascular development were more abundant in H6 and H9 (Figure 1-8). More specifically, members of the

HB, bHLH, C2H2 and MYB families were abundantly expressed in H6 and H9.

The transcriptome of the older hemp hypocotyl was characterized, as expected, by a more abundant expression of TFs related to SCW biosynthesis and vascular development. Three master regulators of SCW deposition, namely *SND2*, *VND1* and *NST1* displayed the highest expression in H15, together with a second-layer master switch, *MYB46*, involved in SCW biogenesis and xylan biosynthesis (Zhong *et al., 2008*). *XND1* was more abundant in H20. As far as the TFs regulating cell wall composition are concerned, a higher expression of TFs involved in hemicellulose and lignin biosynthesis were found in H20, i.e. *KNAT7* and *MYB85*. An ortholog of *Arabidopsis WLIM1*, involved in fibre extension and lignification, was more abundant in H20.

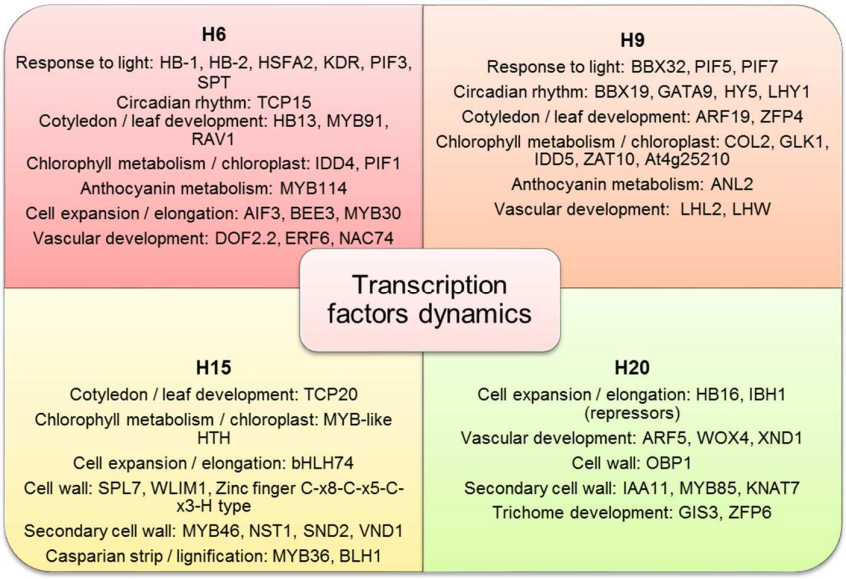


Figure 1-8: Representative functions and genes of TFs differentially regulated

d. Hemp hypocotyl transcriptome dynamics as seen from a cell wall angle

Based on the general analysis of their transcriptomes, one can separate H6 and H9 from H15 and H20. For further characterisation of the hypocotyl transcriptome dynamics, a focused comparison between H6 and H20 will hereby be made.

As a result, from the 3622 transcripts used to build the ICA and clustering (Figure 1-7), those associated with an *Arabidopsis* locus ID and showing a difference corresponding to $-1 > \log_2$ FoldChange (H20/H6) > 1 were used to compare H6 and H20. MapMan functional annotations highlighted the differences between these two time points (Figure 1-9A). Seven categories showed significant differences (*p*-value < 0.05): Photosynthesis, Cell, Cell wall, Gluconeogenesis/glyoxylate cycle, Glycolysis, Nucleotide metabolism, Fermentation. In addition, a subset of the most differentially expressed contigs ($-2 > \log_2$ FC (H20/H6) > 2) have been represented using ClueGO and CluePEDIA within Cytoscape (Figure 1-9B).

As the hypocotyl develops and builds secondary structures, its cellular organisation undergoes dramatic changes, as shown by the expression changes of those genes involved in cell wall biosynthesis and assembly (Supplementary data 1-4). In H6, elongating cells are surrounded by the primary cell wall (PCW), while later vascular and sclerenchyma tissue development is characterised by secondary growth and associated SCW deposition. To highlight the dynamism existing at the cell wall level, a description of the evolution of five families of enzymes playing a predominant role in matrix polysaccharide biosynthesis and remodelling is hereby made following the CAZy nomenclature (Lombard *et al.*, 2014). The families described

are the cellulose synthase-like (CSL, GT2 family), the β-galactosidase (GH35 family), pectin methylesterase (PME, CE8 family), expansin (CBM 63 family), and xyloglucan endotransglucosylase/hydrolase (XTH, GH16 family).

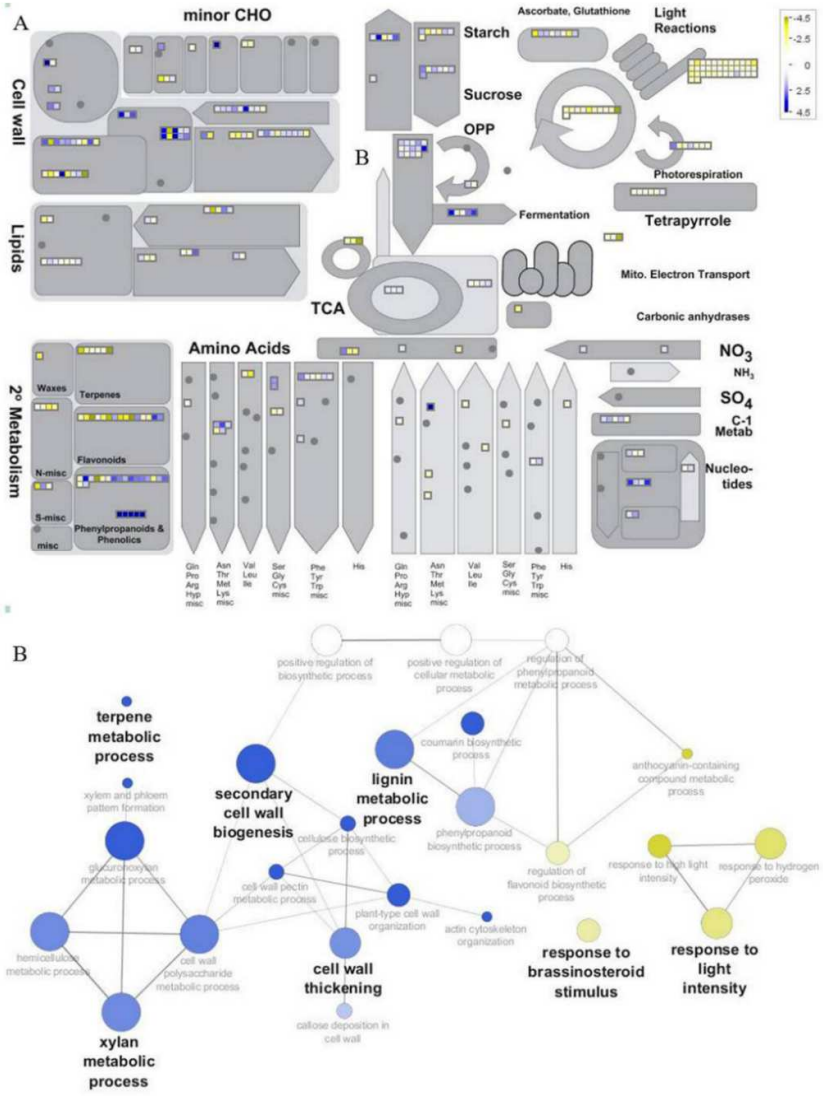


Figure 1-9: Overview of the differentially expressed genes between H6 and H20. A, top: Metabolism overview provided by MapMan functional

annotation. Genes up-regulated in H6 and H20 are represented by yellow and blue squares, respectively. B, bottom: GOE performed with the subset of genes showing -2 > log2FC (H20/H6) >2. Biological processes with genes overrepresented in H6 and H20 are shown in *yellow* and *blue*, respectively. The larger the bubble, the higher the statistical significance.

At the transcriptional level, PCW deposition and organisation depend on a large number of factors. PCW is mainly composed of cellulose deposited in parallel arrays, hemicellulose (predominantly xyloglucan) and pectins (Park and Cosgrove, 2015). The modification of the CW during plant development relies on the active remodelling of the matrix polysaccharides. The RNA-Seq analysis has shown that in H6 genes belonging to the CSL C and E families were more abundantly expressed with respect to H20. More specifically, *CSLC4* and *CSLE1* were more abundant in H6. In H20 different members of the CSL family were highly expressed, namely *CSLD3* and *CSLD5*.

Members of the GH35 and CE8 families were also differentially expressed between H6 and H20. The RNA-Seq analysis showed a differential expression of β-galactosidase (GH35) members in H6 and H20. In H6, the orthologs of *Arabidopsis βGAL2* and *βGAL10* were more abundant. In H20, the ortholog of *βGAL3* was instead more abundant. Pectin methyl esterases (PMEs) and polygalacturonases (GH28 family) were essentially more expressed in H20 (e.g., orthologs of At5g47500 and At5g63180), with only a few being more expressed in H6 (e.g., orthologs of At5g20860 and At4g23820). DUF642, an activator of PME activity (Zúñiga-Sánchez *et al.*, 2014), was strongly up-regulated in H6. Two orthologs of *Arabidopsis* expansins, *EXPA5* and *EXPA8*, were more abundant in the young hypocotyl, as well as three XTHs. *EXPA4* was instead more expressed in H20.

XTH are divided into two categories: xyloglucan endohydrolases (XEHs) and xyloglucan endotransglucosylases (XETs). XEHs irreversibly shorten the length of the chain, while XETs ligate and cleave xyloglucan chain coming from the Golgi (Eklöf and Brumer, 2010). The only XEH (*XTH31*) retrieved in our study was more abundant in H6; some XTHs were more abundant in H6 (*XTH5*, *XTH8*), while others (*XTH15*, *XTH22*) were preferentially expressed in H20.

e. Secondary cell wall-related genes

SCW biogenesis requires the fine tuning of numerous metabolite biosynthesis which are deposited in specific tissues (e.g., xylem vessels and sclerenchyma) undergoing lignification. As expected, several genes linked to SCW formation were more abundant in H20 as compared to H6 (Supplementary data 1-4).

Among the SCW-related genes up-regulated at later stages of development were the orthologs of *A. thaliana CesA4*, *CesA7* and *CesA8*, responsible for the deposition of cellulose (Figure 1-10). Several genes related to the cellulose synthase complex trafficking were more abundant in H20. Such genes included orthologs of *MAP70-5* and *MAP65-8*, *tubulin β-2*, *tubulin β-8* and several genes related to the microtubule motor activity (kinesin and microtubule motor domain).

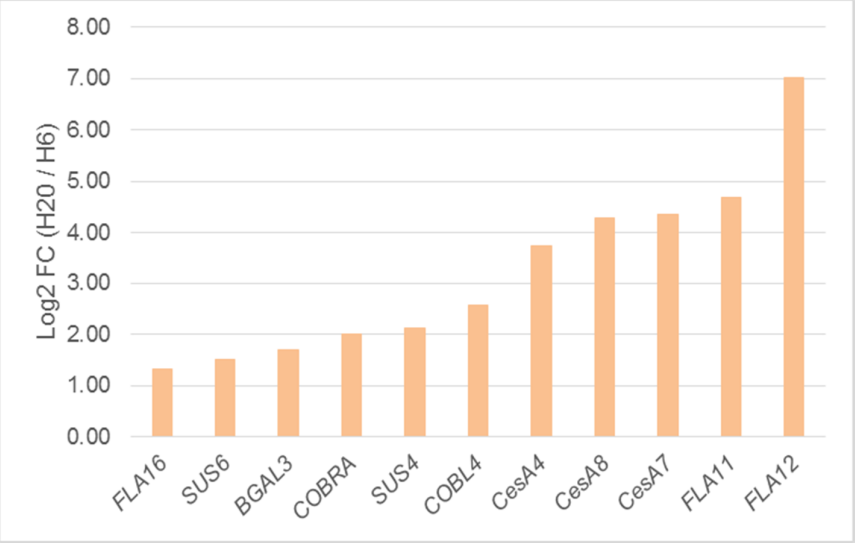


Figure 1-10: Differentially expressed genes involved in cellulose deposition.

Several transcripts related to xyloglucan and xylan biosyntheses and xylan export were more abundant in H20, notably two isoforms of UDP-xylose synthase, *UXS2* and *UXS5*, which provide the building blocks for polymerisation into the xylan backbone. Genes involved in glucuronoxylan synthesis (the main xylan in dicots; Zhong and Ye, 2015) were more abundantly expressed at later stages of development. Other transcripts involved in xylan biosynthesis and more expressed in H20 include *CSLD5*, *TBL34*, *DUF547*, *ESK1* and *GXM3* (Figure 1-11). A more pronounced *trans*-Golgi network activity at later growth stages was highlighted by the higher expression of *DUF707*, *DUF821* and a clathrin light chain protein at H20.

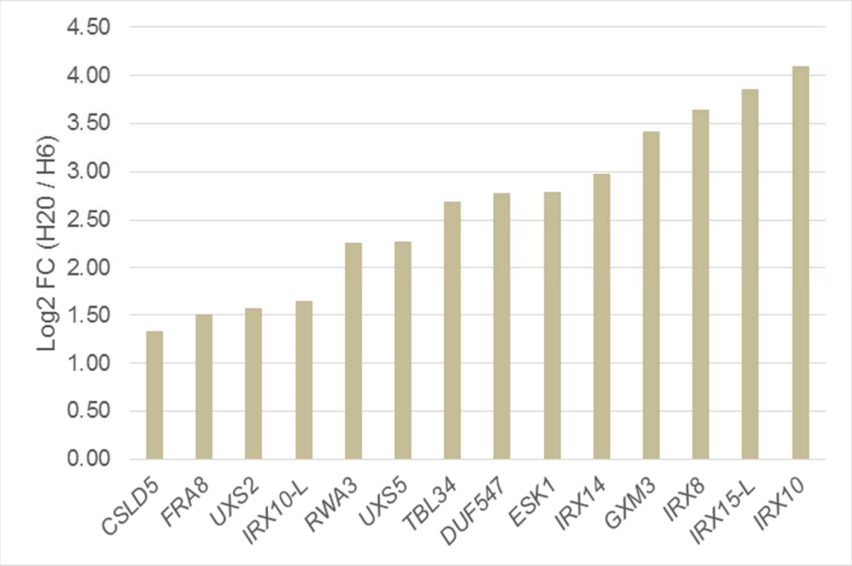


Figure 1-11: Differentially expressed genes involved in xylan biosynthesis.

H20 expression pattern exhibited a substantial up-regulation of the genes associated with lignin biosynthesis. In the general phenylpropanoid pathway, orthologs of *Arabidopsis PAL1*, *PAL2*, *C4H*, *4CL1* and *4CL2* were 1.88 to 2.98 fold more expressed (log_2FC) in H20 than in H6. In the monolignol pathway, orthologs of *Arabidopsis HCT*, *CSE*, *CCR1*, *CCoAOMT* and *CAD4* were up-regulated from 1.46 to 2.79 (log_2FC) (Figure 1-12).

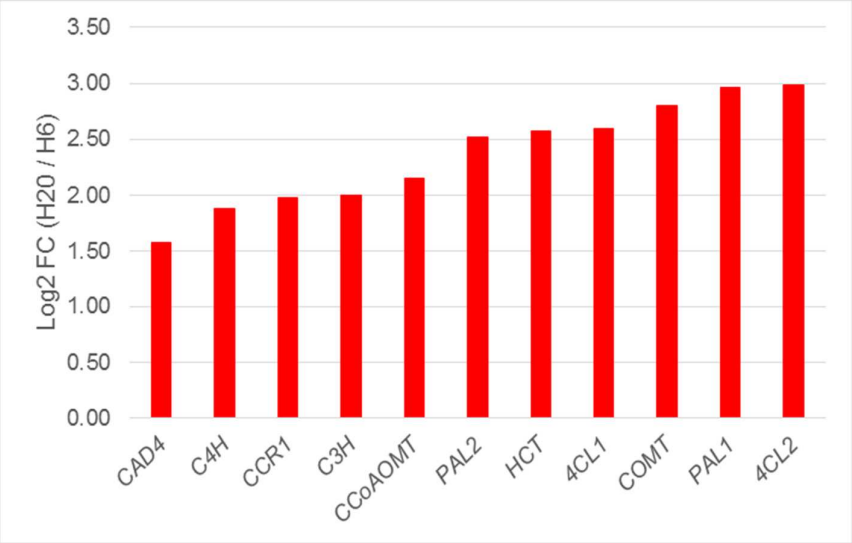


Figure 1-12: Differentially expressed genes involved in phenylpropanoids and monolignols biosyntheses.

In relation to the polymerisation of the monolignol subunits, five orthologs of *Arabidopsis* laccase isoforms (LAC) were dramatically induced in H20: *LAC4* (log_2FC 5.69), *LAC5* (log_2 FC 4.85), *LAC12* (log_2 FC 4.32), *LAC13* (log_2 FC 5.02) and *LAC17* (log_2 FC 5.75). Complementary to the laccase-driven lignin polymerisation, class III peroxidases were also differentially expressed between H6 and H20 (Figure 1-13). The ortholog of *AtPRX52*,which is involved in the synthesis of S-lignin in fibres (Fernández-Pérez *et al.*, 2014), was more abundant in H20. The hemp ortholog of *AtPRX64*, another peroxidase involved in fibre lignification (Tokunaga *et al.*, 2009), was also highly expressed in H20.

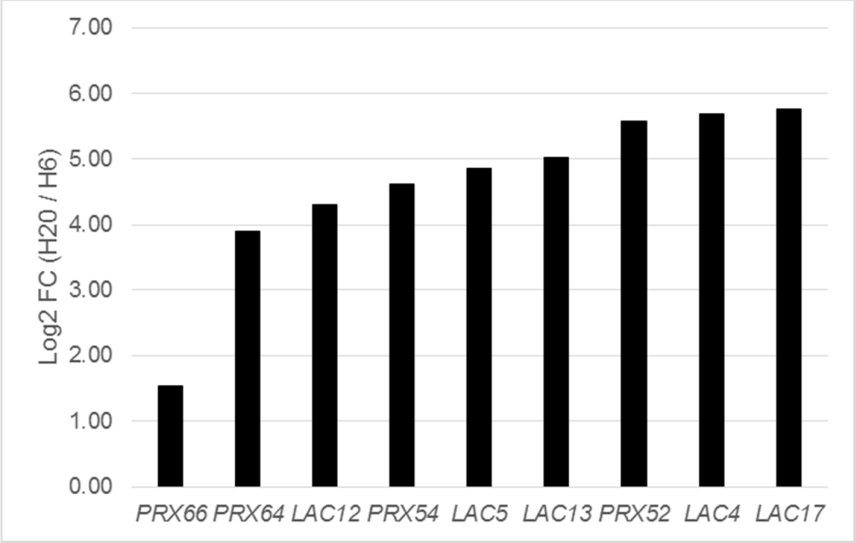


Figure 1-13: Differentially expressed genes involved in lignin polymerisation.

f. Endogenous phytohormones content at different stages of development

Several phytohormones belonging to the auxin, abscisic acid (ABA), cytokinin (CK) and jasmonate (JA) families were quantified (Figure 1-14). The indole-3-acetic acid (IAA) content was not statistically different among the time-points, while the phenylacetic acid (PAA) content was significantly higher in H6 and H9 as compared to H15 and H20.

ABA content was remarkably higher in H20 as compared to H6, H9 and H15. When considering the total content of bioactive CKs, no significant differences were found. The *cis*-zeatin and *cis*-zeatin riboside content was higher in H6. According to Havlová and colleagues (2008), *cis*-zeatin and its riboside lack physiological activity in most CK bioassays and was therefore not considered as an active form of CK. The CK N-glucoconjugates (deactivation forms) and O-glucoconjugates (storage forms) levels were almost equal in H6, H9

and H15 with a decrease in H20. By contrast, the CK-phosphates (CK biosynthesis intermediates) content was higher in H15 and H20. The different forms of JAs include jasmonic acid (JA), jasmonoyl-isoleucine (JA-Ile) and *cis*-OPDA. Both JA and *cis*-OPDA contents were significantly higher in H15 than in H6.

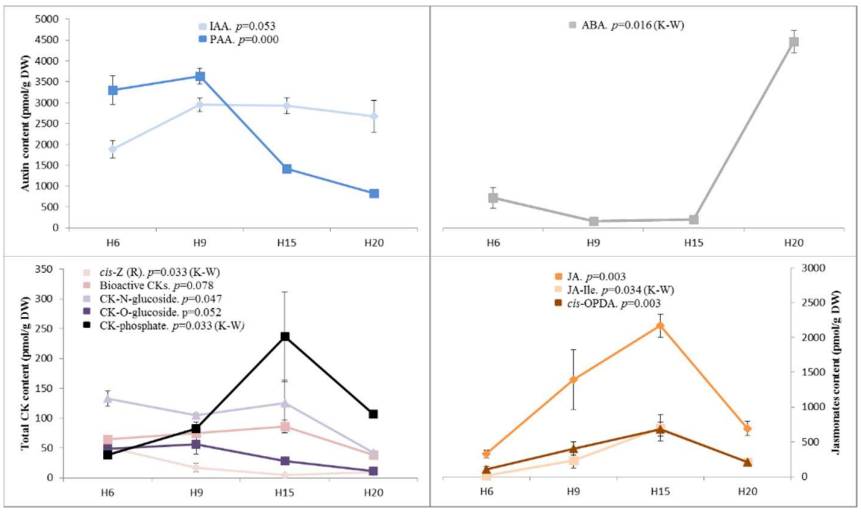


Figure 1-14: Endogenous phytohormone levels in hemp hypocotyls at H6, H9, H15 and H20. *Top left*: bioactive auxins; IAA indole-3-acetic acid, PAA phenylacetic acid. *Top right*: ABA, abscisic acid. *Bottom left*: Total cytokinin. *Bottom right*: bioactive JAs; JA jasmonic acid, JA-Ile jasmonoyl-isoleucine, *cis*-OPDA *cis*-(+)-12-oxo-phytodienoic acid. Mean values ± SE are given. ANOVA or Kruskal-Wallis (K-W) results are indicated depending on the normality of data distribution (Shapiro-Wilk $p>0.05$ or $p<0.05$, respectively). n=3 biological replicates.

The bioactive portion of CK pool is shown in Figure 1-15. The variation described in Figure 1-9C was due to the significant difference observed in the content of *trans*-zeatin (tZ) and its riboside. The *trans*-zeatin/riboside content was higher in H9 and H15 with a minimum in H20 (ANOVA and Tukey post-hoc test, $p=0.014$). No significant

differences were found in the contents of dihydrozeatin (DZ) and its riboside, which were lower as compared to the other forms of bioactive CKs. Even if this trend was not statistically significant, the N^6-(Δ^2-isopentenyl)adenine (iP) pool decreased constantly between H6 and H20. The corresponding riboside iPR transiently peaked in H15. The ratios tZ/tZR, DZ/DZR and iP/iPR were higher in the young stages of development and decreased in older hypocotyls. However, only the tZ/tZR and iP/iPR ratios were found to be significantly different between the time-points (ANOVA *p*-values of 0.035 and 0.008, respectively).

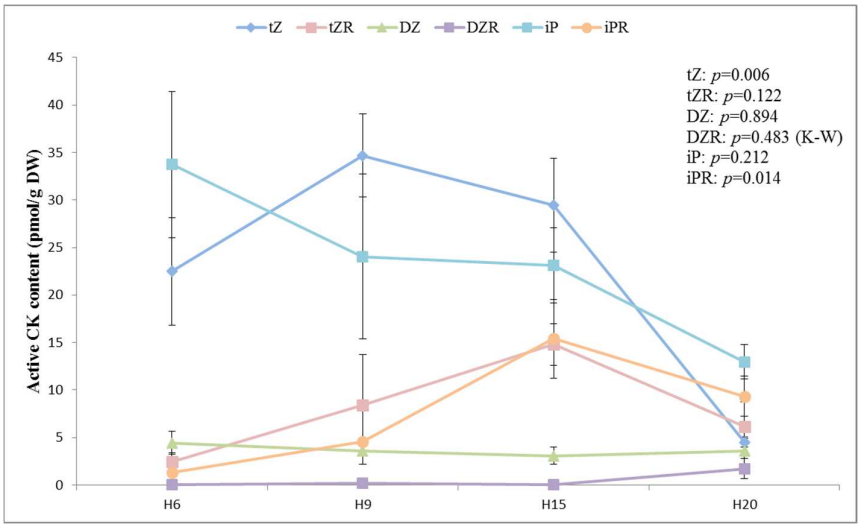


Figure 1-15: Endogenous bioactive cytokinin levels in hemp hypocotyls at H6, H9, H15 and H20. tZ *trans*-zeatin, tZR *trans*-zeatin riboside, DZ dihydrozeatin, DZR dihydrozeatin riboside, iP N6-(Δ2-isopentenyl)adenine, iPR N6-(Δ2-isopentenyl)adenosine. ANOVA or Kruskal-Wallis (K-W) results are indicated depending on the normality of data distribution (Shapiro-Wilk $p>0.05$ or $p<0.05$, respectively).

III. Discussion

a. Regulation of the elongation in the hemp hypocotyl

The hypocotyl elongation is under the control of TFs and genes involved in both cell wall deposition and modification. The bHLHs TFs *KDR* and *AIF3* have antagonistic actions in cell elongation (Ikeda *et al.*, 2013) and were both more expressed in H6. Another elongation repressor of the bHLH system, *IBH1*, was more expressed in H15 and H20 as compared to H6 and H9. *LHW* is related to vascular development (De Rybel *et al.*, 2016) and was more abundant in H9. A TF with a dual function on cell wall dynamics showed an interesting expression pattern, namely *WLIM1*. In cotton fibres, WLIM1 bundles the actin filaments, which favours fibre extension by activating intracellular transport. It also binds to the PAL-box in the promoters of lignin/lignin-like biosynthetic genes, thereby contributing to the lignification of the fibre SCW (Han *et al.*, 2013). The role of WLIM1 in fibre extension can be proposed in young hemp hypocotyls, since a significant increase in expression was observed between H6 and H9.

Xyloglucan, a hemicellulose of the PCW, plays a crucial role in the regulation of cell elongation (Park and Cosgrove, 2015). It was detected in the cell wall at all the time points studied (Figure 1-4) and *CSLC4*, which is involved in its biosynthesis, was more expressed in H6. CSLC4 acts in the Golgi to form the xylosylated glucan backbone during xyloglucan biosynthesis (Park and Cosgrove, 2015). Several transcripts related to xyloglucan remodelling were detected in H6, including expansins, XTH and β-galactosidases. XTH and expansins transcripts were similarly found in elongating flax hypocotyls (Roach and Deyholos, 2008). Expansins loosening the cell wall are thought to

be active in ‘biomechanical hotspots’ where xyloglucan and cellulose are closely intertwined (Park and Cosgrove, 2015). *XTH31* was more expressed in the actively growing hypocotyl (H6), suggesting its potential role in cell expansion (Franková and Fry, 2013). In elongating tissues, *βGAL10* may have an activity on xyloglucan (Sampedro *et al.*, 2012). Cell wall loosening is also controlled by peroxidases (Francoz *et al.*, 2014), which break covalent bonds in cell wall polymers. Notably, in our dataset, the ortholog of *AtPRX57* was strongly upregulated at H6. The deduced protein sequence showed the presence of a signal peptide, a feature supporting the potential role of this peroxidase in cell wall loosening during the growth of the young hemp hypocotyl.

b. Differential expression of genes in the phenylpropanoid pathway among time points

Anthocyanin biosynthesis and monolignol production share the phenylpropanoid pathway from phenylalanine to *p*-coumaroyl CoA. The latter is either directed to the flavonoid pathway through CHS, or through the monolignol pathway via CCR- or HCT- mediated reactions. Several genes implicated in anthocyanin biosynthesis were more expressed in H6 (*MYB114*, *CHS*, *TT18)* while those involved in monolignol biosynthesis were more expressed in H15 or H20 (*HCT*, *CCoAOMT*). Anthocyanins modify the quantity and quality of light trapped by the chloroplasts, acting as a photoprotectant (Steyn *et al.*, 2002). This photoprotection is a required adaptation when the hypocotyl emerges from the soil (Wang *et al.*, 2012). Monolignols are incorporated into the lignin polymer at older developmental stages to provide strength and resistance to the conducting and supporting tissues. A recent work has shown that anthocyanin and monolignol biosynthesis may be interdependent. Indeed, the downregulation of *CCoAOMT*, a core enzyme of the monolignol biosynthetic reactions,

activates the anthocyanin pathway through a MYB transcription factor in petunia (Shaipulah *et al.*, 2016). In addition, silencing of *HCT* in *Arabidopsis* was shown to decrease the S units and increase the H units and flavonoid production through CHS activity (Hao and Mohnen, 2014).

c. Markers of the hypocotyl secondary growth

The cambial activity marking the onset of secondary growth is controlled by phytohormones and, notably, active JAs (Sehr *et al.*, 2010). The JAs content peaked at H15 (Figure 1-14), where cambial division is clearly visible (Figure 1D). JAs stimulate secondary growth by enhancing cambial activity (Sehr *et al.*, 2010) and favour both phloem fibre formation (Sehr *et al.*, 2010) and lignification (Pauwels *et al.*, 2008). At the transcriptional level, the expression of the master regulator of fibre differentiation *NST1* (Zhong and Ye, 2015) increased 6-fold in H15 as compared to H6. XND1 is another NAC TF which acts antagonistically to NST1 to ensure xylem extensibility (Zhao *et al.*, 2008) and its expression increased 23-fold in H20 as compared to H6. PXC1 regulates the xylem fibre formation in stem secondary vasculature (Nieminen *et al.*, 2015) and its expression increased 3-fold in H15 as compared to H6. *WAT1*, encoding a vacuolar auxin transporter essential for fibre differentiation and secondary wall thickening (Ranocha *et al.*, 2013), showed increased expression (2.6-fold) in H20 as compared to H6. *WOX4*, which codes for a homeobox module involved in hypocotyl vascular patterning and cambium proliferation (Ragni and Hardtke, 2014), was more expressed in H20. Genes involved in wall remodelling were also differentially regulated during the development of the hypocotyl. For example, the late expression of the ortholog of *AtEXPA4* might be involved in secondary vascular tissue development by e.g. regulating intrusive growth (Gray-

Mitsumune *et al.*, 2004). Since hemp bast fibres also grow intrusively (Snegireva *et al.*, 2015), a role for this protein in this process might be hypothesised. This gene was found to be more expressed in the phloem-cambium zone of the *Arabidopsis* hypocotyl undergoing secondary growth (Zhao *et al.*, 2005), where active division and cellular growth occur. The hemp ortholog of *TCH4* (*XTH22*) was strongly upregulated in H20 and is known to be inducible by *cis*-OPDA (Taki *et al.*, 2005). The presence of *TCH4* in tissues undergoing secondary wall deposition may be explained by the function of xyloglucan in the biogenesis of the cell wall. Indeed, xyloglucan, after transit through the SCW, may reinforce the links between PCW and SCW in the S1 layer junction (Bourquin *et al.*, 2002). This may explain the strong LM15 signal observed within the cell wall of H15 bast fibres (Figure 1-3C) and the late up-regulation of *TCH4*.

d. Transcripts involved in the SCW biogenesis

Our data showed that the genes involved in the biosynthesis of xylan, cellulose, monolignols and in lignin polymerisation were co-expressed in H15 and H20. Such a co-expression may result from the up-regulation of SCW deposition by TFs, i.e. NST1 for the fibre fate (Hao and Mohnen, 2014) and VND1 for vessels (Zhong and Ye, 2014).

In our dataset, the cellulose synthases which were more expressed during secondary growth were the *Arabidopsis* orthologs of *CesA4*, *CesA7* and *CesA8*, as already previously observed in adult hemp (De Pauw *et al.*, 2007; van den Broeck *et al.*, 2008). *CesA*s require sucrose synthase (SUS) activity for the provision of the UDP-glucose (Doblin *et al.*, 2002). The *Arabidopsis* orthologs of *SUS4* and *SUS6* were found to be differentially expressed between young and old hypocotyls. In *Arabidopsis*, SUS6 is present in the vascular tissues of the cotyledons, leaves, petals, anthers and roots (Bieniawska *et al.*,

2007). The function of SUS4 and SUS6 possibly overlaps, but they may have different patterns of tissular or cellular localisation. These enzymes may be soluble to provide a pool of hexose phosphate for the cellular metabolism or anchored to the plasma membrane to deliver UDP-glucose to synthetize cellulose (Bieniawska *et al.*, 2007). Two genes of the COBRA family, *COB* and *COBL4*, are also involved in cellulose biosynthesis (Brown *et al.*, 2005). Further characterisation of the genes belonging to this family is needed to understand their contribution and the way they act in the cell wall. As their expression was higher in H20, a role in SCW and gelatinous layer (G-layer) deposition may be postulated. Supporting this hypothesis, Hobson and colleagues (2010) have described a COBRA-like gene being strongly up-regulated in the bast fibres of the middle part of adult hemp stems as compared to the top and De Pauw and colleagues (2007) found that *COBL4* was more expressed in the outer tissue of the middle and bottom parts of the stem as compared to the top, where the G-layer is appearing in phloem fibres. In flax it was demonstrated that, after its deposition, the bast fibre galactan layer progressively matures via the assembly of cellulose microfibrils to a crystalline structure (Gorshkova and Morvan, 2006). In flax, this maturation between the galactan-layer and the G-layer requires βGAL1 (Roach *et al.*, 2011). We have found out that a *βGAL* (contig 1317), ortholog of the *Arabidopsis βGAL3*, was specifically expressed in hemp hypocotyls undergoing SCW deposition. This gene clusters with flax *LuβGAL1*, which is involved in SCW deposition in flax hypocotyls and stems (Roach and Deyholos, 2007; Roach *et al.*, 2011). It should however be noted that in hemp we did not observe an LM5 signal in the bast fibres (Figure 1-2), as previously shown in adult plants by Blake and colleagues (2008). This is an important difference with respect to flax fibre composition. However, the presence of rhamnogalacturonan-I in the matrix

polysaccharide of the bast fibre is unambiguously demonstrated by the strong binding of the RU2 antibody. Another gene which may play an important role in hemp cell wall maturation is a class IV chitinase (contig 13996). It was highly expressed in H9, H15 and H20. In flax, three isoforms of chitinase-like genes, most similar to chitinases of class IV, were highly expressed in the phloem fibres (Mokshina *et al.*, 2014), where they may be involved in the development of the G-layer. A signal peptide was found in the deduced hemp chitinase protein sequence, a finding corroborating its action in the cell wall. Three isoforms of FLAs (*FLA11*, *FLA12* and *FLA16*) were more expressed in H15 and H20. The biological functions of FLAs are still debated (Roach and Deyholos, 2007), however several lines of evidence point to a role in SCW deposition (MacMillan *et al.*, 2015; Ito *et al.*, 2005). Our data suggest they may indeed be involved in the deposition of the SCW of the fibres, since they were more expressed at older developmental stages. AtFLA11 and AtFLA12 are highly active in xylem vessels and fibres and impact cellulose, arabinose and galactose content in the cell wall (MacMillan *et al.*, 2015). The group containing AtFLA16 has not yet been characterised. In flax hypocotyls, two FLAs were more expressed at 9 and 15 days as compared to 7 days (Roach and Deyholos, 2008).

Xylan is the chief hemicellulose in hemp hypocotyl secondary tissues (Figure 1-5). Indeed, in accordance with Blake and colleagues (2008), the LM10 epitope was detected in the cell wall of xylem cells and bast fibres. This finding was supported at the transcriptome level by Cytoscape and MapMan analyses (Figure 1-8). In xylem cells, xylan is largely present in the lignifying SCW while in the bast fibres it was localized in the outermost layer enveloping the G-layer.

Glucuronoxylan contains a xylan reducing-end glycosyl sequence (XRES) which is likely produced by FRA8 and IRX8 (Hao and Mohnen, 2014). Some genes involved in the elongation of the glucuronoxylan backbone (*IRX10*, *IRX10L*, *IRX14* and *IRX15L*), backbone acetylation (*ESK1*, *RWA3*) and substituent methylation (*GXM3*) have been found in late stages of the hypocotyl development. Determining the molecular structure and the precise deposition site of xylan in bast fibres, e.g. by immunogold labelling coupled to TEM, may shed light on its function (Mikshina *et al.*, 2013). Indeed, Mortimer and colleagues (2015) have described in *Arabidopsis* a subset of genes involved in the biosynthesis of a specific type of xylan with pentosylated side chains (*PUX5*) which is present in young stems and roots. Interestingly, *IRX10L* and *IRX14* are part of this subset of genes. The *PUX5*-xylan is able to interact with different wall polymers, but with a reduced cross-linking as compared to the xylan of the SCW, to ensure PCW extensibility. It should be noted that IRX10 and IRX15L seem to be implicated in the backbone elongation of SCW xylan (Jensen *et al.*, 2011; Mortimer *et al.*, 2015). Such a differentiation of genes involved in xylan backbone elongation may point to different functions of the xylan deposited in the xylem cells and bast fibres. Further studies are needed to verify this hypothesis.

Most of the genes involved in the transcriptional regulation of lignification, *i.e.* the methyl donors, genes involved in monolignol biosynthesis and lignin polymerisation were up-regulated either in H15 or H20. *WLIM1*, *MYB85* and *NST1* regulate the transcription of the lignin biosynthesizing genes (Zhong and Ye, 2009). These transcription factors may consequently have an important role with respect to the hypolignification of hemp bast fibres. It remains to be determined if they are specifically expressed in lignifying tissues, or if

other regulatory mechanisms exist (such as down regulation of the genes of the monolignol biosynthetic pathway or post-translational control over the transcripts or proteins). C1 metabolism was clearly more intense as the hypocotyls develop, as shown by the up-regulation of one isoform of methylenetetrahydrofolate reductase and methionine synthase and three isoforms of S-adenosylmethionine synthase (orthologs of *Arabidopsis MTHFR2*, *ATMS1*, *SAMS2*, *SAMS3* and *SAMS4*). *Arabidopsis* seedlings deficient in SAMS3 activity display significantly lower content of lignin in the aerial part (Shen *et al.*, 2002), thus highlighting the possible regulation of lignification by the pool of available methyl donors. In adult hemp stem, the core tissue mainly composed of xylem cells also shows a higher rate of C1 recycling (van den Broeck *et al.*, 2008). The genes putatively involved in phenylpropanoid and monolignol biosynthesis were all more expressed in H15 or H20, with the exception of *CAD9*. In *Arabidopsis*, CAD9 is involved in the biosynthesis of coniferyl alcohol in the young developing stem (Eudes *et al.*, 2006). It is therefore plausible to observe a higher expression of CAD9 in H6. The other CAD isoform, *CAD4*, reached its highest expression in H20. None of these two CAD isoforms have been described in the studies of van den Broeck and colleagues (2008) and De Pauw and colleagues (2007). This may be explained by the higher sensitivity of the RNA-Seq for low abundance transcripts (highest RPKM of *CAD9* = 29). Indeed, other genes involved in these pathways display higher RPKM values (>500) and were detected by these studies based on microarrays.

Laccases and peroxidases are enzymes known to polymerise monolignols in *Arabidopsis* (Hao and Mohnen, 2014). The changes associated with the expression of these genes were among the highest observed in our study. Interestingly, laccases were more expressed in

H15 than in H20 (Figure 1-16), suggesting a role in the early lignification. As in *Arabidopsis* (Turlapati *et al.*, 2011), a signal peptide was detected in the orthologs of AtLAC5 and AtLAC13, however their involvement in lignification has not yet been proven. AtLAC4, AtLAC12 and AtLAC17 have a putative role in stem lignification in *Arabidopsis* (Berthet *et al.*, 2012). AtLAC5 and AtLAC13 cluster together, while AtLAC4, AtLAC12 and AtLAC17 are in three separate clusters according to Turlapati *et al.* (2011).

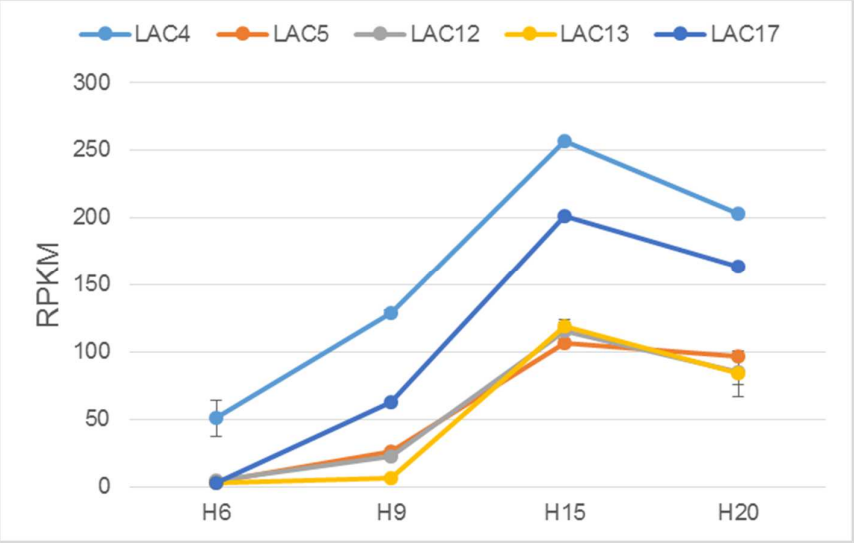


Figure 1-16: Expression of laccases hypothetically involved in lignification.

The orthologs of *AtPRX52* and *AtPRX64*, which are involved in *Arabidopsis* fibre lignification (Fernández-Pérez *et al.*, 2014; Tokunaga *et al.*, 2009), were more expressed in H20 (Figure 1-17). This is an important difference with respect to laccases, which will be broadly discussed in Chapter 2. Interestingly, a signal peptide is present in the deduced protein sequences of PRX52 and PRX64, suggesting their secretion to the cell wall.

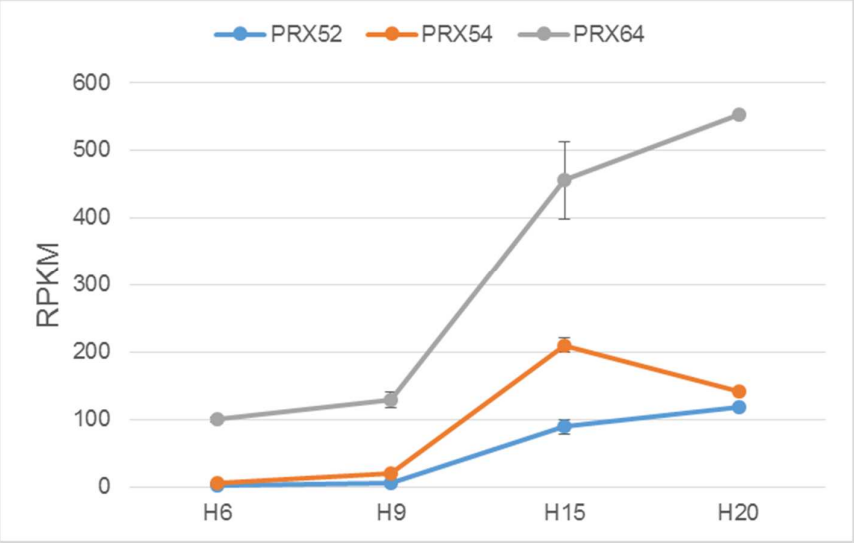


Figure 1-17: Expression of class III peroxidases involved in lignification.

e. Hormonal control of cell wall biogenesis in the hypocotyl of *C. sativa*

Several phytohormones are essential for cell wall biogenesis, from the elongation phase to the strengthening of various tissues by the secondary cell wall. Many studies have pointed to their roles in the expression of cell wall related genes, either by direct addition or by influencing the genes involved in their biosynthesis or signalling (reviewed by Didi *et al.*, 2015; Nieminen *et al.*, 2015; Sánchez-Rodríguez *et al.*, 2010). We have thus combined our transcriptomics data with the quantification of phytohormones and looked for genes involved in cell wall biogenesis known to be regulated by those molecules.

IAA and PAA are two distinct bioactive auxins, but they share the same signalling pathway (e.g., TIR1-IAA3 and TIR1-IAA6) and regulate the same auxin-responsive genes (Sugawara *et al.*, 2015).

For instance, auxins are known to induce cell elongation via the expansin and XTH activities (Sánchez-Rodríguez *et al.*, 2010). In our dataset, *EXPA5*, *EXPA8*, *XTH5* and *XTH8* were more expressed in H6 and H9, where the endogenous bioactive auxin content (IAA3 and PAA) is higher. Another property of auxin is to prevent lignification by promoting the degradation of IAA3 (Nanao *et al.*, 2014; Oh *et al.*, 2014). In *Arabidopsis* seedlings, the gain-of-function *iaa3/shy2-2* mutant shows higher expression of several genes belonging to the phenylpropanoid/lignin pathway, such as *PAL1*, *PAL2, C4H*, *4CL1*, *CCoAOMT* and *PRX52* (Oh *et al.*, 2014). Indeed, the lowest bioactive auxin levels, recovered in H15 and H20, correspond to the highest abundance of these genes. However, auxin may also induce the expression of some genes involved in lignin biosynthesis such as *COMT* and *CCoAOMT* through the binding of the TF BREVIPEDICELLUS to their promoter region (Sánchez-Rodríguez *et al.*, 2010).

ABA acts antagonistically to gibberellic acid to inhibit stem elongation by arranging longitudinally the cortical microtubule (Shibaoka, 1994). This may explain the high ABA content observed in H20, a time-point where elongation ceases. Genes coding for the microtubule-associated proteins, *MAP65*-8 and *MAP70-5*, were found to be highly expressed in H20. MAP65 bundles the microtubules (Soyano *et al.*, 2008), while MAP70-5 determines the cortical patterning of the cell walls of tracheary elements (Derbyshire *et al.*, 2015). ABA also down-regulates the expression of the fasciclin-like arabinogalactan protein *AtFLA2*, whose hemp ortholog was found to be more abundant in H6. AtFLA2 may induce shoot development (Johnson *et al.*, 2003) in an auxin-dependant manner, as it forms a

complex with ABCB19, a regulator of auxin efflux (Titapiwatanakun *et al.*, 2009).

We were not able to significantly discriminate the bioactive cytokinin content from one time-point to another, but the lowest concentration was observed in H20 (Figure 1-14). It has been shown that the concentration of bioactive CKs is higher in young tissues and declines as the tissue senesces (Havlová *et al.*, 2008). However, when considering only tZ / tZR, the highest level was observed in H9 and H15 while the minimum was found in H20. It has been suggested that roots are a major site of tZ production and that tZR is the major CK present in the xylem sap (Hirose *et al.*, 2008; Sakakibara, 2006). After translocation, tZR is metabolised to tZ, leading to signal transduction. Given the decrease of the ratios between CK nucleobases and CK nucleosides (Figure 1-15), we may suppose that there is a shift in bioactive CK allocation along the time-course. As a consequence, the fraction of bioactive CKs able to induce a signal, i.e. the nucleobases, might be higher in younger hypocotyls. At the cell wall level, CKs, and particularly tZ, are involved in the lignification of the fibres and xylem vessels (Didi *et al.*, 2015). From our data, we may speculate that CKs regulate the initiation of lignin deposition in secondary tissues, while later lignification is tuned by other factors. *AtPRX52* (whose hemp ortholog was more expressed in H20) is an example of these genes regulated by several factors: it contains a high number of *Arabidopsis* response regulator 1 (ARR1)-cytokinin-binding element in its promoter region but is also responsive to *bZIP9*, *bZIP50*, *NST1*, *MYB46* or *MYB85* (Herrero *et al.*, 2014). All these TFs were, notably, more abundant either in H15 or in H20.

Methyl jasmonate and other JAs are also factors regulating the biosynthesis of cell wall building components (Men *et al.*, 2013;

Pauwels *et al.*, 2008). This is notably the case for lignin: we have found that the expression of several methyl jasmonate-up-regulated genes peaks simultaneously with the active jasmonates content, namely orthologs of *Arabidopsis MetS1*, *PAL2*, *C3H*, *CCoAOMT*, *HCT*, *CAD4* and *4CL2*. *IRX15-L* (glucuronoxylan backbone biosynthesis) is also up-regulated by methyl jasmonate and its expression peaked in H20.

In conclusion, this study has highlighted the suitability of the hemp hypocotyl for the study of the main events accompanying secondary growth. Phytohormones were here shown to be tightly linked with secondary growth and cell wall deposition. Polysaccharide fingerprinting via immunohistochemistry has been linked to transcriptomic data to emphasize the role of the transcription factors and genes involved in the biosynthesis of the building blocks (polysaccharides, lignin) composing the cell wall, a structure that is crucial for the development of the plant. A concluding figure summarises the genes taking part in this process which have been found in the transcriptomics analysis (Figure 1-18).

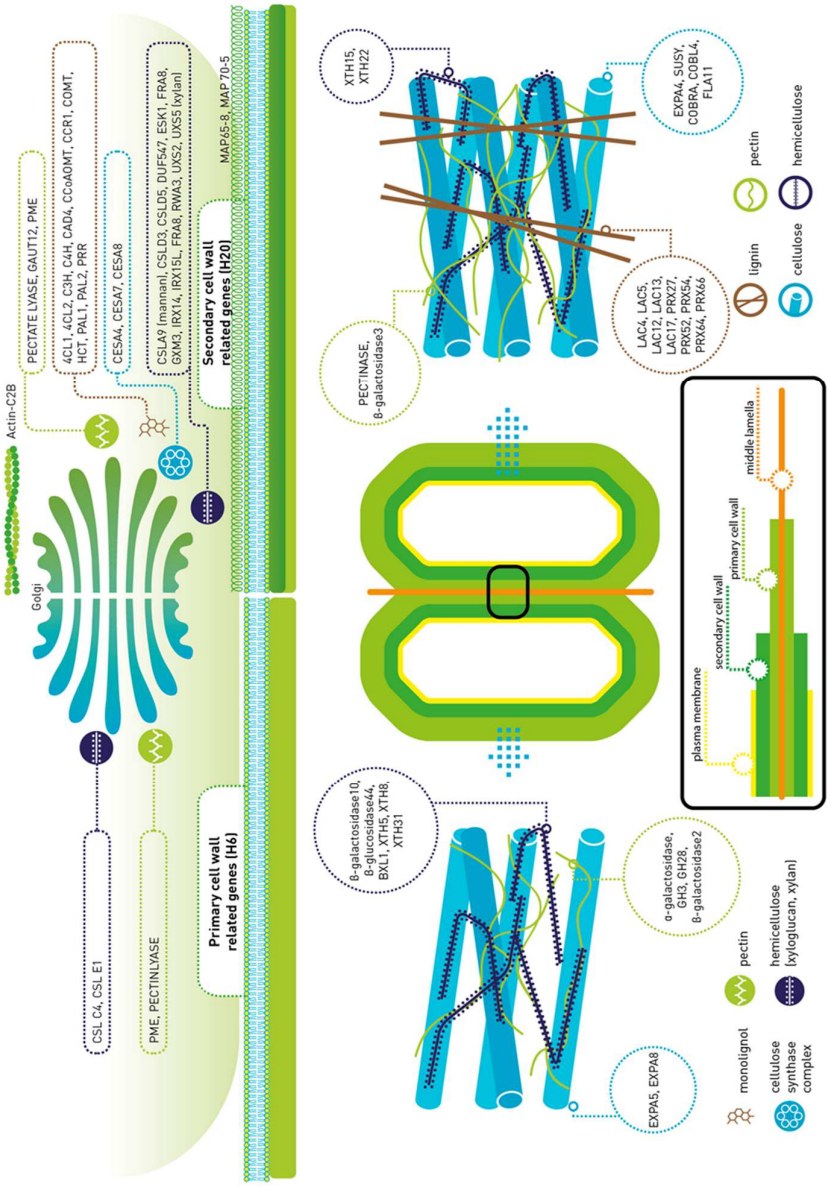


Figure 1-18: Cellular localisation of the activity of proteins involved in cell wall biogenesis whose mRNA have been differentially expressed in H6 (primary cell wall) or in H20 (secondary cell wall)

References

Andre, C.M., Hausman, J.-F., and Guerriero, G. (2016). *Cannabis sativa*: The Plant of the Thousand and One Molecules. *Front Plant Sci* 7, 19.

Berthet, S., Thevenin, J., Baratiny, D., Mont-Caulet, N., Debeaujon, I., Bidzinski, P., *et al.* (2012). Chapter 5 - Role of Plant Laccases in Lignin Polymerization, in Advances in Botanical Research. Lignins Biosynthesis, Biodegradation and Bioengineering, ed. Lise Jouanin and Catherine Lapierre (Academic Press), 145-172.

Bieniawska, Z., Paul Barratt, D.H., Garlick, A.P., Thole, V., Kruger, N.J., Martin, C., et al (2007). Analysis of the sucrose synthase gene family in *Arabidopsis*. *Plant J* 49, 810-828.

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., *et al.* (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 25, 1091-1093.

Bindea, G., Galon, J., and Mlecnik, B. (2013). CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. *Bioinformatics* 29, 661-663.

Blake, A.W., Marcus, S.E., Copeland, J.E., Blackburn, R.S., and Knox, J.P. (2008). In situ analysis of cell wall polymers associated with phloem fiber cells in stems of hemp, *Cannabis sativa* L. *Planta* 228, 1-13.

Bourquin, V., Nishikubo, N., Abe, H., Brumer, H., Denman, S., Eklund, M., et al (2002). Xyloglucan Endotransglycosylases Have a Function during the Formation of Secondary Cell Walls of Vascular Tissues. *Plant Cell* 14, 3073-3088.

Brown, D.M., Zeef, L.A.H., Ellis, J., Goodacre, R., and Turner, S.R. (2005). Identification of Novel Genes in Arabidopsis Involved in Secondary Cell Wall Formation Using Expression Profiling and Reverse Genetics. *Plant Cell* 17, 2281-2295.

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.

Daub, C.O., Kloska, S., and Selbig, J. (2003). MetaGeneAlyse: analysis of integrated transcriptional and metabolite data. *Bioinformatics* 19, 2332-2333.

De Pauw MA, Vidmar J.J., Collins J, Bennett R.A., and Deyholos M.K. (2007). Microarray analysis of bast fiber producing tissues of *Cannabis sativa* identifies transcripts associated with conserved and specialised processes of secondary wall development. *Funct. Plant Biol.* 34, 737-749.

De Rybel, B., Mahonen, A.P., Helariutta, Y., and Weijers, D. (2016). Plant vascular development: from early specification to differentiation. *Nat Rev Mol Cell Biol* 17, 30-40.

Derbyshire, P., Ménard, D., Green, P., Saalbach, G., Buschmann, H., Lloyd, C.W., *et al.* (2015). Proteomic Analysis of Microtubule Interacting Proteins over the Course of Xylem Tracheary Element Formation in *Arabidopsis*. *Plant Cell* 27, 2709-2726.

Didi, V., Jackson, P., and Hejátko, J. (2015). Hormonal regulation of secondary cell wall formation. *J Exp Bot* 66, 5015-5027.

Djilianov, D.L., Dobrev, P.I., Moyankova, D.P., Vankova, R., Georgieva, D.T., Gajdosova, S, *et al.* (2013). Dynamics of endogenous phytohormones during desiccation and recovery of the resurrection plant species *Haberlearhodopensis*. *J Plant Growth Regul* 32, 564-574.

Doblin, M.S., Kurek, I., Jacob-Wilk, D., and Delmer, D.P. (2002). Cellulose Biosynthesis in Plants: from Genes to Rosettes. *Plant Cell Physiol* 43, 1407-1420.

Dobrev, P.I., and Vankova, R. (2012}. Quantification of Abscisic Acid, Cytokinin, and Auxin Content in Salt-Stressed Plant Tissues. in Plant Salt

Tolerance: Methods and Protocols, eds. S. Shabala and A. T. Cuin (Totowa, NJ, Humana Press), 251-261.

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.

Eudes, A., Pollet, B., Sibout, R., Do, C.T., Séguin, A., Lapierre, C., *et al.* (2006). Evidence for a role of AtCAD 1 in lignification of elongating stems of *Arabidopsis thaliana*. *Planta* 225, 23-39.

Fernández-Pérez, F., Pomar, F., Pedreno, M.A., and Novo-Uzal, E. (2014). The suppression of AtPrx52 affects fibers but not xylem lignification in *Arabidopsis* by altering the proportion of syringyl units. *Physiol Plantarum* 154, 395-406.

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

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.

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

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

Gray-Mitsumune, M., Mellerowicz, E.J., Abe, H., Schrader, J., Winzéll, A., Sterky, F., *et al.* (2004). Expansins Abundant in Secondary Xylem Belong to Subgroup A of the α-Expansin Gene Family. *Plant Physiol* 135, 1552-1564.

Guerriero, G., Hausman, J.F., Strauss, J., Ertan, H., and Siddiqui, K.S. (2016). Lignocellulosic biomass: Biosynthesis, degradation, and industrial utilization. *Eng.Life Sci.* 16, 1-16.

Guerriero, G., Sergeant, K., and Hausman, J.F. (2013). Integrated -Omics: A Powerful Approach to Understanding the Heterogeneous Lignification of Fiber Crops. *Int J MolSci* 14, 10958-10978.

Guerriero, G., Sergeant, K., and Hausman, J.F. (2014). Wood biosynthesis and typologies: a molecular rhapsody. *Tree Physiol* 34, 839-855.

Han, L.B., Li, Y.B., Wang, H.Y., Wu, X.M., Li, C.L., Luo, M., *et al.* (2013). The Dual Functions of WLIM1a in Cell Elongation and Secondary Wall Formation in Developing Cotton Fibers. *Plant Cell* 25, 4421-4438.

Hao, Z., and Mohnen, D. (2014). A review of xylan and lignin biosynthesis: Foundation for studying Arabidopsis irregular xylem mutants with pleiotropic phenotypes. *Crit Rev Biochem Mol Biol* 49, 212-241.

Havlová, M., Dobrev, P.I., Motyka, V., Storchová, H., Libus, J., Dobrá, J., *et al.* (2008). The role of cytokinins in responses to water deficit in tobacco plants over-expressing trans-zeatin O-glucosyltransferase gene under 35S or SAG12 promoters. *Plant Cell Environ* 31, 341-353.

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. Bioch.* 80, 192-202.

Hirose, N., Takei, K., Kuroha, T., Kamada-Nobusada, T., Hayashi, H., and Sakakibara, H. (2008). Regulation of cytokinin biosynthesis, compartmentalization and translocation. *J Exp Bot* 59, 75-83.

Hobson, N., Roach, M.J., and Deyholos, M.K. (2010). Gene expression in tension wood and bast fibers. *Russ. J. Plant Physl.* 57, 321-327.

Ikeda, M., Mitsuda, N., and Ohme-Takagi, M. (2013). ATBS1 INTERACTING FACTORs negatively regulate *Arabidopsis* cell elongation in the triantagonistic bHLH system. *Plant Signal Behav* 8, e23448.

Ito, S., Suzuki, Y., Miyamoto, K., Ueda, J., and Yamaguchi, I. (2005). AtFLA11, a Fasciclin-Like Arabinogalactan-Protein, Specifically Localized in Screlenchyma Cells. *Biosci. Biotechnol. Biochem.* 69, 1963-1969.

Jensen, J.K., Kim, H., Cocuron, J.C., Orler, R., Ralph, J., and Wilkerson, C.G. (2011). The DUF579 domain containing proteins IRX15 and IRX15-L affect xylan synthesis in *Arabidopsis*. *Plant J* 66, 387-400.

Jin, J., Zhang, H., Kong, L., Gao, G., and Luo, J. (2014). PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. *Nucleic Acids Res* 42, D1182-D1187.

Johnson, K.L., Jones, B.J., Bacic, A., and Schultz, C.J. (2003). The Fasciclin-Like Arabinogalactan Proteins of Arabidopsis. A Multigene Family of Putative Cell Adhesion Molecules. *Plant Physiol* 133, 1911-1925.

Lombard, V., GolacondaRamulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42, D490-D495.

MacMillan, C.P., Taylor, L., Bi, Y., Southerton, S.G., Evans, R., and Spokevicius, A. (2015). The fasciclin-like arabinogalactan protein family of *Eucalyptus grandis* contains members that impact wood biology and biomechanics. *New Phytol* 206, 1314-1327.

Men, L., Yan, S., and Liu, G. (2013). De novo characterization of *Larixgmelinii* (Rupr.) Rupr. transcriptome and analysis of its gene expression induced by jasmonates. *BMC Genomics* 14, 1-14.

Mikshina, P., Chernova, T., Chemikosova, S.B., Ibragimova, N., Mokshina, N., and Gorshkova, T. (2013). Cellulosic Fibers: Role of Matrix Polysaccharides in Structure and Function, in Cellulose - Fundamental Aspects, eds. T. van de Ven and L. Godbout (InTech), 91-112.

Mokshina, N., Gorshkova, T., and Deyholos, M.K. (2014). Chitinase-Like and Cellulose Synthase Gene Expression in Gelatinous-Type Cellulosic Walls of Flax (*Linum usitatissimum* L.) Bast Fibers. *PLOS ONE* 9, e97949.

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5, 621-628.

Mortimer, J.C., Faria-Blanc, N., Yu, X., Tryfona, T., Sorieul, M., Ng, Y.Z., *et al.* (2015). An unusual xylan in *Arabidopsis* primary cell walls is synthesised by GUX3, IRX9L, IRX10L and IRX14. *Plant J* 83, 413-426.

Moura, J.C.M.S., Bonine, C.A.V., De Oliveira FernandesViana, J., Dornelas, M.C., and Mazzafera, P. (2010). Abiotic and Biotic Stresses and Changes in the Lignin Content and Composition in Plants. *J Integr Plant Biol* 52, 360-376.

Nanao, M.H., Vinos-Poyo, T., Brunoud, G.+., Thévenon, E., Mazzoleni, M., Mast, D., *et al.* (2014). Structural basis for oligomerization of auxin transcriptional regulators. *Nat. Commun.* 5, 3617.

Nieminen, K., Blomster, T., Helariutta, Y., andMähönen, A.P. (2015). Vascular Cambium Development. *The Arabidopsis Book*, e0177.

Oh, E., Zhu, J.Y., Bai, M.Y., Arenhart, R.A., Sun, Y., and Wang, Z.Y. (2014). Cell elongation is regulated through a central circuit of interacting transcription factors in the *Arabidopsis* hypocotyl. *eLife* 3, e03031.

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.

Pauwels, L., Morreel, K., De Witte, E., Lammertyn, F., Van Montagu, M., Boerjan, W., *et al.* (2008). Mapping methyl jasmonate-mediated transcriptional reprogramming of metabolism and cell cycle progression in cultured *Arabidopsis* cells. *P NatlAcadSciUsa* 105, 1380-1385.

Ragni, L., Nieminen, K., Pacheco-Villalobos, D., Sibout, R., Schwechheimer, C., and Hardtke, C.S. (2011). Mobile Gibberellin Directly Stimulates Arabidopsis Hypocotyl Xylem Expansion. *Plant Cell* 23, 1322-1336.

Ragni, L., and Hardtke, C.S. (2014). Small but thick enough-the *Arabidopsis* hypocotyl as a model to study secondary growth. *Physiol Plantarum* 151, 164-171.

Rajkumar, A., Qvist, P., Lazarus, R., Lescai, F., Ju, J., Nyegaard, M. *et al.* (2015). Experimental validation of methods for differential gene expression analysis and sample pooling in RNA-seq. *BMC Genomics* 16, 1-8.

Ranocha, P., Dima, O., Nagy, R., Felten, J., Corratgé-Faillie, C., Novák, O., *et al.* (2013). *Arabidopsis* WAT1 is a vacuolar auxin transport facilitator required for auxin homoeostasis. *Nat. Commun.* 4, 2625.

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

Roach, M.J., and Deyholos, M.K. (2008). Microarray Analysis of Developing Flax Hypocotyls Identifies Novel Transcripts Correlated with Specific Stages of Phloem Fiber Differentiation. *Ann Bot* 102, 317-330.

Roach, M.J., Mokshina, N.Y., Badhan, A., Snegireva, A.V., Hobson, N., Deyholos, M.K., *et al.* (2011). Development of Cellulosic Secondary Walls in Flax Fibers Requires ß-Galactosidase. *Plant Physiol* 156, 1351-1363.

Sakakibara, H. (2006). CYTOKININS: Activity, Biosynthesis, and Translocation. *Annu. Rev. Plant Biol* 57, 431-449.

Sampedro, J., Gianzo, C., Iglesias, N., Guitián, E., Revilla, G., and Zarra, I. (2012). AtBGAL10 Is the Main Xyloglucan ß-Galactosidase in *Arabidopsis*, and Its Absence Results in Unusual Xyloglucan Subunits and Growth Defects. *Plant Physiol* 158, 1146-1157.

Sánchez-Rodríguez, C., Rubio-Somoza, I., Sibout, R., and Persson, S. (2010). Phytohormones and the cell wall in Arabidopsis during seedling growth. *Trends in Plant Science* 15, 291-301.

Sehr, E.M., Agusti, J., Lehner, R., Farmer, E.E., Schwarz, M., and Greb, T. (2010). Analysis of secondary growth in the *Arabidopsis* shoot reveals a positive role of jasmonate signalling in cambium formation. *Plant J* 63, 811-822.

Shaipulah, N.F., Muhlemann, J.K., Woodworth, B.D., Van Moerkercke, A., Verdonk, J.C., Ramirez, A.A., *et al.* (2016). CCoAOMT Down-Regulation Activates Anthocyanin Biosynthesis in *Petunia*. *Plant Physiol* 170, 717-731.

Shen, B., Li, C., and Tarczynski, M.C. (2002). High free-methionine and decreased lignin content result from a mutation in the *Arabidopsis* S-adenosyl-L-methionine synthetase 3 gene. *Plant J* 29, 371-380.

Shibaoka, H. (1994). Plant Hormone-Induced Changes in the Orientation of Cortical Microtubules: Alterations in the Cross-linking Between Microtubules and the Plasma Membrane. *Annu. Rev. Plant Phys.*45, 527-544.

Sibout, R., Plantegenet, S., and Hardtke, C.S. (2008). Flowering as a Condition for Xylem Expansion in *Arabidopsis* Hypocotyl and Root. *CurrBiol* 18, 458-463.

Snegireva, A., Chernova, T., Ageeva, M., Lev-Yadun, S., and Gorshkova, T. (2015). Intrusive growth of primary and secondary phloem fibers in hemp stem determines fiber-bundle formation and structure. *AoB Plants* 7, plv061.

Soyano, T., Thitamadee, S., Machida, Y., and Chua, N.H. (2008). ASYMMETRIC LEAVES2-LIKE19/LATERAL ORGAN BOUNDARIES DOMAIN30 and ASL20/LBD18 Regulate Tracheary Element Differentiation in *Arabidopsis*. *Plant Cell* 20, 3359-3373.

Steyn, W.J., Wand, S.J.E., Holcroft, D.M., and Jacobs, G. (2002). Anthocyanins in vegetative tissues: a proposed unified function in photoprotection. *New Phytol* 155, 349-361.

Sugawara, S., Mashiguchi, K., Tanaka, K., Hishiyama, S., Sakai, T., Hanada, K., *et al.* (2015). Distinct Characteristics of Indole-3-Acetic Acid and Phenylacetic Acid, Two Common Auxins in Plants. *Plant Cell Physiol* 56, 1641-1654.

Taki, N., Sasaki-Sekimoto, Y., Obayashi, T., Kikuta, A., Kobayashi, K., Ainai, T., *et al.* (2005). 12-Oxo-Phytodienoic Acid Triggers Expression of a Distinct

Set of Genes and Plays a Role in Wound-Induced Gene Expression in *Arabidopsis*. *Plant Physiol* 139, 1268-1283.

Taylor-Teeples, M., Lin, L., de Lucas, M., Turco, G., Toal, T.W., Gaudinier, A., *et al.* (2015). An Arabidopsis Gene Regulatory Network for Secondary Cell Wall Synthesis. *Nature* 517, 571-575.

Titapiwatanakun, B., Blakeslee, J.J., Bandyopadhyay, A., Yang, H., Mravec, J., Sauer, M., *et al.* (2009). ABCB19/PGP19 stabilises PIN1 in membrane microdomains in *Arabidopsis*. *Plant J* 57, 27-44.

Tocquard, K., Lafon-Placette, C., Auguin, D., Muries, B., Bronner, G., Lopez, D., *et al.* (2014). In silico study of wall-associated kinase family reveals large-scale genomic expansion potentially connected with functional diversification in *Populus*. *Tree Genet Genomes* 10, 1135-1147.

Tokunaga, N., Kaneta, T., Sato, S., and Sato, Y. (2009). Analysis of expression profiles of three peroxidase genes associated with lignification in *Arabidopsis thaliana*. *Physiol Plantarum* 136, 237-249.

Turlapati, P., Kim, K.W., Davin, L., and Lewis, N. (2011). The laccase multigene family in *Arabidopsis thaliana*: towards addressing the mystery of their gene function(s). *Planta* 233, 439-470.

van Bakel, H., Stout, J.M., Cote, A.G., Tallon, C.M., Sharpe, A.G., Hughes, T.R., *et al.* (2011). The draft genome and transcriptome of *Cannabis sativa*. *Genome Biol* 12, R102.

van den Broeck, H.C., Maliepaard, C., Ebskamp, M.J.M., Toonen, M.A.J., and Koops, A.J. (2008). Differential expression of genes involved in C1 metabolism and lignin biosynthesis in wooden core and bast tissues of fiber hemp (*Cannabis sativa* L.). *Plant Sci* 174, 205-220.

van Raemdonck, D., Pesquet, E., Cloquet, S., Beeckman, H., Boerjan, W., Goffner, D., *et al.* (2005). Molecular changes associated with the setting up of secondary growth in aspen. *J Exp Bot* 56, 2211-2227.

Wang, Y., Chantreau, M., Sibout, R., and Hawkins, S. (2013). Plant cell wall lignification and monolignol metabolism. *Front Plant Sci* 4.

Wang, Y., Zhou, B., Sun, M., Li, Y., and Kawabata, S. (2012). UV-A Light Induces Anthocyanin Biosynthesis in a Manner Distinct from Synergistic Blue + UV-B Light and UV-A/Blue Light Responses in Different Parts of the Hypocotyls in Turnip Seedlings. *Plant Cell Physiol* 53, 1470-1480.

Zhang, Q., Cheetamun, R., Dhugga, K., Rafalski, J., Tingey, S., Shirley, N., *et al.* (2014). Spatial gradients in cell wall composition and transcriptional profiles along elongating maize internodes. *BMC Plant Biol.* 14, 27.

Zhao, C., Craig, J.C., Petzold, H.E., Dickerman, A.W., and Beers, E.P. (2005). The Xylem and Phloem Transcriptomes from Secondary Tissues of the *Arabidopsis* Root-Hypocotyl. *Plant Physiol* 138, 803-818.

Zhao, C., Avci, U., Grant, E.H., Haigler, C.H., and Beers, E.P. (2008). XND1, a member of the NAC domain family in *Arabidopsis thaliana*, negatively regulates lignocellulose synthesis and programmed cell death in xylem. *Plant J* 53, 425-436.

Zhong, R., and Ye, Z.H. (2007). Regulation of cell wall biosynthesis. *Curr Opin Plant Biol* 10, 564-572.

Zhong, R., Lee, C., Zhou, J., McCarthy, R.L., and Ye, Z.H. (2008). A Battery of Transcription Factors Involved in the Regulation of Secondary Cell Wall Biosynthesis in *Arabidopsis*. *Plant Cell* 20, 2763-2782.

Zhong, R., and Ye, Z.H. (2009). Transcriptional regulation of lignin biosynthesis. *Plant Signal Behav* 4, 1028-1034.

Zhong, R., and Ye, Z.H. (2014). Complexity of the transcriptional network controlling secondary wall biosynthesis. *Plant Sci* 229, 193-207.

Zhong, R., and Ye, Z.H. (2015). Secondary Cell Walls: Biosynthesis, Patterned Deposition and Transcriptional Regulation. *Plant Cell Physiol* 56, 195-214.

Zuniga-Sanchez, E., Soriano, D., Martinez-Barajas, E., Orozco-Segovia, A., and Gamboa-deBuen, A. (2014). BIIDXI, the At4g32460 DUF642 gene, is involved in pectin methyl esterase regulation during Arabidopsis thaliana seed germination and plant development. *BMC Plant Biol.* 14, 338.

Chapter 1: Highlights and Perspectives

Secondary growth is an essential aspect of hemp development, as it provides mechanical strength to the stem, allowing this organ to reach considerable heights at maturity. In plants at the beginning of seed maturation, the number of secondary bast fibres exceeds by several orders of magnitude the number of primary bast fibres. Secondary growth is therefore a crucial parameter with respect to the fibre yield of hemp. In this Chapter, we have therefore used a system suitable for the study of secondary growth, i.e. the hemp hypocotyl. Four time-points (6, 9, 15 and 20 days after sowing), corresponding to different stages of the hemp hypocotyl development and spanning elongation, transition to secondary growth and development of secondary tissues, have been investigated via microscopy, transcriptomics and phytohormone analyses, with a special emphasis on cell wall. Schematically, three key developmental stages may be drawn:

1. H6 and H9 undergo elongation. This stage is characterised by primary cell wall deposition and remodelling. Several genes involved in xyloglucan biosynthesis (CSL) and modification (class III peroxidases, expansins, XTHs and β-galactosidases) were up-regulated. The higher bioactive auxin content in these time-points may explain this expression pattern by regulating the expression of genes involved in elongation, such as expansins and XTHs.
2. The onset of secondary growth occurs between H9 and H15. In H15, a substantial increase of the bioactive jasmonates has been observed, together with a higher expression of genes involved in secondary growth such as *WAT1* and *WOX4*. Secondary bast fibres start to thicken.

3. The last stage analysed, H20, corresponds to further thickening of the bast fibres (accompanied by cellulose, rhamnogalacturonan-I and xylan deposition), concomitantly with a higher expression of the secondary cell wall *CesA*s and of the genes involved in monolignol biosynthesis and lignin polymerisation (laccases and class III peroxidases). The high content of abscisic acid at that point may inhibit hypocotyl elongation by modifying the organisation of the cytoskeleton, especially the orientation of the microtubules.

Most of the genes involved in secondary cell wall and G-layer deposition in hemp have thus been identified with this transcriptomic study. Genes involved in lignin biosynthesis were up-regulated at later stages of the hypocotyl development (H15 and H20). Considering that, i. lignin is an important parameter with respect to the quality of hemp fibres (such as suitability for retting), and ii. the link between lignification pattern and enzymes involved in lignin polymerisation (laccases and class III peroxidases) is still fragmentary, Chapter 2 will investigate, more in depth, the main features of the lignification during hypocotyl development. In this respect, the lignin content and monomeric composition of the four time-points will be investigated by chemical analyses and microscopic observations. These results will be put in perspective with the soluble proteome and a gene expression dataset targeting lignin polymerisation and genes which may be involved in lignan biosynthesis.

Chapter 2

Insights into the molecular regulation of monolignol-derived product biosynthesis in the growing hemp hypocotyl

Adapted from the published article: Marc Behr, Kjell Sergeant, Céline C. Leclercq, Sébastien Planchon, Cédric Guignard, Audrey Lenouvel, Jenny Renaut, Jean-Francois Hausman, Stanley Lutts, Gea Guerriero. Insights into the molecular regulation of monolignol-derived product biosynthesis in the growing hemp hypocotyl. 2018. *BMC Plant Biology*, 18:1.

Author contributions:

MB, JFH, SL and GG designed the experiments; MB performed RNA and protein extractions, qPCR, phylogenetic analysis, laccase and peroxidase activities assessment, imaging, lignin quantification, nitrobenzene oxidation, analysed all the data and wrote the manuscript; KS identified peptides and wrote the methods for protein identification (database and MS analysis); CCL and SP performed protein extraction, gel-free and gel-based proteomics experiments and wrote the methods for gel-free and gel-based proteomics; CG and AL run the GC-MS for lignin characterisation and wrote the methods for GC-MS analysis; JR supervised the proteomics analysis.

Abstract

Lignin and lignans both originate from the monolignol pathway. Despite the similarity of their building blocks, they fulfil different functions *in planta*. Lignin strengthens the tissues of the plant, while lignans are involved in plant defence and growth regulation. We propose to study the general molecular events related to monolignol-derived product biosynthesis, especially lignin, in the growing hemp hypocotyl (between 6 and 20 days after sowing). The present work confirms the validity of this system to study the regulation of lignin and lignan biosyntheses. The quantity and the monomeric composition of lignin, as well as the activity of enzymes involved in lignin polymerisation, are shown to be age-dependent. Based on the phylogenetic analysis and targeted gene expression, we suggest a role for the hemp dirigent and dirigent-like proteins in lignan biosynthesis. The transdisciplinary approach adopted results in the gene- and protein-level quantification of the main enzymes involved in the biosynthesis of monolignols and their oxidative coupling (laccases and class III peroxidases), in lignin deposition (dirigent-like proteins) and in the determination of the stereoconformation of lignans (dirigent proteins). Our work sheds light on how the provision of the precursors needed to synthesize the aromatic biomolecules lignin and lignans is regulated in the growing hemp hypocotyl at the transcriptional and proteomic level.

Introduction

The monolignol-derived products lignin and lignans are important plant specialized (secondary) metabolites. They are involved in crucial events related to plant development, such as plant defence, growth regulation, sap conduction and erect growth habit. Lignin strengthens mechanically the xylem by impregnating the secondary cell wall of both tracheary elements and fibres. This phenomenon is particularly obvious in trees, but herbaceous plants also undergo lignification, particularly in the xylem of roots and the hypocotyl (Herrero *et al.,* 2013). Lignans are formed by the stereospecific oxidative coupling of hydroxycinammyl alcohols mediated by a dirigent protein and an oxidase, typically a laccase (Davin *et al.*, 1997; Dixon *et al.*, 2002). Their role in plant defence is known since more than a decade (Dixon *et al.*, 2002), but their plant growth regulatory activity is still under investigation (Yamauchi *et al.*, 2015). Very few studies have addressed how this growth regulatory activity is mediated *in planta* (Binns *et al.*, 1987; Bonawitz and Chapple, 2013). Lignin and lignans originate from the same building blocks, i.e. the monolignols, but have very different functions. Therefore, the allocation of the building blocks to either biosyntheses has to be precisely tuned, both temporally and spatially. As for the majority of genes involved in secondary cell wall deposition, the expression of lignin biosynthetic genes is regulated by master regulators belonging to the NAC and MYB transcription factor families (Zhao and Dixon, 2011). By contrast, the expression of pinoresinol lariciresinol reductase (PLR), a key gene of lignan biosynthesis, was found to be higher in the young stem as compared to older stems in *Forsythia* x *intermedia* (Kwon *et al.*, 2001), suggesting that lignan biosynthesis may be somehow independent from secondary cell wall deposition. However, a role has been ascribed to lignans in

secondary cell wall-forming tissues, where they may participate in the maintenance of the cell redox homeostasis during lignification (Huis *et al.*, 2012; Zhao *et al.*, 2015). The genes coding for enzymes involved in lignan biosynthesis may thus be different according to the stage of development and tissue. This is illustrated by the expression patterns of the pinoresinol reductases *AtPRR1* and *AtPRR2*. The former is coexpressed with several secondary cell wall genes in the lignified internodes, while the latter is highly expressed in the growing hypocotyl (Zhao and Dixon, 2011). Lignan and lignin biosyntheses may be intertwined, as a triple laccase mutant which displays strongly reduced lignin content shows a higher transcript level of *PRR2* (Zhao *et al.*, 2013).

Some dirigent proteins (DIR) are putatively involved in lignan biosynthesis, while others are related to lignin deposition. For example, AtDIR10/ESB1 plays an essential role in the formation of the Casparian strip in *Arabidopsis* by targeting lignin polymerization at specific extracellular sites (Hosmani *et al.*, 2013) and AtDIR6 confers the (-) stereoconformation to pinoresinol (Kim *et al.*, 2012). The phylogenetic analysis of DIR helps to differentiate between those DIR involved in lignan biosynthesis (e.g. AtDIR6) and those to which other functions, such as lignin deposition, are assigned and referred to as DIR-like. According to Ralph (Ralph *et al.*, 2006), proteins of the DIR-a subgroup (AtDIR5, AtDIR6, AtDIR12 and AtDIR13) are wound- or insect-induced while proteins from the DIR-b subgroup are part of the constitutive defence of the plant. Since most of DIR have not yet been functionally characterised, phylogenetic analysis is a helpful tool that complements gene coexpression analysis for functional prediction (Behr *et al.*, 2015).

Because lignin is racemic and since there is no mass spectral evidence about the occurrence of oligolignol stereoisomers, the

proposed model of protein-driven control and template replication of lignin polymerisation (Davin and Lewis, 2000), has failed to prove its robustness (Ralph *et al.*, 2008). Therefore, the currently known functions of DIR are restricted to confer stereospecificity to lignans and to influence lignin deposition. Data from gene expression analysis during specific stages of development (stem elongation, secondary growth, xylem lignification) may contribute to get further details about their functions.

Following monolignol biosynthesis, lignin polymerisation is performed through end-wise radical coupling of phenols to the free-phenolic end of the growing polymer (Ralph *et al.*, 2008). The formation of the radical is catalysed by either laccase or class III peroxidase and mainly takes place in the apoplast. In Angiosperms, the relative proportion of *p*-coumaryl (H), guaiacyl (G) and syringyl (S) units of the lignin polymer depends on the tissue, the stage of development and the subcellular compartment. Starting from phenylalanine, the biosynthesis of H, G and S units requires the activities of 5, 8 and 10 enzymes, respectively (Figure 2-1). The monolignols are interconnected by either carbon-carbon (condensed), such as resinol (*β*-*β*) or phenylcoumaran (*β*-5), or ether (uncondensed) linkages such as *β*-ether (*β*-O-4) (Ralph *et al.*, 2008).

In Angiosperms, the water-conducting cells of the xylem are enriched in G-lignin, while structural fibres (both from xylem and phloem) have a high S-lignin content (Bonawitz and Chapple, 2010). Lignification begins with the deposition of the G units, notably in the secondary cell wall of xylem cells (Berthet *et al.*, 2011; Schuetz *et al.*, 2014), under the activity of laccases (LAC). H units are targeted to the middle lamella, while G units are initially deposited in the S1 sub-layer of both Gymnosperms and Angiosperms (Davin and Lewis, 2000).

Less is known about the role of enzymes partaking in the oxidation and deposition of the lignin macromolecule. It has been suggested that laccases (LAC) and peroxidases (PRX) do not function redundantly in lignin polymerisation in the vascular tissues of *A. thaliana* (Zhao *et al.*, 2013). Indeed, a knock-out mutant of *AtPRX52* has shown a decrease in the S units in the interfascicular fibres (Fernández-Pérez *et al.*, 2014), suggesting that laccases do not compensate for the loss of peroxidase activity. By contrast, the lignin of *AtPRX2*, *AtPRX25* and *AtPRX71* knock-out mutants is richer in S units than wild type (Shigeto *et al.*, 2013), showing that a lower peroxidase activity is not synonymous to a decrease in S unit. The silencing of two genes coding for laccases (*AtLAC4*, *AtLAC17*) induces an increase in the fibre S/G ratio (Berthet *et al.*, 2011). AtLAC17 is specifically involved in the deposition of G lignin in fibres, while the specific activity of AtLAC4 is less clear. More recently, the regulation of flax laccase expression by the microRNA *miR397* has been shown (Le Roy *et al.*, 2017). The question thus arising relates to the activity of these enzymes towards the oxidation of specific lignin subunits.

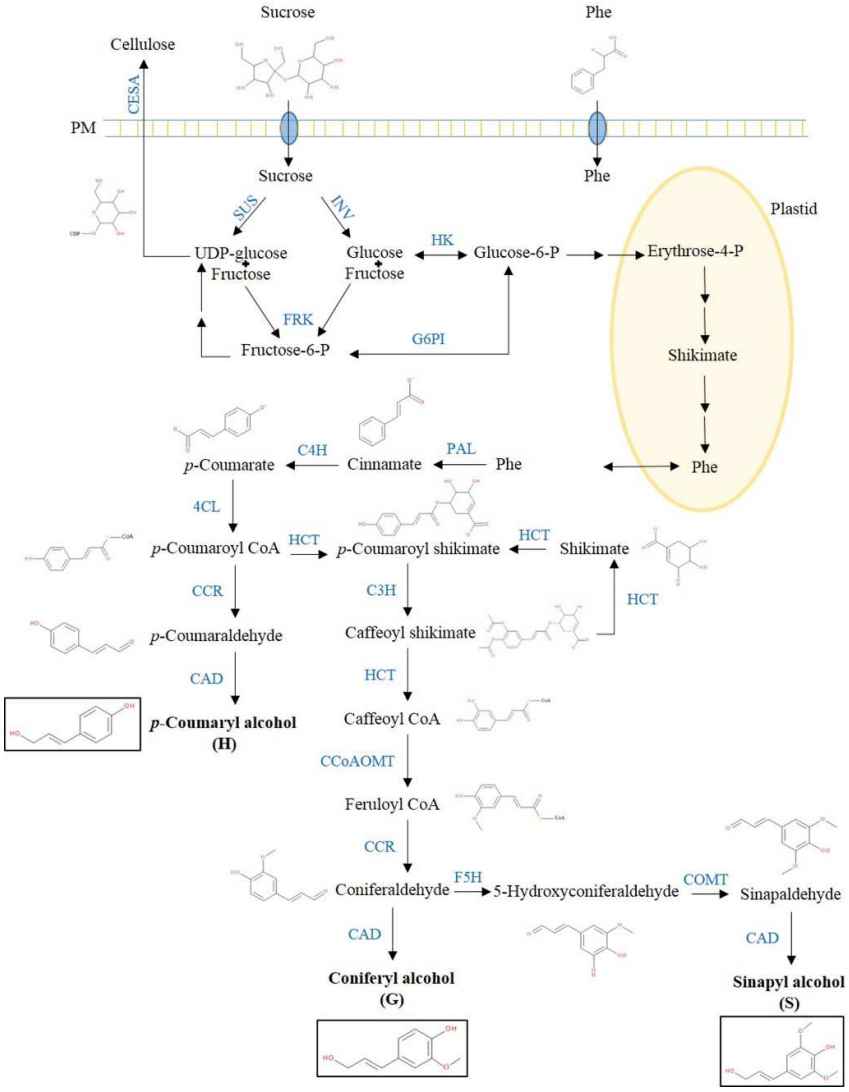


Figure 2-1: Monolignol and cellulose pathways. The molecules (in black) and enzymes (in blue) of the two pathways are indicated. Cellulose synthase (CESA) is membrane bound. Cinnamate-4-hydroxylase (C4H) and coumarate 3-hydroxylase (C3H) localise in the endoplasmic reticulum; hydroxycinnamoyl transferase (HCT) is partially associated with the endoplasmic reticulum (Bassard *et al.*, 2012). All the other enzymes are active in the cytosol. 4CL 4-coumarate ligase, *(continued next page),*

(continued legend Figure 2-1) CAD cinnamyl alcohol dehydrogenase, CCoAOMT caffeoyl-CoA 3-O-methyltransferase, CCR cinnamoyl CoA reductase, COMT caffeate O-methyltransferase, F5H ferulate 5-hydroxylase, FRK fructokinase, G6PI glucose-6-phosphate isomerase, HK hexokinase, INV invertase, PAL phenylalanine ammonia lyase, Phe phenylalanine, SUS sucrose synthase.

Between 6 and 20 days after sowing, the hemp hypocotyl was shown to be a suitable system to study the molecular events underlying secondary growth and secondary cell wall deposition (Chapter 1). The late stages of hypocotyl development are characterised by the up-regulation of transcription factors and genes involved in the synthesis of precursors needed for secondary cell wall deposition, the biosynthesis of monolignols and lignin polymerization. The same experimental set up is here used to study lignin biosynthesis: after chemical characterisation of lignin, the laccase and peroxidase activities are assessed. These results are put in perspective with gene expression and proteomics data. We also provide preliminary results dealing with putative orthologs involved in lignan biosynthesis, based on phylogenetic and targeted gene expression analyses.

I. Material and Methods

a. Plant growth and sampling

The hypocotyls have been grown and sampled following the conditions described in Chapter 1. Each biological replicate consisted in 20 hypocotyls.

b. Lignin analysis

Lignin quantification. Lignin content has been assessed on preparations of cell wall residue (CWR) of 4 or 5 biological replicates (Chantreau *et al.*, 2014). CWR was obtained by washing the powdered plant material first with methanol (80%) under agitation for 4 hours, followed by five additional vortexing / centrifugation cycles with ethanol (80 %). After drying, 5 mg of CWR were digested with 2.6 mL of 25 % acetyl bromide in glacial acetic acid for 2 hours at 50°C using a Hach LT200 system. This method is based on the formation of acetyl derivatives in unsubstituted –OH groups, with the α-carbon OH being replaced by bromide, resulting in solubilisation of lignin in acidic conditions (Moreira-Vilar *et al.*, 2014). After digestion, the solution was transferred to a 50 mL Falcon tube containing 10 mL of 2 M sodium hydroxide and 12 mL of glacial acetic acid. The reaction tube was rinsed with glacial acetic acid and 1.75 mL of 0.5 M hydroxylammonium chloride was added. Finally, the total volume was adjusted to 30 mL with glacial acetic acid. The absorbance of the solution was read at 280 nm in a spectrophotometer, with an extinction coefficient of 22.9 g^{-1} L cm^{-1} for lignin determination.

Lignin characterisation. Lignin was characterised on 3 or 4 biological replicates using the nitrobenzene oxidation method (Billa *et al.*, 1996). This method, in alkaline condition, cleaves the propane

chain, resulting in the formation of aromatic aldehydes (Tarabanko and Tarabanko, 2017). 10 mg of CWR were digested with 2 mL of 2 M NaOH and 30 µL nitrobenzene at 165°C for one hour (Hach LT200 system). After centrifugation, ca. 1500 µL of supernatant was collected and 10 µl of vanillin-D3 (Sigma-Aldrich) at 10 mg/mL in 1,4-dioxan were added as a surrogate standard. Nitrobenzene was removed by four washing steps with ethyl acetate (1 mL, vortexing / centrifugation cycle). The pH of the solution was adjusted to 2-3 by adding approximately 200 µL of 6 N HCl solution. The oxidation products were recovered by two successive extractions with 1 mL ethyl acetate (vortexing / centrifugation cycle) followed by cleaning with 500 µl of saturated NaCl solution and drying with Na_2SO_4. The GC-MS analysis was performed after trimethylsilylation, realized by addition of 50 µl of bis(trimethylsilyl)trifluoroacetamide (BSTFA) to 50 µL of dried extract and derivatization at 60°C for 30 min. Quantitative analyses were performed using a HP-5MS column (30 m x 0.25 mm, 0.25 µm, Agilent) installed in a 7890B-5977A GC-MS system (Agilent). Injection was done at 250°C in splitless mode. The oven program started at 40°C for 5 minutes, increased to 230°C at 10°C/min, then to 320°C at 40°C/min and was kept at 320°C for 10 min. Salicylic acid-D4 was used as internal standard.

c. Gene expression analysis

The gene expression analysis was carried out on biological triplicates using RT-qPCR as described in Chapter 1. The RNAs displayed a RIN value between 7.9 and 10. The specificity of the products was checked at the end of each run with the melt curve. Relative gene expressions were determined using the qBase^{PLUS} software v2.5 (Biogazelle). *CsaETIF3e* and *CsaETIF3h* were the most stable reference genes

among *ETIF3e, ETIF3h*, *Tubulin* and *ETIF4e*. The genes are named based on the putative orthology with the genes from *Arabidopsis*.

d. Proteomics

Hemp hypocotyl soluble protein extraction. The proteomics experiments were carried out with five biological replicates using both gel-based and gel-free methods. Approximately 300 mg of material were treated with ice-cold extraction buffer (TCA 20 %, DTT 0.1 % in acetone) and allowed to precipitate overnight at -20°C. After centrifugation (30,000 g; 45 min at 4°C), the pellet was washed three times in ice-cold acetone and dried *in vacuo*. The dried extract was solubilised in 500 µl labelling buffer (7 M urea, 2 M thiourea, 2% CHAPS, 30 mM Tris) for 30 mg and incubated at room temperature under agitation (900 rpm) for 1h. After centrifugation (15,000 g, 15 min), the supernatant was transferred to a fresh tube and pH was adjusted at 8.5 with sodium hydroxide (50 mM). Protein concentration was determined using the 2-D Quant Kit (GE Healthcare) with BSA for the standard curve according to the protocol defined by the manufacturer.

Gel-based proteome study. 50 µg of protein were labelled with Cy-dyes. Following the labelling, the samples were handled as described in (Szopinska *et al.*, 2016). The analysis of the gel images was performed with the DeCyder™ software (GE Healthcare, v. 7.0.8.53). Spots were considered as significantly different when detected on at least 75% of analysed gel images, protein abundance with a minimum fold change of 1.5 with a Student's t-test *p*-value below 0.05. Following MALDI analysis, the mass spectra of digested peptides were identified by carrying out a MASCOT database search against our in-house hemp transcriptome database (170,598 sequences; 64,508,806 residues) annotated using Blast2GO PRO version 3.0

against the *A. thaliana* non-redundant database and the NCBI Viridiplantae database, with the following parameters: fixed modifications: carbamidomethyl (C); variable modifications: dioxidation (W), oxidation (HW), oxidation (M), Trp→kynurenin (W); mass values: monoisotopic; peptide mass tolerance: ± 100 ppm; fragment mass tolerance: ± 0.5 Da and Max number of missed cleavages: 2. Individual ions scores greater than 42 indicate identity or extensive homology ($p < 0.05$), protein scores greater than 65 are significant ($p < 0.05$). A protein was identified with only one peptide if the individual ion score was higher than 84. Principal Component Analysis (PCA) was performed with the DeCyder™ software.

Gel-free proteome study. 25 µg of proteins were loaded in a Criterion™ XT precast 1D gel 4-12% Bis-Tris (1.0 mm X 12 wells, Bio-Rad). After denaturation of the sample, migration was performed at 200 V during 8 min. The gel was stained with Instant Blue (Gentaur BVBA) for 1 hour. Bands were excised and cut into 1-2 mm cubes at 4°C. The proteins in each band were subsequently reduced (NH_4HCO_3 100 mM + DTT 10 mM, 30 minutes at 56°C), alkylated (NH_4HCO_3 100 mM + IAA 55 mM, 20 min at room temperature), de-stained and digested with trypsin (5 ng/µl in NH_4HCO_3 50 mM, 30 min on ice and 37°C overnight). Peptides were extracted from the gel with ACN 50% / TFA 0.1%, dried with a Speedvac and stored at -20°C until LC-MS analysis. Peptides were analysed with a NanoLC-2D System (Eksigent) coupled to a TripleTOF 5600+ MS (AB Sciex). After desalting and enrichment (C18 PepMap™, 5 µm, 5 mm * 300 µm, Thermo scientific), the peptides were separated with a C18 reverse phase column (PepMap™ 100, 3 µm, 100 Å, 75 µm x 15 cm, Thermo scientific) using a linear binary gradient (solvent A: 0.1% formic acid; solvent B: 80% ACN, 0.1% formic acid) at a flow rate of 300 nl/min. Peptides were eluted from 5% to 55%

solvent B over 40 min, afterwards eluent B increased to 100% to wash the column and the column was re-equilibrated. The LC was coupled to the mass spectrometer with a NanoSpray III source. CID fragmentations for MS/MS spectra acquisitions used the automatically adjusted system of rolling collision energy voltage. A full MS scan was performed (scan range: 300–1250 m/z, accumulation time: 250 ms) and the 20 most intense precursors selected for fragmentation. The CID spectra was analysed with Mascot-Daemon using the hemp database as for gel-based proteomics using the following parameters: 2 missed cleavages, mass accuracy precursor: 20 ppm, mass accuracy fragments: ± 0.3 Da, fixed modifications: carbamidomethyl (C), dynamic modifications: Oxidation (M), Acetyl (protein N-term). Only the contigs where at least one time point has 3 (out of 5) or more than 1 spectral count were considered for further analysis. This filter being applied, a value of 0.5 was added to all the spectral count to compensate for null values and allowing logarithmic transformation (George *et al.*, 2015). The relative quantities of the proteins have been calculated using the Normalized Spectral Abundance Factor (NSAF).

$$NSAF = (100*SpC/MW)/\Sigma(SpC/MW)N$$

where SpC = Spectral Counts, MW = Protein molecular weight in Da and N = Total Number of Proteins (Fanutza *et al.*, 2015).

The Independent Component Analysis (ICA) has been calculated with MetaGeneAlyse (Daub *et al.*, 2003) using the NSAF values. Student's T-tests on the fold-change between the time-points have been calculated with ln-transformed values based on the five biological replicates.

Finally, the list of the differentially abundant proteins was obtained using the same parameters as for 2D-DiGE (fold-change

(NSAF) > 1.5 and Student's t-test *p*-value below 0.05). The NSAF values have been displayed using PermutMatrix (Caraux and Pinloche, 2005) with the following parameters: dissimilarity assessed by Pearson distance, clustering in complete linkage, seriation and tree seriation in multiple-fragment heuristic (MF), rows normalized by Z-score scaling. The proteins are named following the *Arabidopsis* nomenclature.

e. Imaging

FASGA and Mäule staining. For FASGA imaging, hemp hypocotyls were embedded in resin as described in Chapter 1, and cut at a thickness of 10 µm with a microtome. Cross sections deposited on microscopic slides were incubated for 15 min in the FASGA solution at 55°C, rinsed with pure water and observed with an optical microscope (Leica DMR). FASGA is a mix of Alcian blue (affinity for cellulose, stained in blue) and safranin (affinity for phenolics / lignin, stained in orange-red). Mäule staining was performed on fresh hand-cut sections. Sections were incubated in permanganate solution at 1% (w/v) for 5 min, rinsed with pure water, washed with 3.6% hydrochloric acid, mounted in saturated ammonia solution and immediately observed with an optical microscope (Leica DMR).

Peroxidase and laccase activities. Fresh hand cross-sections of H15 and H20 were treated with metal enhanced 3,3'-diaminobenzidine (DAB) substrate kit (ThermoFisher, number 34065), which reacts with horseradish peroxidase in presence of peroxide. Before the incubation for laccase activity, sections were incubated with catalase (100 $\mu g.mL^{-1}$ in TBS pH 7) for 3 h at room temperature. The DAB solution was diluted 10-fold in stable peroxide buffer (Nakagawa *et al.*, 2014) for peroxidase activity and in TBS (pH 6) containing 100 $\mu g.mL^{-1}$ of catalase for laccase activity. Sections were incubated at room temperature for 15 to 30 min, rinsed twice with water, mounted

in water and immediately observed with an optical microscope (Leica DMR). A negative control without DAB solution was included for each time-point, resulting in the absence of the brown signal for both laccase and peroxidase assays. Salicylhydroxamic acid (5 mM) was used as an inhibitor of peroxidase activity (Dunand *et al.*, 2007) and sodium azide (1 mM) as an inhibitor of laccase activity (Johannes and Majcherczyk, 2000).

II. Results

a. Time-course analysis of lignification in the hypocotyl

Hypocotyl lignification was followed between 6 (H6) and 20 (H20) days after sowing. The development was monitored by staining the cross sections with a FASGA solution. In H6 and H9, the xylem cells are stained purple (Figure 2-2A-B). In H15 and H20, the cambial cells display a strong blue coloration, together with a thin layer of the bast fibres (Figure 2-2C-D, insets). A weak red edging is also visible in the cell wall of the bast fibres, indicating the presence of a small amount of lignin (Figure 2-2C-D). The Mäule staining confirms the lignification of the bast fibres, with a red coloration in both primary (H15, Figure 2-2E) and secondary bast fibres (H20, Figure 2-2F). As expected, lignin is detected in all the xylem cells, both with FASGA and Mäule. The Mäule staining differentiates primary from secondary xylem cells based on the differential presence of syringyl (S) units in the lignin of these two tissues. In H15 and H20, primary xylem cells are stained brown, indicating that the lignin polymer lacks the S units; the fibres of the secondary xylem are instead stained red indicating that they are rich in S lignin (Figure 2-2E-F-G-H).

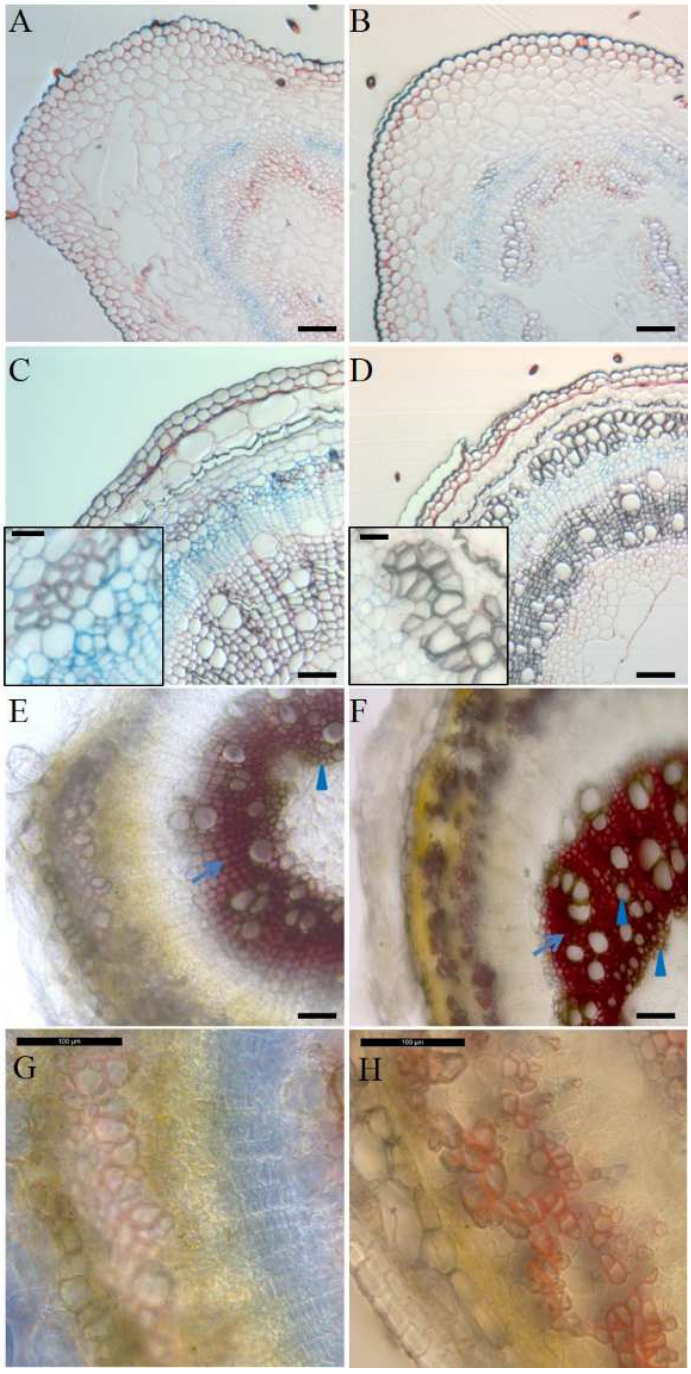

A
B
C
D
E
F
G
H

Figure 2-2: Lignification of the hemp hypocotyl between 6 and 20 days. A to D, fixed cross sections of H6 (A), H9 (B), H15 (C) and H20 (D) stained with FASGA. Lignification of the bast fibres is illustrated in the insets of C and D. E to H, fresh cross-sections of H15 (E and G) and H20 (F and H) stained with Mäule reagent. Higher magnifications of the bast fibres are shown in G and H. The primary xylem cells and tracheary elements of the secondary xylem are indicated with arrow heads, while fibres of the secondary xylem are indicated with an arrow in E and F. Scale bar = 100µm in the main pictures and 25µm in the insets.

Lignin content and composition were determined using the acetyl bromide method and by quantification of products after nitrobenzene oxidation (NBO) of cell wall residue, respectively (Table 2-1).

Table 2-1: Lignin content and monomer composition of the hemp hypocotyl

	H6	H9	H15	H20
Lignin (% CWR)	1.94 (0.39) a	2.71 (0.48) a	4.54 (0.20) b	4.70 (0.19) b
H (µmol / g CWR)	9.72 (0.94) b	7.82 (1.40) ab	4.79 (0.73) a	5.24 (0.90) a
V (µmol / g CWR)	54.12 (8.19) a	73.07 (15.08) a	41.72 (7.14) a	84.93 (22.82) a
S (µmol / g CWR)	9.69 (0.95) a	17.41 (2.00) a	36.09 (7.98) ab	87.29 (23.52) b
S / V	0.18 (0.01) a	0.26 (0.03) a	0.85 (0.04) b	1.04 (0.03) c
H / (H+V+S)	0.13 (0.01) d	0.08 (0.00) c	0.06 (0.00) b	0.03 (0.00) a

Within a row, values with different letters are significantly different (Tukey *p*-value < 0.05). The values indicate the average (± SEM) of 3-5 biological replicates. CWR cell wall residue, H *p*-hydroxybenzaldehyde, V vanillin, S syringaldehyde.

Significant differences were found across development, from 1.94% of the cell wall residue (CWR) in H6 to 4.70% in H20. The three main lignin degradation products (*p*-hydroxybenzaldehyde, vanillin and syringaldehyde corresponding to H, G and S monolignol units, respectively) were recovered in all the samples. Since the proportion of H-units is an indicator of lignin condensation (referring to carbon-carbon interunit linkage), the decrease in the H / (H+V+S) ratio (from 13% to 3%) may indicate that lignin is less condensed in older hypocotyls. The S / V ratio, by contrast, increased with the hypocotyl age (from 0.18 to 1.04).

b. Gene expression analysis during hypocotyl lignification

DIR have been divided into two subfamilies, namely DIR-a (DIR) and DIR-like (DLP, (Ralph *et al.*, 2006)), based on their sequences and possible biochemical functions. In order to predict a function for the dirigent proteins of *C. sativa*, a phylogenetic tree was built with DIR and DLP from *Arabidopsis*, *Linum usitatissimum, Forsythia* x *intermedia*, *Schisandra chinensis* and poplar (Hertzberg *et al.*, 2001;Ralph *et al.*, 2006;Kim *et al.*, 2012;Dalisay *et al.*, 2015). Two main clades appeared (Figure 2-3). Clade I contains pinoresinol forming DIR from *Arabidopsis*, *L. usitatissimum*, *F. intermedia* and *S. chinensis* as well as two DIR from *C. sativa*, i.e. CsaDIR6A and CsaDIR6B. Furthermore, two sub-clades (I-a and I-b) respectively group (-)- and (+)- pinoresinol forming DP. Interestingly, CsaDIR6A belongs to subclade I-a while CsaDIR6B belongs to subclade I-b. Clade II contains the DIR-like proteins from *Arabidopsis*. The subclades II-a and II-b correspond to the DIR-d and DIR-e family described by Ralph and colleagues (Ralph *et al.*, 2006), respectively.

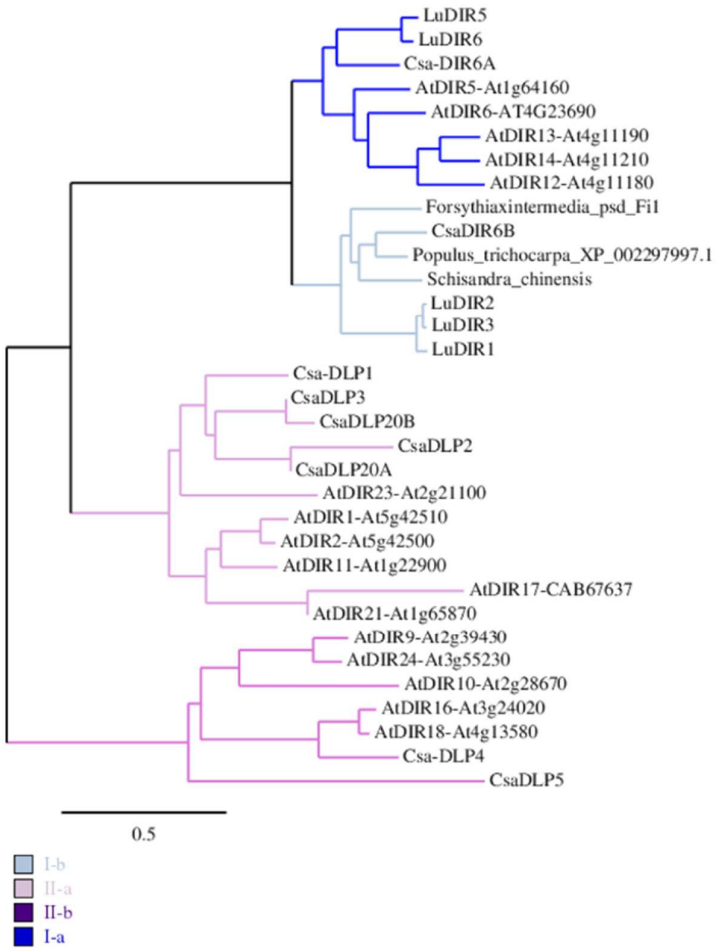


Figure 2-3: Phylogenetic analysis of DIR and DIR-like proteins (DLP). *Cannabis sativa* (Csa), *A. thaliana* (At), *Forsythia* x *intermedia*, *Populus trichocarpa, Schisandra chinensis* and *Linum usitatissimum* (Lu). Neighbour-joining tree calculated with 1000 bootstraps replicates with bioNJ algorithm (phylogeny.fr; (Dereeper *et al.*, 2008)). Scale bar: expected numbers of amino acid substitutions per site. Sequences in Supplementary data 2-1.

Two patterns of gene expression are observed based on the hierarchical clustering (Figure 2-4). With the exception of *DLP20A*, all

the DIR-like genes branched in the left cluster (*DLP1*, *DLP20B*, *DLP3*, *DLP2*, *DLP5* and *DLP4*). Their corresponding proteins belong to clade II of the phylogenetic analysis (Figure 2-3). *DLP20A*, together with *DIR6A* and *DIR6B* (clade I of the phylogenetic tree), were part of the bottom cluster. Genes of the left cluster were either more expressed in H6 or H9 (*DLP1*, *DLP4*, *PRR1*, *LAC17*), or showed no major changes in their expression. PLR was most expressed in H6 and H9. The opposite trend was observed for the right cluster: genes involved in secondary cell wall biogenesis (*NST1*, *MET1*, *SAM*, *PRX49*, *PRX52* and *PRX72*), *DIR6A*, *DIR6B* (based on sequence clustered with proteins involved in pinoresinol biosynthesis) and *DLP20A* were more expressed in H15 or H20. Two trends of expression were thus observed for the DIR and DLP, which may point to different physiological roles.

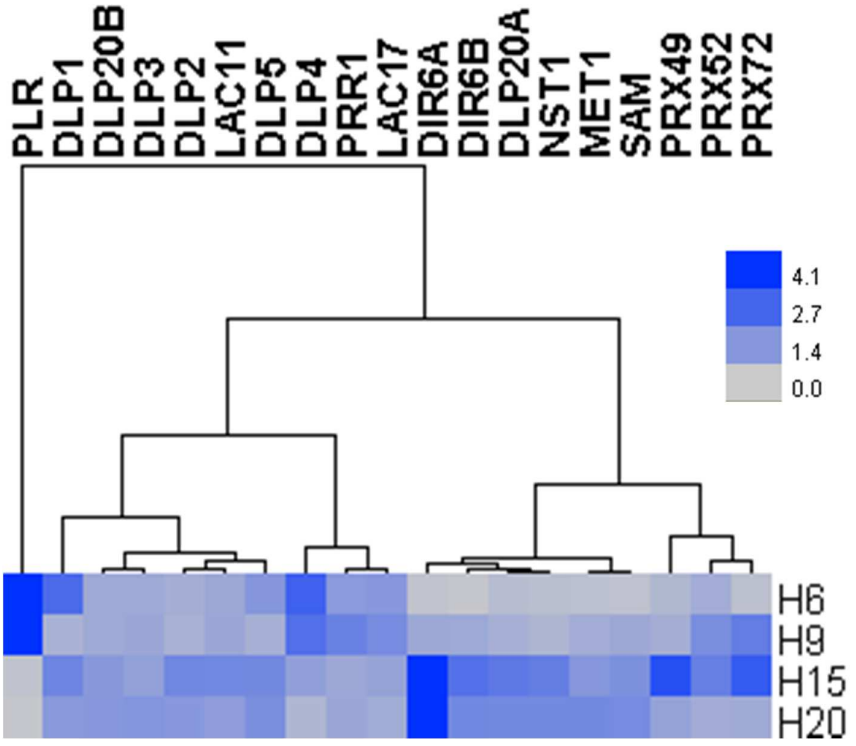


Figure 2-4: Heatmap hierarchical clustering showing the expression of genes assessed by RT-qPCR (Supplementary data 2-2). Values represent Calibrated Normalized Relative Quantities (CNRQ) calculated with qbase+. DLP dirigent-like protein, LAC laccase, PRR1 pinoresinol reductase 1, DIR dirigent protein, NST1 NAC secondary cell wall thickening 1, MET1 methionine synthase 1, SAM *S*-adenosylmethionine synthase, PRX peroxidase, PLR pinoresinol-lariciresinol reductase. The colour bar indicates the expression values represented as an increasing intensity gradient.

c. Proteomics analysis by gel-based and gel-free methods

Soluble proteomes from H6, H9, H15 and H20 have been analysed in 5 biological replicates by gel-based and gel-free methods (Supplementary data 2-3).

In the two-dimensional difference gel electrophoresis (2D-DiGE) experiment, 433 spots were reproducibly matched across the gels and used in a Principal Component Analysis (PCA, Figure 2-5A). In the LC-MS experiment, only peptides detected in at least three replicates (*i.e.*, at least one spectral count) were considered for the Independent Component Analysis (ICA, Figure 2-5B), resulting in 404 variables. Both approaches allowed to discriminate young (H6 and H9) from more mature (H15 and H20) hemp hypocotyls. Using LC-MS/MS, more proteins involved in cell wall formation and, more specifically, in monolignol biosynthesis, were identified and quantified. The patterns of abundance of these proteins are shown using the NSAF values (Figure 2-6).

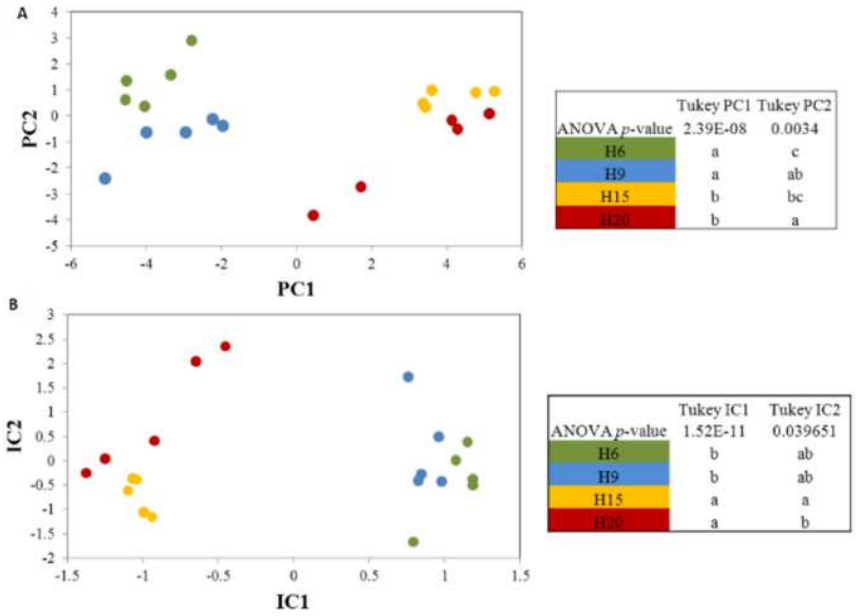


Figure 2-5: Clustering of the proteome profiles of hypocotyls at different ages. H6 green dots, H9 blue dots, H15 orange dots, H20 red dots. A: Principal component analysis based on the gel-based proteome study (total variance explained: 67.2+10.8=78.0%). B: Independent component analysis of LC-MS/MS based proteome profiles (total variance explained: 62.3%). In both panels, the significance of the coordinates in the two main axes was assessed using a Tukey post-hoc test, different letters within one column indicate that the proteome profiles are significantly different .

When the proteins involved in cell wall development and monolignol synthesis are used for a hierarchical clustering (Figure 2-6), three groups can be discerned. Group I (in yellow) includes proteins which are more abundant in H15 or H20 as compared to younger hypocotyls, namely UDP-xylose synthase 5 (UXS5), fructokinase-3 (FRK3), fasciclin-like arabinogalactan 12 (FLA12), PAL1, cellulase, SUS4, UDP-glucose dehydrogenase 4 (UDG4), methionine synthase 2 (METS2), *S*-adenosylmethionine synthases 3-4 (SAMS3-4), *S*-ADENOSYL-L-HOMOCYSTEINE HYDROLASE 2 (SAHH2),

methylenetetrahydrofolate reductase 2 (MTHFR2), annexin 2 and one invertase/PME. The patterns of gene expression of *SAMS* and *METS* (Figure 2-4) are closely related to the abundance of their respective proteins.

The proteins of the group II (in magenta) are more abundant in H20. They include the majority of the proteins involved in the phenylpropanoid and monolignol biosynthetic pathways, as well as in lignin polymerisation, such as PAL1, 4CL2, HCT, CAD4, COMT, CCoAOMT, PHENYLCOUMARAN BENZYLIC ETHER REDUCTASE (PCBER), and orthologs of AtPRX52, AtPRX54 and AtPRX3. Some proteins having a role in cell wall biosynthesis and modification are also present in this second group: PATELLIN-3, PHOSPHOGLUCOMUTASE 3 (PGM3), FLA11, POLYGALACTURONASE INHIBITING PROTEIN 1 (PGIP1), METHYLESTERASE, CLASS V CHITINASE, VOLTAGE DEPENDENT ANION CHANNEL 1 (VDAC1). The gene expression of other peroxidases involved in lignin polymerisation (orthologs of AtPRX49, AtPRX52 and AtPRX72) are similar to those detected in the LC-MS experiment, highlighting the ongoing lignification in the old hypocotyls. PATELLIN-3 is involved in the protein transport to the plasma membrane (Ito *et al.*, 2011; Nikolovski *et al.*, 2012). VDAC1 is involved in the regulation of hydrogen peroxide generation (Tateda *et al.*, 2011) and thus may play a role in lignin polymerisation by the peroxidases.

Finally, the proteins present in the group III (in blue) are more abundant in young hypocotyls. They are mainly devoted to the modification of the cell wall to ensure the extensibility of the hypocotyl: FLA1, PRXs, PMEs, XYLOGLUCAN ENDOTRANSGLUCOSYLASE/HYDROLASE 5 (XTH5), ALPHA-

XYLOSIDASE 1 (XYL1), BETA-XYLOSIDASE 1-4 (BXL1, XYL4), and ALPHA-L-ARABINOFURANOSIDASE 1 (ASD1). 10-FORMYLTETRAHYDROFOLATE SYNTHETASE (THFS) is involved in the metabolism of folic acid.

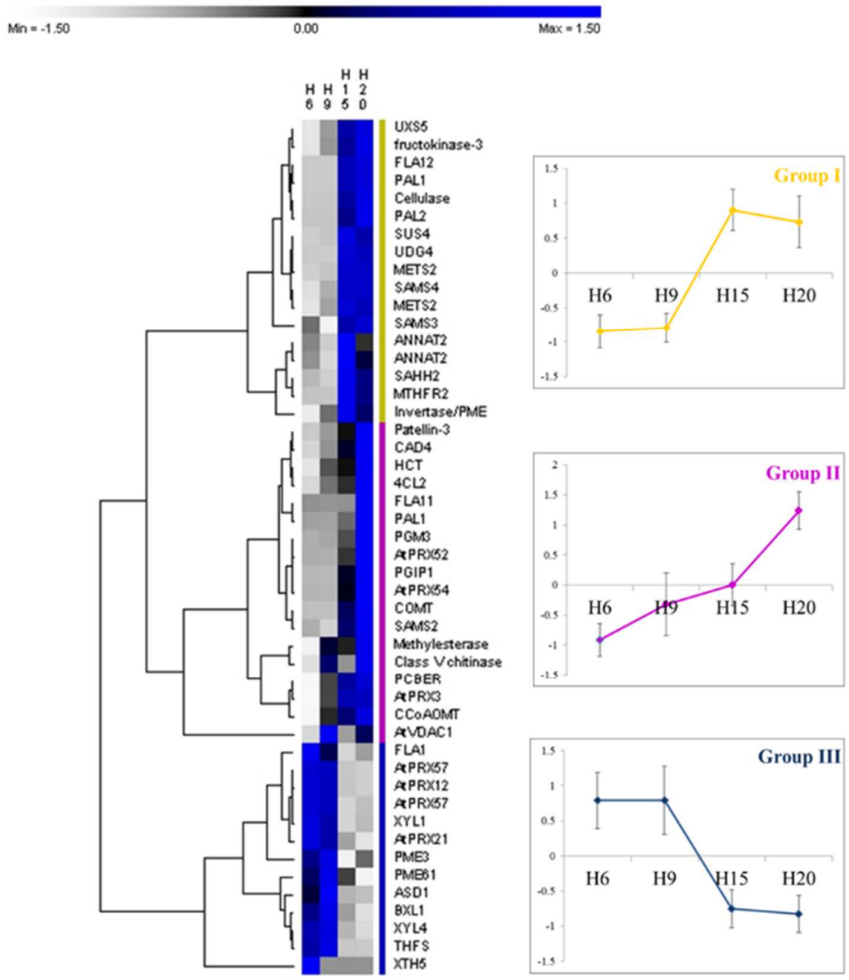


Figure 2-6: NSAF relative quantities of proteins involved in cell wall biogenesis assessed by LC-MS. Abbreviations are as in the text. For each group, the average of the abundances as calculated for the hierarchical clustering was plotted (± standard deviation).

d. Peroxidase activity

The assessment of the peroxidase activity, as determined by the oxidation of 3,3'-diaminobenzidine, was done to complement the qualitative data related to lignin as visualised by the Mäule staining in H15 and H20. Overall, peroxidase activity was detected in the bast fibres, with a brown staining starting in the cell corners and extending to the middle lamella (Figure 2-7 A,B,E,F). In xylem cells, the activity was stronger in developing cells adjacent to the cambial region (Figure 2-7 C,D,G,H). The xylem fibres showed a more homogenous signal, from the cambium to the pith, while the vessels displayed a stronger staining in the young developing xylem (Figure 2-7 G). Application of salicylhydroxamic acid, an inhibitor of peroxidase activity (Dunand *et al.*, 2007), resulted in a decreased peroxidase activity in the cambial zone (Figure 2-7 J). This decrease was not so obvious in xylem and bast fibres, presumably because of the presence of secondary cell walls and lignin, hindering the penetration of salicylhydroxamic acid.

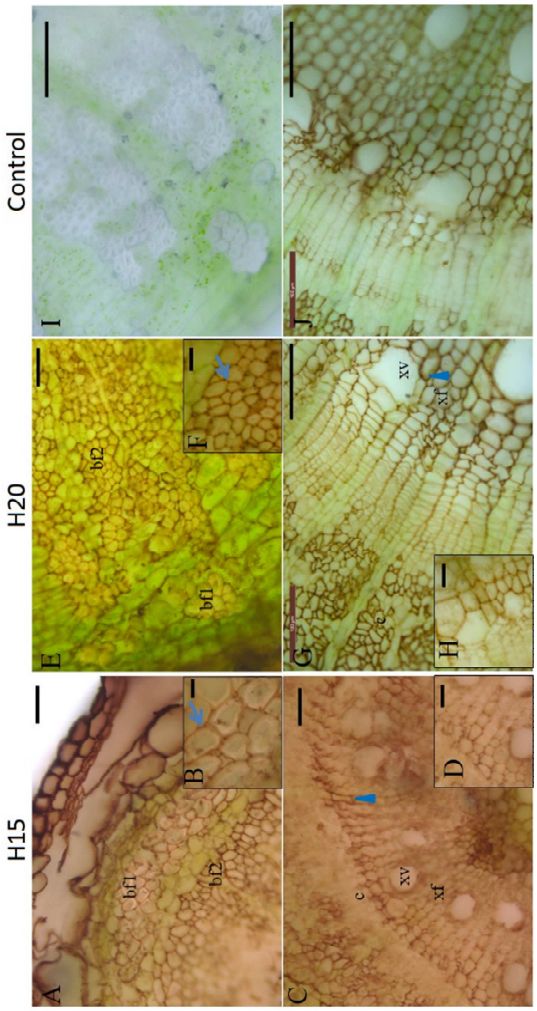


Figure 2-7: Peroxidase activity in H15 and H20. A to D, H15; E to H, H20. Details of the bast fibres and xylem regions are shown in B and F and D and H, respectively. Blue arrows indicate peroxidase activity in the middle lamella and cell corners of the bast fibres; blue arrowheads indicate peroxidase activity in the xylem vessels and fibres. I: negative control without DAB in H20. J: negative control with salicylhydroxamic acid as inhibitor of peroxidase activity in H20. bf1 primary bast fibre, bf2 secondary bast fibre, c cambial zone, xf xylem fibre, xv xylem vessel. Scale bar: 100µm (G, I, J), 50µm (A, C, E); 25µm (D, F, H); 10µm (B).

e. Laccase activity

A second activity stain, consisting in the oxidation of 3,3'-diaminobenzidine in the optimal pH range of laccase activity, was applied in hemp hypocotyls (Figure 2-8). The signal was higher in the secondary xylem than in the bast fibres, both in H15 and H20 (Figure 2-8 C-G). We did not detect any staining in the primary xylem at any time point (Figure 2-8 A-B-F). In H15, the xylem vessels displayed a stronger staining with respect to the xylem fibres (Figure 2-8 C). In H20, the secondary bast fibres displayed a stronger laccase activity than primary fibres (Figure 2-8 G). Sodium azide was used as an inhibitor of laccase activity (Johannes and Majcherczyk, 2000), and resulted in a decreased orange coloration (Figure 2-8 J).

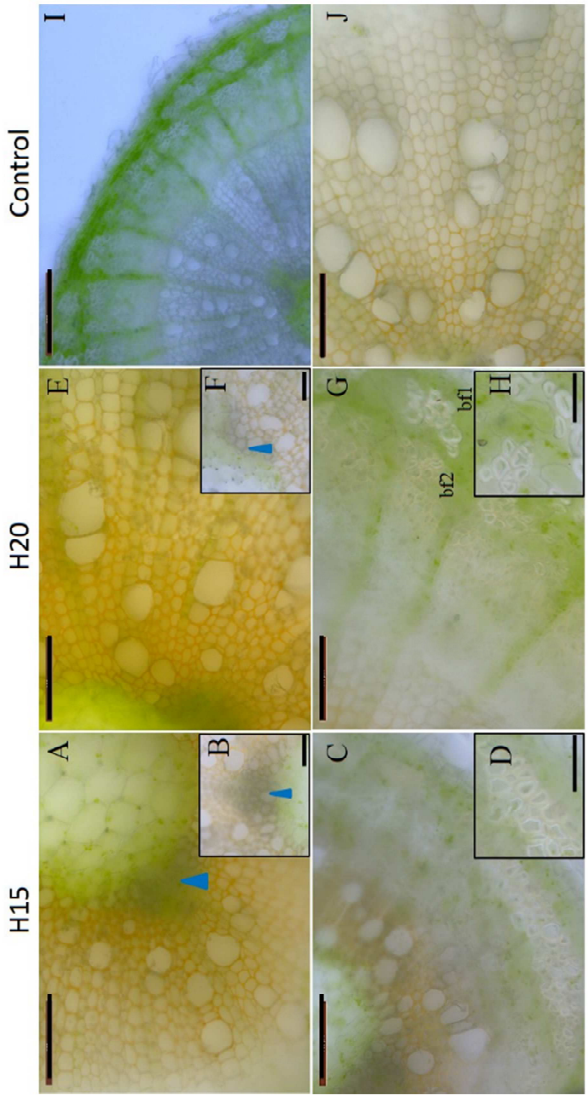


Figure 2-8: Laccase activity in H15 and H20. A to D, H15; E to H, H20. Details of the xylem regions and bast fibres are shown in B and F and D and H, respectively. Orange colour indicates the presence of laccase activity. Blue arrowheads indicate the absence of laccase activity in the lignified primary xylem. I: negative control without DAB in H20. J: negative control with sodium azide as inhibitor of laccase activity in H20. bf1 primary bast fibre, bf2 secondary bast fibre. Scale bar: 100µm (50 µm in the insets).

III. Discussion

a. Lignans are related to various developmental stages

In this study, we have focused on two families of lignan-related genes, namely the *DIR* and *PLR*/*PRR*. In *Arabidopsis*, 25 DIR have been found. Since we have retrieved only 9 DIR genes in hemp, several other DIR-genes may be missing in our analysis. Two main clades of hemp DIR are clearly distinct from the phylogenetic analysis (Figure 2-3). CsaDIR6A and CsaDIR6B belong to the same clade as the *Arabidopsis*, *Schisandra chinensis*, *Forsythia* x *intermedia* and flax dirigent proteins (Clade I). In those species, these DIRs are involved in the stereoselective initial coupling step of coniferyl alcohol to yield either (+)-pinoresinol or (-)-pinoresinol, a G(8-8)G lignan, in the presence of laccases (Davin and Lewis, 2000). We found that these hemp DIRs were more expressed at later stages of the hypocotyl development (H15 and H20) compared to H6 and H9 (Figure 2-4). In *Arabidopsis*, PRR reduces (+)-pinoresinol to lariciresinol (Nakatsubo *et al.*, 2008). However, the patterns of expression of hemp *DIR6A*, *DIR6B*, *PLR* and *PRR1* did not overlap, since *PLR* and *PRR1* were more expressed in H6 and H9. It is therefore plausible that, in hemp, DIR6A and DIR6B may be involved in other biochemical reactions, such as those occurring in the ellagitannin or sesquiterpenoid pathways (Gasper *et al.*, 2016). Various terpenes, including sesquiterpenes, have been detected in both medicinal and fibre varieties of *C. sativa* (Andre *et al.*, 2016). Several squalene epoxidases, key-enzymes of terpenoids biosyntheses, are expressed in the thale cress hypocotyl (Rasbery *et al.*, 2007). Therefore, the higher abundance of hemp *DIR6A* and *DIR6B* might be related to the

biosynthesis of terpenes. Further analyses are required to determine the DIR involved in the stereoconformation of the lignans in hemp. Candidates may be found in the DIR-like proteins subfamilies, which regroup proteins whose functions are not yet clearly established (Ralph *et al.*, 2006). The expression pattern of *CsaDLP4* is compatible with a role in lignan biosynthesis, as it is coexpressed with *PLR* and *PRR1*. The comparison with the co-expression data of genes related to *AtPRR1* may support this hypothesis. Indeed, the closest ortholog of CsaDLP4 (AtDIR18) according to the phylogenetic analysis (Figure 2-3) is co-expressed with *AtPRR1* (*p*-value < 1E-03) (atted.jp). AtPRR1 is solely responsible for lignan biosynthesis in the stem (Zhao *et al.*, 2015). However, in contrast with the expression of hemp *PRR1*, *AtPRR1* is co-expressed with many genes involved in secondary cell wall deposition such as *MYB46*, *SND1*, *CESA7* or *LAC17*. As described in Chapter 1, the genes involved in secondary cell wall biosynthesis were more expressed either in H15 or in H20. It remains to be investigated whether the expression of *PLR* / *PRR* in our hemp system is transient and linked to the young developmental stage, or if it also applies to older phases.

Lignans positively (e.g., syringaresinol) or negatively (e.g., sesamin) regulate root and shoot lengths (Cutillo *et al.*, 2003; Yamauchi *et al.*, 2015). A disruption in (+)-secoisolariciresinol diglucoside (SDG) biosynthesis by RNAi of flax *PLR* resulted in higher concentrations of dehydrodiconiferyl alcohol glucoside (DCG) and dihydro-dehydrodiconiferyl alcohol glucoside (DDCG) without compromising growth (Renouard *et al.*, 2014). DCG promotes cell division, possibly by transducing the cytokinin signal (Binns *et al.*, 1987). This may indicate that the plant compensates the lack of SDG induced by the down-regulation of *PLR* by the biosynthesis of DCG. A

detailed analysis of lignans present at the different time points of the hemp hypocotyl system may shed additional light on the exact functions of this class of molecules.

PLR expression and thus the synthesis and accumulation of lignans may also be involved in the response to oxidative stress. In *F. intermedia*, *PLR* mRNA is more abundant in young stems as compared to more mature stems and is localised in the vascular cambium and developing xylem (Kwon *et al.*, 2001). During the early stages of xylem parenchyma cell development, the activity of PLR results in the synthesis of lignans, while lignification occurs in tracheary elements and fibres (Kwon *et al.*, 2001). Xylem parenchyma cells, tracheary elements and fibres which did not yet complete their programmed cell death have to cope with oxidative intracellular conditions because of the production of H_2O_2 by xylem parenchyma cells (Ros Barceló, 2005). H_2O_2 is able to diffuse from cell to cell, and may be used as a substrate for the activity of peroxidases involved in lignification. Niculaes and colleagues (Niculaes *et al.*, 2014) propose that class I peroxidases catalyse the oxidation of monolignols and their subsequent dimerization into dilignols such as pinoresinol to prevent damages to the cells that have not completed their programmed cell death. The suggestion of a PLR protective function against oxidative stress in the developing xylem cells may explain its high expression in young hemp hypocotyls, besides its potential implication in the regulation of plant growth. Later in the development, the abundance of PCBER was higher (H15-H20, Figure 2-6). *In vitro*, PCBER reduces dehydrodiconiferyl alcohol (DDC) to isodihydrodehydrodiconiferyl alcohol (IDDDC), a G(8-5)G neolignan. In living cells, such reduced products are oxidised by a peroxidase consuming H_2O_2, providing a protection against oxidative stress during lignification (Niculaes *et al.*,

2014). The authors suggest that PCBER and pinoresinol reductases prevent oxidative damage by producing radical scavenging molecules such as reduced phenylpropanoid coupling products. At the time points where PCBER was more abundant, one can speculate that neolignans are excreted to the secondary cell walls to cope with the oxidative stress accompanying lignin polymerization (Huis *et al.*, 2012; Zhao *et al.*, 2015). Alternatively, PCBER may also be involved in xylem lignification by reducing the arylglycerol of S(8–5)G glycoside, whose product is finally used for scavenging H_2O_2 in the oxidative conditions found in lignifying tissues (Niculaes *et al.*, 2014).

b. Regulation of lignin content and composition

At the transcriptomic level, the ortholog of *Arabidopsis NST1*, a transcription factor involved in secondary cell wall deposition and lignification (Zhao and Dixon, 2011) was significantly more expressed in H15 and H20 as compared to H6 and H9 (Figure 2-4). Consequently, most of the genes of the monolignol and lignin biosynthetic pathways were more expressed in older hypocotyls. The increased expression of *NST1* and genes of the monolignol-lignin biosynthesis is in accordance with the data retrieved in our previous RNA-Seq experiment (Chapter 1). During the proteome study, the abundance of enzymes involved in the generation of methyl donors was identified as significantly changing: METS, SAMS, SAHH and MTHFR were differentially abundant within the time course experiment (Table 2-2). The two proteins identified as the orthologs of AtMETS2 displayed different trends: the first was highly abundant at all the time points, with small but significant fold-change, while the second was following the same trend as the proteins of the monolignol pathway, with massive fold-change variations. This may point to different functions of these two isoforms. The same observation applies to the abundance of three

isoforms of SAM synthases, the orthologs of AtSAMS2, AtSAMS3 and AtSAMS4, with SAMS2 and SAMS3 abundances remaining almost constant and SAMS4 being more abundant as the hypocotyl ages. Such changes in the abundance / expression profiles of the proteins / genes involved in the biosynthesis of methyl donors have also been documented in the flax hypocotyl, where genes of the *S*-adenosylmethionine pathway were more expressed at 15 days as compared to 6 days (Roach and Deyholos, 2008). The methyl donors are involved in many biochemical reactions, including the methylation of G and S monolignol (Shen *et al.*, 2002; Tang *et al.*, 2014).

Table 2-2: Identification and NSAF fold-change of the proteins involved in methyl donors metabolism. Abbreviations are as in the text. The ratio of the average NSAF for each time point is calculated. (-) means that the second term of the comparison is more abundant (*e.g.*, H6 in the comparison H9/H6). Statistical significance with Student T-test at $p < 0.05$, 0.01 and 0.001 is indicated with *, ** and ***, respectively.

Protein	TAIR ID	H9/H6	H15/H6	H20/H6	H20/H15
MTHFR2	AT2G44160	-1.07	10.25*	6.62	-1.55
THFS	AT1G50480	1.31	-11.34*	-12.55*	-1.1
METS2	AT5G17920	30.72	119.93***	108.45***	-1.1
METS2	AT5G17920	1.03	1.85***	1.83**	-1.01
SAHH2	AT3G23810	-1.07	1.76**	1.45*	-1.21
SAMS2	AT4G01850	-1.06	1.23**	1.47***	1.18*
SAMS3	AT2G36880	-1.17	1.15	1.19*	1.04
SAMS4	AT3G17390	8.97	42.03***	42.36**	1

The lignin content and composition depends on the activity of ca. ten enzymes, from PAL to COMT (Wang *et al.*, 2013) (Figure 2-1). Perturbations in the enzymatic activities of those enzymes lead to changes in lignin content, lignin composition, or both (Nakashima *et al.*, 2008). Accordingly, the higher abundances of PAL1, PAL2, 4CL2,

CAD4, HCT, CCoAOMT and COMT in H15 and H20 correlate with the rise in lignin content (Figure 2-6). The abundance pattern of the proteins involved in lignification is consistent with the transcriptomics data obtained with the RNA-Seq analysis of the same time points (Chapter 1). The lignin composition also depends on the cell type (e.g., vessels *versus* fibres) and age (deposition and maturation of the secondary cell wall). H lignin accumulates in the middle lamella of vascular cells (Davin and Lewis, 2000), prior to G and S lignin deposition in the S1 sublayer. H lignin is believed to be a factor determining the shape of the vascular cells (Nakashima *et al.*, 2008). Lignin rich in H units is more condensed because this subunit is capable of forming condensed units at the 3 and 5 positions (Terashima and Fukushima, 1988), resulting in branched polymers. The pectic nature of the middle lamella may explain why lignin is present in a more condensed substructure with a high H unit content (Terashima and Fukushima, 1989; Hao *et al.*, 2014). Indeed, the loose structure of the pectic matrix sterically favours the formation of the condensed lignin and accommodates its bulky and branched organisation (Ruel *et al.*, 2001). Noncellulosic polysaccharides such as galactan and xylan have roles in controlling the cellulose microfibril orientation, resulting in qualitative and quantitative changes in the lignification of the middle lamella and cell walls (Donaldson, 1994; Hao *et al.*, 2014). In alfalfa, the middle lamella lignification occurs in specific spots where pectin, peroxidase activity and H2O2 are present (Wi *et*

al., 2005). The difference in terms of structure between condensed and non-condensed lignin is depicted in Figure 2-9.

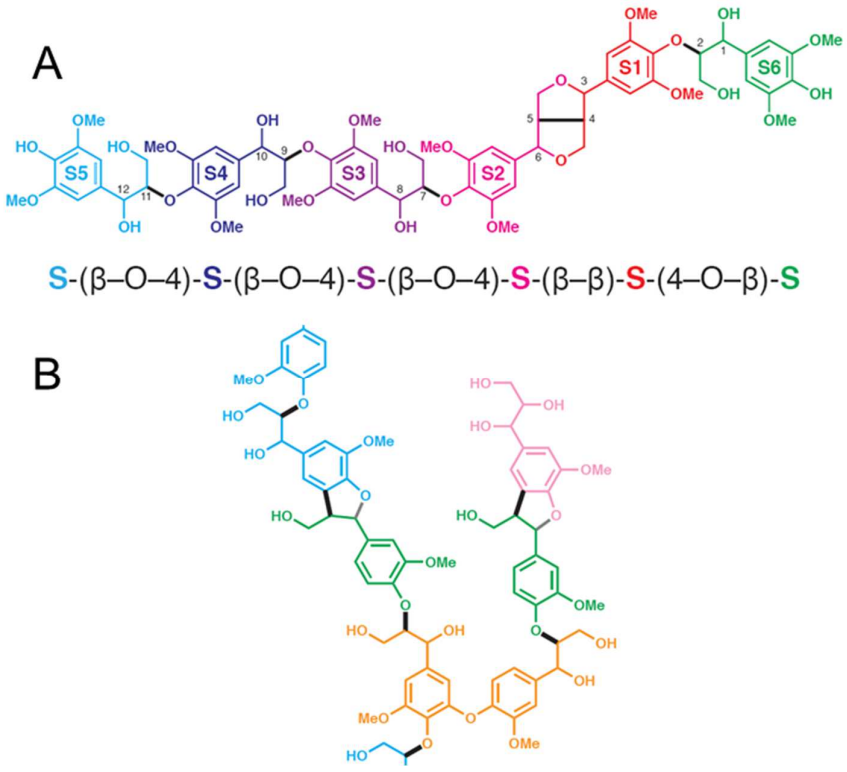


Figure 2-9: Typical structure of non-condensed (A) and condensed, branched (B) lignin. In A, the syringyl units are linked via β-O-4 (β-ether) and β-β (resinol) bonds, conferring the linear structure. In B, the two guaiacyl units in orange are linked via 5-O-4 (biphenyl ether) bond, responsible for the so-called branched structure of bulk lignin. Adapted from Ralph *et al.*, 2008.

During the hypocotyl elongation, the expression of several genes involved in lignification was shown to be under control of the circadian clock (Rogers and Campbell, 2004). Among them, *C4H*,

COMT and *CCoAOMT* show circadian-dependent expression in thale cress. More specifically, their transcripts are more abundant 4-8h before dawn, when cell elongation is slowed down or has stopped. These results may be linked to the availability of metabolizable sugars for lignification (Rogers *et al.*, 2005). In support of this hypothesis, these authors have shown that *sex1*, a mutant impaired in starch turnover resulting in a reduced pool of available carbon, accumulates less lignin than the wild type. In addition, the lignin of the *sex1* mutant was completely depleted of H unit because of the higher C3'H activity relative to other enzymes of the phenylpropanoid/monolignol pathway (Rogers *et al.*, 2005). Moreover, sucrose supplied to dark-grown hypocotyl induces lignification (Rogers *et al.*, 2005). Sucrose may be a signalling molecule to induce the activity of the lignin biosynthetic pathway. As a result, sucrose may be considered both as a source of carbon-rich skeletons for lignification, and as a signalling molecule regulating a suite of developmental processes, including the differentiation of xylem cells (Rogers *et al.*, 2005). One may speculate that the deposition of H lignin in the elongating hypocotyl may be (partially) tuned by sugar availability. Several enzymes involved in sucrose, glucose and fructose metabolism (e.g. SUS4, INV, FRK3; Figure 2-1) were more abundant in H15 and H20. Increased sucrose synthase activity leads to higher cellulose content, as it provides UDP-glucose, the precursor used by the cellulose synthase complex (Coleman *et al.*, 2009). Invertase hydrolyses sucrose into glucose and fructose. Glc-6-P may be converted in erythrose-4-P to be shunted to the shikimate pathway, producing phenylalanine, a precursor for the biosynthesis of lignols (Figure 2-1) (Novaes *et al.*, 2010). Specifically in H20, most of the proteins associated with downstream phenylalanine metabolism, especially monolignol biosynthesis (PAL1, PAL2, 4CL2, CAD4, HCT, COMT and CCoAOMT) and lignin polymerisation

(orthologs of AtPRX3, AtPRX52 and AtPRX54), reached their maximum abundance. Fructose is likely phosphorylated by FRK3 to avoid feedback inhibition of SUS4 and invertase, therefore contributing to cellulose biosynthesis. Fructose may also be converted to UDP-glucose and finally sucrose (Roach *et al.*, 2012).The enzymes involved in lignin biosynthesis and sucrose metabolism are listed in Table 2-3.

Table 2-3: Identification and NSAF fold-change of the proteins involved in sucrose metabolism and lignin biosynthesis. Abbreviations are as in the text. The ratio of the average NSAF for each time point is calculated. (-) means that the second term of the comparison is more abundant (*e.g.*, H6 in the comparison H9/H6). Statistical significance with Student T-test at $p < 0.05$, 0.01 and 0.001 is indicated with *, ** and ***, respectively.

Protein	TAIR ID	H9/H6	H15/H6	H20/H6	H20/H15
SUS4	AT3G43190	1.86*	29.38***	23.22***	-1.27
FRK3	AT1G06020	1.27*	1.73***	2.00***	1.16
INV	AT3G47400	2.55	5.32**	3.66**	-1.45*
PAL1	AT2G37040	-1.07	22.32***	26.03***	1.16
PAL1	AT2G37040	-1.07	3.25	17.36***	5.32**
PAL2	AT3G53260	-1.07	20.58***	31.84***	1.54**
4CL2	AT3G21240	4.84	6.35	15.26***	2.40*
HCT	AT5G48930	9.96**	11.93**	25.46***	2.13
CAD4	AT3G19450	2.78	5.13	10.25***	1.99
CCoAOMT	AT4G34050	8.39*	10.88*	15.04*	1.38
COMT	AT5G54160	1	2.07***	3.19***	1.54***
PCBER	AT1G75280	19.86***	31.21***	37.13***	1.18

As the hypocotyl ages, increased CCoAOMT and COMT abundances (Table 2-3) result in lignin richer in G and S subunits (Table 2-1). We therefore show a consistent link between proteomics data and lignin monomeric composition. The lignification of the secondary cell wall begins with deposition of G units in the S1 sub-

layer (Davin and Lewis, 2000) in discrete domains where LAC4 and LAC17 are present (Schuetz *et al.*, 2014). Despite the lack of a precise localisation of these two enzymes in the hemp hypocotyl system, we may however assume that LAC4 and LAC17 activities lead to a decrease in the relative proportion of H lignin (Table 2-1). This was previously observed in hemp bast fibres: the H lignin proportion of apical fibres was systematically higher than in basal fibres, irrespective of the stage of development (Crônier *et al.*, 2005). Using thioacidolysis, the same authors found out that the S/G molar ratio was also higher in the basal fibres. Primary xylem is almost completely devoid of S lignin, as shown by the brown coloration of the Mäule staining (Figure 2-2) and as previously described (Venverloo, 1971). Lignin in the secondary xylem is progressively enriched in S units (Venverloo, 1971). This rise in the S/G ratio was also observed during lignification of mature secondary cell walls in woody Angiosperms, mainly due to S lignin polymerisation (Lourenço *et al.*, 2016). The development of fibres in both xylem and phloem likewise contributes to this increase: fibres, as compared to tracheary elements, are richer in S lignin to provide mechanical strength. In *Quercus suber*, the xylem lignin is enriched in S units because of the large proportion of fibres (S/G of 1.2, (Lourenço *et al.*, 2016)). The monomeric composition of lignin also depends on the carbohydrate composition of the cell wall where the polymerization occurs (Terashima and Fukushima, 1989). In secondary cell walls, elongated patches of lignin are deposited between the cellulose microfibrils, acting as a template to guide the lignification (Donaldson, 1994; Ruel *et al.*, 2001). Because secondary cell walls are richer in hemicellulose, e.g. xylan in the xylem cells of hemp (Blake *et al.*, 2008; Chapter 1), than the middle lamella, lignification mainly occurs with polymerization of G and S subunits (Terashima and Fukushima, 1989). By controlling the orientation of the cellulose microfibrils in the

secondary cell wall, xylan favours the formation of the microfibril matrix (Ruel *et al.*, 2001). This, in turn, favours the formation of non-condensed lignin (mainly composed of G and S subunits with ether linkage, such as *β-O*-4, as in Figure 2-9A) and its extended conformation adapted to the tight volume available between the cellulose microfibrils (Ruel *et al.*, 2001). For example, the *irx8* mutant, which is disturbed in the xylan architecture of the secondary cell wall, has less lignin because of the lower amount in G subunits (Hao *et al.*, 2014).

In old cell walls, lignin and xylan are covalently linked to form the lignin-carbohydrate complex (LCC) by the addition of nucleophilic groups (hydroxyl or carboxylic groups of hemicelluloses) on the transient quinone methide intermediate, determining the final step of cell wall construction (Barakat *et al.*, 2007). In hemp, primary and secondary phloems are rich in extraxylary fibres, which are already lignified in the hypocotyl aged 20 days (Figure 2-2). However, the rise in the S/G ratio in the hypocotyl time course system differs from the trend observed in the outer tissues of adult plants, where the S/G ratio in the same stem fragment (apical or basal) does not change significantly when it gets older (Crônier *et al.*, 2005). The authors suggest that the high amount of crystalline cellulose of the bast fibres may impair the polymerisation process. Since in our hypocotyl system the vascular tissue of the xylem is more developed than the sclerenchyma bast fibre tissue, we can observe an increase in the S/G ratio throughout time. Laser capture microdissection was found to be a reliable method to study the lignin composition of specific cell types in herbaceous species, where the manual separation of specific tissues is cumbersome (Nakashima *et al.*, 2008) and may provide important data complementing the Mäule staining.

The synthesized monolignols are excreted to the cell wall where they are polymerised into lignin under the activities of laccases and class III peroxidases (Hao *et al.*, 2014), with the possible intervention of DIR. The involvement of peroxidases in the lignification process is obvious when considering the abundances of both transcripts (orthologs of *AtPRX49*, *AtPRX52* and *AtPRX72*) and proteins (orthologs of AtPRX3, AtPRX52 and AtPRX54), which are in general higher at older stages of development (Table 2-4). The three transcripts were more abundant in H15, while the three proteins were more abundant in H20.

Table 2-4: CNRQ (transcripts) and NSAF (proteins) fold-change of the genes involved in lignin polymerisation. Abbreviations are as in the text. The ratio of the average CNRQ or NSAF for each time point is calculated. (-) means that the second term of the comparison is more abundant (*e.g.*, H6 in the comparison H9/H6). Statistical significance with Student T-test at $p < 0.05$, 0.01 and 0.001 is indicated with *, ** and ***, respectively.

Transcript	TAIR ID	H9/H6	H15/H6	H20/H6	H20/H15
AtPRX49	AT4G36430	1.44	6.14*	2.06*	-2.98
AtPRX52	AT5G05340	2.01	2.53	1.04	-2.43
AtPRX72	AT5G66390	7.06	10.08*	2.90**	-3.47
Protein	**TAIR ID**	**H9/H6**	**H15/H6**	**H20/H6**	**H20/H15**
AtPRX3	AT1G05260	1.21*	1.36**	1.38*	1.01
AtPRX52	AT5G05340	-1.07	6.62	22.04***	3.32*
AtPRX54	AT5G06730	-1.07	10.40**	27.75***	2.66*

Peroxidase activity is essential for fibre lignification in several species including *Arabidopsis* (Tokunaga *et al.*, 2009; Fernández-Pérez *et al.*, 2014) and flax (Chantreau *et al.*, 2014) and is required for S lignin polymerisation, since laccases seem not to be able to catalyse this polymerisation (Novo-Uzal *et al.*, 2013). Likewise, S lignin staining and peroxidase activity were overlapping in the bast fibres (Figures 2-2 and 2-7). Moreover, the increase in the peroxidase activity observed during tracheary element differentiation (close to the cambial region) and lignification (Figure 2-7D,H) was previously described in *Zinnia* (Novo-Uzal *et al.*, 2013). During the hemp hypocotyl development, secondary xylem cells are visible 12 days after sowing (Chapter 1) and lignification is ongoing at least until day 20 (Table 2-1). The role of the two laccases LAC11 and LAC17 is more difficult to define, as there is no significant change in their expression when determined by RT-qPCR (Figure 2-4). However, we have previously described the expression pattern of putative orthologs of *A. thaliana LAC17* using RNA-Seq in the same hypocotyl system, showing that one isoform of this gene was strongly upregulated in H20 (Chapter 1). As several isoforms of *LAC17* were found using RNA-Seq, it is plausible that the transcripts detected by the two methods are different isoforms. One may also speculate that there is a basal level of laccase activity until the hypocotyl reaches its final diameter, i.e. when all fibres and vessels have differentiated from the cambium and started their initial lignification with G units. Laccases are involved in secondary cell wall lignification of protoxylem tracheary elements in young elongating tissues (Schuetz *et al.*, 2014). In H15 and H20, no laccase activity signal was observed in the primary xylem (Figure 2-8), but a strong peroxidase signal was instead visible (Figure 2-7C). This may suggest that a peroxidase-driven lignification occurs after the polymerisation of monolignols by laccases. Berthet and colleagues (Berthet *et al.*, 2011)

suggested that laccases are expressed at the beginning of the lignification and involved in the polymerisation of G-rich lignin in fibres. In flax, three orthologs of *AtLAC4* and *AtLAC17* were more abundant in the upper region of inner stem tissues, containing the xylem, while five peroxidases were more expressed in the lower region of inner stem tissues (Huis *et al.*, 2012). The proposed mechanism of lignification suggested by these two studies is compatible with the increase in the S/G ratio along the time course.

The role of the DIR in lignification is limited. Since lignin polymerisation is performed through end-wise radical coupling of phenols to the free-phenolic end of the growing polymer and not by enzymatic control (Ralph *et al.*, 2008), a direct role of DIR in this process is unlikely. However, they may play a role in lignin localisation in specific regions of the cell wall. By silencing *AtDIR10* (*ESB1*), perturbations in the organisation of Casparian strips have been observed (Hosmani *et al.*, 2013). The authors suggest that ESB1 plays a role in the localisation of lignin. In the phylogenetic analysis (Figure 2-3), CsaDLP4 and CsaDLP5 fall into the same cluster as ESB1, possibly suggesting a role in lignification. The expression pattern of *DLP5* is in line with such a function, as it is more expressed in old hypocotyls and differently from *DLP4*. The hemp genes *DLP2*, *DLP20A* and *DLP20B* were more expressed in H15 and H20 (Figure 2-4). These DLPs belong to another subgroup, from which no proteins have been functionally characterised yet. The need of such studies is obvious to understand the precise role of the DIR-like proteins.

The molecular aspects of lignin and lignan biosyntheses in the hemp hypocotyl system were studied. Lignin content and composition were in line with proteomics, RT-qPCR and microscopic observations of laccase and peroxidase activities. These results foster our

understanding of lignification during primary and secondary growth and open venues of functional studies of the mechanisms underlying primary and secondary cell wall lignification.

References

André, C.M., Hausman, J.F., and Guerriero, G. (2016) *Cannabis sativa*: The Plant of the Thousand and One Molecules. *Front. Plant Sci.* 7: 19.

Barakat, A., Winter, H., Rondeau-Mouro, C., Saake, B., Chabbert, B., and Cathala, B. (2007) Studies of xylan interactions and cross-linking to synthetic lignins formed by bulk and end-wise polymerization: a model study of lignin carbohydrate complex formation. *Planta* 226: 267.

Bassard, J.E., Richert, L., Geerink, J., Renault, H., Duval, F., Ullmann, P., *et al.* (2012) Protein-Protein and Protein-Membrane Associations in the Lignin Pathway. *Plant Cell* 24: 4465-4482.

Behr, M., Legay, S., Hausman, J.-F. and Guerriero, G. (2015) Analysis of Cell Wall-Related Genes in Organs of *Medicago sativa* L. under Different Abiotic Stresses*. Int. J. Mol. Sci.* 16: 16104.

Berthet, S., Demont-Caulet, N., Pollet, B., Bidzinski, P., Cézard, L., Le Bris, P., et al (2011) Disruption of LACCASE4 and 17 Results in Tissue-Specific Alterations to Lignification of *Arabidopsis thaliana* Stems. *Plant Cell* 23: 1124-1137.

Billa, E., Tollier, M.T., and Monties, B. (1996) Characterisation of the Monomeric Composition of in situ Wheat Straw Lignins by Alkaline Nitrobenzene Oxidation: Effect of Temperature and Reaction Time. *J. Sci. Food Agric.* 72: 250-256.

Binns, A.N., Chen, R.H., Wood, H.N., and Lynn, D.G. (1987) Cell division promoting activity of naturally occurring dehydrodiconiferyl glucosides: do cell wall components control cell division? *Proc. Natl Acad. Sci. U S A* 84: 980-984.

Blake, A.W., Marcus, S.E., Copeland, J.E., Blackburn, R.S. and Knox, J.P. (2008) In situ analysis of cell wall polymers associated with phloem fibre cells in stems of hemp, *Cannabis sativa* L. *Planta* 228: 1-13.

Bonawitz, N.D., and Chapple, C. (2010) The Genetics of Lignin Biosynthesis: Connecting Genotype to Phenotype. *Annu. Rev. Genet.* 44: 337-363.

Bonawitz, N.D., and Chapple, C. (2013) Can genetic engineering of lignin deposition be accomplished without an unacceptable yield penalty? *Curr. Opin. Biotech.* 24: 336-343.

Caraux, G., and Pinloche, S. (2005) PermutMatrix: a graphical environment to arrange gene expression profiles in optimal linear order. *Bioinformatics* 21: 1280-1281.

Chantreau, M., Portelette, A., Dauwe, R., Kiyoto, S., Crônier, D., Morreel, K., *et al.* (2014) Ectopic Lignification in the Flax lignified bast fibre1 Mutant Stem Is Associated with Tissue-Specific Modifications in Gene Expression and Cell Wall Composition. *Plant Cell* 26: 4462-4482.

Coleman, H.D., Yan, J. and Mansfield, S. D. (2009) Sucrose synthase affects carbon partitioning to increase cellulose production and altered cell wall ultrastructure. *Proc. Natl Acad. Sci. U S A.* 106: 13118-13123.

Crônier, D., Monties, B., and Chabbert, B. (2005) Structure and Chemical Composition of Bast Fibres Isolated from Developing Hemp Stem. *J. Agric. Food Chem.* 53: 8279-8289.

Cutillo, F., D'Abrosca, B., DellaGreca, M., Fiorentino, A., and Zarrelli, A. (2003) Lignans and Neolignans from *Brassica fruticulosa*: Effects on Seed Germination and Plant Growth. *J. Agric. Food Chem.* 51: 6165-6172.

Dalisay, D.S., Kim, K.W., Lee, C., Yang, H., Rübel, O., Bowen, B.P., *et al.* (2015) Dirigent Protein-Mediated Lignan and Cyanogenic Glucoside Formation in Flax Seed: Integrated Omics and MALDI Mass Spectrometry Imaging. *J. Nat. Prod.* 78:1231-1242.

Daub, C.O., Kloska, S., and Selbig, J. (2003) MetaGeneAlyse: analysis of integrated transcriptional and metabolite data. *Bioinformatics* 19: 2332-2333.

Davin, L.B., Wang, H.B., Crowell, A.L., Bedgar, D.L., Martin, D.M., Sarkanen, S., *et al.* (1997) Stereoselective Bimolecular Phenoxy Radical Coupling by an Auxiliary (Dirigent) Protein Without an Active Center. *Science* 275: 362-367.

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

Dereeper A., Guignon V., Blanc G., Audic S., Buffet S., Chevenet F., *et al.* (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res*. 36:465-469.

Dunand, C., Crèvecoeur, M., Penel, C. (2007) Distribution of superoxide and hydrogen peroxide in *Arabidopsis* root and their influence on root development: possible interaction with peroxidases. *New Phytol*. 174(2):332-341.

Dixon, R.A., Achnine, L., Kota, P., Liu, C.J., Reddy, M.S.S., and Wang, L. (2002) The phenylpropanoid pathway and plant defence - a genomics perspective. *Mol. Plant Pathol.* 3: 371-390.

Donaldson, L.A. (1994) Mechanical constraints on lignin deposition during lignification. *Wood Sci. Technol.* 28: 111-118.

Fanutza, T., Del Prete, D., Ford, M.J., Castillo, P.E., and D'Adamio, L. (2015) APP and APLP2 interact with the synaptic release machinery and facilitate transmitter release at hippocampal synapses. *eLife* 4: e09743.

Fernández-Pérez, F., Pomar, F., Pedreno, M.A., and Novo-Uzal, E. (2014) The suppression of AtPrx52 affects fibres but not xylem lignification in *Arabidopsis* by altering the proportion of syringyl units. *Physiol. Plantarum* 154: 395-406.

Gasper, R., Effenberger, I., Kolesinski, P., Terlecka, B., Hofmann, E., and Schaller, A. (2016) Dirigent Protein Mode of Action Revealed by the Crystal Structure of AtDIR6. *Plant Physiol.* 172: 2165-2175.

George, I.S., Pascovici, D., Mirzaei, M., and Haynes, P.A. (2015) Quantitative proteomic analysis of cabernet sauvignon grape cells exposed to thermal stresses reveals alterations in sugar and phenylpropanoid metabolism. *Proteomics* 15: 3048-3060.

Hao, Z., Avci, U., Tan, L., Zhu, X., Glushka, J., Pattathil, S., *et al.* (2014). Loss of *Arabidopsis* GAUT12/IRX8 causes anther indehiscence and leads to reduced G lignin associated with altered matrix polysaccharide deposition. *Front. Plant Sci.* 5: 357.

Herrero, J., Fernández-Pérez, F., Yebra, T., Novo-Uzal, E., Pomar, F., Pedreno, M., *et al.* (2013) Bioinformatic and functional characterization of the basic peroxidase 72 from *Arabidopsis thaliana* involved in lignin biosynthesis. *Planta* 237: 1599-1612.

Hertzberg, M., Aspeborg, H., Schrader, J., Andersson, A., Erlandsson, R., Blomqvist, K., *et al.* (2001) A transcriptional roadmap to wood formation. *Proc. Natl Acad. Sci. U S A* 98: 14732-14737.

Hosmani, P.S., Kamiya, T., Danku, J., Naseer, S., Geldner, N., Guerinot, M.L., and Salt, D.E. (2013) Dirigent domain-containing protein is part of the machinery required for formation of the lignin-based Casparian strip in the root. *Proc. Natl Acad. Sci. U S A* 110: 14498-14503.

Huis, R., Morreel, K., Fliniaux, O., Lucau-Danila, A., Fénart, S., Grec, S., *et al.* (2012) Natural Hypolignification Is Associated with Extensive Oligolignol Accumulation in Flax Stems. *Plant Physiol.* 158: 1893-1915.

Ito, J., Batth, T.S., Petzold, C.J., Redding-Johanson, A.M., Mukhopadhyay, A., Verboom, R., *et al.* (2011) Analysis of the *Arabidopsis* Cytosolic Proteome Highlights Subcellular Partitioning of Central Plant Metabolism. *J. Proteome Res.* 10: 1571-1582.

Johannes, C, and Majcherczyk, A. (2000) Laccase activity tests and laccase inhibitors. *J. Biotechnol.* 78(2):193-199.

Kim, K.W., Moinuddin, S.G.A., Atwell, K.M., Costa, M.A., Davin, L.B., and Lewis, N.G. (2012) Opposite Stereoselectivities of Dirigent Proteins in *Arabidopsis* and *Schizandra* Species. *J. Biol. Chem.* 287: 33957-33972.

Kwon, M., Davin, L.B., and Lewis, N.G. (2001) In situ hybridization and immunolocalization of lignan reductases in woody tissues: implications for heartwood formation and other forms of vascular tissue preservation. *Phytochemistry* 57: 899-914.

Lourenço, A., Rencoret, J., Chemetova, C., Gominho, J., Gutiérrez, A., del Río, J.C., *et al.* (2016) Lignin Composition and Structure Differs between Xylem, Phloem and Phellem in *Quercus suber* L. *Front. Plant Sci.* 7: 1612.

Moreira-Vilar, F.C., Siqueira-Soares, R.D.C., Finger-Teixeira, A., Oliveira, D.M.D., Ferro, A.P., Da Rocha, G.J., *et al.* (2014). The Acetyl Bromide Method Is Faster, Simpler and Presents Best Recovery of Lignin in Different Herbaceous Tissues than Klason and Thioglycolic Acid Methods. *PLOS ONE* 9**,** e110000.

Nakagawa, K., Yoshinaga, A., and Takabe, K. (2014) Xylan deposition and lignification in the multi-layered cell walls of phloem fibres in *Mallotus japonicus* (Euphorbiaceae). *Tree Physiol.* 34: 1018-1029.

Nakashima, J., Chen, F., Jackson, L., Shadle, G., and Dixon, R.A. (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.

Nakatsubo, T., Mizutani, M., Suzuki, S., Hattori, T., and Umezawa, T. (2008) Characterization of *Arabidopsis thaliana* Pinoresinol Reductase, a New Type of Enzyme Involved in Lignan Biosynthesis. *J. Biol. Chem.* 283: 15550-15557.

Niculaes, C., Morreel, K., Kim, H., Lu, F., McKee, L.S., Ivens, B., *et al.* (2014) Phenylcoumaran Benzylic Ether Reductase Prevents Accumulation of Compounds Formed under Oxidative Conditions in Poplar Xylem. *Plant Cell* 26: 3775-3791.

Nikolovski, N., Rubtsov, D., Segura, M.P., Miles, G.P., Stevens, T.J., Dunkley, T.P., *et al.* (2012) Putative Glycosyltransferases and Other Plant Golgi Apparatus Proteins Are Revealed by LOPIT Proteomics. *Plant Physiol.* 160: 1037-1051.

Novaes, E., Kirst, M., Chiang, V., Winter-Sederoff, H. and Sederoff, R. (2010) Lignin and Biomass: A Negative Correlation for Wood Formation and Lignin Content in Trees. *Plant Physiol.* 154: 555-561.

Novo-Uzal, E., Fernández-Pérez, F., Herrero, J., Gutiérrez, J., Gómez-Ros, L.V., Bernal, M.Á., *et al.* (2013). From *Zinnia* to *Arabidopsis*: approaching the involvement of peroxidases in lignification. *J. Exp. Bot.* 64: 3499-3518.

Ralph, J., Brunow, G., Harris, P.J., Dixon, R.A., Schatz, P.F., and Boerjan, W. (2008) Lignification: are Lignins Biosynthesized via simple Combinatorial Chemistry or via Proteinaceous Control and Template Replication? in Recent Advances in Polyphenol Research Wiley-Blackwell, 36-66.

Ralph, S., Park, J.Y., Bohlmann, J., and Mansfield, S. (2006) Dirigent Proteins in Conifer Defense: Gene Discovery, Phylogeny, and Differential Wound- and Insect-induced Expression of a Family of DIR and DIR-like Genes in Spruce (*Picea* spp.). *Plant Mol. Biol.* 60: 21-40.

Rasbery, J. M., Shan, H., LeClair, R. J., Norman, M., Matsuda, S. P. T., and Bartel B. (2007) *Arabidopsis thaliana* Squalene Epoxidase 1 Is Essential for Root and Seed Development. *J. Biol. Chem.* 282: 17002-17013.

Renouard, S., Tribalatc, M.A., Lamblin, F., Mongelard, G., Fliniaux, O., Corbin, C., *et al.* (2014) RNAi-mediated pinoresinol lariciresinol reductase gene silencing in flax (*Linum usitatissimum* L.) seed coat: Consequences on lignans and neolignans accumulation. *J. Plant Physiol.* 171: 1372-1377.

Roach, M.J., and Deyholos, M.K. (2008) Microarray Analysis of Developing Flax Hypocotyls Identifies Novel Transcripts Correlated with Specific Stages of Phloem Fibre Differentiation. *Ann. Bot.-London* 102: 317-330.

Roach, M., Gerber, L., Sandquist, D., Gorzsás, A., Hedenström, M., Kumar, M., *et al.* (2012) Fructokinase is required for carbon partitioning to cellulose in aspen wood. *Plant J.* 70: 967-977.

Rogers, L.A., and Campbell, M.M. (2004) The genetic control of lignin deposition during plant growth and development. *New Phytol.* 164: 17-30.

Rogers, L.A., Dubos, C., Cullis, I.F., Surman, C., Poole, M., Willment, J., *et al.* (2005) Light, the circadian clock, and sugar perception in the control of lignin biosynthesis. *J. Exp. Bot.* 56: 1651-1663.

Ros Barceló, A. (2005) Xylem parenchyma cells deliver the H_2O_2 necessary for lignification in differentiating xylem vessels. *Planta* 220: 747-756.

Ruel, K., Chabannes, M., Boudet, A.M., Legrand, M., and Joseleau, J.P. (2001) Reassessment of qualitative changes in lignification of transgenic tobacco plants and their impact on cell wall assembly. *Phytochemistry* 57: 875-882.

Schuetz, M., Benske, A., Smith, R.A., Watanabe, Y., Tobimatsu, Y., Ralph, J., *et al.* (2014) Laccases Direct Lignification in the Discrete Secondary Cell Wall Domains of Protoxylem. *Plant Physiol.* 166: 798-807.

Shen, B., Li, C., and Tarczynski, M.C. (2002) High free-methionine and decreased lignin content result from a mutation in the *Arabidopsis* S-adenosyl-L-methionine synthetase 3 gene. *Plant J.* 29: 371-380.

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.

Szopinska, A., Christ, E., Planchon, S., König, H., Evers, D., and Renaut, J. (2016) Stuck at work? Quantitative proteomics of environmental wine yeast strains reveals the natural mechanism of overcoming stuck fermentation. *Proteomics* 16: 593-608.

Tang, H.M., Liu, S., Hill-Skinner, S., Wu, W., Reed, D., Yeh, C.T., *et al.* (2014) The maize brown midrib2 (bm2) gene encodes a methylenetetrahydrofolate reductase that contributes to lignin accumulation. *Plant J.* 77: 380-392.

Tarabanko, V., and Tarabanko, N. (2017). Catalytic Oxidation of Lignins into the Aromatic Aldehydes: General Process Trends and Development Prospects. *Int. J. Mol. Sci.* 18**,** 2421.

Tateda, C., Watanabe, K., Kusano, T., and Takahashi, Y. (2011) Molecular and genetic characterization of the gene family encoding the voltage-dependent anion channel in *Arabidopsis*. *J. Exp. Bot.* 62: 4773-4785.

Terashima, N., and Fukushima, K. (1988) Heterogeneity in formation of lignin-XI: An autoradiographic study of the heterogeneous formation and structure of pine lignin. *Wood Sci. Technol.* 22: 259-270.

Terashima, N., and Fukushima, K. (1989) Biogenesis and Structure of Macromolecular Lignin in the Cell Wall of Tree Xylem as Studied by Microautoradiography. in Plant Cell Wall Polymers American Chemical Society, 160-168.

Venverloo, C.J. (1971) The Lignin of *Populus nigra* L. cv. 'Italica' and Some Other Salicaceae. *Holzforschung* 18.

Wang, Y., Chantreau, M., Sibout, R., and Hawkins, S. (2013) Plant cell wall lignification and monolignol metabolism. *Front. Plant Sci.* 4: 220.

Wi, S.G., Singh, A.P., Lee, K.H., and Kim, Y.S. (2005) The Pattern of Distribution of Pectin, Peroxidase and Lignin in the Middle Lamella of Secondary Xylem Fibres in Alfalfa (*Medicago sativa*) *Ann. Bot. London* 95: 863-868.

Yamauchi, S., Ichikawa, H., Nishiwaki, H., and Shuto, Y. (2015) Evaluation of Plant Growth Regulatory Activity of Furofuran Lignan Bearing a 7,9':7',9-Diepoxy Structure Using Optically Pure (+)- and (-)-Enantiomers. *J. Agric. Food Chem.* 63: 5224-5228.

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., *et al.* (2013) LACCASE Is Necessary and Nonredundant with PEROXIDASE for Lignin Polymerization during Vascular Development in *Arabidopsis*. *Plant Cell* 25: 3976-3987.

Zhao, Q., Zeng, Y., Yin, Y., Pu, Y., Jackson, L.A., Engle, N.L., *et al.* (2015) Pinoresinol reductase 1 impacts lignin distribution during secondary cell wall biosynthesis in *Arabidopsis*. *Phytochemistry* 112: 170-178.

Chapter 2: Highlights and Perspectives

The aim of this Chapter was to characterise the lignification of the growing hemp hypocotyl. Lignin plays an important role in plant development, as it provides mechanical strength to the stem and facilitates the transport of water in the xylem by conferring hydrophobicity to vessels. The quantity and the monomeric composition of the lignin deposited in the hemp hypocotyl significantly differ between the developmental stages. The differential abundance of proteins involved in monolignol biosynthesis, as well as the changes in gene expression, explain this trend. The function of lignans and the regulation of their biosyntheses have also been discussed. We may thus draw three stages of lignan and lignin biosyntheses in the growing hemp hypocotyl.

1. During elongation (H6 and H9), a higher proportion of H-lignin was recovered. This lignin is mostly deposited in the middle lamella because the loose structure of its pectic matrix mirrors the bulky and relatively unorganised structure of the condensed lignin. At this stage, monolignol biosynthesis may be driven by the pool of available sugars, in competition with the carbon required for elongation. Lignans may act as regulators of stem elongation.
2. After elongation has ceased, the lignin content raises and is enriched in G-subunits. This constitutes the first step of the lignification of the secondary cell wall. In this compartment, the orientation of the cellulose microfibrils is controlled by xylan, resulting in the deposition of a non-condensed lignin (made of G- and S-lignin) within this tight available volume. Laccases

may be the chief enzymes regulating G-lignin polymerisation and deposition.

3. At the latest stage of the hypocotyl development (H20), lignin is enriched in S-units. This is explained by the higher abundances of enzymes involved in monolignol methylation. S-lignin deposition is favoured in the xylem and phloem fibres to provide mechanical strength. Since laccases are likely not able to polymerise S-lignin, a burst of peroxidase activity has been detected in these tissues. Lignans may participate in the regulation of the redox homeostasis during lignification.

Bringing together the results from Chapter 1 (transcriptomics and phytohormone analysis) and Chapter 2 (pattern of lignification), I demonstrated the suitability of the hemp hypocotyl system for the study of primary and secondary growths, secondary cell wall biogenesis and lignification. Untargeted *–omics* analyses have evidenced the role of several genes and proteins in these processes which are crucial for plant growth and development.

The advantage of the hypocotyl system lies in the fact that the experimenter can study different stages of development within the same developing organ. The biological processes (gene expression, protein biosynthesis, phytohormone contents) mirroring the steps of the hypocotyl development (primary and secondary growths, SCW deposition) are studied by comparing different ages. In the second part of this thesis, the results obtained on the hypocotyl system will be complemented by studying the development of the adult hemp stem. This new focus aims at studying the plant in the context of its downstream applications, especially the production of bast fibres. The hemp stem is characterized by two main types of fibres. Those of the inner tissues belong to the xylem and have xylan-type SCW

impregnated with lignin. The fibres of the phloem are markedly different from those of the xylem: they are poorly lignified, longer and have a G-layer rich in crystalline cellulose. The separation of the two stem tissues is easy in the adult hemp plants and is essential to accurately determine the factors involved in their different properties.

The histology (i.e. primary *vs* secondary tissues) and the chemical composition of the stem follow a basipetal gradient. Within this gradient of development, the snap point marks the transition between elongation and thickening of the bast fibres. By analysing cortical and core tissues located above, at or below the snap point, we will shed light on the factors involved in the development of these two tissues.

In the next Chapter, the major differences between the lignified core tissue and the fibre-bearing cortical tissue on one hand, and the position relatively to the snap point on the other, are discussed using a transcriptomic angle. A special emphasis will be given to the genes involved in lignin biosynthesis, cellulose deposition and stem biomechanics.

Chapter 3

Gene expression analysis in hemp identifies transcripts important for bast fibre development

Partially adapted from the published article: Gea Guerriero, Lauralie Mangeot-Peter, Sylvain Legay, Marc Behr, Stanley Lutts, Khawar Sohail Siddiqui and Jean-Francois Hausman. Identification of fasciclin-like arabinogalactan proteins in textile hemp (*Cannabis sativa* L.): *in silico* analyses and gene expression patterns in different tissues. 2017. *BMC Genomics* 18:741.

Author contributions:

Gea Guerriero conceived and designed the experiments and performed promoter analysis, bioinformatics and gene expression analysis. Lauralie Mangeot-Peter performed the gene expression, cloning and sequencing. Marc Behr performed the gene expression analysis in hypocotyls and contributed to the RT-qPCR in adult stem and bioinformatics analyses. Khawar Sohail Siddiqui performed protein modelling. All the authors analysed the data and wrote the article.

Abstract

Cortical and core tissues of the hemp stem have contrasting composition. While the xylem tissue found in the core is highly lignified, the phloem fibres have a gelatinous-type secondary cell wall poor in lignin. A gene expression analysis performed on stem tissues (isolated bast fibres and shivs sampled at five heights) and hypocotyls (6-9-12-15-17-20 days-old) of hemp, indicates that most of the genes involved in cellulose deposition, monolignol biosynthesis and lignin polymerisation are more expressed in the basal core tissue. Moreover, some *FLA*s are more expressed during early phases of fibre growth (elongation), while others are more expressed at the middle and base of the stem and thus potentially involved in secondary cell wall formation (fibre thickening). We mention that the *N*-linked glycosyl moieties of FLAs may trigger a signal for the formation and thickening of the secondary cell wall. The bioinformatic analysis of the promoter regions shows that the *FLA*s upregulated in the younger regions of the stem share a conserved motif related to flowering control and regulation of photoperiod perception. The promoters of the *FLA* genes expressed at higher levels in the older stem regions, instead, share a motif putatively recognized by MYB3, a transcriptional repressor of phenylpropanoid biosynthesis-related genes belonging to the MYB family subgroup S4.

Introduction

Textile hemp is an economically important bast fibre-producing crop, with several applications in industry, namely the biocomposite, textile and construction sector (Fike, 2016). This plant is not only important as a multi-purpose crop, but also useful for fundamental studies centered on cell wall biosynthesis and remodelling (Chapter 1), because its stem tissues show strong differences in cell wall composition (Crônier *et al.*, 2005). The core of hemp stems is indeed woody, while the cortex harbors the bast fibres, with a high content in crystalline cellulose and poor in lignin (Guerriero *et al.*, 2013). The different stem heights correspond to distinctive stages of bast fibre development (from intrusive growth to thickening; Figure 3-1). It is hence possible to study the mechanisms involved in the development of gelatinous- and xylan-type fibres by separating the stem tissues of the same plant. Using this method, van den Broeck and colleagues (2008) have observed that the genes involved in lignin biosynthesis were more expressed in core tissue. The bast fibres can also be separated from the surrounding parenchymatic cells from the cortex with the use of 80% ethanol, a mortar and a pestle (Gorshkov *et al.*, 2017).

Arabinogalactan proteins (AGPs) are cell surface glycoproteins belonging to the hydroxyproline-rich glycoprotein superfamily (Showalter and Basu, 2016) which are involved in many aspects of plant development, i.e. pattern formation, phytohormone interaction, tissue differentiation, reproduction, response to (a)biotic stresses, cell expansion and secondary cell wall deposition (Seifert and Roberts, 2007; Nguema-Ona *et al.*, 2014). These heavily glycosylated proteins are subdivided into four main classes: classical AGPs, AG peptides, Lys-rich AGPs and FLAs (Johnson *et al.*, 2003).

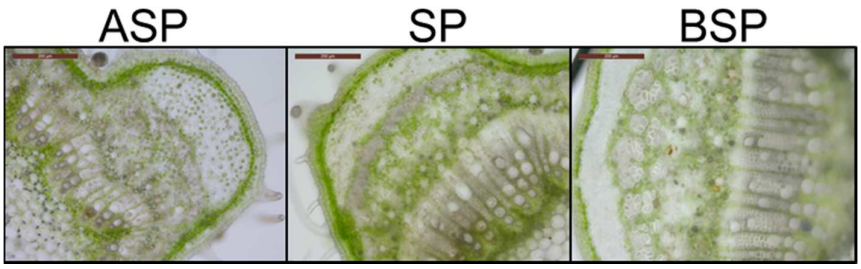


Figure 3-1: Stem cross-sections of the internodes above the snap point (ASP), containing the snap point (SP) and below the snap point (BSP). Scale bar: 200 µm.

FLAs are characterized by the occurrence of one or two AGP regions, as well as one or two fasciclin (FAS) domains (Johnson *et al.*, 2003). Although a consensus sequence for the FAS domains is lacking, two regions are highly conserved, named H1 and H2 (of *ca*. 10 amino acids) (Johnson *et al.*, 2003). Additionally, most FLAs show an N-terminal signal peptide and a C-terminal glycosylphosphatidylinositol (GPI) membrane anchor (Johnson *et al.*, 2003), mediating attachment to the cell surface.

*FLA*s constitute multigene families in plants: for example, 21 *FLA*s have been identified in thale cress (Johnson *et al.*, 2003), 35 in poplar (Zang *et al.*, 2015) and 18 in eucalypt (MacMillan *et al.*, 2015). Molecular studies focused on FLAs are important, since they increase our understanding of the molecular functions of this protein family: the available literature on the topic has shown that FLAs in plants are not only related to tissue-specific functions, but also involved in generalised responses to environmental constraints, both biotic and abiotic (Johnson *et al.*, 2003; Nguema-Ona *et al.*, 2014; Ihsan *et al.*, 2017). Additionally, a strong body of evidence in the literature has highlighted the importance of FLAs in regulating aspects linked to cell

wall biosynthesis and, more generally, to stem mechanics in both herbaceous and woody species, as well as fibre growth. For instance, in *Arabidopsis*, insertional mutants of *Atfla11* and *Atfla12* and *Atfla11*/*fla12* double mutants show modified stem mechanics, due to a decrease in cellulose, arabinose and galactose in secondary cell walls (MacMillan *et al.*, 2010). Likewise, in *Eucalyptus*, FLAs belonging to the subclade A are involved in stem mechanics (MacMillan *et al.*, 2015): in particular EgrFLA2 is linked to cellulose microfibril angle. In poplar, antisense expression of *PtFLA6* alters secondary cell wall composition in the xylem, by affecting the biosynthesis of lignin and cellulose (Wang *et al.*, 2015). In flax, some *FLA*s were shown to be up-regulated at the snap point, hence confirming the potential function of these genes in the regulation of fibre development (Roach and Deyholos, 2007).

In Chapter 1, we have shown that key aspects of the secondary cell wall development in the hypocotyl, such as lignification or cellulose deposition, are regulated at the gene expression level. In Chapter 2, we have demonstrated that the up-regulation of the genes involved in monolignol biosynthesis or coding for FLAs results in a higher abundance of the corresponding enzymes. In this Chapter, we sought to identify and study the expression patterns of the genes involved in cell wall biogenesis in the different stem tissues. By using bioinformatics coupled to RT-qPCR, we additionally show that some *FLA* genes are highly expressed in bast fibres. Moreover, we identify groups of *FLA*s upregulated either at the top or the bottom of the stem, which share putative conserved elements in their promoters.

I. Material and Methods

a. Plant material and growth conditions

After six weeks of growth in controlled chambers, samples were taken along five stem regions localized at different heights with respect to the "snap point" (an empirically defined reference region marking the transition from elongation to secondary cell wall thickening; Gorshkova *et al.*, 2003). The "ASP" (Above snap point) segment corresponds to the region right below the apex, the "SP" (Snap point) segment is the internode containing the snap point, the "BSP" (Below snap point) segment is located one internode below the snap point and the "BBSP" (Below-below snap point) segment is located two internodes below the snap point. Finally, the "BOT" region used for the expression analysis of *FLA* genes is located three internodes below the snap point (Figure 3-2). A segment of 2.5 cm was collected in the middle of each internode to avoid too much variation in gene expression, due to the varying developmental stages of the cell types.

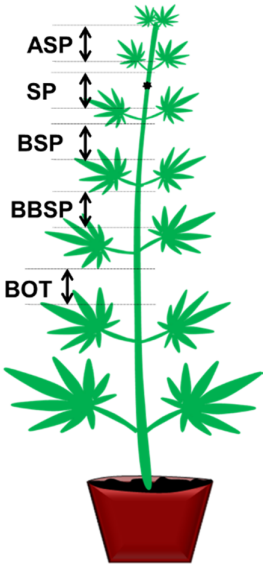


Figure 3-2: Sampling strategy of the stem tissues.

Each sample is 2.5 cm long.

The cortex was separated from the shivs by peeling the cortical tissues. For the expression analysis of *FLA* genes, the fibres were separated in a mortar, by repeatedly pressing the cortex with a pestle. The shivs were directly plunged in liquid nitrogen and stored at −80 °C. The number of independent biological replicates is four, with the exception of the BOT inner tissues, for which the biological replicates are three.

Cross sections of the ASP, SP and BSP sections were prepared with a vibratome. Fresh sections were embedded without fixation in 5 % low gelling point agarose, immediately cut at a thickness of 100 µm with a vibratome (Leica VT1200 S) and observed with a light microscope.

The hypocotyls, aged from 6 to 20 days after sowing, were grown and sampled as described in Chapter 1. Three biological replicates, each consisting of a pool of 20 hypocotyls, were used.

b. RNA extraction and RT-qPCR

Total RNA was extracted using a modified CTAB extraction protocol combined with a Qiagen RNeasy Plant Mini Kit in adult plants (Guerriero *et al.*, 2016a) and with the protocol described in Chapter 1 for the hypocotyls. The quality check of total RNA and the retro-transcription have been performed as in Chapter 1. A tissue maximization design was used to prepare the microplates for *CsaFLA* gene expression analysis (Bustin *et al.*, 2009). The expression of each *CsaFLA* was normalised using five reference genes (*Tubulin*, *CDPK*, *RAN*, *Clathrin* and *F-box*, which geNORM^{PLUS} identified as sufficient for appropriate data normalization). The expression of the cell wall-related genes was normalised using *Tubulin* and *Ubiquitin* (among *ETIF3E*, *ETIF3H*, *ETIF4* and *Cyclophilin*). *ETIF4E* and *GAPDH* were the most

suitable genes for normalisation in hypocotyls (among *Histone*, *EF2*, *Actin*, *Cyclophilin*, *Ubiquitin*, *Tubulin*, *ETIF3H* and *ETIF3E*). For statistical analysis, the normalized relative quantities exported from qBasePLUS were log2 transformed. A one-way ANOVA was carried out using IBM SPSS Statistics v19. A Tukey's HSD was performed as post-hoc test. The normal distribution of the data was verified with a Kolmogorov–Smirnov test.

c. Identification of *CsaFLA* genes using bioinformatics

In order to identify the *FLA* genes in *C. sativa* (hereafter referred to as *CsaFLA*s for the genes and CsaFLAs for the corresponding proteins), different databases were searched: the Medicinal Plant Genomics Resource and the *Cannabis sativa* Genome Browser Gateway. *CsaFLA*s were identified by using orthologous FLA protein sequences of *Arabidopsis thaliana* (Johnson *et al.*, 2003) and *Populus trichocarpa* (Zang *et al.*, 2015). These sequences were used to perform a BLAT analysis against the hemp Finola and Purple Kush database (*Cannabis* Genome Browser Gateway; van Bakel *et al.*, 2011) and a BLASTP in the MPGR database. Several incomplete sequences were retrieved when using the MPGR database; however it was possible to deduce their full length sequences either by querying the *Cannabis* Genome Browser Gateway, or the EST database at NCBI.

d. Primer design

Primers were designed using Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3-plus.cgi/) and verified with the OligoAnalyzer 3.1 tool from Integrated DNA technologies (http://eu.idtdna.com/calc/analyzer). Primer efficiencies

were checked via qPCR using a serial five-fold dilution of cDNA (25, 5, 1, 0.2, 0.04, 0.008 ng/µL).

e. In silico and phylogenetic analyses of CsaFLA protein sequences

Putative FAS domains were identified with the Motif Scan algorithm, N terminal signal peptides were identified with SignalP and SignalBlast; the subcellular localization was predicted with TargetP. The big-PI Plant Predictor program (Eisenhaber *et al.*, 2003) was used to identify the glycosylphosphatidylinositol (GPI) anchor. The H1 and H2 conserved regions, motifs and residues implicated in adhesion in both proteins were manually annotated according to Johnson and colleagues (2003). Conserved motifs in the *CsaFLA* promoter sequences (retrieved at the *Cannabis sativa* Genome Browser Gateway) were identified using the MEME Suite 4.11.2 (Bailey and Elkan, 1995). The identified motifs were subsequently analysed with Tomtom (Gupta *et al.*, 2007) for a comparison against the available motifs in the JASPAR CORE plant database 2016 (Mathelier *et al.*, 2016).

For the phylogenetic analysis, full-length sequences of FLAs from thale cress (Johnson *et al.*, 2003), poplar (Zang *et al.*, 2015), eucalypt (MacMillan *et al.*, 2015) and hemp were aligned with Clustal-Omega and the generated alignment submitted to PHYML (Dereeper *et al.*, 2008) to obtain a maximum likelihood phylogenetic tree. The Maximum Likelihood tree was constructed using an aLRT (approximate likelihood ratio test) for non-parametric branch support, based on a Shimodaira-Hasegawa-like procedure. The tree was visualized with iTOL-Interactive Tree Of Life.

f. Sequencing of some representative *CsaFLA* promoters

To determine the homology of the promoter sequences of the variety Santhica 27 with those from the Purple Kush and Finola reference genomes, primers were designed on 3 representative genes (*CsaFLA2–7-16*) using the available sequences at the *Cannabis sativa* Genome Browser to perform nested PCRs. Genomic DNA was extracted from stem tissues (whole internodes) by using a CTAB-based protocol coupled to the NucleoSpin Plant II kit (Macherey-Nagel). Briefly, 500 µl of extraction buffer (2% CTAB, 2.5% PVP-40, 2 M NaCl, 100 mM Tris-HCl pH 8.0, 25 mM EDTA and 10 µl RNase) were added to 100 mg of finely ground sample and the slurry was vortexed vigorously. After an incubation step at 60 °C for 10 min, 20 µl β-ME/ml buffer were added and the samples were further incubated for 20 min at 60 °C. Subsequently, 500 µl chloroform/isoamyl alcohol (24:1 v/v) were added, the samples were vortexed and centrifuged at RT for 10 min at 10000 g. To the aqueous phase, 2/3 cold isopropanol were added and the DNA was precipitated for 1 h at −20 °C. After this stage, the NucleoSpin II columns were used to bind the DNA and the manufacturer's instructions were followed to elute genomic DNA. PCRs were performed using 50 ng DNA and the Q5 Hot Start High-Fidelity 2X Master Mix, following the manufacturer's instructions. The optimal annealing temperatures were computed using the NEB Tm calculator. PCR products were ligated into the pGEM-T Easy vector, following the manufacturer's instructions and cloned into JM109 chemically competent cells. Three positive clones for each gene promoter were grown overnight at 37 °C in LB medium supplemented with ampicillin 100 µg/ml. Plasmids were extracted using the QIAGEN plasmid miniprep kit and sequenced on an Applied Biosystems 3500

Genetic Analyser using the BigDye Terminator v3.1 Cycle Sequencing and the BigDye XTerminator Purification kits, according to the manufacturer's instructions.

II. Results

a. General overview of the processes related to the cell wall in adult plant

An RT-qPCR general overview of the events related to primary and secondary cell wall deposition is provided in Figure 3-3. The hierarchical clustering shows four main expression patterns. In group A the genes associated with primary cell wall cellulose deposition are found (*CesA1*, *CesA3*, *CesA6A* and *CesA6B*), as well as *β-galactosidase 2* (*BGAL2*). These genes are ubiquitously expressed within the targeted tissues. Group B gathers together some major genes involved in lignification: *PAL*, *CAD4* and *4CL1* are part of the monolignol pathway (Figure 2-1), *MET1* and *SAM* are involved in monolignol methylation (Shen *et al.*, 2002) and *LAC4* is one of the enzymes partaking in lignin polymerisation (Berthet *et al.*, 2011). These genes are in general more expressed in the core tissue than in the cortical tissue. In group C are two master transcriptional regulators of secondary cell wall deposition, *NST1* and *NST2* (Sakamoto and Mitsuda, 2014), the three cellulose synthases associated with secondary cell wall biogenesis, *CesA4*, *CesA7* and *CesA8* (Hill *et al.*, 2014), as well as a class III peroxidase which may be involved in lignification, the ortholog of *Arabidopsis PRX49* (Herrero *et al.*, 2013). These genes are more expressed at the snap point and in lower internodes and slightly up-regulated in the inner tissues as compared to outer tissues. Finally, the similar expression patterns of *Pinoresinol lariciresinol reductase* (*PLR)* and *PRX72* form group D. These two

genes are more expressed above the snap point. By contrast with the three other groups, they are up-regulated in the outer tissues, which may point to an important role in bast fibre development.

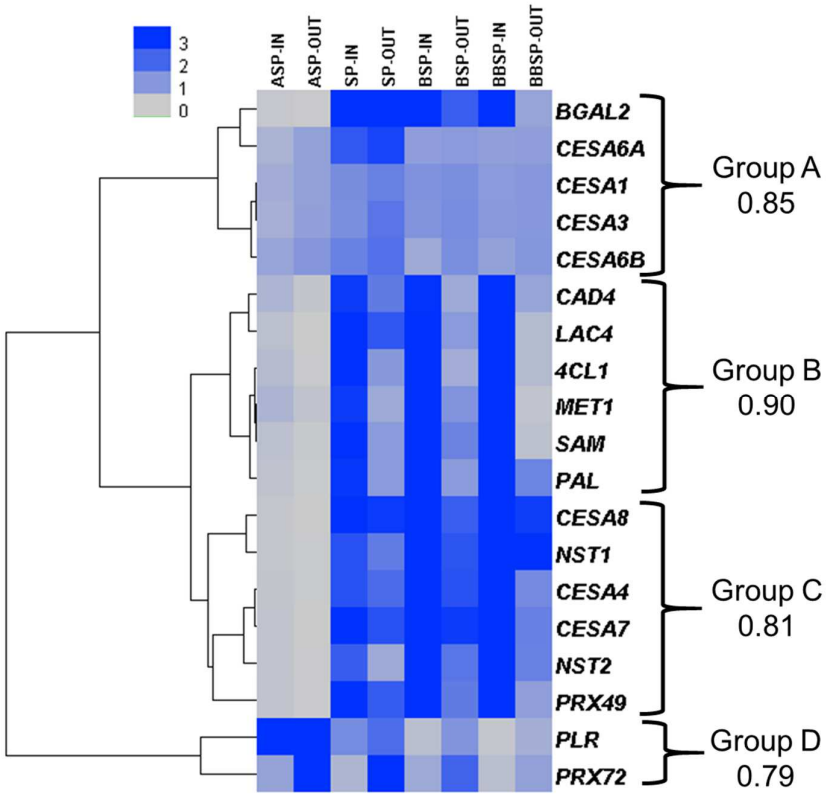


Figure 3-3: Gene expression analysis targeting processes related to cell wall deposition. Hierarchical clustering of profiles from nineteen genes in four stem regions and inner (-IN) or outer (-OUT) part. The higher the value, the higher the gene expression (value normalised for each gene). For each group, the Pearson coefficient is provided. Abbreviations are as in Figure 3-2 and text.

b. Identification of putative FLAs in *C. sativa*: Protein architecture and phylogenetic analysis

BLAST/BLAT analyses of the 21 *A. thaliana* sequences (*AtFLA*s) performed against the Medicinal Plant Genomics Resource, the NCBI EST and the *Cannabis* Genome Browser Gateway databases led to the identification of 23 *CsaFLA*s. It should be noted that, during the database queries, a contig, i.e. csa_locus_44222_iso_1_len_407_ver_2, which was initially called *CsaFLA22* and retrieved at the Medicinal Plant Genomics Resource (MPGR), was also found. However, we believe that this partial gene was erroneously attributed to *C. sativa*, since we never amplified any product with different primers designed on it and the reported FPKM values at the MPGR are 0 for all the tissues examined. We discarded this gene from the analyses, but kept the original nomenclature given to the hemp FLA genes (i.e. *CsaFLA1–24*), as at this stage we cannot rule out the existence of this gene in textile hemp.

To check that the putative CsaFLAs belong to the FLA family, the occurrence of the following features was checked: the presence of at least one FAS domain, a signal peptide at the N-terminus, (in some cases) a GPI anchor at the C terminus and the presence of AGP domains (Figure 3-4).

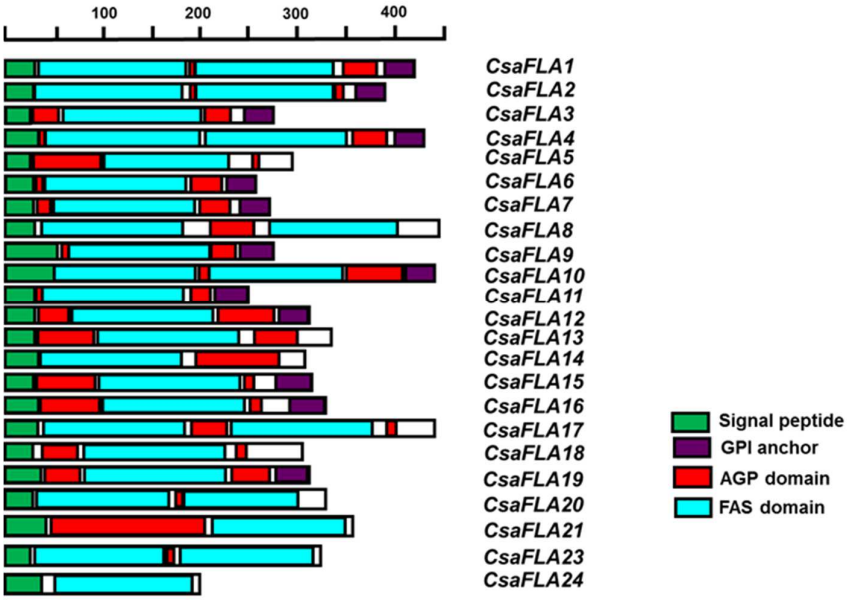


Figure 3-4: Schematic representation of CsaFLA domains. The details relative to the signal peptide, GPI anchor, AGP and FAS domain(s) are indicated for the 23 CsaFLAs.

In order to investigate the evolutionary relationship between known plant FLA proteins, a maximum-likelihood phylogenetic tree was built using the identified 23 CsaFLAs, 21 AtFLAs, 18 Eucalyptus (EgrFLAs) and 35 poplar sequences (PtrFLAs) (Figure 3-5). It should be noted that we chose to perform the phylogenetic analysis using full-length CsaFLA sequences to conform to the previously published tree of poplar FLAs (Zang *et al.*, 2015). Different results may be obtained if the mature protein sequences are used.

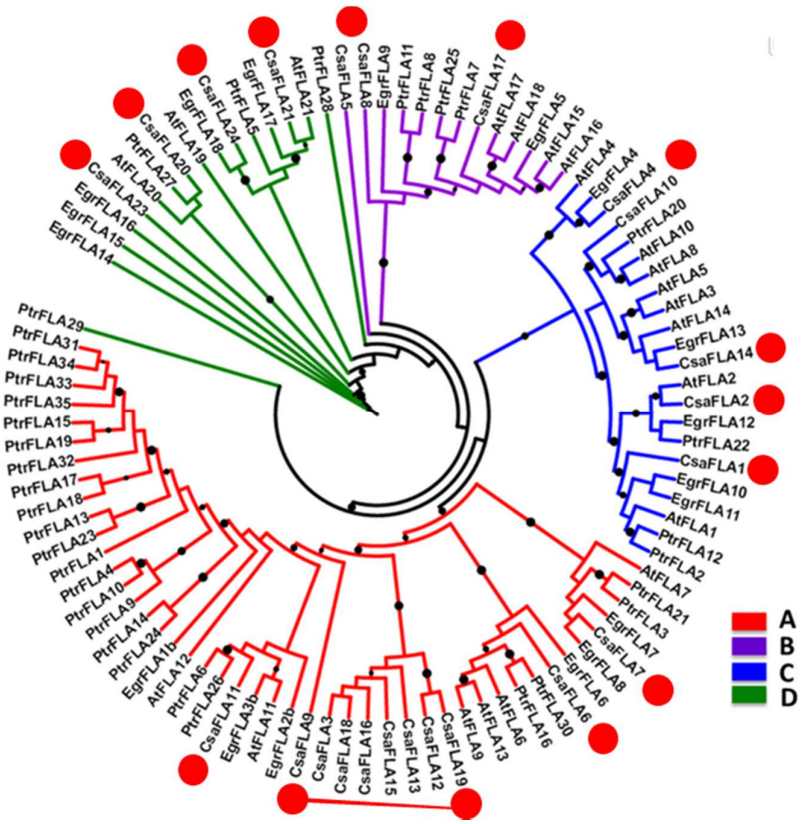


Figure 3-5: The phylogenetic tree of FLA proteins from hemp, poplar, eucalypt and thale cress. The maximum likelihood tree was constructed with 23 FLA proteins in hemp (red dots), 18 in eucalypt, 35 in poplar and 21 in *Arabidopsis*, resulting in four clades (A to D). Only bootstrap values higher than 0.8 are here visualized as black circles (the bigger the circle, the higher the bootstrap value).

CsaFLAs cluster into four major clades (A-D). Clade A is the largest, with 11 hemp members (CsaFLA3/6/7/9/11/12/13/15/16/18/19) containing a single FAS domain flanked by two AGP domain and a GPI anchor at the C-terminus. Clade B includes CsaFLA5/8/17 which contains two FAS domains and a single AGP region. Clade C comprises 5 members

(CsaFLA1/2/4/10/14), characterized by two FAS, two AGP domains and a GPI anchor. The last clade includes 4 proteins (CsaFLA20/21/23/24) with no distinctive protein architecture (although Csa-FLA20 and CsaFLA23 have 2 FAS domains like clade B, however their lengths are smaller). The percentage of CsaFLA identity with the putative orthologs from thale cress is between 32 and 76% (Table 3-1).

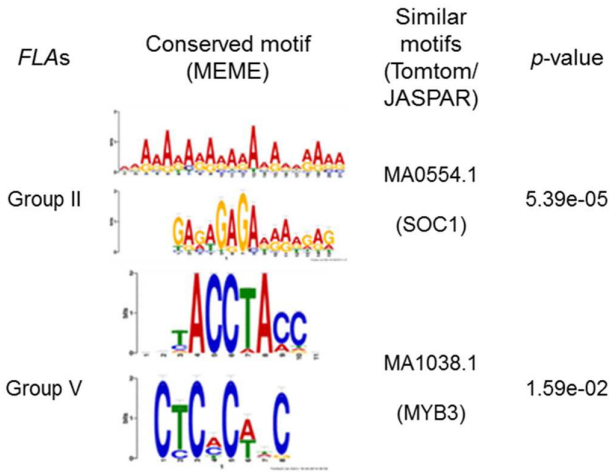
Table 3-1: Identities of CsaFLAs with AtFLAs

Name	Class	Arabidopsis sequence identity		Arabidopsis sequence homology
CsaFLA1	Class C	AtFLA1	AT5G55730	63%
CsaFLA2	Class C	AtFLA2	AT4G12730	60%
CsaFLA3	Class A	AtFLA11	AT5G03170	49%
CsaFLA4	Class C	AtFLA4	AT3G46550	56%
CsaFLA5	Class B	AtFLA19	AT1G15190	32%
CsaFLA6	Class A	AtFLA9	AT1G03870	54%
CsaFLA7	Class A	AtFLA7	AT2G04780	62%
CsaFLA8	Class B	AtFLA17	AT5G06390	71%
CsaFLA9	Class A	AtFLA11	AT5G03170	57%
CsaFLA10	Class C	AtFLA10	AT3G60900	72%
CsaFLA11	Class A	AtFLA11	AT5G03170	64%
CsaFLA12	Class A	AtFLA12	AT5G60490	55%
CsaFLA13	Class A	AtFLA13	AT5G44130	42%
CsaFLA14	Class C	AtFLA14	AT3G12660	45%
CsaFLA15	Class A	AtFLA11	AT5G03170	53%
CsaFLA16	Class A	AtFLA12	AT5G60490	55%
CsaFLA17	Class B	AtFLA17	AT5G06390	76%
CsaFLA18	Class A	AtFLA11	AT5G03170	56%
CsaFLA19	Class A	AtFLA11	AT5G03170	55%
CsaFLA20	Class D	AtFLA20	AT5G40940	36%
CsaFLA21	Class D	AtFLA21	AT5G06920	45%
CsaFLA23	Class D	AtFLA20	AT5G40940	36%
CsaFLA24	Class D	No homology		

c. Identification of conserved motifs in the promoters of some *CsaFLA*s

A bioinformatic analysis was carried out on the available promoter regions (between 116 and 1064 bp retrieved at the *Cannabis* Genome Browser Gateway) of the *CsaFLA* genes showing a distinctive expression pattern in the bast fibres. The promoters of the genes within group I, II and V were selected, in the light of their upregulation at the snap point, at the younger and older stem regions respectively (Figure 3-6). While no conserved motifs could be found with the MEME suite tool for the genes in group I, one motif was found for the genes within group II and V. The search carried out with the conserved sequence of group II and V *FLA*s in the JASPAR CORE 2016 plants database identified SOC1 (lowest p-value among the matches retrieved) and MYB3 as candidates recognizing similar motifs (Table 3-2). SOC1 is a transcription activator controlling flowering time and is also probably involved in photoperiod perception. MYB3 represses the phenylpropanoid biosynthesis-related genes and responds to salt stress, wounding, abscisic acid and salicylic acid.

Table 3-2: Conserved motifs in the promoters of FLAs from group II and V. The first motif in each table cell is the one found in the database, the second motif is the one found in the promoters of hemp FLAs. The computed *p*-value is indicated.



d. *CsaFLA* expression patterns

Of the twenty three *CsaFLA*s identified, twenty two were expressed in the stem tissues (Figure 3-6). *CsaFLA14* was detected at very low levels in the stem tissues (Ct > 32).

The heat-map hierarchical clustering shows five major expression trends (Figure 3-6). These are the following: 1) a group of genes (*CsaFLA2–6-24*) is upregulated at the middle internode containing the snap point (in the core, the expression decreases towards the base of the stem); 2) *CsaFLA1-4-7-8-10-20-23* are expressed at higher levels in the top and decreased towards the bottom internode; 3) two *FLA*s, *CsaFLA5* and *CsaFLA21*, are downregulated at the snap point; 4) three genes, *CsaFLA9-11-17*, show a tendency to upregulation at the snap point, although the pattern is less marked with respect to group I (and in the core the expression increased towards the stem base); 5) the last group comprises *FLA*s upregulated at the middle and bottom (*CsaFLA3-12-13-15-16-18-19*).

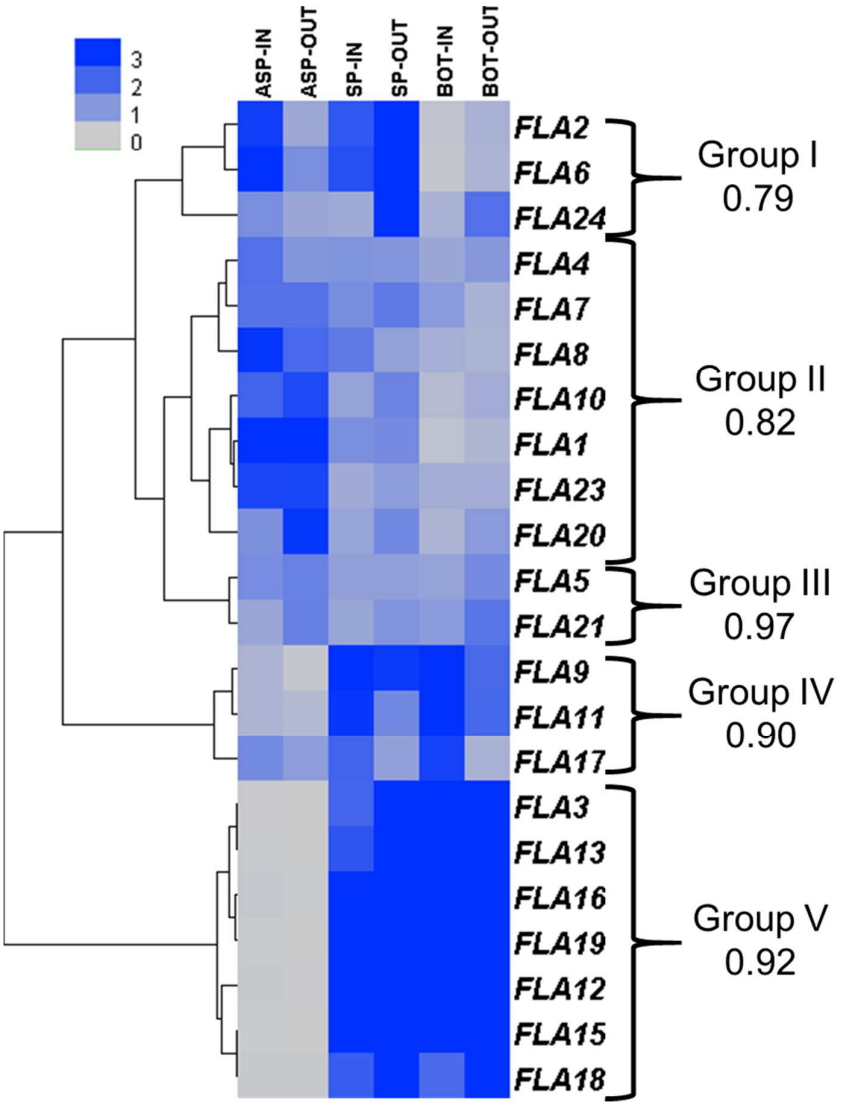


Figure 3-6: Expression profiles of *CsaFLA*s in hemp stems. Hierarchical clustering of profiles from twenty two genes in three stem regions and inner (-IN) or outer (-OUT) part. For each group, the Pearson coefficient is provided. Internode name as in Figure 3-2.

Based on the expression profiles observed in adult plants, the following *FLA*s have been selected for an analysis on the hypocotyls

aged 6-9-12-15-17-20 (H6 to H20): *FLA2* (Group I), *FLA1*, *FLA8* (Group II), *FLA21* (Group III), *FLA9*, *FLA11* (Group IV), and *FLA3* and *FLA13* (Group V). The hemp hypocotyl was proven to be a suitable model to study cell wall-related processes accompanying secondary growth (Chapter 1), therefore the goal was to verify whether their expression pattern highlighted the same trend observed in adult stems. *CsaFLA1-2-8-21* were more expressed in young hypocotyls (H6); *CsaFLA3-9-11-13* were more expressed in H15, H17 and H20 (Figure 3-7).

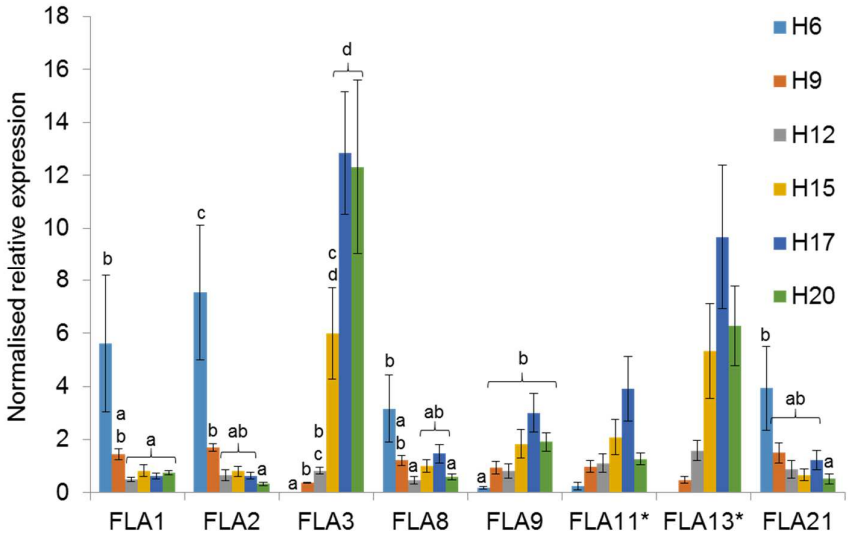


Figure 3-7: Expression analysis of eight *CsaFLA*s in the hypocotyls of *C. sativa*. Error bars indicate the standard error of the mean (n = 3). Different letters indicate statistically significant values at the one-way ANOVA test ($p < 0.05$) with a Tukey post-hoc test. * indicates a non-normal distribution (Kolmogorov-Smirnov test) and a Kruskal-Wallis associated *p*-value < 0.05.

III. Discussion

As stressed in Chapters 1 and 2, the pattern of several *FLA*s is associated either with elongation or secondary growth of the hypocotyl. Considering that FLAs are involved in the regulation of events associated with cell wall biosynthesis and stem biomechanics (Ito *et al.*, 2005; Persson *et al.*, 2005; Huang *et al.*, 2013), we have identified the *CsaFLA*s and determined their expression in tissues at several stages of development, together with genes involved in cell wall formation.

a. Cellulose and lignin associated genes

The targeted gene expression analysis (Figure 3-3) shows that the higher lignin content of the xylem tissues is correlated with an up-regulation of the genes of the phenylpropanoid / monolignol pathway. Those genes are found in the cluster B. The *CesA*s associated with the cellulose synthase complex of the secondary cell wall (*CesA4*, *CesA7* and *CesA8*; Hill *et al.*, 2014) are grouped in the cluster C. They are slightly more expressed in the inner tissues and weakly in the elongating stem (Figure 3-3). In the elongating internode (ASP), the deposition of secondary cell wall is restricted to the metaxylem and protoxylem. In secondary xylem, the secondary cell wall is deposited in a reticulated or pitted pattern, characterised by a massive deposition of cellulose (Schuetz *et al.*, 2014). Fibres and tracheary elements of the xylem have a xylan-type secondary cell wall (i.e., organised through S1, S2 and S3 sub-layers) while phloem fibres have a gelatinous-type secondary cell wall (S1, S2 and G-layer) (Guerriero *et al.*, 2016b). Generally, there is, if any, only slight differences in the *CesA*s expression patterns between xylan-type and gelatinous-type secondary cell walls (Gorshkova *et al.*, 2012). However, a recent study

performed on flax has highlighted a higher expression of both primary and secondary cell wall-related *CesA*s in phloem fibres depositing their G-layer (Mokshina *et al.*, 2017). In addition, a strong bast fibre phenotype (reduced number and irregular cell shape associated with altered cell wall composition) has been observed in flax plants with virus induced silencing of *CesA* genes (*CesA1* or *CesA6*), usually acting in primary cell wall biogenesis (Chantreau *et al.*, 2015). The data presented on Figure 3-3 do not allow to confirm nor invalidate this hypothesis, however it is noteworthy to mention that several *CesA* isoforms may be missing from our analysis. Between fifteen and sixteen predicted *CesA*s have been found in the flax genome (Mokshina *et al.*, 2014; Chantreau *et al.*, 2015). We may thus anticipate additional *CesA* isoforms in hemp. In agreement with the data from Chantreau and Mokshina, *CesA4*, *CesA7* and *CesA8* were more highly expressed in inner tissues and may thus be considered as functional orthologs of *AtCesA4*, *AtCesA7* and *AtCesA8*.

The expression profiles of *PLR* and *PRX72* is in sharp contrast with the genes from cluster C. Both genes are more expressed in the elongating internode (ASP) and in the outer tissues. *PLR* is an entry enzyme for the biosynthesis of lignans. Lignans are formed by enantioselective coupling of two monolignol units (Nakatsubo *et al.*, 2008). This family of molecules is involved in plant growth (Cutillo *et al.*, 2003), lignin distribution during secondary cell wall biosynthesis (Zhao *et al.*, 2015) and redox homeostasis during lignification (Niculaes *et al.*, 2014). From our data, we can hypothesise that *PLR* regulates stem elongation via the biosynthesis of specific lignans (Yamauchi *et al.*, 2015). The expression of *PLR* was also higher in the elongating hypocotyl (H6 and H9), as explained in Chapter 2. A functional analysis of this gene, as well as a detailed chemical characterisation of lignans

present in elongating- and non-elongating tissues will validate this hypothesis. The expression of *PLR* and *PRX72* is biased towards the cortical tissue and may thus be important for the development of bast fibres. The *Arabidopsis* ortholog of *PRX72* is involved in lignin biosynthesis (Fernández-Pérez *et al.*, 2015). A mutant defective in AtPRX72 shows thinner secondary cell walls only in interfascicular fibres and a lower lignin S/G ratio. Based on its expression profile in hemp, a role in bast fibre lignification is unlikely. Indeed, bast fibres lignify mostly after they reach their final size and S-lignin is deposited at the latest stage of lignification (Fernández-Pérez *et al.*, 2015).

b. Characterisation of *CsaFLA*s

The multigenic nature of the FLA family suggests that these proteins have evolved to fit rather specific roles in plant development. The variability in the peptide sequence and the number of FAS and AGP domains between FLAs point to an interaction with various cell wall polysaccharides, such as pectins for FAS and hemicelluloses and pectins for AGP (Ito *et al.*, 2005; Huang *et al.*, 2013). The evolution of the composition and spatial organisation of the polysaccharide extracellular matrix during development (from elongation to secondary cell wall formation) may explain why some *FLA*s are more expressed in elongating tissues, while others are up-regulated during the thickening stage. Moreover, some of these proteins have a GPI region anchoring them to the plasma membrane. Such a domain is found in lipid rafts, contributing to the stabilisation and activity of the embedded proteins (Langhans *et al.*, 2017) while interacting with the cytoskeleton (Johnson *et al.*, 2003).

The FLAs identified in *C. sativa* group into the previously described four phylogenetic classes (Figure 3-5) (Johnson *et al.*, 2003). A nomenclature of CsaFLAs is hereby also proposed which

follows the *Arabidopsis* classification (i.e. when the phylogenetic tree highlighted clustering of a CsaFLA proteins with a specific AtFLA, the same number was assigned to *C. sativa*).

Within class A, the largest (Figure 3-5), it is possible to observe a separate clade represented by *CsaFLA3-12-13-15-16-18-19* which is highly expressed at the snap point and in the older stem regions, both in the bast fibres and the shivs (Figure 3-6). A subset of class A genes (*CsaFLA3-9-11-13*) was more expressed in the old hypocotyls (peaking at H17 with high values at H15 and H20). As shown in Chapter 1, the hypocotyl undergoes secondary growth from H9 onwards. The phylogenetic position of this cluster of FLAs, together with their common expression pattern, indicates a specific role in secondary growth. This group of genes may indeed represent hemp-specific single FAS domain FLAs specialized in secondary growth, in a manner analogous to what was previously shown in eucalypt and thale cress (MacMillan *et al*, 2015). Hemp is unfortunately recalcitrant to transformation, therefore homologous testing, as previously performed on e.g. eucalypt *FLA*s (MacMillan *et al*, 2015), is cumbersome. However heterologous testing in a more amenable system, e.g. *Nicotiana tabacum*, can confirm or refute the hypothesis.

It is here worth discussing also the phylogenetic position of CsaFLA11 in a clade grouping AtFLA11, EgrFLA2b and EgrFLA3b (Figure 3-5). These genes were shown to affect stem mechanics, as well as cell wall architecture (MacMillan *et al*, 2010; MacMillan *et al.*, 2015). The *AtFLA11* transcript was detected in the xylem and interfascicular fibres in inflorescence stem, preceding the lignification of those two tissues (Ito *et al.*, 2005); *CsaFLA11* also shows a gradual increase in expression towards the older regions of the stem and it is slightly more expressed in the older hypocotyl too (Figure 3-7). This

FLA represents another interesting candidate putatively involved in cell wall-related processes in textile hemp.

Within class C, CsaFLA4 and CsaFLA1 group together with the characterized orthologs from thale cress (Figure 3-5). AtFLA4 (SOS5) is involved in cell expansion (Shi *et al.*, 2003) and AtFLA1 was shown to regulate root and shoot development in tissue culture (Johnson *et al.*, 2011). *CsaFLA8* was more expressed in the TOP region of the stem, as well as in H6, suggesting a role in elongating tissue. However, it remains to be shown whether the hemp genes are involved in the same regulatory networks as in *Arabidopsis*.

The first group of *FLA* genes, represented by *CsaFLA2-6-24*, shows a different expression profile in the bast fibres and the shivs (Figure 3-6). The expression in the shivs shows a decrease from the top to the bottom of the stem, while in the bast fibres their expression peaks at the snap point. This is quite interesting if we consider that the snap point is the region marking a shift in the stem mechanical properties, as it determines the transition from cell elongation to thickening (Gorshkova *et al.*, 2003). It was shown that the young stem regions of hemp at the vegetative stage of growth are characterized by the presence of *ca*. 66 % glucose, while older regions have about 82 %: this result confirms that during their transition from elongation to thickening, bast fibres require great amounts of glucose for the synthesis of cellulose (Crônier *et al.*, 2005). These three FLAs may therefore be involved in cell wall-related processes occurring during this transition. Additionally, this is in agreement with the flax microarray data showing upregulation of certain *FLA*s around the snap point (Roach and Deyholos, 2007) and with the increased expression of poplar *FLA*s in tension wood, which, like bast fibres, is composed of a cellulosic G-layer (Lafarguette *et al.*, 2004; Gritsch *et al.*, 2015). As

previously discussed for poplar tension wood, specific FLAs with a GPI-anchor might be involved in the cytoskeleton-cell wall connections during fibre expansion / elongation (Lafarguette *et al.*, 2004). This would be the case of CsaFLA2 and CsaFLA6, which possess a GPI-anchor (Figure 3-4). In the hypocotyl, *CsaFLA2* was significantly more expressed in H6 (Figure 3-7).

FLAs might also be involved in triggering a cellular signal inducing the formation of the G-layer, via the cleavage of their *N*-acetylglucosamine oligosaccharides [$(GlcNAc)_n$] by the action of chitinases (Mokshina *et al.*, 2014). It was shown that in flax stems, specific chitinases are highly expressed in bast fibres and may regulate G-layer formation in these cell types (Mokshina *et al.*, 2014). AGPs contain GlcNAc and can be activated by enzymes such as endochitinases, resulting in the production of various oligosaccharides (van Hengel *et al.*, 2001). The accumulation of short $(GlcNAc)_n$ triggered by the heterologous expression of *N-ACETYLGLUCOSAMINYLTRANSFERASE* (*NodC*) of *Azorhizobium caulinodans* alters internode length and leads to a general loss of cell wall strength in *Arabidopsis* (Vanholme *et al.*, 2014). These oligosaccharides may compete in carbohydrate-carbohydrate and carbohydrate-protein interactions, leading to perturbation in the noncovalent interactions in the cell wall or at the cell wall-plasma membrane interface (Vanholme *et al.*, 2014). Therefore, it is reasonable to assume that the concerted action of specific FLAs, chitinases and specific oligosaccharides may be involved in the transition from elongation to G-layer formation in hemp.

In group II and V are *FLA*s which, in the bast fibres, show a gradual decreased expression from the apical to the basal part of the stem and an increase in expression, respectively. A similar trend was

observed in the hypocotyls: *CsaFLA8* (belonging to the stem group II) was more expressed in H6; *CsaFLA13* (belonging to the stem group V) was more expressed in H15, H17 and H20 (Figure 3-7). In addition, the hypocotyl expression pattern of *CsaFLA3* was similar to the one of *CsaFLA13*. Our study therefore identified specific FLAs likely involved in bast fibre elongation during intrusive growth (*CsaFLA1-4-7-8-10-20-23*) and others involved in secondary cell wall deposition during the thickening stage (*CsaFLA3-12-13-15-16-18-19*).

c. *FLA*s promoter analysis

In order to investigate whether specific regulatory elements occurred in the promoters of the genes showing specific expression patterns in the stem tissues, we analysed the genes from group I, II and V (Figure 3-6). While for group I no conserved motifs could be obtained, two conserved sequences were found for group II and V (Table 3-2). A conserved motif recognised by the MADS box transcription factor SOC1 could be identified in the promoters of the genes upregulated in the apical stem regions: this finding suggests that they may be involved in a developmental program regulating the transition from vegetative to reproductive growth and/or the response to hormonal regulation (e.g. via gibberellin). In this respect it is noteworthy that in *A. thaliana* SOC1 was shown to control the annual growth habit: *soc1 ful* mutants show indeed woody growth reminiscent of the perennial lifestyle (Melzer *et al.*, 2008). Hence the *FLA*s upregulated at the top of the stems might belong to a regulatory circuit controlling elongation and suppressing secondary growth.

The genes in group V (Figure 3-6) show the presence of a conserved motif putatively recognized by MYB3, which is an R2R3 MYB transcriptional repressor belonging to subgroup S4 together with the well-characterized AtMYB4 (Jin *et al.*, 2000). MYB4 negatively

regulates phenylpropanoid biosynthesis and is consistently down-regulated in several mutants (*pom1*, *eli1* and *det3*) in which lignin highly accumulates (Rogers *et al.*, 2005). It is therefore possible that the identified element is involved in the coordination of phenylpropanoid biosynthesis in bast fibres and might regulate the hypolignification observed in these cells (Day *et al.*, 2005; Huis *et al.*, 2012). As extensively presented in Chapter 4, we observed an upregulation of the *SOC1* gene at the top (4-fold induction with respect to the bottom and 1.3-fold induction with respect to the middle) and *MYB4* at the bottom (1.7-fold induction with respect to the top and 4.6-fold induction with respect to the middle). This result therefore strengthens the existence of a putative regulatory circuit at the top and bottom of adult hemp plants.

IV. Conclusion

In conclusion, we have identified (at least) 23 genes coding for FLAs in textile hemp, some of which specific to distinct stages of bast fibre development, and put them in perspective with the expression of genes involved in cellulose and lignin deposition. Bioinformatics has highlighted the occurrence of conserved motifs in the promoters of *FLA* genes upregulated either at the top or at the bottom of the stem. This finding points to the existence of a fine regulatory network controlling bast fibre elongation and cell wall composition. Future functional analyses carried out on heterologous systems will shed more light on the functions of the identified genes. The differential cell wall composition of inner and outer stem tissues is at least partially regulated at the gene expression level, especially for lignin. Additional studies should be performed to determine the CESA isoforms responsible for the deposition of cellulose in the gelatinous layer of the bast fibres.

References

Bailey, T.L., and Elkan, C. (1995). "Fitting a mixture model by expectation maximization to discover motifs in biopolymers", in: *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology.* (Menlo Park, California,: AAAI Press).

Berthet, S., Demont-Caulet, N., Pollet, B., Bidzinski, P., Cézard, L., Le Bris, P., *et al.* (2011). Disruption of LACCASE4 and 17 Results in Tissue-Specific Alterations to Lignification of *Arabidopsis thaliana* Stems. *Plant Cell* 23**,** 1124-1137.

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., *et al.* (2009). The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR Experiments. *Clin. Chem.* 55**,** 611-622.

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.***,** 1312-1324.

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.

Cutillo, F., D'abrosca, B., Dellagreca, M., Fiorentino, A., and Zarrelli, A. (2003). Lignans and Neolignans from *Brassica fruticulosa*: Effects on Seed Germination and Plant Growth. *J. Agric. Food Chem.* 51**,** 6165-6172.

Day, A., Addi, M., Kim, W., David, H., Bert, F., Mesnage, P., *et al.* (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.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., *et al.* (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* 36**,** W465-W469.

Eisenhaber, B., Wildpaner, M., Schultz, C.J., Borner, G.H.H., Dupree, P., and Eisenhaber, F. (2003). Glycosylphosphatidylinositol Lipid Anchoring of Plant Proteins. Sensitive Prediction from Sequence- and Genome-Wide Studies for *Arabidopsis* and Rice. *Plant Physiol.* 133**,** 1691-1701.

Fernández-Pérez, F., Pomar, F., Pedreño, M.A., and Novo-Uzal, E. (2015). Suppression of *Arabidopsis* peroxidase 72 alters cell wall and phenylpropanoid metabolism. *Plant Sci.* 239**,** 192-199.

Fike, J. (2016). Industrial Hemp: Renewed Opportunities for an Ancient Crop. *Crit. Rev. Plant Sci.* 35**,** 406-424.

Gorshkov, O., Mokshina, N., Gorshkov, V., Chemikosova, S., Gogolev, Y., and Gorshkova, T. (2017). Transcriptome portrait of cellulose-enriched flax fibres at advanced stage of specialization. *Plant Mol. Biol.* 93**,** 431-449.

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

Gorshkova, T., Brutch, N., Chabbert, B., Deyholos, M., Hayashi, T., Lev-Yadun, S., *et al.* (2012). Plant Fiber Formation: State of the Art, Recent and Expected Progress, and Open Questions. *Crit. Rev. Plant Sci.* 31**,** 201-228.

Gritsch, C., Wan, Y., Mitchell, R.a.C., Shewry, P.R., Hanley, S.J., *et al.* (2015). G-fibre cell wall development in willow stems during tension wood induction. *J. Exp. Bot.* 66**,** 6447-6459.

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.

Guerriero, G., Mangeot-Peter, L., Hausman, J.F., and Legay, S. (2016a). Extraction of High Quality RNA from *Cannabis sativa* Bast Fibres: A Vademecum for Molecular Biologists. *Fibers* 4, 23.

Guerriero, G., Hausman, J.-F., Strauss, J., Ertan, H., and Siddiqui, K.S. (2016b). Lignocellulosic biomass: Biosynthesis, degradation, and industrial utilization. *Eng. Life Sci.* 16**,** 1-16.

Gupta, S., Stamatoyannopoulos, J.A., Bailey, T.L., and Noble, W.S. (2007). Quantifying similarity between motifs. *Genome Biol.* 8**,** R24.

Herrero, J., Fernández-Pérez, F., Yebra, T., Novo-Uzal, E., Pomar, F., Pedreno, M., *et al.* (2013). Bioinformatic and functional characterization of the basic peroxidase 72 from *Arabidopsis thaliana* involved in lignin biosynthesis. *Planta* 237**,** 1599-1612.

Hill, J.L., Hammudi, M.B., and Tien, M. (2014). The *Arabidopsis* Cellulose Synthase Complex: A Proposed Hexamer of CESA Trimers in an Equimolar Stoichiometry. *Plant Cell* 26**,** 4834-4842.

Huang, G.Q., Gong, S.Y., Xu, W.L., Li, W., Li, P., Zhang, C.J., *et al.* (2013). A Fasciclin-Like Arabinogalactan Protein, GhFLA1, Is Involved in Fiber Initiation and Elongation of Cotton. *Plant Physiol.* 161**,** 1278-1290.

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

Ihsan, M.Z., Ahmad, S.J.N., Shah, Z.H., Rehman, H.M., Aslam, Z., Ahuja, I., *et al.* (2017). Gene Mining for Proline Based Signaling Proteins in Cell Wall of

Arabidopsis thaliana. *Front. Plant Sci.* 8-233.

Ito, S., Suzuki, Y., Miyamoto, K., Ueda, J., and Yamaguchi, I. (2005). AtFLA11, a Fasciclin-Like Arabinogalactan-Protein, Specifically Localized in Screlenchyma Cells. *Biosci., Biotechnol., Biochem.* 69**,** 1963-1969.

Jin, H., Cominelli, E., Bailey, P., Parr, A., Mehrtens, F., Jones, J., *et al.* (2000). Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in *Arabidopsis*. *EMBO J.* 19**,** 6150-6161.

Johnson, K.L., Jones, B.J., Bacic, A., and Schultz, C.J. (2003). The Fasciclin-Like Arabinogalactan Proteins of *Arabidopsis*. A Multigene Family of Putative Cell Adhesion Molecules. *Plant Physiol.* 133**,** 1911-1925.

Johnson, K.L., Kibble, N.J., Bacic, A., and Schultz, C.J. (2011). A Fasciclin-Like Arabinogalactan-Protein (FLA) Mutant of *Arabidopsis thaliana*, *fla1*, Shows Defects in Shoot Regeneration. *PLoS One* 6**,** e25154.

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

Langhans, M., Weber, W., Babel, L., Grunewald, M., and Meckel, T. (2017). The right motifs for plant cell adhesion: what makes an adhesive site? *Protoplasma* 254, 95-108.

Macmillan, C.P., Mansfield, S.D., Stachurski, Z.H., Evans, R., and Southerton, S.G. (2010). Fasciclin-like arabinogalactan proteins: specialization for stem biomechanics and cell wall architecture in *Arabidopsis* and *Eucalyptus*. *Plant J.* 62**,** 689-703.

Macmillan, C.P., Taylor, L., Bi, Y., Southerton, S.G., Evans, R., and

Spokevicius, A. (2015). The fasciclin-like arabinogalactan protein family of *Eucalyptus grandis* contains members that impact wood biology and biomechanics. *New Phytol.* 206**,** 1314-1327.

Mathelier, A., Fornes, O., Arenillas, D.J., Chen, C.Y., Denay, G., Lee, J., *et al.* (2016). JASPAR 2016: a major expansion and update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res.* 44**,** D110-D115.

Melzer, S., Lens, F., Gennen, J., Vanneste, S., Rohde, A., and Beeckman, T. (2008). Flowering-time genes modulate meristem determinacy and growth form in *Arabidopsis thaliana*. *Nat. Genet.* 40**,** 1489-1492.

Mokshina, N., Gorshkova, T., and Deyholos, M.K. (2014). Chitinase-Like and Cellulose Synthase Gene Expression in Gelatinous-Type Cellulosic Walls of Flax (*Linum usitatissimum* L.) Bast Fibers. *PLoS One* 9**,** e97949.

Mokshina, N., Gorshkov, O., Ibragimova, N., Chernova, T., and Gorshkova, T. (2017). Cellulosic fibres of flax recruit both primary and secondary cell wall cellulose synthases during deposition of thick tertiary cell walls and in the course of graviresponse. *Funct. Plant Biol.* 44**,** 820-831.

Nakatsubo, T., Mizutani, M., Suzuki, S., Hattori, T., and Umezawa, T. (2008). Characterization of *Arabidopsis thaliana* Pinoresinol Reductase, a New Type of Enzyme Involved in Lignan Biosynthesis. *J. Biol. Chem.* 283**,** 15550-15557.

Nguema-Ona, E., Vicré, M., Gotté, M., Plancot, B., Lerouge, P., Bardor, M., and *et al.* (2014). Cell wall *O*-glycoproteins and *N*-glycoproteins: biosynthesis and some functional aspects. *Front. Plant Sci.* 5-499.

Niculaes, C., Morreel, K., Kim, H., Lu, F., Mckee, L.S., Ivens, B., *et al.* (2014). Phenylcoumaran Benzylic Ether Reductase Prevents Accumulation of Compounds Formed under Oxidative Conditions in Poplar Xylem. *Plant Cell*

26**,** 3775-3791.

Persson, S., Wei, H., Milne, J., Page, G.P., and Somerville, C.R. (2005). Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proc. Natl. Acad. Sci. U. S. A.* 102**,** 8633-8638.

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.

Rogers, L.A., Dubos, C., Surman, C., Willment, J., Cullis, I.F., Mansfield, S.D., and Campbell, M.M. (2005). Comparison of lignin deposition in three ectopic lignification mutants. *New Phytol.* 168**,** 123-140.

Sakamoto, S., and Mitsuda, N. (2015). Reconstitution of a Secondary Cell Wall in a Secondary Cell Wall-Deficient *Arabidopsis* Mutant. *Plant Cell Physiol.* 56**,** 299-310.

Schuetz, M., Benske, A., Smith, R.A., Watanabe, Y., Tobimatsu, Y., Ralph, J., *et al.* (2014). Laccases Direct Lignification in the Discrete Secondary Cell Wall Domains of Protoxylem. *Plant Physiol.* 166**,** 798-807.

Seifert, G.J., and Roberts, K. (2007). The Biology of Arabinogalactan Proteins. *Annu. Rev. Plant Biol.* 58**,** 137-161.

Shen, B., Li, C., and Tarczynski, M.C. (2002). High free-methionine and decreased lignin content result from a mutation in the *Arabidopsis* S-adenosyl-L-methionine synthetase 3 gene. *Plant J.* 29**,** 371-380.

Shi, H., Kim, Y., Guo, Y., Stevenson, B., and Zhu, J.K. (2003). The *Arabidopsis* SOS5 Locus Encodes a Putative Cell Surface Adhesion Protein and Is Required for Normal Cell Expansion. *Plant Cell* 15**,** 19-32.

Van Bakel, H., Stout, J.M., Cote, A.G., Tallon, C.M., Sharpe, A.G., Hughes, T.R., *et al.* (2011). The draft genome and transcriptome of *Cannabis sativa*. *Genome Biol.* 12**,** R102-R102.

Van Den Broeck, H.C., Maliepaard, C., Ebskamp, M.J.M., Toonen, M.a.J., and Koops, A.J. (2008). Differential expression of genes involved in C1 metabolism and lignin biosynthesis in wooden core and bast tissues of fibre hemp (*Cannabis sativa* L.). *Plant Sci.* 174**,** 205-220.

Van Hengel, A.J., Tadesse, Z., Immerzeel, P., Schols, H., Van Kammen, A., and De Vries, S.C. (2001). *N*-Acetylglucosamine and Glucosamine-Containing Arabinogalactan Proteins Control Somatic Embryogenesis. *Plant Physiol.* 125**,** 1880-1890.

Vanholme, B., Vanholme, R., Turumtay, H., Goeminne, G., Cesarino, I., Goubet, F., *et al.* (2014). Accumulation of *N*-Acetylglucosamine Oligomers in the Plant Cell Wall Affects Plant Architecture in a Dose-Dependent and Conditional Manner. *Plant Physiol.* 165**,** 290-308.

Wang, H., Jiang, C., Wang, C., Yang, Y., Yang, L., Gao, X., *et al.* (2015). Antisense expression of the fasciclin-like arabinogalactan protein *FLA6* gene in *Populus* inhibits expression of its homologous genes and alters stem biomechanics and cell wall composition in transgenic trees. *J. Exp. Bot.* 66**,** 1291-1302.

Yamauchi, S., Ichikawa, H., Nishiwaki, H., and Shuto, Y. (2015). Evaluation of Plant Growth Regulatory Activity of Furofuran Lignan Bearing a 7,9':7',9-Diepoxy Structure Using Optically Pure (+)- and (-)-Enantiomers. *J. Agric. Food Chem.* 63**,** 5224-5228.

Zang, L., Zheng, T., Chu, Y., Ding, C., Zhang, W., Huang, Q., *et al.* (2015). Genome-Wide Analysis of the Fasciclin-Like Arabinogalactan Protein Gene Family Reveals Differential Expression Patterns, Localization, and Salt Stress Response in *Populus*. *Front. Plant Sci.* 6-1140.

Zhao, Q., Zeng, Y., Yin, Y., Pu, Y., Jackson, L.A., Engle, N.L., *et al.* (2015). Pinoresinol reductase 1 impacts lignin distribution during secondary cell wall biosynthesis in *Arabidopsis*. *Phytochemistry* 112**,** 170-178.

Chapter 3: Highlights and Perspectives

This Chapter aimed at providing a molecular perspective of the genes involved in the differential cell wall composition of cortical and core tissues of adult hemp plant, with a special emphasis on *FLA*s. The cortical tissue carries the bast fibres, which are characterised by a thick gelatinous-type secondary cell wall. By contrast, xylem cells constituting the core tissue are richer in lignin and xylan. The analysis of internodes at different developmental stages complemented this approach. By doing so, we were able to shed light on some genes important for the characteristics of the hemp stem.

1. The contrasting lignin content between xylem and bast fibres is regulated at the gene expression level. Several genes of the monolignol pathway (*PAL*, *4CL1*, *CAD4*) and involved in lignin polymerisation (*LAC4*, *PRX49*) were more expressed in the core tissue.
2. Master transcription factors of the secondary cell wall (*NST1*, *NST2*) were up-regulated at the internode of the snap point downwards, together with *CesA4*, *CesA7* and *CesA8*. These genes were slightly more expressed in the core than in the cortex. They are therefore involved in the deposition of the secondary cell wall in both tissues. Further experiments are required to identify the CESAs depositing the cellulose of the gelatinous layer in the bast fibres.
3. The phylogeny of FLAs from hemp and those originating from species where those proteins have been characterised (thale cress, poplar and eucalypt) was investigated. Hemp FLAs are distributed within the four clades previously described in those species.

4. We identified different gene expression patterns within the twenty three *FLA* genes. Specific subsets have been attributed to elongating stems, stem undergoing secondary growth or maturating bast fibres, pointing to different functions and opening avenues for functional analyses of this yet enigmatic class of protein.
5. These specific functions may be linked to the presence of conserved motifs in their promoter region. In the *FLA*s which were more expressed in elongating tissues, we found a conserved motif recognised by the transcription factor SOC1 (probably involved in flowering and photoperiod perception). The *FLA*s more expressed in older tissues shared the presence of a conserved motif recognised by the transcription factor MYB3, which may repress the phenylpropanoid biosynthesis-related genes.

Key genes involved in stem growth and patterning have been analysed with a targeted method. In this respect, striking differences have been observed between bast fibre and xylem tissues. With the objective to understand more deeply the molecular mechanism of bast fibre development, from intrusive growth to gelatinous layer maturation, a RNA-Seq experiment carried out on enriched bast fibre fractions will be presented in Chapter 4. As in Chapter 1, we will pay a special attention to the genes related to cell wall (cellulose, hemicellulose and lignin), phytohormones (gibberellins, cytokinins, ethylene and jasmonic acid) and transcriptional regulation.

Chapter 4

Transcriptomic profiling of hemp bast fibres at different developmental stages

Partially adapted from the published article: Gea Guerriero, Marc Behr, Sylvain Legay, Lauralie Mangeot-Peter, Simone Zorzan, Mohammad Ghoniem, Jean-Francois Hausman. Transcriptomic profiling of hemp bast fibres at different developmental stages. 2017. *Scientific reports* 7:4961.

Author contributions:

Gea Guerriero conceived the idea of the study and designed the experiments. Marc Behr performed the RT-qPCR validation. Lauralie Mangeot-Peter grew the plants, collected the fibres and extracted the RNA. Gea Guerriero and Lauralie Mangeot-Peter performed library preparation, quantification and designed the primers for the RT-qPCR analysis. Gea Guerriero, Lauralie Mangeot-Peter, Sylvain Legay and Simone Zorzan carried out the bioinformatics analyses. All the authors interpreted the data and wrote the manuscript.

Abstract

Bast fibres are long extraxylary cells which mechanically support the phloem. They are divided into xylan- and gelatinous-type, depending on the composition of their secondary cell walls. The former, typical of jute/kenaf bast fibres, are characterized by the presence of xylan and a high degree of lignification, while the latter, found in tension wood, as well as flax, ramie and hemp bast fibres, have a high abundance of crystalline cellulose. During their differentiation, bast fibres undergo specific developmental stages: the cells initially elongate rapidly by intrusive growth, subsequently they cease elongation and start to thicken. The goal of the present study is to provide a transcriptomic close-up of the key events accompanying bast fibre development in textile hemp (*Cannabis sativa* L.), a fibre crop of great importance. Bast fibres have been sampled from different stem regions. The developmental stages corresponding to active elongation and cell wall thickening have been studied using RNA-Seq. The results show that the fibres sampled at each stem region are characterized by a specific transcriptomic signature and that the major changes in cell wall-related processes take place at the internode containing the snap point. The data generated also identify several interesting candidates for future functional analysis.

Introduction

Fibre crops are important bioresources as they provide strong and long fibres (up to 100 mm in some cases; Guerriero *et al.*, 2013), also known as bast fibres. These extraxylary cells belong to the sclerenchyma, they support mechanically the phloem and are differentiated into xylan- and gelatinous-type (Mikshina *et al.*, 2013). The cell walls of xylan-type fibres are lignified, contain predominantly xylan as hemicellulose and show a typical layered structure (S1–S3) because of the different orientation of the cellulose microfibrils (Mikshina *et al.*, 2013). The gelatinous fibres, typically found in hemp bast fibres, are characterized by a thick cellulosic cell wall (referred to as G-layer; Neutelings, 2011; Guerriero *et al.*, 2013). Bast fibre G-layer is reminiscent of the cell walls occurring in tension wood. However, the former does not exert the same contractile function as the latter (Mellerowicz *et al.*, 2008). Fibre crops like textile hemp or flax (*Cannabis sativa* L. and *Linum usitatissimum* L.) are very attractive models to carry out investigations on cell wall processes, because their stems are characterized by tissues displaying remarkable differences in cell wall composition. The cortical tissues, which can be easily peeled off and separated, harbour the cellulosic bast fibres and are characterized by the occurrence of low amounts of lignin (ca. 2–7%; Guerriero *et al.*, 2013). The core, also referred to as shivs or hurds, is instead woody. Along the stem axis it is possible to identify an empirically-determined region, called the “snap point” (Gorshkova *et al.*, 2003), which marks the transition from elongation to fibre thickening (and resulting in changes in fibre mechanical properties). The fibres in the younger regions of the stem (at the top) first grow symplastically with the surrounding tissues (Guerriero *et al.*, 2014; Snegireva *et al.*, 2015), then they start to elongate actively by a mechanism known as intrusive growth, where

the tip of the fibres invades the middle lamella of neighbouring cells (Gorshkova *et al.*, 2012). This growth mechanism ensures that the number of fibres in a given transverse section of the stem increases, without changing the total number of cells (Figure 2). This gradient of fibre developmental stages is accompanied by a basipetal lignification gradient in the stem tissues, where genes involved in the production of phenylpropanoids, and more generally in the provision of metabolic precursors needed for lignin synthesis, are expressed at higher levels. In this respect, in hemp it was recently shown that genes involved in the non-oxidative phase of the pentose phosphate pathway and in the first reaction of the shikimate pathway were expressed at higher levels in the core tissues at the bottom of the stem (Mangeot-Peter *et al.*, 2016). The stem of fibre crops is therefore ideal to carry out high throughput molecular analyses focusing on the cell wall, because its tissue polarity and spatial lignification gradient enable the study of sequential developmental stages. Several studies have indeed been published on fibre crops, namely flax (Roach and Deyholos, 2007; Gorshkov *et al.*, 2017), jute (Chakraborty *et al.*, 2015; Samanta *et al.*, 2015; Islam *et al.*, 2017), ramie (Chen *et al.*, 2014), kenaf (Li *et al.*, 2016) and hemp (de Pauw *et al.*, 2007; van den Broeck *et al.*, 2008), where a molecular approach was adopted to shed light on the mechanisms underlying bast fibre differentiation and development. These studies have identified important genes involved in bast fibre development, notably chitinases and cellulose synthases (Mokshina *et al.*, 2014), as well as transcription factors (Chakraborty *et al.*, 2015; Samanta *et al.*, 2015) and genes involved in secondary metabolism and monolignol-associated pathway (Chakraborty *et al.*, 2015; van den Broeck *et al.*, 2008). The advent of high-throughput techniques like transcriptomics has enabled huge steps forward in the study of fibre crops. For example, a very recent molecular study on flax, has shed

light on the molecular mechanisms underlying advanced phases of bast fibre development, by identifying several transcription factors, as well as glycosyltransferases and unknown/not fully annotated genes (Gorshkov *et al.*, 2017). Another recent study using transcriptomics/genomics has compared two varieties of jute differing in the cellulose/lignin fibre content and has demonstrated the expansion of lignin-biosynthetic genes with respect to flax (Islam *et al.*, 2017).

In the light of the industrial importance that gelatinous bast fibres are receiving, we investigated, via RNA-Seq, the molecular events accompanying their development in textile hemp. By sampling bast fibres from the top (TOP), middle (MID) and bottom (BOT) internodes of hemp stems (Figure 4-1), we show that the transcriptional signature at each stem region is unique. These results are useful to identify and characterize candidate genes involved in bast fibre elongation/thickening which can be further studied functionally and used for future biotechnological applications.

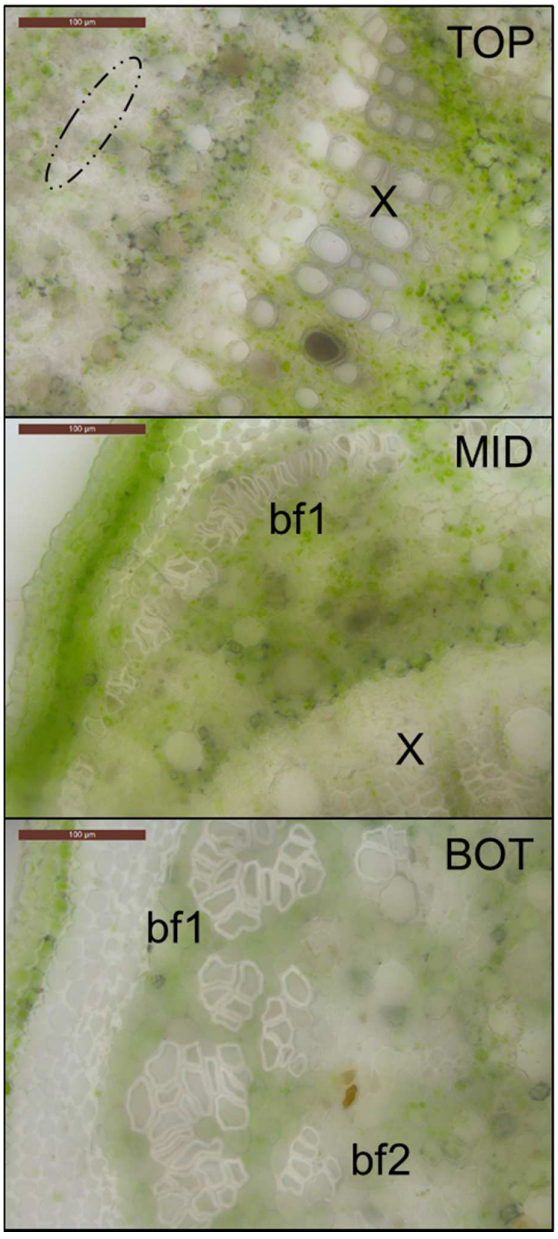


Figure 4-1. Optical microscope pictures of the different hemp stem regions (TOP, MID, BOT) analysed in this study. In the TOP, primary bast fibres are localised within the circle. bf1 primary bast fibres, bf2 secondary bast fibres, X xylem. Scale bars are 100 µm.

I. Material and Methods

a. Plant material, growth conditions and optical microscopy

A hemp monoecious fibre variety (*C. sativa* cv. Santhica 27) was studied in this work. After six weeks of growth in controlled chambers, samples were taken along three stem regions localized at different heights with respect to the snap point (determined empirically by gently tilting the stem apex until a kink could be observed). The top corresponds to the internode right below the apex, the middle (MID) is the internode containing the snap point and the bottom (BOT) is located three internodes below the middle sample (Figure 4-2).

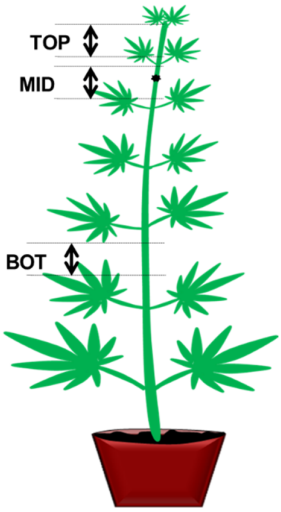


Figure 4-2: Sampling strategy.

Each sample is 2.5 cm long.

Black star at the MID: snap point

At the time of sampling, the plants had ca. 6–7 internodes. A segment of 2.5 cm was collected in the middle of each internode to avoid too much variation in gene expression, because of the varying developmental stages of the cell types. Fibres were separated from the shivs by peeling the cortical tissues and by quickly processing them with ethanol 80 % as described in Chapter 3. The number of biological

replicates is four with 13 plants in each replicate. Cross sections with the vibratome was performed as described in Chapter 3.

b. RNA extraction and preparation of the libraries

Total RNA was extracted using a modified CTAB extraction protocol combined with an RNeasy Plant Mini Kit (Qiagen) (Guerriero *et al.*, 2016). The RNA concentration and quality were measured by using a Nanodrop ND-1000 (Thermo Scientific) and a 2100 Bioanalyzer (Agilent), respectively. All the RNAs had a RIN value between 8 and 9. Libraries were prepared, quantified and their average size analysed as described in Chapter 1. The libraries were pooled at the concentration of 20 pM and sequenced on an Illumina MiSeq in five consecutive runs (MiSeq reagent kit V3, 150 cycles). Raw sequences have been deposited at the NCBI Gene Expression Omnibus (GEO) with the accession number GSE94156.

c. Processing of the reads, mapping and RNA-Seq analysis

The raw sequences obtained were uploaded in CLC Genomics Workbench 9.0.1. Sequences were filtered as follows: sequences > 35 bps, the sequence quality score was left as default value (0.05), the maximum number of ambiguities was set to 0. After two consecutive trimmings, the sequences reached an average length of 59 bps. We had previously published a *de novo* assembly for the variety Santhica 27 (Chapter 1) and proven its validity by comparing the results generated with our *de novo* assembly and with the Finola transcriptome (van Bakel *et al.*, 2011). We decided to merge the reads generated in this study with those previously obtained on the hemp hypocotyl (Chapter 1) to get a better assembly of the transcriptome of the variety under study. We therefore uploaded in CLC Genomics

Workbench 9.0.1 the reads obtained previously for the hypocotyls and those obtained in the present study for the fibres from adult plants. The parameters used are: wording size was set to 20, the bubble size to 50 and minimum contig length of 300. The reads were mapped back to the assembly with a mismatch, insertion and deletion cost of 3 (stringent criteria), and a length and similarity fraction of 0.95. The assembly was then annotated using Blast2GO PRO version 3.0 against the Viridiplantae and *A. thaliana* non-redundant database. The annotation against the *Arabidopsis* database was used for the subsequent Gene Ontology term Enrichment analysis (GOE) in Cytoscape (*vide infra*). For each library, the mapping was performed with a maximum hits per read of 3, a similarity and length fraction of 0.95, a mismatch, insertion and deletion cost of 3. Mapping was also performed using the transcriptome of the variety Finola (van Bakel *et al.*, 2011). The expression values were then calculated using the RPKM method (Mortazavi *et al.*, 2008).

The expression values were subjected to an ANOVA statistical test with three groups (TOP, MID, BOT), each composed of four biological replicates and, subsequently, to a false discovery rate (FDR) correction. Only the genes showing a corrected *p*-value < 0.05 were retained for downstream analysis. The obtained data were further filtered by removing those genes showing a maximum value of the means < 1 RPKM (this was done with the purpose of removing those contigs showing negligible changes in gene expression) and a maximum FC < 4 in absolute value. A total of 3268 contigs was obtained.

d. Primer design

Primers were designed using Primer3Plus and verified with the OligoAnalyzer 3.1 tool from Integrated DNA technologies. Primer

efficiencies were checked via qPCR using a serial dilution of cDNA (from 10 ng to 0.0032 ng/µl).

e. RNA-Seq validation with RT-qPCR

The RNA extracted using the above-mentioned protocol was retrotranscribed into cDNA using the ProtoScript II RTase (NEB) and random primers, according to the manufacturer's instructions. The cDNA was diluted to 2 ng/µl and 2 µl were used for the RT-qPCR analysis in 384-wells microplates (10 µl final volume). An automated liquid handling robot (epMotion 5073) was used to prepare the microplates. The expression of each target gene was normalized using two reference genes (*clathrin* and *F-Box*), after screening the twelve reference genes described previously (Mangeot-Peter *et al.*, 2016). To check the specificity of the amplicons, a melt curve analysis was performed. The expression of the genes was calculated using qBasePLUS by using the above-mentioned two reference genes.

f. Bioinformatic analysis

The annotation of the putative transcription factors (TFs) in the *de novo* assembly was carried out with PlantTFcat (Dai *et al.*, 2013), which gave a total of 2484 TFs. The ICA was performed with the on-line program MetaGeneAlyse as described in Chapter 2. The Gene Ontology term Enrichment analysis (GOE) was performed as previously described (Legay *et al.*, 2015) using Cytoscape (v3.4.0) with the ClueGO v2.3.2 plugin (Bindea *et al.*, 2009) (p-value < 0.05, Benjamini-Hochberg enrichment, gene ontology from level 3 to 8, kappa score set at 0.6). RNA-Seq RPKMs were log2 transformed and loaded for clustering and expression profile analysis in a data analytics software developed in-house. The software includes a Web-based user interface providing interactive data visualisation in the form of a parallel coordinates plot

synchronised with 2D scatter plots of PCA projections; the user interface is backed by an R server providing the necessary statistical analyses, in particular correlation clustering and PCA projection of multidimensional data. The software allowed us to configure, execute and visually analyse the RNA-Seq RPKMs.

II. Results and Discussion

a. RNA-Seq of hemp bast fibres

To analyse the transcriptional signature of hemp bast fibres at different developmental stages, RNA-Seq was performed on fibres sampled at three stem heights, i.e. top (TOP), middle (MID, containing the snap point) and bottom (BOT). Although fibres located at the top are difficult to separate because of the lack of a well-developed tertiary cell wall (G-layer), we carefully peeled the cortical tissues from the core and got rid of the majority of epidermal/parenchymatic and xylem cells by using a mortar with pestle and ethanol 80%. We reasoned that this procedure would enable us to minimize the “contamination” from the other tissues: inspection at the microscope showed that some non-glandular trichomes, parenchyma and xylem cells were still present, however those elements were sporadic.

A *de novo* assembly of the Santhica transcriptome was performed by merging the reads obtained from the hypocotyl with those obtained here for the bast fibres of adult plants: we reasoned that this would enable us to better capture the cell wall-related dynamism in the isolated phloem fibres of the variety at the developmental stages under investigation. We nevertheless validated the data by comparing the results obtained after mapping using our *de novo* assembly with those generated after mapping against the Finola transcriptome (van Bakel

et al., 2011). As discussed in the next paragraphs, the two mapping strategies gave comparable results.

A total of 3268 differentially expressed contigs ranging from 283 to 7095 bps was obtained after data processing; of these, 2317 are annotated (Supplementary data 4-1). The Independent Component Analysis (computed with the FastICA algorithm in MetaGeneAlyse) revealed a good separation of the different stem regions when two components were used (percentage of variance explained: 98.79%) (Figure 4-3). This indicates that the bast fibres sampled for the analysis were in different developmental stages and therefore characterized by distinct transcriptomic signatures.

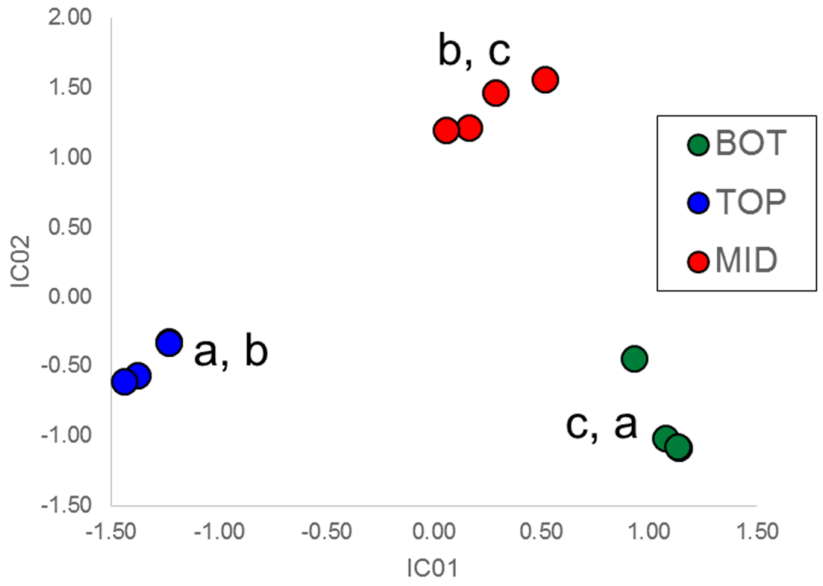


Figure 4-3: Independent component analysis (ICA) of transcriptomic profiles based on the RPKM of the 3268 differentially expressed contigs. The significance of the coordinates in the two IC was assessed using a Tukey post-hoc test (first letter for IC01, second letter for IC02), different letters indicate that the profiles are significantly different.

The RNA-Seq data were validated using targeted RT-qPCR on a subset of 12 genes: the calculated coefficient is 0.9284, which indicates a very good correlation between the RNA-Seq and RT-qPCR data.

To get insights into the gene expression patterns, data clustering was performed using an arbitrary Pearson correlation coefficient threshold of 0.75. The analysis resulted in eleven clusters (Figure 4-4); among these, cluster 1 and 2 account for the highest number of contigs (843 and 703 annotated contigs, respectively). The hemp contigs can be assigned to three major expression patterns: progressive decrease in expression from the top to the bottom of the stem (clusters 1, 3, 4), progressive increase in expression along the stem axis (clusters 2, 5, 7) and maximum expression at the internode containing the snap point (clusters 6, 9, 10). Two additional trends were revealed with the clustering analysis: cluster 11 groups contigs showing a tendency towards decreased expression at the snap point, cluster 8 comprises genes showing no major changes along the stem axis.

Since the bast fibres undergo progressive cell wall thickening from the top to the bottom of the hemp stem axis, we reasoned that the bulk of cell wall-related information would be obtained by focusing on the above-mentioned three major expression patterns (clusters 1-3-4, clusters 2-5-7, clusters 6-9-10).

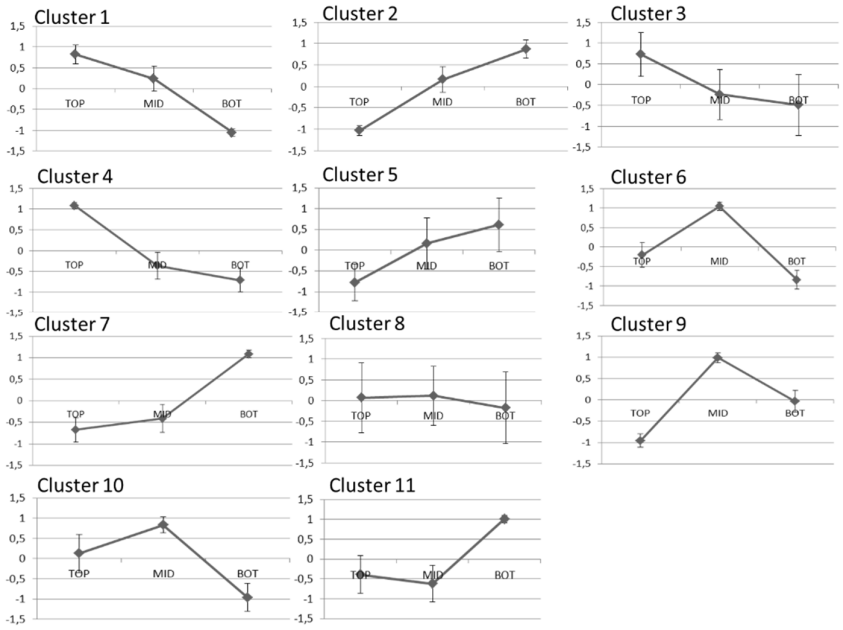


Figure 4-4: Profiles of the eleven clusters of genes obtained using a Pearson coefficient threshold of 0.75. The data represent the log_2RPKM rescaled values ± standard deviation (the rescaled values were calculated by subtracting to each contig expression value the average among the three stem regions and dividing by the standard deviation).

b. Gene ontology enrichment analysis of the TOP region

The transcriptomic landscape of the TOP region is dominated by genes belonging to the DNA replication and cell cycle ontologies (Figure 4-5), a result confirmed also by the analysis performed after mapping using the Finola transcriptome.

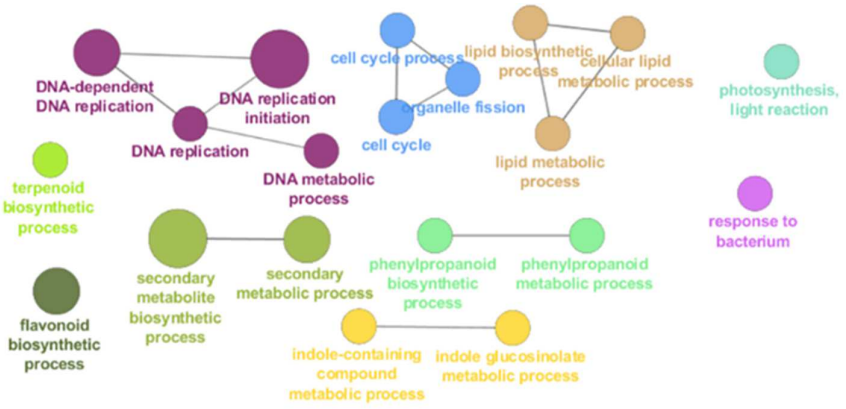


Figure 4-5: GOE analysis of the TOP region. Different colours indicate different ontologies. The bigger the circle, the higher the statistical significance.

These findings are indicative of active nuclear division and are in agreement with what was previously reported in flax bast fibres (Roach and Deyholos, 2007). Bast fibres are indeed multinucleate and during intrusive growth the number of nuclei increases, as previously documented in flax (Ageeva *et al.*, 2005). Compared to the MID and BOT respectively, *CDC6* (cell division control 6) was expressed ca. 2 and 10 times more, *DMC1* (encoding a meiotic recombination protein) was upregulated 2.4 and 7 times, *MCM2*, *4* and *5* (minichromosome maintenance protein 2, 4 and 5) were ca. 3 and 5 times more abundantly expressed, the *ORC1A* and *ORC6* genes (origin recognition complex subunit 1 and 6) were between 2.5 and 5‑7 times more expressed. Additionally, the key gene *PCNA2* (proliferating cell nuclear antigen 2) involved in DNA replication (Strzalka and Ziemienowisz, 2011) was highly expressed at the TOP (ca. 99 RPKM) with respect to the BOT (18 RPKM); finally, *DRT100* (coding for a DNA-damage-repair/toleration protein) was expressed ca. 12 times more at the TOP as compared to the BOT.

The enrichment of genes involved in cell division was accompanied by the over-representation of candidates partaking in the lipid biosynthetic process: several 3-ketoacyl-CoA synthase isoforms (*KCS10*, *11*, *19*) were highly expressed at the TOP, together with the β-ketoacyl reductase 1 (*KCR1*) and the fatty acid desaturases *FAD5* and *8*. The enrichment of these genes can be explained by the diffuse (intrusive) growth mechanism of bast fibres (Ageeva *et al.*, 2005): the elongation of fibres is indeed ensured by the growth of the entire surface of the cell. Notably, among the genes belonging to the lipid biosynthetic process, there are also homogentisate phytyltransferases (*HPT1*), which are involved in the biosynthesis of tocopherols. These are lipid antioxidants protecting against oxidative stress (e.g. light stress; Collakova *et al.*, 2003).

The growth of the cell surface during intrusive growth was coupled to the strong up-regulation (4355 RPKM at the TOP and 89 at the BOT) of a gene coding for a non-specific lipid transfer protein (*LTP3*). LTPs are cell wall proteins with diversified roles in plant growth and development (Chae *et al.*, 2010) and *LTP3* expression has been detected in the cambium, xylem and phloem tissues of *Arabidopsis* inflorescence stem (Wenzel *et al.*, 2008). In addition, several other *LTP*s were more expressed in the TOP. Several genes of this family are strongly expressed during *Arabidopsis* shoot development (Che *et al.*, 2002). Considering the expression profiles of *LTP*s in hemp, cotton (Deng *et al.*, 2016) and flax (Roach and Deyholos, 2007; Gorshkov *et al.*, 2017) and their high RPKM values, an important role of this gene family can be anticipated for (bast) fibre development. As highlighted in cotton, *LTP*s expressed during fibre elongation likely mediate the transport of phosphatidylinositol monophosphates (Deng *et al.*, 2016), polar lipids important for plasma membrane biosynthesis. A

glycosylphosphatidylinositol (GPI)-anchor can therefore be added to LTPs (LTPG). The cleavage of the LTPG GPI-anchor is performed by a GPI-specific phospholipase. In our dataset, a gene annotated as a patatin-related phospholipase A III-δ (*PLP9*) might be involved in this cleavage. Interestingly, a meta-analysis of RNA-Seq experiments performed on *Arabidopsis thaliana* (ATTED-II) has shown that *PLP9* is co-expressed with several fasciclin-like arabinogalactans (*FLA1*, *FLA7*, *FLA8*, *FLA15* and *FLA16*). As explained in Chapter 3, FLAs have a GPI-anchor. *PLP9* is 9-fold more expressed in the TOP than in BOT and *FLA1*, *FLA7* and *FLA16* are all more expressed in the TOP than in the BOT. After cleavage, LTPG could be released from the plasma membrane to the cell wall and fulfil a different function, such as deliver alkanes in the cuticle of stem epidermis (DeBono *et al.*, 2009). Such a GPI-predicted anchored LTP (annotated as *XYP2*) was more expressed in the TOP / MID as compared to the BOT (10-fold higher expression). XYP2 (xylogen protein 2) contains one LTP domain, one AGP domain and a GPI-anchor (Kobayashi *et al.*, 2011), together with a signal peptide (signal-P 4.1). It is co-expressed with *PLP9*, both in our dataset and in the ATTED-II co-expression network. *XYP2* is preferentially expressed in vascular tissues and has been involved in pattern formation of procambial tissues (Kobayashi *et al.*, 2011; Gujas and Rodriguez-Villalon, 2016). We suppose that XYP2 plays a role during hemp bast fibre elongation, as suggested for several AGPs (Chapter 3).

A transcription factor (TF) annotated as *IAA19* (indole-3-acetic acid inducible 19) was highly expressed at the TOP and MID (251 and 216 RPKM, respectively) with respect to BOT (60 RPKM). IAA19 mediates the auxin- and brassinosteroid-induced growth in the hypocotyl (Zhou *et al.*, 2013). Cell elongation is regulated by a complex

circuit involving auxin, brassinosteroids and helix-loop-helix (HLH), bHLH and IAA TFs (Figures 14 and 4-6). Importantly, several bHLH TFs belonging to this complex are up-regulated at the TOP. Among them, *KIDARI* mediates light-mediated cell elongation within the antagonising HLH / bHLH module (Zhang *et al.*, 2009) and was 13-fold more expressed at the TOP as compared to the BOT.

Genes involved in photosynthetic light reaction were likewise enriched at the TOP, notably *CAB1* (chlorophyll A/B binding protein 1), *LHCB4.3* and *LHCA5* (light harvesting complex), expressed between 5 and 8 times more as compared to the BOT. The TOP region was also characterized by genes involved in the response to pathogens: among them it is worth mentioning here *UVI4* (*UV-B-Insensitive 4*, encoding the protein POLYCHOME), whose overexpression was shown to induce increased resistance to a bacterial pathogen via the activation of disease resistance genes (Bao *et al.*, 2013). In our dataset, several pathogenesis-related (PR) genes were also upregulated at the TOP, namely *PRB1*, *PR-1-like*, *PR4*.

Interestingly, the Gene Ontology term Enrichment (GOE) analysis highlighted the enrichment of genes involved in secondary metabolism biosynthetic processes too, namely terpenoids, flavonoids, indole-containing compounds and phenylpropanoids (Figure 4-5). In the terpenoid biosynthetic process there are genes coding for three cytochrome P450s, i.e. *CYP76C1* and *C2* and *CYP82G1*. CYP76C1 and C2 are involved in floral linalool metabolism (Boachon *et al.*, 2015), while CYP82G1 is responsible for the synthesis of homoterpene volatiles in *Arabidopsis* (Lee *et al.*, 2010). Hemp trichomes are factories producing several terpenes contributing to the plant peculiar aroma (André *et al.*, 2016), however these metabolites can also be found in resin ducts. Since we observed the presence of non-glandular

trichomes in our fibre samples, we speculate that these genes may be expressed in "contaminating" resin ducts present in our bast fibre samples.

Three genes encoding UDP glucosyltransferases were present in the "flavonoid biosynthetic process" ontology: *UGT73C1* and *C7* and *UGT78D2*. The first two belong to group D, whose members are related to stress-inducible response (Langlois-Meurinne *et al.*, 2005); notably, UGT73C1 was shown to glucosylate cytokinins (*trans*-zeatin and dihydrozeatin; Hou *et al.*, 2004) and may therefore play a role in hormone homeostasis. UGT78D2 glucosylates instead flavonols in the C3 position and is strongly co-regulated with flavanone 3-hydroxylase (F3H, participating in the conversion of *p*-coumaroyl CoA to kaempferol/quercetin; Le Roy *et al.*, 2016), whose gene was also upregulated at the TOP in our dataset.

Two genes encoding cytochrome P450s, i.e. *CYP81D1* and *D8* and two MYB TFs, *MYB34* and *MYB122*, were present in the indole-containing compound metabolic process: *MYB34* and *MYB122* are two of the three TFs reported to control indole glucosinolate biosynthesis (Frerigmann and Gigolashvili, 2014). These TFs respond to phytohormones in different manners: MYB34 responds to both abscisic acid (ABA) and jasmonate (JA), while MYB122 plays a minor role in indole glucosinolate biosynthesis upon ethylene and JA signalling (Frerigmann and Gigolashvili, 2014). Given the role of indole glucosinolates in defence responses upon mechanical damage (e.g. herbivore attack; War *et al.*, 2012) and since the phytohormone JA is considered "the wound hormone" (Koo and Howe, 2009), it is tempting to speculate that a mechanism involving the synthesis of JA and the subsequent activation of indole glucosinolates may be involved in the intrusive growth phase of bast fibres. Among the TFs up-

regulated at the TOP, four of them are ethylene-responsive factors (*ERF3*, *ERF12*, *ERF34* and *CRF4*). Importantly, the expression of *ERF3* is induced by methyl jasmonate (Brown *et al.*, 2003) and by wounding or touching (Jeter *et al.*, 2004). ERFs also regulate jasmonate-responsive gene expression (Brown *et al.*, 2003). In addition to the roles of auxin and brassinosteroids in cell elongation, ethylene was also found to induce hypocotyl elongation through an auxin signalling route (De Grauwe *et al.*, 2005; Pierik *et al.*, 2009). Taken together, we may hypothesise that a cross-talk between auxin, brassinosteroids, ethylene and JA may play an important role in intrusive growth (Figure 4-6). The role of phytohormones in bast fibre intrusive growth remains however to be confirmed.

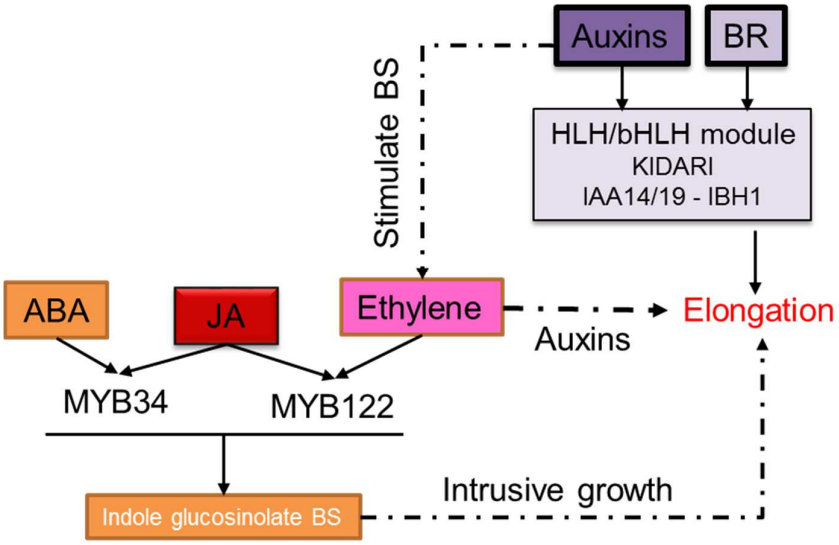


Figure 4-6: Hypothetical model of bast fibre elongation. Auxins stimulate the biosynthesis of ethylene (Abel *et al.*, 1995). BR brassinosteroids, BS biosynthesis. Other abbreviations as in the text.

We cannot exclude that the presence of genes belonging to the phenylpropanoid biosynthetic ontology (i.e. the laccases *LAC11* and

17, the cinnamyl alcohol dehydrogenase *CAD9*) may be (partly) due to the traces of contaminating xylem tissues during the separation of fibre-rich peels from the TOP stem region. The xylem tissue present is however in lower amount compared to the fibres; hence these transcripts may reflect an actual gene network linked to the phenylpropanoid metabolism in the TOP bast fibres. In this respect it should be noted that hemp bast fibres contain ca. 4% lignin (van den Broeck *et al.*, 2008) and that a previous study highlighted the presence of transcripts associated with the secondary metabolism (namely peroxidases, methyltransferases) in the bast fibres sampled at the TOP (de Pauw *et al.*, 2007).

Among the genes grouped in the "phenylpropanoid biosynthetic process" ontology, there are five contigs annotated as *PRR1* (pinoresinol reductase 1), whose expression was upregulated at the TOP. In flax, lignomics unveiled a complex monolignol metabolism associated with bast fibre hypolignification, with the accumulation of both aglycone forms and glycosides of (neo)lignans (Huis *et al.*, 2012); hence *PRR1* may be involved in a similar rich monolignol metabolism in hemp bast fibres. Additionally, some lignans, as dehydrodiconiferyl alcohol, are known to regulate cell division (Binns *et al.*, 1987); therefore PRR1 may also partake in the synthesis of specific lignans contributing to the regulation of cell division. Recently a transcriptome analysis performed on a wild-type and mutant jute (referred to as *deficient lignified phloem fibre*) identified the presence of many isoforms (more than other bast fibre-producing plants) belonging to the monolignol and shikimate-aromatic amino acid metabolism (Chakraborty *et al.*, 2015). In particular, the identification of several shikimate *O*-hydroxycinnamoyl transferases demonstrated that in jute fibres there is a shunting of the phenylpropanoid metabolism from *p*-

coumaroyl-CoA towards the production of monolignols, instead of flavonoids. Hemp bast fibres are of gelatinous-type and therefore the presence in the TOP bast fibres of transcripts belonging to both the flavonoid and phenylpropanoid pathways (Figure 4-5) may reflect a difference with respect to the xylan-type fibres present in jute.

Among the transcripts showing the highest FC at the TOP with respect to the MID, we would like to draw the reader's attention on two genes in particular, i.e. a jacalin-like lectin domain-containing protein orthologous to *AT3G16460* and a protodermal factor 1 (*PDF1*). In maize, two jacalin-like lectin domain-containing proteins were shown to be involved in cell wall-related processes (Bosch *et al.*, 2011); in cotton, a *PDF1* gene was found to be associated with fibre initiation and early growth phases (Deng *et al.*, 2012). The exact role of these genes remains to be elucidated via functional studies; however their expression pattern strongly points towards an involvement during hemp bast fibre early growth stages.

c. Gene ontology enrichment analysis of the MID region

The transcriptome of the MID region is mainly characterized by processes related to secondary cell wall biogenesis (Figure 4-7); this result was confirmed with the mapping against the Finola transcriptome.

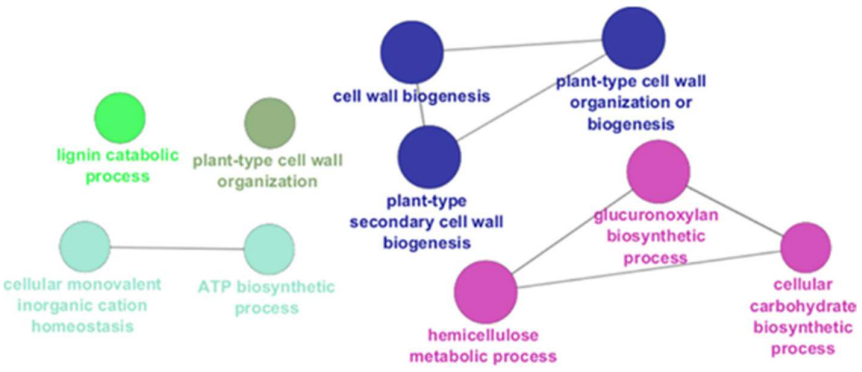


Figure 4-7: GOE analysis of the MID region. Different colours indicate different ontologies. The bigger the circle, the higher the statistical significance.

These data are of particular relevance because they provide an overview of the key cell wall-related events responsible for the change in mechanical properties of the bast fibres observed at the snap point. In this study the MID region corresponds to the internode containing the snap point (Figure 4-2); hence the fibres separated are progressively shifting from a stage of elongation to a phase of tertiary cell wall formation. According to our data, the transcriptome of the bast fibres at the MID region is characterized by genes involved in cell expansion and cell wall loosening, as well as by an active transcriptional dynamics of candidates involved in hemicellulose biosynthesis and transcriptional regulators orchestrating secondary cell wall biogenesis. Alpha expansin genes (*EXPA8*, *10*, *11* and *12*) and a xyloglucan endotransglycosylase/hydrolase (*XTH33*) peaked at the MID region (*EXPA10* and *11* showed an increase in FC > 5 with respect to the TOP); their higher expression is probably due to the phase of elongation characterizing the heterogeneous fibre stages at the MID. The master regulator *MYB46* and its downstream target *MYB63* were expressed 4.8 and 1.4 times more at the MID as

compared to the TOP (and 3.1 and 4.3 times more at the MID as compared to the BOT): these TFs activate genes involved in xylan and lignin biosynthesis and therefore the secondary cell wall biosynthetic program (Zhong and Ye, 2012). Three laccases, i.e. *LAC4* (*IRX1*; contig_10035), *LAC12* (contig_15462) and two contigs annotated as *LAC17* (contig_19910 and contig_17371) were also highly upregulated at the MID with respect to the TOP: these same genes were shown to be upregulated at older stages of development in the hemp hypocotyl (Chapter 1) and in the stem of adult hemp plants they may be associated with the peculiar lignification of the bast fibres. Bast fibres are hypolignified, however condensed guaiacyl (G) lignins were immunologically detected in the middle lamellas, cell wall junctions and S1 layer of phloem fibres in another fibre crop, flax (Day *et al.*, 2005). In hemp bast fibres the presence of a similar condensed lignin was reported (Crônier *et al.*, 2005); hence the laccases here detected may participate in the formation of condensed lignins in the cell wall of bast fibres.

In the "hemicellulose metabolic process" ontogeny there are genes involved in glucuronoxylan biosynthesis (the principal hemicellulose in dicots' secondary cell walls; Zhong and Ye, 2015). *IRX8*/*GAUT12* (irregular xylem 8/galacturonosyl transferase 12)*, IRX9, IRX10*/*GUT2* (glucuronosyl transferase 2) and *PGSIP3*/*GUX2* (plant glycogenin-like starch initiation protein 3/glucuronic acid substitution of xylan 2) were upregulated 3.2, 4.5, 4.8 and 5.4 times respectively at the MID as compared to the TOP. IRX8 belongs to the glycosyltransferase (GT) family 8 and is involved in the synthesis of the reducing end tetrasaccharide of glucuronoxylans, while IRX9 is a GT43 involved in xylan backbone elongation (Penã *et al.*, 2007); IRX10 is a GT47 also involved in the elongation of the xylan backbone (Brown *et*

al., 2009). GUX2 is a Golgi-localized GT8 with xylan glucuronosyltransferase activity. At the MID, the upregulation of two contigs, annotated as *ASD1*/*ARAF1* (which codes for a bifunctional α -L-arabinofuranosidase/β -D-xylosidase), was also observed (FC ca. 10 between MID and TOP). ARAF1 was shown to act *in vivo* on arabinan-containing pectins (Chávez Montes *et al.*, 2008) and its upregulation at the MID suggests cell wall remodelling associated with the deposition of secondary cell walls. Taking into account the upregulation of genes involved in xylan biosynthesis together with ARAF1 and considering the proof for the existence of a covalent link between hemicelluloses and pectins via an arabinogalactan protein (APAP1; Tan *et al.*, 2013), it is reasonable to hypothesize the presence of a similar association in the cell walls of bast fibres. ARAF1 may therefore participate in the remodelling of pectins at the onset of secondary cell wall deposition.

A contig annotated as *IRX1*/*CesA8* showed upregulation at the MID (FC MID *vs* TOP ca. 11): this result indicates that at the internode containing the snap point major molecular events related to cell wall biosynthesis take place with, notably, the upregulation of transcripts related to both cellulosic and non-cellulosic polysaccharide deposition.

Three genes coding for H+-ATPases (*HA3*, *5*, *6*) were upregulated at the MID: these genes may be associated with fibre elongation via the accumulation of osmotically-active compounds maintaining turgor pressure, or the acidification of the apoplast, in a manner analogous to the mechanism invoked for cotton fibre elongation (Ruan *et al.*, 2001).

Among the contigs encoding TFs upregulated at the MID there are *IBH1* (ILI1 binding bHLH 1 protein), with the highest FC increase

in expression as compared to the TOP (ca. 34), and *IAA14*. IBH1 is a negative regulator of cell elongation (Ikeda *et al.*, 2012) and its higher expression at the MID may indicate a function in the shift from elongation to thickening in bast fibres. IAA14 is a transcriptional repressor of auxin signalling (Liscum and Reed, 2002) and may attenuate the auxin-driven elongation in the MID segment.

Among the contigs showing the highest FC increase between MID and TOP there is a putative acid phosphatase (*AT1G04040*): notably, the corresponding protein in thale cress was found associated to the cell wall fraction (Bayer *et al.*, 2006). This phosphatase may be involved in cell wall-related processes in bast fibres during the transition from elongation to thickening and is therefore an interesting candidate for further analysis.

d. Gene ontology enrichment analysis of the BOT region

The BOT region is characterized by ontologies related to phytohormone and non-cellulosic polysaccharide biosynthesis, as well as secondary metabolic processes (Figure 4-8). Mapping against the Finola transcriptome confirmed these data and additionally showed the enrichment of auxin metabolism-related genes, among which, *WAT1* (*WALLS ARE THIN1*), an auxin efflux transporter involved in secondary cell wall deposition (Ranocha *et al.*, 2010; Ranocha *et al.*, 2013). In our assembly, two contigs annotated as a homeodomain TF (*BLH1*) were significantly up-regulated at the BOT as compared to the TOP (379 *vs* 87 and 83 *vs* 16 RPKM, respectively). Interestingly, *BLH1* is up-regulated in wood-forming *Arabidopsis* stem and has auxin responsive *cis*-elements in its promoter region (Ko *et al.*, 2004). Auxin

may consequently play a significant role during the thickening and the maturation of hemp bast fibres.

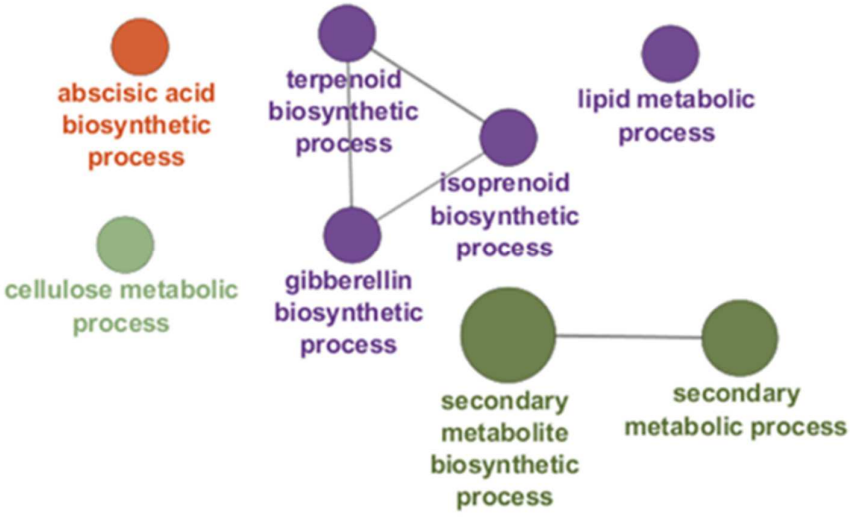


Figure 4-8: GOE analysis of the BOT region. Different colours indicate different ontologies. The bigger the circle, the higher the statistical significance.

In the “gibberellin biosynthetic process” ontogeny there are genes involved in the three steps required for gibberellin (GA) synthesis, i.e. formation of *ent*-kaurene from geranyl geranyl diphosphate, conversion of *ent*-kaurene to GA12 and synthesis of C20- and C19-Gas (Sun, 2008). *GA1*/*CPS* (*ent*-copalyl diphosphate synthase) involved in the formation of *ent*-kaurene was expressed at the BOT > 7 times with respect to the TOP; *KAO2* (*ent*-kaurenoic acid oxidase 2), which codes for a CYP88A member catalysing the conversion of *ent*-kaurenoic acid to GA12 (Helliwell *et al.*, 2001), showed an increase in FC > 4 as compared to the TOP; the genes encoding 2-oxoglutarate-dependent dioxygenases *GA20OX2* and *GA3OX1* showed a FC > 4 as compared to the TOP. The expression

of *GA2OX2* also peaked at the BOT (expressed 12-times more as compared to the TOP): this gene encodes a gibberellin 2-β-dioxygenase responsible for the conversion of GA to an inactive form and participates in GA homeostasis. In addition to that, a contig annotated as a CYP714A1 (contig_19239), which participates in GA deactivation (Zhang *et al.*, 2011), was also expressed at higher levels at the BOT with respect to the TOP (FC > 7). In the BOT region of the hemp stem, characterized by bast fibres with a thick tertiary cell wall, elongation has ceased and the homeostasis of GA levels by *GA2OX2* may represent an important regulatory mechanism. Transgenic tobacco plants in which the gene *GA2OX2* had been silenced were taller than *GA20OX1* overexpressors (the GA 20-oxidase catalyses the rate limiting step in GA biosynthesis); these results therefore show that deactivation is the key factor in the maintenance of GA homeostasis (Dayan *et al.*, 2010). Concomitantly with the activation of *GA2OX2*, genes involved in abscisic acid (ABA) biosynthesis increased in expression at the BOT. ABA inhibits stem elongation and in our previous study it was found in higher abundance in the hemp hypocotyl at older developmental stages (Chapter 1). In our dataset, the ABA biosynthetic genes *NCED3* and *5* were both upregulated at the BOT (FC > 5 and >11 respectively, as compared to the TOP); the gene *XERICO* encoding a RING-H2 zinc finger protein and involved in ABA homeostasis (Ko *et al.*, 2006) was also upregulated at the BOT (FC > 4 with respect to the TOP). Additionally, three contigs annotated as CYP704A2, which was proposed to be a candidate for ABA 8′-hydroxylation (i.e. catabolism; Kushiro *et al.*, 2004), were expressed at higher levels at the BOT (FC BOT *vs* TOP between 28 and 135).

A contig annotated as a CYP82C2 was upregulated at the BOT (FC difference between BOT and TOP > 16): the *A. thaliana* ortholog

is involved in the modulation of tryptophan-derived secondary metabolites under conditions of high JA levels (Liu *et al.*, 2010).

Genes encoding putative cytochrome P450 94 members (CYP94D2 and B2) were also upregulated at the BOT (FC BOT *vs* TOP between 4.6 and 38.9): the turnover of jasmonoyl-L-isoleucine (the major bioactive form of JA) is regulated by members of this class (Koo *et al.*, 2012). Our results show that a molecular control over the levels GA, ABA, JA is present at the BOT of adult hemp stems.

In the "cellulose metabolic process" ontogeny it is here worth mentioning the cellulose synthase-like genes *CSLG1* and *G3* (FC > 12 and >5, respectively, as compared to the TOP); recently, in fibres of flax with tertiary cell walls, the upregulation of a *CSLG* member was observed and it was proposed that the corresponding enzyme may catalyse the synthesis of β-1,4-galactans (Gorshkov *et al.*, 2017). Although the LM5 antibody recognizing β-1,4-galactans does not label hemp bast fibres (Chapter 1), it may be possible that pectic galactans in hemp fibres are masked and therefore do not react with the LM5 antibody. Similarly to flax (Gorshkov *et al.*, 2017), several LTPs were up-regulated during the maturation of hemp bast fibres. Among them, LTP4 shows a signal peptide (signalP 4.1), so may be secreted to the cell wall. *LTP3* and *LTP4* have opposite expression trends, suggesting that *LTP*s are expressed during specific stages of cell wall development. In fully maturated epidermal cells of thale cress, LTPG shows a fibrous cellulose microfibril-like pattern which resembles the gelatinous wall of bast fibres (Ambrose *et al.*, 2013). This indicates that LTPs, in addition to their role during intrusive growth, may also be involved in the latest stages of bast fibre development.

Genes involved in lignin biosynthesis, i.e. *PAL1* (phenylalanine ammonia lyase 1) and one contig annotated as *OMT1* (caffeic acid/5-hydroxyferulic acid *O*-methyltransferase; contig_2172) showed, respectively, a FC > 11 and >9 as compared to the TOP. Likewise, one contig encoding IRX12/LAC4 (contig_13683) and a cytochrome P450 involved in lignification (CYP71A20; van de Mortel *et al.*, 2008) showed a FC increase in expression > 28 and >35 at the BOT with respect to the TOP. A β-glucosidase with an activity on monolignol glucosides (*BGLU45*; Chapelle *et al.*, 2012) was 12-fold more expressed in the BOT as compared to the TOP. In addition, the UDP-glycosyltransferase 72B1, whose knock-out mutant is characterised by a higher lignin content (Lin *et al.*, 2016), was 11-fold more expressed in the TOP than in the BOT. UGT72B1 catalyses the glycosylation of monolignols (Lin *et al.*, 2016). From the expression profiles of these two genes, it is tempting to speculate that the balance between monolignols and their glucosides is an important parameter regulating lignification. As suggested by several recent experiments, monolignol glucosides are not used to synthesise lignin (Le Roy *et al.*, 2016). As shown in flax for the lignan secoisolariciresinol diglucoside, they may be protectant molecules against reactive oxygen species, viral, fungal and bacterial diseases or abiotic stresses (Le Roy *et al.*, 2016). For instance, the glycoconjugate scopolin is converted into aglycone scopoletin which shows antiviral activity (Le Roy *et al.*, 2016). The pool of monolignol glucosides may also cause a global feedback regulation on lignin biosynthesis (Lin *et al.*, 2016), leading to the lignification of the bast fibres in the BOT. The fibres in the basal region show a positive signal at the level of the middle lamella after phloroglucinol staining (Crônier *et al.*, 2005).

As in flax, the bast fibres of hemp contain minor amount of lignin, especially as compared with xylem cells. In the Chapter 3, we showed that some genes involved in monolignol biosynthesis (*PAL*, *4CL1* and *CAD4*) and lignin polymerisation (*LAC4* and *PRX49*) were less expressed in the cortical tissue than in the core. As suggested by its expression pattern, it is possible that *MYB4* participates in the regulation of this tissue-specific regulation, since it is a negative regulator of lignification (Huis *et al.*, 2012; Zhou *et al.*, 2017). We have also demonstrated in Chapter 3 that several fasciclin-like arabinogalactans that are more expressed at the BOT share a motif putatively recognized by MYB3 (belonging to the same R2R3 MYB subgroup S4 as MYB4), pointing to a possible function of MYB3 / MYB4 in hemp bast fibre hypolignification. A second gene regulatory network may explain this pattern. The expression of the homeobox (HD) protein knotted-1-like 3 (*KNAT3*) peaked at the BOT (4.4-fold more expressed than at the TOP). KNAT3 is a class II TALE (three amino acid loop extension) homeodomain from the KNOX protein family (Hake *et al.*, 2004). KNOX proteins function as heterodimers with BELL proteins co-factors, which are another class of TALE proteins. BLH1 is a BELL protein interacting with KNAT3 (Hay and Tsiantis, 2010) and is more expressed at the BOT. In addition, the genes encoding BEL1 and BLH2, which are in the same BELL phylum as BLH1 (Hake *et al.*, 2004), show the same expression pattern (11-fold and 8-fold more expressed in the BOT *vs* the TOP, respectively). Importantly, the KNOX-BELL heterodimers repress the expression of some lignin biosynthetic genes (*COMT1*, *CCoAOMT* and *AtPRX2*) and modulate the abundance of gibberellins and cytokinins (Hay and Tsiantis, 2010). This process activates cell division and represses cell differentiation during the leaf primary morphogenesis to allow leaflet initiation (Hay and Tsiantis, 2010). Considering the co-expression of KNAT3, BEL1,

BLH1 and BLH2, we may anticipate that the interaction between these proteins and the downstream consequences on lignin-related gene expression play an important role for the hypolignification of hemp bast fibres (Figure 4-9).

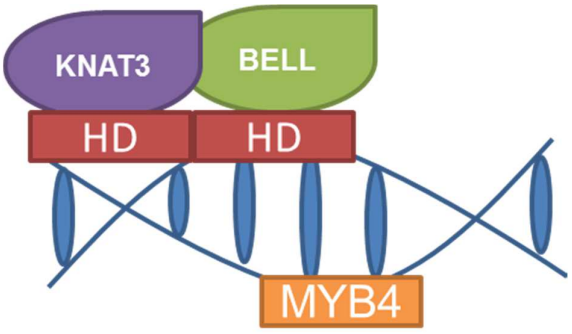


Figure 4-9: Suggested genetic mechanism of hypolignification in hemp bast fibres. KNOX protein (here KNAT3) forms a heterodimer with a BELL protein (BEL1, BLH1 or BLH2). This heterodimer or MYB4 may prevent the expression of lignin-related genes by binding to DNA.

Secondary growth (cambial activity) and cell wall thickening (G-layer deposition) are two major events occurring at the BOT. In this framework, it is worth mentioning the genes encoding TFs upregulated at this developmental stage. A LOB-domain containing protein 4 (*LBD4*) showed a BOT *vs* TOP increase in expression of ca. 218. LOBs are key regulators of plant organ polarity which play also a role in plant secondary growth (Xu *et al.*, 2016); in *Populus*, LBD1 positively regulates secondary phloem development in the stem (Yordanov *et al.*, 2010). Consequently, LBDs represent interesting targets in phloem fibre development. A contig annotated as *NAC047 / SPEEDY HYPONASTIC GROWTH* (*SHYG*) peaked at the BOT (BOT *vs* TOP FC > 78): the ortholog in thale cress was shown to be associated with hyponastic leaf movements upon waterlogging via modulation of an

enzyme involved in ethylene biosynthesis, i.e. ACC OXIDASE5 (Rauf *et al.*, 2013). Interestingly SHYG acts on expansin and xyloglucan endotransglycosylase/hydrolase genes to induce expansion of the petiole abaxial side. Its role in bast fibre thickening is yet to be unveiled, but it may entail a suite of cell wall-related events. Ethylene may play a role during fibre thickening: several genes encoding ethylene-responsive genes or TFs were indeed upregulated at the BOT, notably *EIN3*, *ERF1*, *WRI1* and *ERF106*. Similarly, several genes involved in ethylene biosynthesis and signalling (*EIN3*) were significantly up-regulated in *Populus* tension wood (Andersson-Gunneråsa *et al.*, 2006). A gene encoding a Zim17-type zinc finger protein (AT1G68730) was more expressed at the BOT (118 *vs* 18 RPKM at the TOP); this gene is also highly expressed in tension wood of poplar (Andersson-Gunneråsa *et al.*, 2006). As in hemp, fibres with thick gelatinous cell wall are found in tension wood. These TFs may therefore play a conserved role in the formation of the gelatinous-type secondary cell wall.

Several contigs coding for HB, MYB, NAC, bHLH, bZIP and WRKY TFs were found enriched in the bast fibres at the BOT. Among the 62 TFs upregulated at the BOT, eleven of them belong to the Homeobox family. Two Zinc finger proteins (*ZFP4* and *ZAT5*) were also more expressed in this region. Importantly, one HB and one ZFP TF were found to be differentially expressed between wild type jute and a bast fibre-deficient mutant (Samanta *et al.*, 2015). Additionally, in jute, bHLH, MYB-related, WRKY and NAC TFs were among the eight most abundant families (Chakraborty *et al.*, 2015). These TFs may therefore have overlapping functions in the development of both xylan- and gelatinous-type bast fibres.

A contig encoding a RADIALIS-LIKE 6 TF was upregulated at the BOT ca. 18 fold with respect to the TOP: very recently a RADIALIS-

LIKE TF was found upregulated in flax bast fibres and a role in phloem fibre development was therefore suggested (Gorshkov *et al.*, 2017). RADIALIS-LIKE TFs play a role in organ (floral) symmetry; hence, together with LBDs, they may regulate aspects related with positional cues. Three contigs coding for a PLATZ TF also peaked at the BOT (BOT *vs* TOP FC of 8, 12 and 20, respectively): PLATZ TFs were proposed to act as negative regulators of cell proliferation (Nagano *et al.*, 2001) and may contribute to the transition of bast fibres from a phase of active division to secondary growth.

III. Conclusions

We have here provided a transcriptional fingerprinting of bast fibres from textile hemp sampled at different stem heights, which correspond to different developmental stages. We have discussed the results using, predominantly, a cell wall angle, since the development of phloem fibres requires major modifications of the cell wall. Our study has shown that each region of the stem is characterized by distinct gene expression profiles. Young stem regions are dominated by cell cycle- and photosynthesis-related genes, together with candidates involved in the biosynthesis of specific secondary metabolites, notably indole-containing compounds and oligolignols; older internodes show enrichment of phytohormone-related genes, together with genes involved in non-cellulosic polysaccharides deposition and lignification. According to our results, the bulk of cell wall-related gene dynamics in hemp bast fibres is localized at the internode containing the snap point, where the fibres shift from a phase of elongation to thickening. The data here shown contribute to the understanding of the molecular events accompanying hemp bast fibre development and identify several genes deserving further functional study.

References

Abel, S., Nguyen, M.D., Chow, W., and Theologis, A. (1995). ASC4, a Primary Indoleacetic Acid-responsive Gene Encoding 1-Aminocyclopropane-1-carboxylate Synthase in *Arabidopsis thaliana*: STRUCTURAL CHARACTERIZATION, EXPRESSION IN *ESCHERICHIA COLI,* AND EXPRESSION CHARACTERISTICS IN RESPONSE TO AUXIN. *J. Biol. Chem.* 270**,** 19093-19099.

Ageeva, M.V., Petrovská, B., Kieft, H., Sal'nikov, V.V., Snegireva, A.V., Van Dam, J.E.G., *et al.* (2005). Intrusive growth of flax phloem fibers is of intercalary type. *Planta* 222**,** 565-574.

Ambrose, C., Debono, A., and Wasteneys, G. (2013). Cell Geometry Guides the Dynamic Targeting of Apoplastic GPI-Linked Lipid Transfer Protein to Cell Wall Elements and Cell Borders in *Arabidopsis thaliana*. *PLoS One* 8**,** e81215.

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

Andre, C.M., Hausman, J.F., and Guerriero, G. (2016). *Cannabis sativa*: The plant of the thousand and one molecules. *Front. Plant Sci.* 7.

Bao, Z., Yang, H., and Hua, J. (2013). Perturbation of cell cycle regulation triggers plant immune response via activation of disease resistance genes. *Proc. Natl. Acad. Sci. U. S. A.* 110**,** 2407-2412.

Bayer, E.M., Bottrill, A.R., Walshaw, J., Vigouroux, M., Naldrett, M.J., Thomas, C.L., *et al.* (2006). *Arabidopsis* cell wall proteome defined using multidimensional protein identification technology. *Proteomics* 6**,** 301-311.

Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., *et al.* (2009). ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 25**,** 1091-1093.

Binns, A.N., Chen, R.H., Wood, H.N., and Lynn, D.G. (1987). Cell division promoting activity of naturally occurring dehydrodiconiferyl glucosides: Do cell wall components control cell division? *Proc. Natl. Acad. Sci. U. S. A.* 84**,** 980-984.

Boachon, B., Junker, R.R., Miesch, L., Bassard, J.E., Höfer, R., Caillieaudeaux, R., *et al.* (2015). CYP76C1 (Cytochrome P450)-mediated linalool metabolism and the formation of volatile and soluble linalool oxides in arabidopsis flowers: A strategy for defense against floral antagonists. *Plant Cell* 27**,** 2972-2990.

Bosch, M., Mayer, C.D., Cookson, A., and Donnison, I.S. (2011). Identification of genes involved in cell wall biogenesis in grasses by differential gene expression profiling of elongating and non-elongating maize internodes. *J. Exp. Bot.* 62**,** 3545-3561.

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

Brown, R.L., Kazan, K., Mcgrath, K.C., Maclean, D.J., and Manners, J.M. (2003). A Role for the GCC-Box in Jasmonate-Mediated Activation of the *PDF1.2* Gene of *Arabidopsis*. *Plant Physiol.* 132**,** 1020-1032.

Chae, K., Gonong, B.J., Kim, S.-C., Kieslich, C.A., Morikis, D., Balasubramanian, S., *et al.* (2010). A multifaceted study of stigma/style cysteine-rich adhesin (SCA)-like *Arabidopsis* lipid transfer proteins (LTPs)

suggests diversified roles for these LTPs in plant growth and reproduction. *J. Exp. Bot.* 61**,** 4277-4290.

Chakraborty, A., Sarkar, D., Satya, P., Karmakar, P.G., and Singh, N.K. (2015). Pathways associated with lignin biosynthesis in lignomaniac jute fibres. *Mol. Genet. Genomics* 290**,** 1523-1542.

Chapelle, A., Morreel, K., Vanholme, R., Le-Bris, P., Morin, H., Lapierre, C., *et al.* (2012). Impact of the Absence of Stem-Specific β-Glucosidases on Lignin and Monolignols. *Plant Physiol.* 160**,** 1204-1217.

Chávez Montes, R.A., Ranocha, P., Martinez, Y., Minic, Z., Jouanin, L., Marquis, M., *et al.* (2008). Cell wall modifications in *Arabidopsis* plants with altered α-l-arabinofuranosidase activity. *Plant Physiol.* 147**,** 63-67.

Che, P., Gingerich, D.J., Lall, S., and Howell, S.H. (2002). Global and Hormone-Induced Gene Expression Changes during Shoot Development in *Arabidopsis*. *Plant Cell* 14**,** 2771-2785.

Chen, J., Pei, Z., Dai, L., Wang, B., Liu, L., An, X., *et al.* (2014). Transcriptome profiling using pyrosequencing shows genes associated with bast fiber development in ramie (*Boehmeria nivea* L.). *BMC Genomics* 15.

Collakova, E., and Dellapenna, D. (2003). The Role of Homogentisate Phytyltransferase and Other Tocopherol Pathway Enzymes in the Regulation of Tocopherol Synthesis during Abiotic Stress. *Plant Physiol.* 133**,** 930-940.

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.

Dai, X., Sinharoy, S., Udvardi, M., and Zhao, P.X. (2013). PlantTFcat: An online plant transcription factor and transcriptional regulator categorization

and analysis tool. *BMC Bioinformatics* 14.

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

Dayan, J., Schwarzkopf, M., Avni, A., and Aloni, R. (2010). Enhancing plant growth and fiber production by silencing *GA 2-oxidase*. *Plant Biotechnol. J.* 8**,** 425-435.

De Grauwe, L., Vandenbussche, F., Tietz, O., Palme, K., and Van Der Straeten, D. (2005). Auxin, Ethylene and Brassinosteroids: Tripartite Control of Growth in the *Arabidopsis* Hypocotyl. *Plant Cell Physiol.* 46**,** 827-836.

De Pauw, M.A., Vidmar, J.J., Collins, J., Bennett, R.A., and Deyholos, M.K. (2007). Microarray analysis of bast fibre producing tissues of *Cannabis sativa* identifies transcripts associated with conserved and specialised processes of secondary wall development. *Funct. Plant Biol.* 34**,** 737-749.

DeBono, A., Yeats, T.H., Rose, J.K.C., Bird, D., Jetter, R., Kunst, L., *et al.* (2009). *Arabidopsis* LTPG Is a Glycosylphosphatidylinositol-Anchored Lipid Transfer Protein Required for Export of Lipids to the Plant Surface. *Plant Cell* 21**,** 1230-1238.

Deng, F., Tu, L., Tan, J., Li, Y., Nie, Y., and Zhang, X. (2012). GbPDF1 is involved in cotton fiber initiation via the core cis-element HDZIP2ATATHB2. *Plant Physiol.* 158**,** 890-904.

Deng, T., Yao, H., Wang, J., Wang, J., Xue, H., and Zuo, K. (2016). GhLTPG1, a cotton GPI-anchored lipid transfer protein, regulates the transport of phosphatidylinositol monophosphates and cotton fiber elongation. *Sci. Rep.* 6**,** 26829.

Frerigmann, H., and Gigolashvili, T. (2014). MYB34, MYB51, and MYB122 distinctly regulate indolic glucosinolate biosynthesis in *Arabidopsis thaliana*. *Mol. Plant* 7**,** 814-828.

Gorshkov, O., Mokshina, N., Gorshkov, V., Chemikosova, S., Gogolev, Y., and Gorshkova, T. (2017). Transcriptome portrait of cellulose-enriched flax fibres at advanced stage of specialization. *Plant Mol. Biol.* 93**,** 431-449.

Gorshkova, T., Brutch, N., Chabbert, B., Deyholos, M., Hayashi, T., Lev-Yadun, S., *et al.* (2012). Plant Fiber Formation: State of the Art, Recent and Expected Progress, and Open Questions. *Crit. Rev. Plant Sci.* 31**,** 201-228.

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

Guerriero, G., Hausman, J.F., and Cai, G. (2014). No stress! relax! mechanisms governing growth and shape in plant cells. *Int. J. Mol. Sci.* 15**,** 5094-5114.

Guerriero, G., Mangeot-Peter, L., Hausman, J.F., and Legay, S. (2016). Extraction of high quality RNA from *Cannabis sativa* bast fibres: A Vademecum for Molecular Biologists. *Fibers* 4**,** 23.

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.

Gujas, B., and Rodriguez-Villalon, A. (2016). Plant Phosphoglycerolipids: The Gatekeepers of Vascular Cell Differentiation. *Front. Plant Sci.* 7**,** 103.

Hake, S., Smith, H.M.S., Holtan, H., Magnani, E., Mele, G., and Ramirez, J.

(2004). THE ROLE OF KNOX GENES IN PLANT DEVELOPMENT. *Annu. Rev. Cell Dev. Biol.* 20**,** 125-151.

Hay, A., and Tsiantis, M. (2010). KNOX genes: versatile regulators of plant development and diversity. *Development* 137**,** 3153-3165.

Helliwell, C.A., Chandler, P.M., Poole, A., Dennis, E.S., and Peacock, W.J. (2001). The CYP88A cytochrome P450, ent-kaurenoic acid oxidase, catalyzes three steps of the gibberellin biosynthesis pathway. *Proc. Natl. Acad. Sci. U. S. A.* 98**,** 2065-2070.

Hou, B., Lim, E.K., Higgins, G.S., and Bowles, D.J. (2004). *N*-glucosylation of cytokinins by glycosyltransferases of *Arabidopsis thaliana*. *J. Biol. Chem.* 279**,** 47822-47832.

Huis, R., Morreel, K., Fliniaux, O., Lucau-Danila, A., Fénart, S., Grec, S., *et al.* (2012). Natural hypolignification is associated with extensive oligolignol accumulation in flax stems. *Plant Physiol.* 158**,** 1893-1915.

Ikeda, M., Fujiwara, S., Mitsuda, N., and Ohme-Takagi, M. (2012). A triantagonistic basic helix-loop-helix system regulates cell elongation in *Arabidopsis*. *Plant Cell* 24**,** 4483-4497.

Islam, M.S., Saito, J.A., Emdad, E.M., Ahmed, B., Islam, M.M., Halim, A., *et al.* (2017). Comparative genomics of two jute species and insight into fibre biogenesis. *Nature Plants* 3.

Jeter, C.R., Tang, W., Henaff, E., Butterfield, T., and Roux, S.J. (2004). Evidence of a Novel Cell Signaling Role for Extracellular Adenosine Triphosphates and Diphosphates in *Arabidopsis*. *Plant Cell* 16**,** 2652-2664.

Ko, J.H., Yang, S.H., and Han, K.H. (2006). Upregulation of an *Arabidopsis* RING-H2 gene, *XERICO*, confers drought tolerance through increased

abscisic acid biosynthesis. *Plant J.* 47**,** 343-355.

Ko, J.-H., Han, K.-H., Park, S., and Yang, J. (2004). Plant Body Weight-Induced Secondary Growth in *Arabidopsis* and Its Transcription Phenotype Revealed by Whole-Transcriptome Profiling. *Plant Physiol.* 135**,** 1069-1083.

Kobayashi, Y., Motose, H., Iwamoto, K., and Fukuda, H. (2011). Expression and Genome-Wide Analysis of the Xylogen-Type Gene Family. *Plant Cell Physiol.* 52**,** 1095-1106.

Koo, A.J.K., and Howe, G.A. (2009). The wound hormone jasmonate. *Phytochemistry* 70**,** 1571-1580.

Koo, A.J.K., and Howe, G.A. (2012). Catabolism and deactivation of the lipid-derived hormone jasmonoyl-isoleucine. *Front. Plant Sci.* 3.

Kushiro, T., Okamoto, M., Nakabayashi, K., Yamagishi, K., Kitamura, S., Asami, T., *et al.* (2004). The *Arabidopsis* cytochrome P450 CYP707A encodes ABA 8′-hydroxylases: Key enzymes in ABA catabolism. *EMBO J.* 23**,** 1647-1656.

Langlois-Meurinne, M., Gachon, C.M.M., and Saindrenan, P. (2005). Pathogen-responsive expression of glycosyltransferase genes *UGT73B3* and *UGT73B5* is necessary for resistance to *Pseudomonas syringae* pv tomato in *Arabidopsis*. *Plant Physiol.* 139**,** 1890-1901.

Le Roy, J., Huss, B., Creach, A., Hawkins, S., and Neutelings, G. (2016). Glycosylation is a major regulator of phenylpropanoid availability and biological activity in plants. *Front. Plant Sci.* 7.

Lee, S., Badieyan, S., Bevan, D.R., Herde, M., Gatz, C., and Tholl, D. (2010). Herbivore-induced and floral homoterpene volatiles are biosynthesized by a single P450 enzyme (CYP82G1) in *Arabidopsis*. *Proc. Natl. Acad. Sci. U. S.*

A. 107**,** 21205-21210.

Legay, S., Guerriero, G., Deleruelle, A., Lateur, M., Evers, D., André, C.M., *et al.* (2015). Apple russeting as seen through the RNA-seq lens: strong alterations in the exocarp cell wall. *Plant Mol. Biol.* 88**,** 21-40.

Li, H., Li, D., Chen, A., Tang, H., Li, J., and Huang, S. (2016). Characterization of the kenaf (*Hibiscus cannabinus*) global transcriptome using Illumina paired-end sequencing and development of EST-SSR markers. *PLoS One* 11.

Lin, J.-S., Huang, X.-X., Li, Q., Cao, Y., Bao, Y., Meng, X.-F., *et al.* (2016). UDP-glycosyltransferase 72B1 catalyzes the glucose conjugation of monolignols and is essential for the normal cell wall lignification in *Arabidopsis thaliana*. *Plant J.* 88**,** 26-42.

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

Liu, F., Jiang, H., Ye, S., Chen, W.P., Liang, W., Xu, Y., *et al.* (2010). The *Arabidopsis* P450 protein CYP82C2 modulates jasmonate-induced root growth inhibition, defense gene expression and indole glucosinolate biosynthesis. *Cell Res.* 20**,** 539-552.

Mangeot-Peter, L., Legay, S., Hausman, J.F., Esposito, S., and Guerriero, G. (2016). Identification of reference genes for RT-qPCR data normalization in *Cannabis sativa* stem tissues. *Int. J. Mol. Sci.* 17.

Mellerowicz, E.J., Immerzeel, P., and Hayashi, T. (2008). Xyloglucan: The molecular muscle of trees. *Ann. Bot.* 102**,** 659-665.

Mikshina, P. (2013). Cellulosic Fibers: Role of Matrix Polysaccharides in Structure and Function. Cellulose - Fundamental Aspects. Theo G.M. Van De Ven (Ed.), *InTech*, DOI: 10.5772/51941.

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.

Mortazavi, A., Williams, B.A., Mccue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* 5**,** 621-628.

Nagano, Y., Furuhashi, H., Inaba, T., and Sasaki, Y. (2001). A novel class of plant-specific zinc-dependent DNA-binding protein that binds to A/T-rich DNA sequences. *Nucleic Acids Res.* 29**,** 4097-4105.

Neutelings, G. (2011). Lignin variability in plant cell walls: Contribution of new models. *Plant Sci.* 181**,** 379-386.

Peña, M.J., Zhong, R., Zhou, G.K., Richardson, E.A., O'neill, M.A., Darvill, A.G., *et al.* (2007). *Arabidopsis* irregular xylem8 and irregular xylem9: Implications for the complexity of glucuronoxylan biosynthesis. *Plant Cell* 19**,** 549-563.

Pierik, R., Djakovic-Petrovic, T., Keuskamp, D.H., De Wit, M., and Voesenek, L.a.C.J. (2009). Auxin and Ethylene Regulate Elongation Responses to Neighbor Proximity Signals Independent of Gibberellin and DELLA Proteins in *Arabidopsis*. *Plant Physiol.* 149**,** 1701-1712.

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

Ranocha, P., Dima, O., Nagy, R., Felten, J., Corratge-Faillie, C., Novak, O., *et al.* (2013). *Arabidopsis* WAT1 is a vacuolar auxin transport facilitator

required for auxin homoeostasis. *Nature Commun.* 4.

Rauf, M., Arif, M., Fisahn, J., Xue, G.P., Balazadeh, S., and Mueller-Roeber, B. (2013). NAC transcription factor SPEEDY HYPONASTIC GROWTH regulates flooding-induced leaf movement in *Arabidopsis*. *Plant Cell* 25**,** 4941-4955.

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.

Ruan, Y.L., Llewellyn, D.J., and Furbank, R.T. (2001). The control of single-celled cotton fiber elongation by developmentally reversible gating of plasmodesmata and coordinated expression of sucrose and K^+ transporters and expansin. *Plant Cell* 13**,** 47-60.

Samanta, P., Sadhukhan, S., and Basu, A. (2015). Identification of differentially expressed transcripts associated with bast fibre development in *Corchorus capsularis* by suppression subtractive hybridization. *Planta* 241**,** 371-385.

Snegireva, A., Chernova, T., Ageeva, M., Lev-Yadun, S., and Gorshkova, T. (2015). Intrusive growth of primary and secondary phloem fibres in hemp stem determines fibre-bundle formation and structure. *AoB PLANTS* 7.

Strzalka, W., and Ziemienowicz, A. (2011). Proliferating cell nuclear antigen (PCNA): A key factor in DNA replication and cell cycle regulation. *Ann. Bot.* 107**,** 1127-1140.

Sun, T.P. (2008). Gibberellin metabolism, perception and signaling pathways in Arabidopsis. *Arabidopsis Book* 6.

Tan, L., Eberhard, S., Pattathil, S., Warder, C., Glushka, J., Yuan, C., *et al.*

(2013). An *Arabidopsis* cell wall proteoglycan consists of pectin and arabinoxylan covalently linked to an arabinogalactan protein. *Plant Cell* 25**,** 270-287.

van Bakel, H., Stout, J.M., Cote, A.G., Tallon, C.M., Sharpe, A.G., Hughes, T.R., *et al.* (2011). The draft genome and transcriptome of *Cannabis sativa*. *Genome Biol.* 12.

van de Mortel, J.E., Schat, H., Moerland, P.D., van Themaat, E.V.L., van der Ent, S., Blankestijn, H., *et al.* (2008). Expression differences for genes involved in lignin, glutathione and sulphate metabolism in response to cadmium in *Arabidopsis thaliana* and the related Zn/Cd-hyperaccumulator *Thlaspi caerulescens*. *Plant Cell Environ.* 31**,** 301-324.

van den Broeck, H.C., Maliepaard, C., Ebskamp, M.J.M., Toonen, M.a.J., and Koops, A.J. (2008). Differential expression of genes involved in C1 metabolism and lignin biosynthesis in wooden core and bast tissues of fibre hemp (*Cannabis sativa* L.). *Plant Sci.* 174**,** 205-220.

War, A.R., Paulraj, M.G., Ahmad, T., Buhroo, A.A., Hussain, B., Ignacimuthu, S., *et al.* (2012). Mechanisms of plant defense against insect herbivores. *Plant signal. behav.* 7**,** 1306-1320.

Wenzel, C.L., Hester, Q., and Mattsson, J. (2008). Identification of Genes Expressed in Vascular Tissues Using NPA-Induced Vascular Overgrowth in *Arabidopsis*. *Plant Cell Physiol.* 49**,** 457-468.

Xu, C., Luo, F., and Hochholdinger, F. (2016). LOB Domain Proteins: Beyond Lateral Organ Boundaries. *Trends Plant Sci.* 21**,** 159-167.

Yordanov, Y.S., Regan, S., and Busov, V. (2010). Members of the LATERAL ORGAN BOUNDARIES DOMAIN Transcription Factor Family Are Involved in the Regulation of Secondary Growth in *Populus*. *Plant Cell* 22**,** 3662-3677.

Zhang, L.-Y., Bai, M.-Y., Wu, J., Zhu, J.-Y., Wang, H., Zhang, Z., *et al.* (2009). Antagonistic HLH/bHLH Transcription Factors Mediate Brassinosteroid Regulation of Cell Elongation and Plant Development in Rice and *Arabidopsis*. *Plant Cell* 21**,** 3767-3780.

Zhang, Y., Zhang, B., Yan, D., Dong, W., Yang, W., Li, Q., *et al.* (2011). Two *Arabidopsis* cytochrome P450 monooxygenases, CYP714A1 and CYP714A2, function redundantly in plant development through gibberellin deactivation. *Plant J.* 67**,** 342-353.

Zhong, R., and Ye, Z.H. (2012). MYB46 and MYB83 bind to the SMRE sites and directly activate a suite of transcription factors and secondary wall biosynthetic genes. *Plant Cell Physiol.* 53**,** 368-380.

Zhong, R., and Ye, Z.H. (2015). Secondary cell walls: Biosynthesis, patterned deposition and transcriptional regulation. *Plant Cell Physiol.* 56**,** 195-214.

Zhou, X.-Y., Song, L., and Xue, H.-W. (2013). Brassinosteroids Regulate the Differential Growth of *Arabidopsis* Hypocotyls through Auxin Signaling Components IAA19 and ARF7. *Mol. Plant* 6**,** 887-904.

Zhou, M., Zhang, K., Sun, Z., Yan, M., Chen, C., Zhang, X., *et al.* (2017). LNK1 and LNK2 Corepressors Interact with the MYB3 Transcription Factor in Phenylpropanoid Biosynthesis. *Plant Physiol.* 174**,** 1348-1358.

Chapter 4: Highlights and Perspectives

With the objective to understand more deeply the molecular mechanism of bast fibre development, from intrusive growth to gelatinous layer maturation, an RNA-Seq experiment was carried out on enriched bast fibre fractions. Three stages of development have been investigated: at the TOP, the bast fibres are elongating; the MID sample marks the transition to secondary cell wall deposition; the BOT is characterised by a thick gelatinous cell wall. As in Chapter 1, we have paid a special attention to the genes related to cell wall, phytohormones and transcriptional regulation. Each sample was characterised by a specific transcriptional signature:

1. At the TOP, the bast fibres undergo intrusive growth. Genes associated with (phospho)lipid metabolism (*KCS*, *LTP*) were strongly up-regulated to follow cell elongation. We have suggested that intrusive growth is regulated by multiple phytohormones (auxins, BR, ethylene, jasmonic and abscisic acids) and transcription factors (*KIDARI*, *IBH1*, *MYB34* and *MYB122*). Similarly to what was observed in Chapter 3, several *FLA*s which may be involved in cell elongation were highly expressed in this sample.
2. The MID sample is characterised by the up-regulation of genes involved in SCW deposition. Several genes involved in xylan biosynthesis and lignification (*MYB46*, *MYB63*, and laccases) were highly up-regulated. The pectic fraction of the cell wall seems to be modified by expansins, XTHs and arabinoxylosidase (ARAF1). The up-regulation of the secondary cell wall-associated *CesA8* and of several cellulose

synthase-like genes indicate both cellulosic and non-cellulosic polysaccharide deposition.

3. Cell wall thickening and secondary growth occur at the BOT. Auxin is a major actor of SCW deposition, as shown by the strong up-regulation of the auxin transporter WAT1. The homeostasis of gibberellins and abscisic acid may play an important role for stopping elongation. The hypolignification of bast fibres is possibly related with the monolignol glycosylation balance, as well as with the strong expression of transcription factors (*MYB3-4* and *KNOX* homeodomain) which negatively regulate the expression of lignin-related genes.

The transcriptomic data confirm that the major changes related to cell wall occur at the internode containing the snap point (MID). In the next chapter a molecular analysis of the snap point will be performed at a high spatial resolution. This study will aim at investigating whether the phenotypic transition occurring in this peculiar region is accompanied by a measurable change in the gene expression profile. To this end, some genes important for cell wall deposition and modification will be targeted by RT-qPCR. We will provide additional data related to the role of FLA proteins and auxin in the framework of cellular elongation and cell wall thickening. The gene expression pattern of lignin-related genes confirms that hypolignification is tuned at the transcriptional level.

Chapter 5

Molecular investigation of the stem snap point in textile hemp

From the published article: Marc Behr, Sylvain Legay, Jean-Francois Hausman, Stanley Lutts and Gea Guerriero. Molecular investigation of the stem snap point in textile hemp. 2017. *Genes* 8 (12): 363.

Author contributions:

Marc Behr, Sylvain Legay and Gea Guerriero designed the experiment, collected the samples and analysed the data. Marc Behr performed the RT-qPCR analysis. Marc Behr and Gea Guerriero wrote the manuscript. Sylvain Legay, Jean-Francois Hausman and Stanley Lutts critically revised the manuscript.

Abstract

Among fibre crops, textile hemp (*Cannabis sativa* L.) is appreciated for its long and strong gelatinous bast fibres. The stem of fibre crops is a useful system for cell wall-oriented studies, because it shows a strong tissue polarity with a lignified inner core and a cellulosic hypolignified cortex, as well as a basipetal lignification gradient. Along the stem axis of fibre crops, a specific region, denoted snap point, marks the transition from elongation (above it) to fibre thickening (below it). After empirically determining the snap point by tilting the plant, we divided the stem segment containing it into three non-overlapping consecutive regions of 1 cm each and carried out targeted RT-qPCR on cell wall-related genes in outer and inner tissues, separately. Different gene clusters can be observed: two are the major gene groups, i.e. one group with members expressed at higher levels in the inner tissues and one whose genes are more expressed in the cortex. In the core tissue, genes associated with cellulose deposition (*CesA4*, *CesA7* and *CesA8*), lignin biosynthesis (such as *PAL*, methyl donors and *LAC4*) and stem biomechanics (*FLA2*, *FLA6*) were up-regulated. The outer tissues were characterised by the higher expression of *FLA3* and *PLR*. The present results provide a molecular validation that the snap point is characterized by a gradient in the expression of cell wall-related genes associated with the shift from fibre elongation to thickening.

Introduction

The stems of fibre crops like hemp and flax are characterized by the occurrence of both cellulosic and woody fibres, localized at the cortex and core, respectively. The cellulosic fibres are long cells developing a thick tertiary cell wall, which is rich in crystalline cellulose and similar to the gelatinous (G) layer forming in tension wood (TW) (Mellerowicz and Gorshkova, 2012; Guerriero *et al.*, 2013; Guerriero *et al.*, 2014).

Along the stem axis of fibre crops, a gradient accompanying the development of bast fibres and vascular tissues is present: young internodes elongate rapidly, while, at the base, older internodes cease elongation and thicken. Bast fibres elongate mostly by intrusive growth and undergo secondary and tertiary cell wall deposition only when elongation has stopped (Snegireva *et al.*, 2015). The shift to secondary and tertiary cell wall deposition results in changes in the mechanical properties of the stem; the region physically marking the boundaries of this transition is referred to as snap point (Gorshkova *et al.*, 2003; Koziel, 2010). In Chapter 4, we showed that the majority of secondary cell wall biosynthetic processes in the bast fibres takes place at the snap point, where genes involved in both hemicellulose and lignin metabolism are significantly upregulated. We have here focused our attention on the snap point and provide evidence for the existence of a gradient of cell wall-related processes in this crucial stem region. We performed a molecular « dissection » of the stem internode comprising the snap point region by dividing it into three 1-cm non-overlapping consecutive disks; we separated the cortical and core tissues of each disk. We monitored the expression of key genes involved in cellulose and lignin biosynthesis previously identified as being highly dynamic. Our results provide further insight into the regulation of cell wall-related genes in the stem tissues of textile hemp.

I. Material and Methods

a. Plant material and growth conditions

The monoecious hemp fibre variety (*C. sativa* cv. Santhica 27) was analysed in this work. Plants were grown as described in Chapter 3 and sampled at the age of six weeks (at a height of 100-120 cm). The snap point was located 5.5 to 6.5 cm below the stem apex. Disks of 1 cm were excised from a 3 cm internode segment containing the snap point. SP1, SP2 and SP3 refer to the top, medium and basal disks of the excised segment, respectively. The snap point was determined by gently tilting the plant and by identifying the region of the stem showing a kink (Koziel, 2010). Subsequently, the stem region containing the kink was quickly excised (1.5 cm above and below the kink). The cortical tissues containing epidermal, parenchymatic and bast fibre cells (labelled OUT) were peeled from the shivs (labelled IN) and quickly frozen in liquid nitrogen. The number of biological replicates was four, with one plant in each replicate. A representative microscopic observation of the sampled region, prepared as described in Chapter 3, is depicted in Figure 5-1.

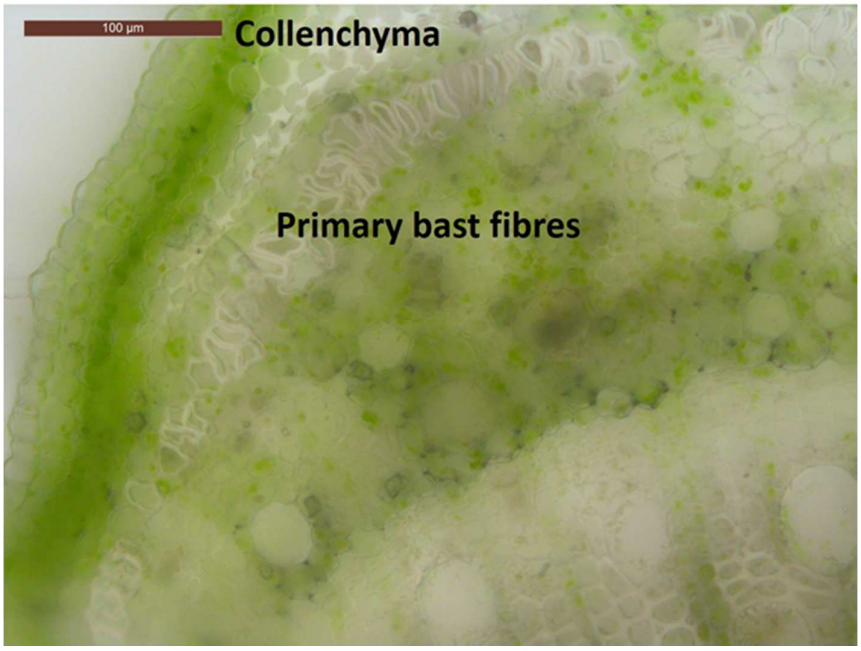


Figure 5-1: Cross-section of the snap point showing the thickening of primary bast fibres. Scale bar is 100 µm.

b. RNA extraction and RT-qPCR

Total RNA was extracted using the RNeasy Plant Mini Kit (Qiagen) coupled to the on-column DNase treatment. The RNA concentration was measured with a Nanodrop ND-1000 (Thermo Scientific). One microgram of total RNA was retrotranscribed into cDNA using the ProtoScript II Reverse transcriptase (NEB) and random primers oligonucleotides (Invitrogen), according to the manufacturer's instructions. The cDNA (4 ng) was used for the RT-qPCR analysis with Takyon SYBR Green low ROX (Eurogentec) for a total reaction volume of 10 µl in 384-wells microplates. An automated liquid handling robot (epMotion 5073) was used to prepare the microplates. The expression of each target gene was normalized using two reference genes (*Cyclophilin* and *Etif3H*) selected among six (*Clathrin*, *F-Box*, *Etif3E*

and *Ubiquitin*). The specificity of the amplicons was checked with a melt curve analysis. The normalised relative expression of the genes was calculated in qBasePLUS (Hellemans *et al.*, 2007). The primers used in this study have been previously described and validated either in the previous chapters or by Mangeot-Peter and colleagues (2016) or designed based on the transcript sequences retrieved from the Medicinal Plant Genomics Resource database (http://medicinalplantgenomics.msu.edu/index.shtml). The normalised relative expression values are indicated as a heat map hierarchical clustering performed with the software PermutMatrix (Caraux and Pinloche, 2005) using the following parameters: dissimilarity assessed by Pearson distance, clustering in complete linkage, seriation and tree seriation in multiple-fragment heuristic (MF), rows normalized by Z-score scaling. The genes were named according to the *Arabidopsis thaliana* nomenclature. Two statistical tests have been performed, i. an ANOVA or a Kruskall-Wallis on the whole set of samples (SP1, SP2 and SP3 IN and OUT), and ii. a Student *t*-test between the different positions of the inner and outer tissues, to follow their individual developments.

II. Results and Discussion

A total of 22 cell wall-related genes, investigated in the previous chapters and related to cellulose (Persson *et al.*, 2005) and lignin (Zhong and Ye, 2015) were targeted using RT-qPCR. The heat map hierarchical clustering of the normalised relative expressions indicates the presence of five major groups (Figure 5-2).

The first group is composed of eleven genes belonging to the secondary cell wall (SCW) and lignin biosynthetic process and shows increased expression in the inner tissues (with upregulation in SP2 and

SP3). The following genes are found in this cluster: *4-coumarate ligase* (*4CL*), *Methionine synthase 2* (*MET2*), *Phenylalanine ammonia lyase* (*PAL*), *S-adenosylmethionine synthetase 1* (*SAM1*), *Caffeoyl-CoA 3-O-methyltransferase* (*CCoAOMT*), *Laccase 4* (*LAC4*), *MYB46-1*, *Cellulose synthase 7* (*CesA7*), *CesA8*, *NAC secondary cell wall thickening 1* (*NST1*) and *CesA4. CesA4* was slightly less expressed than the other genes of this cluster in SP1-IN, explaining its specific position in this group.

The second group shows higher expression in the core tissue of the different disks with respect to the cortex (with, in the cortex, a gradual decrease in expression from SP1 to SP3). This group gathers the primary cell wall (PCW)-related genes *CesA6A*, *Fasciclin-like arabinogalactan 2* (*FLA2*) and *FLA6*.

Genes clustered in the third group are the elongation-related genes *α-expansin 8* (*EXPA8*) and *FLA8*. Their expression is high in the cortex of SP1 and lower in the cortical tissues of SP2 and SP3. The genes of the fourth group *Peroxidase 4* (*PRX4*), *PRX72*, and *Pinoresinol lariciresinol reductase* (*PLR*) show upregulation in the cortex of SP1 and SP2. Finally, the expression of the genes of the fifth group *FLA3* and *Walls are thin 1* (*WAT1*) increases in the cortex of SP2 and SP3.

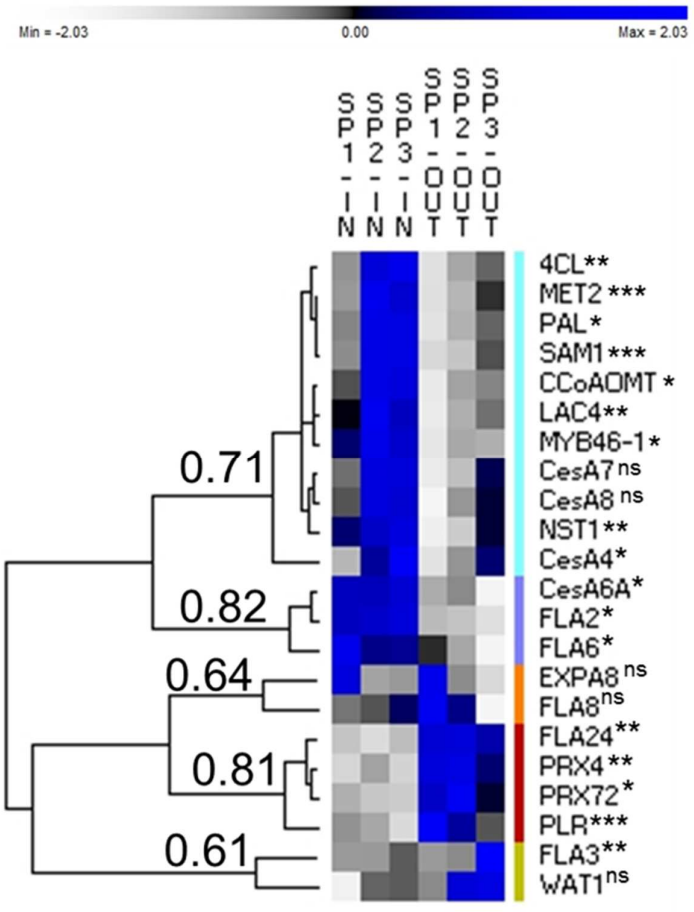


Figure 5-2: Normalised relative expression values of the six samples for each gene as calculated by qBASE^{+} and PermutMatrix. n = 4. First cluster in turquoise; second cluster in violet; third cluster in orange; fourth cluster in red and fifth cluster in dark yellow. *p*-value < 0.05*, 0.01**, 0.001*** and > 0.05^{ns} from ANOVA or Kruskall-Wallis test. The Pearson correlation coefficient is indicated for each group. The statistical tests on normalised relative expression values are in Tables 5-1 to 5-5. Abbreviations as in the text. Values and statistics in Supplementary data 5-1.

a. Genes preferentially expressed in the inner tissues

The genes of the first group belongs to the SCW-biosynthetic program. Their expressions are regulated by the two transcription factors (TFs) NST1 and MYB46-1. NST1 is the master regulator of fibre differentiation (Zhong and Ye, 2015) and MYB46-1 is a tier 2 transcription factor involved in SCW biogenesis and xylan biosynthesis (Zhong *et al.*, 2008). It is therefore not surprising that the genes involved in cellulose deposition in the SCW (*CesA4*, *CesA7* and *CesA8*) and in monolignol biosynthesis (*PAL*, *4CL*, *CCoAOMT* and the methyl donors *SAM1* and *MET2*) display a similar expression pattern. These genes are more expressed in the inner tissue of the lower part stem, containing the xylem cells with xylan-type SCW. These data are in agreement with the results of van den Broeck and colleagues (2008) showing, in the core tissue, an up-regulation of *PAL*, *4CL* and *CCoAOMT* for monolignol biosynthesis, *SAM* and *MET* for the methyl donor metabolism. Hemp bast fibres are weakly lignified, but we nevertheless observed a significantly increased expression of the genes belonging to the first group in the basal segment (SP3-OUT) as compared to SP1-OUT (Table 5-1). *PAL1* and *COMT* were also significantly more expressed in the bast fibres located below the snap point of hemp stem (Chapter 4). Methylation of monolignols is catalysed by COMT and CCoAOMT and *S*-adenosylmethionine is a methyl donor for this reaction, explaining the higher transcription of genes related to this metabolism such as *SAM1* and *MET2* in the core tissue. The expression of *CesA4*, *CesA7* and *CesA8* slightly increases in SP2 and SP3 as compared to SP1, both in the core and the cortex. *CesA8* shows a significant upregulation at SP3-OUT (fold change SP3-OUT *vs* SP2-OUT and SP1-OUT of 1.3 and 4.4). The onset of bast

fibre SCW deposition occurs at the snap point. Even though the histology of the three consecutive segments is very similar because of their vicinity, our sampling strategy enabled the detection of (even small) differences in the expression of genes related to SCW deposition. When observed in different internodes, the trend is stronger as the fold change of *CesA8* between the internode of the snap point and the internode above the snap point is around 11 (Chapter 4).

Table 5-1: Student's *t*-test on normalised relative expression values from the first group. The 3 samples from inner or outer tissues are compared. A green cell means that the first term of the comparison is up-regulated (*e.g.*, SP3 in the comparison SP3/SP1).

	IN			**OUT**		
	SP2/SP1	SP3/SP1	SP3/SP2	SP2/SP1	SP3/SP1	SP3/SP2
4CL	0.189	0.073	0.438	0.079	0.002	0.217
MET2	0.163	0.026	0.408	0.114	0.000	0.074
PAL	0.220	0.071	0.480	0.064	0.000	0.109
SAM1	0.198	0.047	0.491	0.046	0.005	0.020
CCoAOMT	0.175	0.073	0.452	0.067	0.003	0.355
LAC4	0.215	0.203	0.345	0.054	0.002	0.209
MYB46-1	0.178	0.222	0.357	0.070	0.049	0.457
CESA7	0.175	0.082	0.478	0.083	0.005	0.044
CESA8	0.253	0.197	0.457	0.046	0.003	0.199
NST1	0.240	0.198	0.389	0.105	0.001	0.032
CESA4	0.143	0.036	0.184	0.066	0.016	0.139

CesA6A, *FLA2* and *FLA6* are the genes of the second group (Figure 5-2). These genes are more expressed in the core as compared to the cortex (Figure 5-2). In the outer tissue, *FLA6* is significantly more expressed in SP1 than in SP3 (Table 5-2). CESA6 is involved in the deposition of cellulose in the PCW (Gonneau *et al.*, 2014), but is also present in flax fibres depositing the G-layer (Mokshina *et al.*, 2017). CesA6 has different isoforms (5 in flax

according to Mokshina *et al.*, 2014). Additional characterisations of hemp *CesA* genes are required to define the exact number of isoforms for each CESA and to clearly define their respective expression patterns in the different tissues / stem regions. *FLA2* and *FLA6* are more expressed in the core and cortical tissues of younger internodes (Chapter 3). The closest ortholog of CsaFLA6 in *Arabidopsis* is FLA9 (Chapter 3). Expression of *AtFLA9* is significantly repressed in a mutant with altered auxin signalling, caused by loss of function of several VIER F-BOX PROTEINES (VFBs). Loss of VFB function by RNAi causes changes in gene expression and delays plant growth (Schwager *et al.*, 2007). The *FLA9* expression level is also negatively affected in the *Arabidopsis* auxin-insensitive gain-of-function mutant *iaa3* but is higher in seedlings exposed to brassinolide, a brassinosteroid (Oh *et al.*, 2014). Interestingly, the expression of *EXPA8* is also induced by auxin (Sánchez-Rodríguez *et al.*, 2010) and brassinolide (Oh *et al.*, 2014) and shows a pattern similar to *CsaFLA6* (Figure 5-2). Considering the significantly higher expression of *CsaFLA6* and *EXPA8* in outer SP1 *vs* SP3 (Tables 5-2 and 5-3), we may speculate that these proteins play a role in the elongation of the bast fibres in relation with auxin and brassinosteroids.

Table 5-2: Student's *t*-test on normalised relative expression values from the second group. The 3 samples from inner or outer tissues are compared. A red cell means that the second term of the comparison is up-regulated (*e.g.*, SP1 in the comparison SP3/SP1).

	IN			**OUT**		
	SP2/SP1	SP3/SP1	SP3/SP2	SP2/SP1	SP3/SP1	SP3/SP2
CESA6A	0.473	0.432	0.428	0.400	0.109	0.104
FLA2	0.455	0.426	0.459	0.433	0.189	0.208
FLA6	0.208	0.275	0.481	0.217	0.004	0.063

b. Genes preferentially expressed in the outer tissues

Expression of the genes of the third cluster, *EXPA8* and *FLA8*, does not change significantly across the disk sections according to the ANOVA test. However, the expression of *EXPA8* decreases from the top to the bottom, both in inner and outer tissues, resulting in a significant difference between SP1-OUT and SP3-OUT (Table 5-3). Expansins are involved in cell elongation by mediating the degradation of the biomechanical hotspots where xyloglucan and cellulose are closely intertwined (Park and Cosgrove, 2015). In this respect, the expression of *EXPA8* was higher in elongating hemp hypocotyls, as compared to thickening hypocotyls (Chapter 1) and in the bast fibres of younger internodes (Chapter 4). Using a microarrays approach, a transcript for a putative expansin was up-regulated in the bast fibres of the upper part of hemp stem (de Pauw *et al.*, 2007). One may therefore suppose that the expression of *EXPA8* is not drastically changing because stem elongation around the snap point is very limited, as compared to the stem part closer to the shoot apical meristem. Similarly, the relatively constant expression of *CsaFLA8*, which was shown to be upregulated during bast fibre elongation (Chapter 3), may be due to the limited elongation in the sampled region.

Table 5-3: Student's *t*-test on normalised relative expression values from the third group. The 3 samples from inner or outer tissues are compared. A red cell means that the second term of the comparison is up-regulated (e.g., SP1 in the comparison SP3/SP1).

	IN			**OUT**		
	SP2/SP1	SP3/SP1	SP3/SP2	SP2/SP1	SP3/SP1	SP3/SP2
EXPA8	0.148	0.197	0.474	0.068	0.022	0.226
FLA8	0.426	0.368	0.406	0.281	0.060	0.058

The genes of the fourth group are more expressed in the outer tissue (Figure 5-2). *PLR* transcript abundance significantly decreases from the top to the bottom in both tissues (Table 5-4), similarly to what was observed previously (Chapter 4). PLR is involved in the biosynthesis of lignans such as pinoresinol. Lignans are monolignol-derived molecules which are highly accumulated in the hypolignified bast fibres of flax (Huis *et al.*, 2012) and may thus be final acceptors of the products of the monolignol pathway in hemp bast fibre. Stem and root elongation is regulated by specific lignans, such as syringaresinol and sesamin, in lettuce and ryegrass (Yamauchi *et al.*, 2015). We can thus suggest that lignans may partake in organ elongation in hemp. Interestingly, the hemp genes annotated as *PRX4* and *PRX72* share a similar expression pattern with *PLR*. Alignment of hemp and thale cress protein sequences of PRX4 and PRX72 with the EMBOSS Needle suite has shown identity / similarity rates of 62.8% / 78.0% for PRX4 and 77.2% / 87.8% for PRX72. AtPRX5 and AtPRX72 are the closest orthologs of CsaPRX4 and CsaPRX72, respectively, according to BLASTP (E values of 3e-149 and 0, respectively). In *Arabidopsis*, PRX4 and PRX72 are involved in the formation of syringyl lignin (Herrero *et al.*, 2013; Fernández-Pérez *et al.*, 2015). Considering that these two genes are more expressed in the outer tissue, which is poor in lignin, and follow an acropetal gradient of expression (Table 5-4), we suggest that, in hemp, they are not involved in lignin polymerisation, and that they are not functional orthologs of AtPRX4 and AtPRX72. They may be involved in other cellular processes occurring in the apoplast (e.g., protein bridges). *FLA24* is the last gene of this cluster, no significant change in its expression was observed, except that it was slightly more expressed in the cortical tissue.

Table 5-4: Student's *t*-test on normalised relative expression values from the fourth group. The 3 samples from inner or outer tissues are compared. A red cell means that the second term of the comparison is up-regulated (*e.g.*, SP1 in the comparison SP3/SP1).

	IN			**OUT**		
	SP2/SP1	SP3/SP1	SP3/SP2	SP2/SP1	SP3/SP1	SP3/SP2
FLA24	0.296	0.448	0.297	0.457	0.270	0.221
PRX4	0.089	0.473	0.045	0.409	0.058	0.060
PRX72	0.153	0.078	0.254	0.342	0.262	0.129
PLR	0.287	0.003	0.026	0.034	0.014	0.166

The last cluster is composed of *FLA3* and *WAT1*. Both genes are more expressed in the outer tissue and in the section undergoing SCW deposition (SP3). The expression of *FLA3* is significantly higher in SP3 than in SP1 and SP2, in both inner and outer tissues (Table 5-5).

Table 5-5: Student's *t*-test on normalised relative expression values from the fifth group. The 3 samples from inner or outer tissues are compared. A green cell means that the first term of the comparison is up-regulated (*e.g.*, SP3 in the comparison SP3/SP1).

	IN			**OUT**		
	SP2/SP1	SP3/SP1	SP3/SP2	SP2/SP1	SP3/SP1	SP3/SP2
FLA3	0.097	0.012	0.014	0.117	0.012	0.014
WAT1	0.220	0.132	0.493	0.004	0.059	0.432

This result confirms the data obtained i. with the time-course analysis of the hemp hypocotyl, where *CsaFLA3* was more expressed in hypocotyl undergoing secondary growth, and ii. with hemp internodes at different developmental stages, where the expression of *CsaFLA3* peaked in the internode below the snap point (Chapter 3).

Hemp FLA3 shows a single-fasciclin (FAS) domain, an N-terminal signal peptide and a C-terminal glycosylphosphatidylinositol (GPI) membrane anchor at the C-terminus mediating attachment to the cell surface and is closely related to AtFLA11, EgrFLA2 and PtrFLA6 (Chapter 3). Importantly, the transcripts of several *FLA* are more abundant during the formation of TW in poplar (Andersson-Gunneras *et al.*, 2006). Recently, it has been demonstrated in poplar that gibberellins positively regulate TW formation by inducing the expression of several *FLA*s including *PtrFLA6* (Wang *et al.*, 2017). It has been shown in trees and thale cress that the expression of single-FAS domain FLA is correlated with cellulose microfibril angle (MFA) and thus wood stiffness (MacMillan *et al.*, 2015). In this respect, eucalypt *35S*:FLA2 fibres have a reduced MFA and increased stiffness and cellulose crystallinity. In *Arabidopsis*, *fla11* insertion line shows a subtle irregular xylem phenotype (IRX13), as well as alterations in non-cellulosic polymers and a cellulose reduction, which may be linked to *FLA11* strong coexpression with *CesA4*, *CesA7* and *CesA8* (Persson *et al.*, 2005). We therefore strengthen the hypothesis that CsaFLA3 may be involved in SCW and G-layer deposition during the thickening stage. Functional analyses of *CsaFLA3* or closely related *FLA* genes in other fibre crops, such as flax, may shed light on their roles in G-layer assembly. The expression of *WAT1* is slightly higher at the bottom part of the snap point internode, both in the inner and outer tissue (Figure 5-2). WAT1 is a vacuolar auxin efflux transporter playing a key role in SCW thickening of fibres (Ranocha *et al.*, 2013). The *wat1* mutant shows altered expression in several transcription factors (*NST3*, *MYB46* and *MYB85*) and structural genes (*CesA4*, *CesA7*, *CesA8* and *IRX9*) important for SCW development (Ranocha *et al.*, 2010; Cassan-Wang *et al.*, 2013). The authors also showed that the early differentiation of fibres was not impacted in the *wat1* mutant,

contrasting with later stage of SCW deposition, when the monolignol profile of the mutant is significantly altered. A role in fibre lignification is also possible, as *wat1* shows a lower amount of mono/oligolignols, but higher quantity of lignans (Ranocha *et al.*, 2010). It is therefore plausible that the gene expression of the *WAT1* hemp ortholog does not change dramatically in our experimental set-up, because the fibres are not yet in their late stage of development. In support of this hypothesis, *WAT1* was found to be significantly more expressed in the internode below the snap point as compared to the internode above the snap point (Chapter 4).

III. Conclusion

The expression profile of genes known to be involved in elongation (expansin or specific *FLA*s) and SCW deposition (*CesA*s, lignin biosynthesis) highlights the shift from fibre elongation to fibre thickening. Outer and inner tissues have contrasting patterns of gene expression, especially those genes involved in the monolignol / lignin biosynthetic pathway. It is also noteworthy to observe that some *FLA* genes, whose roles are just starting to be understood in woody and herbaceous species, are more expressed in elongating tissues (e.g., *FLA6*), while others are up-regulated in tissues undergoing SCW deposition (e.g., *FLA3*). It will be interesting to study how the expression of the genes involved in the transition from elongation to thickening at the snap point differs in similar fibre crops, such as flax or ramie, but also crops with xylan-type bast fibres (jute and kenaf). FLAs may play an important role *viz.* concerning the different composition of gelatinous-type and xylan-type cell wall. Such comparisons may help decipher the functions of some genes whose roles are still enigmatic.

References

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

Caraux, G., and Pinloche, S. (2005). PermutMatrix: a graphical environment to arrange gene expression profiles in optimal linear order. *Bioinformatics* 21**,** 1280-1281.

Cassan-Wang, H., Goué, N., Saidi, M.N., Legay, S., Sivadon, P., Goffner, D., *et al.* (2013). Identification of novel transcription factors regulating secondary cell wall formation in *Arabidopsis*. *Front. Plant Sci.* 4.

de Pauw, M.A., Vidmar, J.J., Collins, J., Bennett, R.A., and Deyholos M.K. (2007). Microarray analysis of bast fibre producing tissues of *Cannabis sativa* identifies transcripts associated with conserved and specialised processes of secondary wall development. *Funct. Plant Biol.* 34**,** 737-749.

Fernández-Pérez, F., Vivar, T., Pomar, F., Pedreno, M.Á., and Novo-Uzal, E. (2015). Peroxidase 4 is involved in syringyl lignin formation in *Arabidopsis thaliana*. *J. Plant Physiol.* 175**,** 86-94.

Gonneau, M., Desprez, T., Guillot, A., Vernhettes, S., and Höfte, H. (2014). Catalytic Subunit Stoichiometry within the Cellulose Synthase Complex. *Plant Physiol.* 166**,** 1709-1712.

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

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.

Guerriero, G., Sergeant, K., and Hausman, J.F. (2014). Wood biosynthesis and typologies: a molecular rhapsody. *Tree Physiol.* 34**,** 839-855.

Hellemans, J., Mortier, G., De Paepe, A., Speleman, F., and Vandesompele, J. (2007). qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol.* 8**,** R19.

Herrero, J., Fernández-Pérez, F., Yebra, T., Novo-Uzal, E., Pomar, F., Pedreno, M., *et al.* (2013). Bioinformatic and functional characterization of the basic peroxidase 72 from *Arabidopsis thaliana* involved in lignin biosynthesis. *Planta* 237**,** 1599-1612.

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

Koziel, S.P. (2010). Genetic Analysis of Lignification and Secondary Wall Development in Bast Fibers of Industrial Hemp (*Cannabis sativa*)*. Master of Science, University of Alberta*.

Macmillan, C.P., Taylor, L., Bi, Y., Southerton, S.G., Evans, R., and Spokevicius, A. (2015). The fasciclin-like arabinogalactan protein family of *Eucalyptus grandis* contains members that impact wood biology and biomechanics. *New Phytol.* 206**,** 1314-1327.

Mangeot-Peter, L., Legay, S., Hausman, J.F., Esposito, S., and Guerriero, G. (2016). Identification of Reference Genes for RT-qPCR Data Normalization in *Cannabis sativa* Stem Tissues. *Int. J. Mol. Sci.* 17.

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.

Mokshina, N., Gorshkov, O., Ibragimova, N., Chernova, T., and Gorshkova, T. (2017). Cellulosic fibres of flax recruit both primary and secondary cell wall

cellulose synthases during deposition of thick tertiary cell walls and in the course of graviresponse. *Funct. Plant Biol.* 44**,** 820-831.

Mokshina, N., Gorshkova, T., and Deyholos, M.K. (2014). Chitinase-Like and Cellulose Synthase Gene Expression in Gelatinous-Type Cellulosic Walls of Flax (*Linum usitatissimum* L.) Bast Fibers. *PLoS One* 9**,** e97949.

Oh, E., Zhu, J.Y., Bai, M.Y., Arenhart, R.A., Sun, Y., and Wang, Z.Y. (2014). Cell elongation is regulated through a central circuit of interacting transcription factors in the *Arabidopsis* hypocotyl. *eLife* 3**,** e03031.

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.

Persson, S., Wei, H., Milne, J., Page, G.P., and Somerville, C.R. (2005). Identification of genes required for cellulose synthesis by regression analysis of public microarray data sets. *Proc. Natl. Acad. Sci. U. S. A.* 102**,** 8633-8638.

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

Ranocha, P., Dima, O., Nagy, R., Felten, J., Corratgé-Faillie, C., Novák, O., *et al.* (2013). *Arabidopsis* WAT1 is a vacuolar auxin transport facilitator required for auxin homeostasis. *Nature Commun.* 4.

Sánchez-Rodríguez, C., Rubio-Somoza, I., Sibout, R., and Persson, S. (2010). Phytohormones and the cell wall in *Arabidopsis* during seedling growth. *Trends Plant Sci.* 15**,** 291-301.

Schwager, K.M., Calderon-Villalobos, L.I.A., Dohmann, E.M.N., Willige, B.C., Knierer, S., Nill, C., *et al.* (2007). Characterization of the *VIER F-BOX PROTEINE* Genes from *Arabidopsis* Reveals Their Importance for Plant Growth and Development. *Plant Cell* 19**,** 1163-1178.

Snegireva, A., Chernova, T., Ageeva, M., Lev-Yadun, S., and Gorshkova, T. (2015). Intrusive growth of primary and secondary phloem fibres in hemp stem determines fibre-bundle formation and structure. *AoB Plants* 7.

van den Broeck, H.C., Maliepaard, C., Ebskamp, M.J.M., Toonen, M.a.J., and Koops, A.J. (2008). Differential expression of genes involved in C1 metabolism and lignin biosynthesis in wooden core and bast tissues of fibre hemp (*Cannabis sativa* L.). *Plant Sci.* 174**,** 205-220.

Wang, H., Jin, Y., Wang, C., Li, B., Jiang, C., Sun, Z., *et al.* (2017). Fasciclin-like arabinogalactan proteins, PtFLAs, play important roles in GA-mediated tension wood formation in *Populus*. *Sci. Rep.* 7**,** 6182.

Yamauchi, S., Ichikawa, H., Nishiwaki, H., and Shuto, Y. (2015). Evaluation of Plant Growth Regulatory Activity of Furofuran Lignan Bearing a 7,9':7',9-Diepoxy Structure Using Optically Pure (+)- and (-)-Enantiomers. *J. Agric. Food Chem.* 63**,** 5224-5228.

Zhong, R., Lee, C., Zhou, J., Mccarthy, R.L., and Ye, Z.H. (2008). A Battery of Transcription Factors Involved in the Regulation of Secondary Cell Wall Biosynthesis in *Arabidopsis*. *Plant Cell* 20**,** 2763-2782.

Zhong, R., and Ye, Z.H. (2015). Secondary Cell Walls: Biosynthesis, Patterned Deposition and Transcriptional Regulation. *Plant Cell Physiol.* 56**,** 195-214.

Chapter 5: Highlights

In this chapter, we demonstrate that the phenotypic transition occurring in the snap point is accompanied by a measurable change in the gene expression profile. Some differences have also been observed between the cortical tissue harbouring the bast fibres and the core tissue containing the xylem. Four main groups can be observed.

1. The first group is composed of genes involved in the transcriptional regulation and biosynthesis of the secondary cell wall (lignin, cellulose). These genes are more expressed in the inner tissue and follow a basipetal gradient. These data are in accordance with the gene expression analysis performed in Chapter 3, suggesting that the hypolignification of the bast fibres is at least partially regulated at this level.
2. The genes found in the second group are also up-regulated in the core tissue, but their expression is higher in the upper segment. The transcription of some of these genes (e.g., *CsaFLA6*) may be induced by auxin and brassinosteroids to regulate the elongation of the stem.
3. The expression of the fourth group genes is higher in the cortical tissue and decreases from the top to the bottom. They participate in the biosynthesis of monolignol-derived molecules, such as lignans or lignin. Lignans are known to regulate stem elongation. They may also contribute to the hypolignification of bast fibres, by subtracting monolignols from the lignin polymerisation process.
4. The last group is composed of two genes, *CsaFLA3* and *WAT1*. They are more expressed in the cortical tissue undergoing secondary cell wall deposition. The thale cress and eucalypt

orthologs of *CsaFLA3* are involved in cellulose deposition, while in poplar this gene is up-regulated during tension wood formation. *WAT1* is involved in the late stage of secondary cell wall deposition. These two genes are therefore interesting candidates for functional analysis.

In this chapter, we have demonstrated that, despite the vicinity of the studied samples, the snap point is characterised by discernible modifications in the expression of genes associated with cell wall biosynthesis / remodelling. The wider amplitude of expression was recorded for *CsaFLA3*, suggesting that FLAs, in hemp as in other species, are important proteins for the transition to secondary growth and secondary cell wall thickening.

Discussion and Perspectives

Global warming is a serious threat to humankind as it is presently impacting biogeochemical cycles (such as water and carbon) with direct or indirect impacts in several countries worldwide. But global warming also fuels the opportunity to develop or rediscover tools to take benefit of renewable resources. Among them, lignocellulosics have a promising future. As petroleum, lignocellulosics serve as a source of energy (heating systems, bioethanol and biofuel) and of raw material for a broad range of products (bioplastic, paper, construction and composite materials and biochemicals). Nature and evolution have designed a very large panel of plant metabolites (e.g., polysaccharides, specialised metabolites) which are, or may be, directly used by humankind. However, some specific applications require modifications of the biomass, either directly in the plant system and / or by a pre-treatment. The production of bioethanol, which results from the fermentation of polysaccharides, has been improved by these methods. In woody species for instance, lignin hampers the conversion of polysaccharides (such as cellulose) into ethanol, and a massive effort has been undertaken i. to reduce the lignin content or alter its monomeric composition without strong penalties for the development of the trees, and ii. to optimise the pre-treatment step before the fermentation (Guerriero *et al.*, 2016). Such fine-tuning of the composition of a biological system is made easier by the understanding of its underlying network. To this end, the research effort from the last decades has significantly strengthened our knowledge of plant cell walls in terms of molecular biology, composition and structure (Keegstra, 2010).

The cell walls are much related to the interesting characteristics of hemp stem: both xylem tissue and phloem fibres have thick secondary cell walls with interesting properties. The xylan-type secondary cell wall found in the xylem is rich in lignin and is used to manufacture concrete-like material. By contrast, the gelatinous-type secondary cell wall found in the bast fibres is rich in crystalline cellulose and poor in lignin, opening applications in the production of composite materials and speciality papers (Fike, 2016). From the researcher's perspective, this offers the opportunity to study the biosynthesis of these two different cell walls on the same plant system. In addition to be tissue-dependent, the bulk deposition of the secondary cell wall (SCW) starts at a determined stage of development: the end of elongation. The organisation of the SCW prevents further cellular elongation. After completion of primary growth, the stem diameter increases as a result of cambial activity. Secondary growth is responsible for the majority of plant biomass and results in the production of secondary bast fibres and secondary xylem in hemp.

This thesis aims at answering to two questions related to these aspects:

1) What are the transcriptional and hormonal signatures of primary and secondary growth in hemp?

2) Which molecular factors are associated with the development of hemp bast fibres?

We have studied the mechanisms related to primary and secondary growth using the hypocotyl system. Different transcriptomic signatures characterise these two stages of development. The composition of the cell walls (cellulose, hemicelluloses, rhamnogalacturonan I and lignin), as well as the phytohormone profiles

(such as auxins, jasmonic acid and abscisic acid), markedly changed throughout time. Next, we investigated the genes involved in the differential cell wall composition of cortical and core tissues of adult hemp plants, demonstrating that the hypolignification of the bast fibres is, at least partially, regulated at the transcriptional level. The molecular actors involved in the development of the bast fibres were thoroughly studied using transcriptomics, highlighting the role of phytohormones and transcription factors during their whole life cycle, from elongation to gelatinous layer maturation. We have demonstrated that the empirically-determined snap point displays a gene expression signature typical of SCW deposition.

Across this Discussion chapter, we will bring together the key elements dealing with vascular patterning, phytohormones, lignification and fasciclin-like arabinogalactan proteins in the framework of hemp fibres.

How can the data from the two systems explain the vascular patterning of hemp stem?

By combining the hypocotyl and the adult stem systems, we contributed to the global comprehension of the events related to stem development and bast fibre formation. The hypocotyl system has shed light on the transcriptional regulation of secondary growth and on the general events related to the cell walls, while the specific development of bast fibres was investigated at the transcriptomic level in the adult stem system.

In fully-developed internodes, secondary tissues are far more abundant than primary structures. Considering that hemp provides xylem and bast fibres, understanding the factors of secondary growth

is of considerable interest. To this end, we took advantage of a time-course analysis on the hypocotyl system (i.e., at 6, 9, 15 and 20 days after sowing), with a special emphasis on cell wall. In this organ, elongation and secondary growth are temporally uncoupled. The elongating hypocotyls were characterised by the higher expression of genes related to photosynthesis and response to light (*RubisCO*, chlorophyll metabolism, pigments biosynthesis), primary cell wall (xyloglucan biosynthesis) and cellular elongation (expansins, *XTH*s, *PRX57*). Several genes related to the onset of secondary growth were more expressed in older hypocotyls. It is worth mentioning *WUSCHEL-RELATED HOMEOBOX 4* (*WOX4*), *TDIF receptor TDR / PXY* and *PXY-CORRELATED GENE 1* (*PXC1*), which are all important players in vascular cambium activity and identity of conductive tissues (Nieminen *et al.*, 2015) and may constitute interesting targets for functional analyses of bast fibre development in hemp. Secondary xylem and secondary bast fibres, together with primary bast fibres, undergo SCW deposition after the beginning of secondary growth. As a result, the genes coding for secondary CESAs (CESA4, CESA7 and CESA8), xylan (backbone and substituent polymerisation, acetylation), and lignin biosynthesis (from PAL to CAD, laccases and peroxidases) were more expressed in older hypocotyls.

The cell wall composition of xylem and bast fibres strongly differs, mainly in terms of pectin, xylan and lignin content, as evidenced by immunohistochemical (Chapter 1) and lignin-specific (Chapter 2) microscopic observations. In order to selectively study this patterning, we switched to the adult stem system, in which cortical and core tissues are readily separable. The contrasting profile of the two tissues is (partially) determined at the transcriptional level: the genes involved in lignin biosynthesis (*PAL*, *4CL1*, and *CAD4*) were more expressed in

the core tissue. The xylan-type SCW found in xylem cells has been widely investigated and has a rather conserved structure in many plant species (reviewed in Gorshkova *et al.*, 2012; Zhong and Ye, 2015). By contrast, the organisation and the composition of the gelatinous-type SCW strongly differ between species. In flax, the gelatinous-type cell wall is organised in two substructures, the galactan-layer and the compact G-layer (Gorshkova and Morvan, 2006). Considering this biological diversity, there is a striking interest to decipher the composition and the organisation of the gelatinous layer found in hemp bast fibres. We bring a first insight to this objective by performing a next-generation sequencing analysis of developing bast fibre (Chapter 4). Genes associated with (phospho)lipid metabolism (*KCS*, *LTP*) and cell wall loosening (*α*- and *β-galactosidases*, *α-xylosidases*, *expansins*, *XET*s) were strongly up-regulated in elongating tissues, in relation with intrusive growth. We have also identified, in enriched bast fibre fractions, several transcription factors which may play important roles in phloem fibre formation (*LBD4*, *DOF zinc finger protein 5.6*), cell wall deposition (Zim17-type zinc finger protein) and bast fibre hypolignification (*BEL1* and *BLH*s homeodomain, *MYB4*). However, the general lack of knowledge concerning the enzymes involved in pectin biosynthesis, and most specifically RG-I, prevents a detailed molecular analysis explaining difference of RG-I structure observed between flax and hemp bast fibres, as evidenced by the absence of LM5 signal in hemp bast fibres.

How are phytohormones involved in hemp stem development?

Phytohormones are associated with every aspect of plant development (Crozier *et al.*, 2000). Throughout this thesis, we have investigated the possible links between the main phytohormones and gene expression profiles, with an emphasis on cellular elongation, secondary growth and cell wall formation.

The implication of auxins in cellular elongation was obvious in our two systems. The higher bioactive auxin content of the elongating hypocotyls was concomitant with the up-regulation of genes involved in cell wall loosening (expansins, *XTH*s, *PRX57*) (Chapter 1). In both systems, we have evidenced the presence of several transcription factors involved in the common auxin-brassinosteroids gene regulatory network, such as *KIDARI*, *IBH1* and *PIF7*, which were previously characterised in *A. thaliana* (Oh *et al.*, 2014). These phytohormones therefore regulate the expression of several hundreds of genes involved in cell elongation. In addition to auxins, it seems that other phytohormones are involved in the intrusive growth of hemp bast fibres. Transcription factors responsive to ethylene, jasmonic acid or abscisic acid (*MYB34* and *MYB122*) were up-regulated in the bast fibres of elongating internodes, possibly playing a role *via* the biosynthesis of indole glucosinolate (Chapter 4). It is known that, in thale cress seedlings, auxins stimulate the formation of ethylene (Abel *et al.*, 1995), and that ethylene may trigger elongation through an auxin signalling pathway (Pierik *et al.*, 2009). Auxins and ethylene may therefore regulate, either directly or via a cross-talk, the elongation of the hemp hypocotyl as well as the intrusive growth of bast fibres. The positive role of gibberellins in stem elongation is long-time

acknowledged (Aloni, 1979). GAs from the leaves are implicated in elongation and secondary growth of the stem, but shoots are normally autonomous for the biosynthesis of GAs (Hedden and Sponsel, 2015). We suggest that, in hemp, gibberellins are involved in elongation and secondary growth. *GA20OX2*, which catalyses the rate limiting step in GA biosynthesis, was more expressed in the elongating hypocotyls (H6 and H9), but its expression also peaked in thickening bast fibres (at BOT) (Chapters 1 and 4). A previous study has highlighted that deactivation is the crucial point for gibberellins homeostasis (Dayan *et al.*, 2010), and thus may adjust the biological activity of GAs. The higher expression of two genes involved in GA deactivation (*GA2OX2* and *CYP714A1*) in the thickening bast fibres may trigger a signal causing the end of the elongation. In the hypocotyl, the expression of *CYP714A1* strongly decreased during the time course, suggesting that the pool of bioactive GA is higher in the hypocotyls undergoing secondary growth and secondary cell wall formation. It will therefore be interesting to quantify the pools of bioactive and inactive GAs, as well as their histological repartition, both in the hypocotyls and in internodes spanning elongation and bast fibre maturation. Recently, the formation of poplar tension wood has been shown to be regulated by gibberellin A3, through the up-regulation of *PtrFLA6* (Wang *et al.*, 2017). As shown in Chapter 3, the expression of *FLA* genes is related to the stage of the development of the bast fibres and to the age of the hypocotyls. The hemp ortholog of PtrFLA6 is CsaFLA11, whose coding gene is highly expressed in xylem (Chapter 3), but also in the outer tissue (Chapter 4), from the snap point downwards. Gibberellins may consequently have a role in xylem and bast fibre development.

The hormonal regulation of secondary growth is studied since several decades but the underlying molecular regulation is not yet

completely understood (Ursache *et al.*, 2013). A strong correlation does exist between the cambial auxin content and the initiation of secondary growth: perturbations in the auxin fluxes lead to inhibition of cambial activity (Ursache *et al.*, 2013). Our data agree with this hypothesis, as several genes involved in auxin cellular transport were up-regulated in hypocotyls undergoing secondary growth. Among them, we may cite *AUX1*, *PIN1*, *PIN5* and *ABCB1*, which may direct the flux of auxins towards the cambial zone and induce secondary growth. Many genes with auxin responsive elements in their promoter region and up-regulated during secondary growth (Ko *et al.*, 2004) were more expressed in H15 or H20 (*MYB85*, *CCoAOMT*, *CesA7*) (Chapter 1). In addition, auxin promotes cambial growth through the TDIF receptor-WOX4 signalling route (Nieminen *et al.*, 2015). These two genes were up-regulated in H15 and H20, respectively (Chapter 1). The regulation of cambial activity by cytokinins is also well described. The expression of the transcription factor *AINTEGUMENTA* (*ANT*) and of some D-type cyclins (*CYCD*) are induced by cytokinins; notably, they are expressed in the dividing cambial cells (Bhalerao and Fischer, 2017). The expression of *ANT* and *CYCD*s were higher in hypocotyls undergoing secondary growth (H15 and H20), with a minimum in H6. We have not observed a strong correlation between the cytokinin content and the expression of these genes. Especially, the bioactive cytokinin content reached a minimum in H20 (Chapter 1). It is possible that the concentration is low in the overall hypocotyl, but high in the cambial zone. We also found out that the expression of the GA receptor *GID1C* peaked in H9 (Chapter 1). In aspen, the genes of the *GID1* family stimulate xylogenesis in the cambial zone and fibre elongation in the young xylem (Mauriat and Moritz, 2009). GA are thus likely important for the xylem development in hemp, which may have industrial applications, such as the valorisation in “hempcrete”.

Jasmonates play a role in secondary growth, as recently demonstrated by Sehr and colleagues (2010). Indeed, silencing the negative regulator of jasmonate signalling JAZ10 results in a more pronounced interfascicular cambial activity and jasmonate-treated stem has enhanced cambial activity and additional phloem fibres in *Arabidopsis*. The maximum jasmonate content was found in H15 (Chapter 1), suggesting a role for this class of phytohormone in hemp stem secondary growth. The application of exogenous jasmonic acid results in an increased secondary growth, additional secondary phloem fibres and higher lignin content in the hemp hypocotyl (data presented in Appendix).

Finally, phytohormones regulate several aspects of SCW deposition. As for cell elongation and secondary growth, the role of auxins is well documented. According to its expression profile, its high abundance and its role in SCW thickening, the gene *Walls are thin 1* (*WAT1*) is expected to have an important role in hemp fibres. In both models, the expression of *WAT1* reached several hundreds of RPKM (>300 in H15 and H20; >500 at the BOT; Chapters 1 and 4). In thale cress, the *wat1-1* mutant has reduced SCW deposition, as a result of decreased expression of *CesA4*, *CesA7*, *CesA8*, *IRX9* (glucuronoxylan biosynthesis) and most of the genes of the monolignol biosynthesis pathway (Ranocha *et al.* 2010). In the hypocotyls, the genes involved in cellulose, xylan and lignin deposition were co-expressed with *WAT1* (Chapter 1). Considering the strong cell wall phenotype of the *wat1-1* mutant in *Arabidopsis* and the relationship of *WAT1* with the fibre yield of ramie (Liu *et al.*, 2017), this gene should definitely be more deeply investigated to understand its role in hemp bast fibre development. Cytokinins are involved in the lignification of the fibres and xylem cells (Didi *et al.*, 2015). For instance, the ortholog

of *AtPRX52* has a high number of cytokinin-binding element in its promoter region (Herrero *et al.*, 2014). Abscisic acid (ABA) content in the hypocotyls and the expression of the related genes (biosynthesis and signalling) in both systems suggest that this phytohormone is important for bast fibre thickening. ABA inhibits stem elongation by arranging longitudinally the cortical microtubules (Shibaoka, 1994), leading to the specific pattern of SCW. This may be achieved through the modulation of the expression of the genes involved in microtubule patterning, such as *MAP65-8* and *MAP70-5*, which were highly expressed in H20. Jasmonates regulate the biosynthesis of monolignols and oligolignols (Pauwels *et al.*, 2008). Accordingly, the expression of several methyl jasmonate-inducible genes involved in lignin biosynthesis peaked with the bioactive jasmonate content in hypocotyls (Chapter 1).

What do our two systems (hypocotyl and bast fibre) teach us about hemp stem lignification?

Lignification is important for plant development and affects the mechanical performance of fibre-based composites. For instance, lignin could potentially bind with polypropylene, increasing the interfacial strength of the final polymer, as compared to flax fibres which have less lignin (Pickering *et al.*, 2016). We first investigated how the lignification follows the development of the hypocotyl (Chapter 2). We then tried to unravel, by transcriptomics, the molecular factors involved in the hypolignification of bast fibres (Chapter 4).

The lignin content was significantly lower in elongating hypocotyls (H6 and H9), as compared to thickening hypocotyls. Deposition of lignin was consistent with the up-regulation of genes associated with monolignol biosynthesis and polymerisation (Chapter

1), as well as with the abundance of the corresponding proteins (Chapter 2). Lignin is first deposited within the middle lamella and cell corners (Davin and Lewis, 2000). The polysaccharide composition of the lignifying spots determines the monomeric composition of the polymer. In elongating cells, the apoplast rich in pectins (Phyo *et al.*, 2017) favours the formation of branched lignin, mainly composed of H and G subunits (Terashima and Fukushima, 1989), as shown in Figure 6-1.

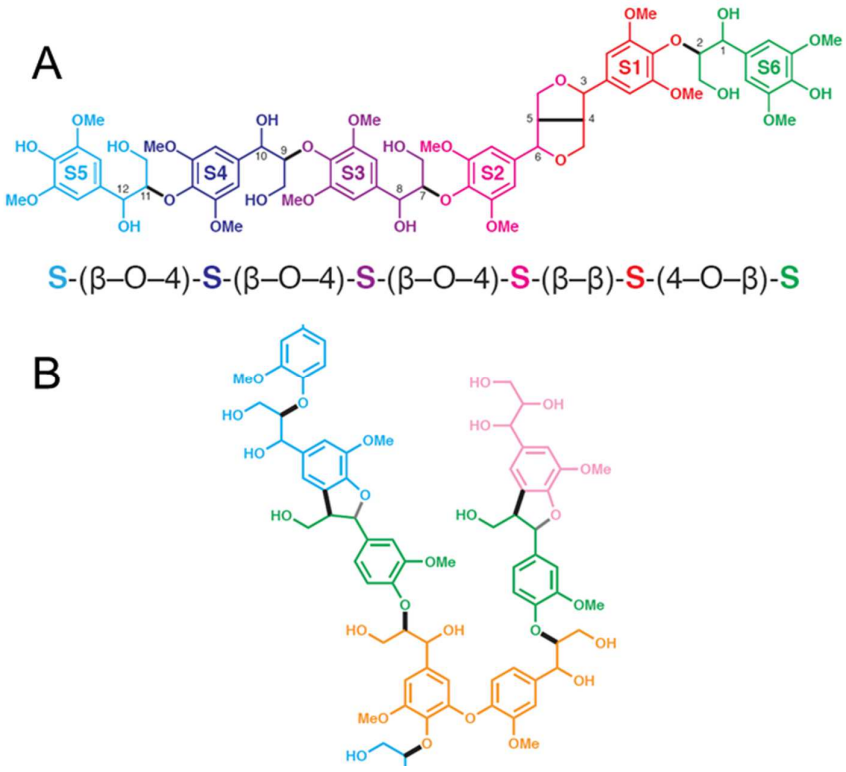


Figure 6-1: Examples of lignin oligomers. In A, a syringyl (S) hexamer, with uncondensed linkages. In B, the lignin is more branched because of the higher proportion of H and G subunits (in orange) and therefore has a higher steric volume. Adapted from Ralph *et al.*, 2008.

The end of elongation marks the onset of SCW deposition and modifications in the monomeric composition of lignin, as observed on the hypocotyls (Chapter 2). The SCW is characterised by the compact structure of the cellulose matrix. The biosynthesis of monolignols is tuned to produce a higher proportion of G and S subunits, which polymerise into long, almost linear non-condensed lignin (Figure 6-1). In this respect, the proportion of H lignin in the cell wall decreases throughout the development of the hypocotyl (Chapter 2). The families of enzymes polymerising monolignols into lignin change during the process. In agreement with the findings of Berthet and colleagues (2011), the *in situ* enzymatic activities suggest that lignin polymerisation is first performed with laccase, and then by peroxidase (Chapter 2). The peroxidase activity was particularly visible in the bast fibres, whose lignin is rich in S-units. These results may be strengthened by comparing the cell wall composition (polysaccharides and specific lignin epitopes) at the ultrastructural level with immunocytochemistry. Lignin signatures assessed by specific antibodies, such as H- or S-enriched epitopes, should be put in perspective with the composition of the extracellular matrix (pectin, xyloglucan, xylan and cellulose) to prove the link between these classes of cell wall molecules.

The hypolignification of bast fibres is clearly associated with a low transcription rate of genes involved in monolignol biosynthesis as compared to the xylem, where xylan-type SCW are found (Chapter 3). With this thesis, we aimed at identifying some molecular actors regulating this differential lignification. A first hypothesis is that the lower lignification of bast fibres is due to the transcriptional repressor *MYB4* (MYB subgroup 4), because this gene was expressed, even though at a low level (< 2 RPKM), in isolated bast fibres (Chapter 4). A

motif putatively recognised by MYB3 (also a MYB subgroup 4) was found in the promoter region of *FLA*s highly expressed in old stem region, which undergoes SCW deposition (Chapter 3). A repressor belonging to the same MYB subgroup 4 was significantly more expressed in flax stem outer tissue *vs* inner tissue (Huis *et al.*, 2012). The second hypothesis suggested by transcriptomics points to the involvement of heterodimers of homeobox transcription factors (Chapter 4). Such heterodimers are built through the interaction between a KNOX protein (such as KNAT3) and a BELL protein (such as BEL1, BLH1 or BLH2). The expression levels of these genes are far higher than *MYB4*, ranging from 4 RPKM (*BLH2*) to > 370 RPKM (*BLH1*). A heterodimer with the same KNOX / BELL structure represses the expression of some lignin biosynthetic genes (*COMT1*, *CCoAOMT* and *AtPRX2)* during leaf morphogenesis (Hay and Tsiantis, 2010). The hypolignification of hemp bast fibres may also be regulated at the post-translational level: we found out that three genes, annotated as F-box proteins kelch-repeat (such as *KFB39*), were up-regulated in young hypocotyls (Chapter 1) and in thickening bast fibres (Chapter 4). In *Arabidopsis*, this group of proteins regulates the ubiquitination and subsequent degradation of the four PAL isozymes, controlling the flux of the phenylpropanoid pathway (Zhang and Liu, 2015). These data need to be strengthened by comparing the expression profile of those genes in cortical and inner tissues and by functional analysis.

Which roles for FLAs during stem elongation and thickening?

FLAs are cell wall proteins linked to cell wall deposition and stem development in many species, from herbaceous (thale cress; MacMillan, 2010) to woody (poplar; Wang *et al.*, 2015). However, their

functions in this context are at the moment not fully known and understood. Cell wall proteins form, together with polysaccharides and lignin, a complex interactive network known as the extracellular matrix (Johnson *et al.*, 2003). The cell adhesion properties conferred by the fasciclin region of these proteins suggest that they play an important role in the organisation of the extracellular matrix. The presence of GPI regions may participate in the formation of membrane rafts and in the molecular mobility at the cell surface / cell wall interface (Johnson *et al.*, 2003). These lipid rafts may interact with the cytoskeleton to stabilise cell-cell adhesion (Johnson *et al.*, 2003). The differential expression of *FLA*s during stem and hypocotyl development may result from the evolution of the development stage of the cell, from elongation to cell wall thickening (Chapter 3). The cell surface is indeed drastically changed during the development of the internodes: for example, the cell wall composition and thickness are strongly modified during development. This organisation is partially controlled by the cytoskeleton (Bashline *et al.*, 2014). We may speculate that, being at the cell surface, FLAs may partake in this process. It is particularly relevant that, in our dataset (Chapter 3), the hierarchical clustering was able to discriminate *FLA*s that are more expressed in a specific region (ASP to BSP, inner or outer part). The ortholog of *AtFLA2* was more expressed in H6 while those from *AtFLA11* and *AtFLA12* were strongly up-regulated in H15 and H20 (Chapter 1). From these observations, it seems that the expression of a specific *FLA* is related to the nature of the cell wall. In addition, a group was transiently up-regulated at the snap point (*CsaFLA2*, *CsaFLA6* and *CsaFLA24*) and thus may play a role in the transition to SCW deposition. At this stage, the cytoskeleton is deeply reorganised to modify the pattern of cellulose deposition (Bashline *et al.*, 2014). Such a transcriptional tuning takes place in the tension wood of woody species, such as poplar, where several *FLA*s

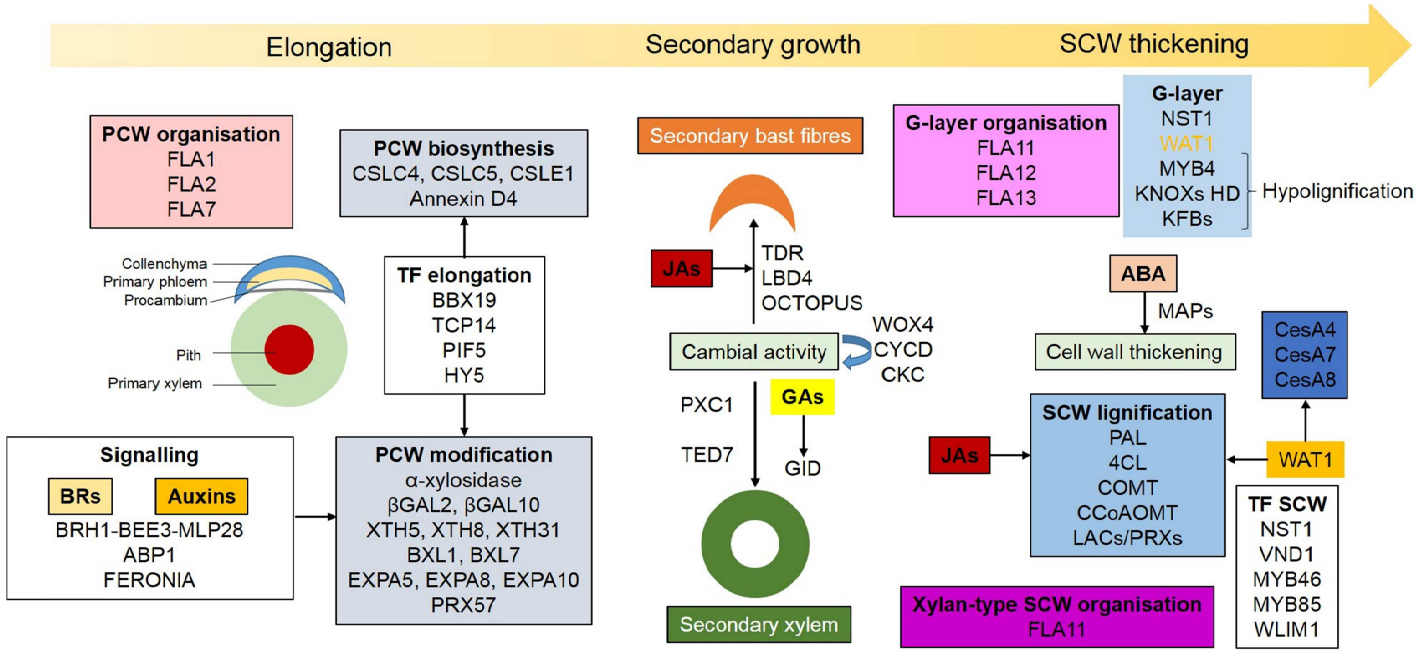
are up-regulated during elongation or thickening of fibre under tensional stress (Lafarguette *et al.*, 2004; Andersson-Gunneråas *et al.*, 2006). Interestingly, FLAs are also associated with the development of bast fibre of other crops, such as flax (Roach and Deyholos, 2007) or jute (Chakraborty *et al.*, 2015). In jute, CcFLA6 (which is closely related to CsaFLA3 and AtFLA11) may coordinate the deposition of the SCW during the bast fibre transition from elongation to thickening (Chakraborty *et al.*, 2015). The presence of FLAs in these distant species indicate that they play a conserved role in the biogenesis of SCW.

FLAs might also be involved in triggering a cellular signal inducing the formation of the G-layer, via the cleavage of their *N*-acetylglucosamine oligosaccharides [$(GlcNAc)_n$] by the action of chitinases (Lafarguette *et al.*, 2004; Mokshina *et al.*, 2014). The expression of specific chitinases is up-regulated during the formation of the G-layer, both in flax (Mokshina *et al.*, 2014) and hemp bast fibres (Chapters 1 and 4) and in poplar tension wood (Lafarguette *et al.*, 2004). These oligosaccharides may lead to perturbations in the noncovalent interactions in the cell wall or at the cell surface interface (Vanholme *et al.*, 2014). They compete in carbohydrate-carbohydrate and carbohydrate-protein interactions, leading to modifications of the cell wall structure (Vanholme *et al.*, 2014). Non-covalent interactions are of crucial importance for the cell wall structure, as demonstrated with xyloglucan-cellulose (Park and Cosgrove, 2015) and xylan-cellulose (Grantham *et al.*, 2017) interactions. It would make sense that proteins located at the cell surface, such as FLAs, trigger, upon the specific expression of *e.g.* chitinases, modifications in the extracellular matrix.

Conclusion

With the aim to provide a global overview of the hemp stem development and of the molecular actors (genes, proteins and phytohormones) partaking in this process, a scheme summarizing the events related to primary and secondary growth, xylem and bast fibre development is hereafter provided.

Figure 6-2: Simplified model of hemp stem development. The main genes, proteins and phytohormones (auxins, brassinosteroids, gibberellic acid, jasmonic acid and abscisic acid) involved in elongation, secondary growth and secondary cell wall thickening in hemp are depicted. Biological processes related to primary (PCW) and xylan-type and gelatinous-type secondary cell wall (SCW) formation, composition and modification are highlighted. The names of genes and proteins follow the *Arabidopsis* nomenclature. For each stage of development, the most important transcription factors (TF) are indicated. Abbreviations as in the text.


Elongation
Secondary growth
SCW thickening
PCW organisation
FLA1
FLA2
FLA7
PCW biosynthesis
CSLC4, CSLC5, CSLE1
Annexin D4
Collenchyma
Primary phloem
Procambium
Pith
Primary xylem
TF elongation
BBX19
TCP14
PIF5
HY5
Signalling
BRs
Auxins
BRH1-BEE3-MLP28
ABP1
FERONIA
PCW modification
α-xylosidase
βGAL2, βGAL10
XTH5, XTH8, XTH31
BXL1, BXL7
EXPA5, EXPA8, EXPA10
PRX57
Secondary bast fibres
JAs
TDR
LBD4
OCTOPUS
Cambial activity
WOX4
CYCD
CKC
PXC1
TED7
GAs
GID
Secondary xylem
G-layer organisation
FLA11
FLA12
FLA13
G-layer
NST1
WAT1
MYB4
KNOXs HD
KFBs
Hypolignification
ABA
MAPs
Cell wall thickening
CesA4
CesA7
CesA8
JAs
SCW lignification
PAL
4CL
COMT
CCoAOMT
LACs/PRXs
WAT1
TF SCW
NST1
VND1
MYB46
MYB85
WLIM1
Xylan-type SCW organisation
FLA11

Perspectives

Several important questions related to bast fibre development have not been addressed by this thesis. Considering the abundance and the peculiar organisation of cellulose in this tissue, it will be of primary interest to identify the cellulose synthases involved in this process, especially during G-layer deposition. The transcriptomic profiles presented in Chapter 4 provided a partial picture on the role of *CesA* genes. In flax, two recent articles have provided different methods to reach this objective. Mokshina and colleagues (2017) have performed RNA-Seq and RT-qPCR analyses in cellulosic fibres at different stages of development or subjected to gravistimulation. In both cases, the bast fibres were undergoing G-layer deposition. By comparing the gene expression profiles between the samples, the authors concluded that the G-layer cellulose may be deposited by two functional cellulose synthase complexes. In the second study, Chantreau and colleagues (2015) have performed functional analyses of cellulose synthases in stem tissues. They have investigated the impact of silencing the expression of the cellulose synthase genes and found out that *LuCesA1* and *LuCesA6* silencing have marked effects on the formation and thickening of bast fibres. It is also highlighted that the expression of *CesA*s is coregulated, as the silencing of a specific *CesA* results in a down-regulation of other *CesA*s, in accordance with the structure of the cellulose synthase complexes. In hemp, there is to date no detailed analysis of the genes of the cellulose synthase family. We can take advantage of our transcriptome assembly to obtain preliminary data about the identity of these genes. Additional investigation in the draft genome of hemp released by van Bakel and coworkers (2011) may further deepen this analysis. As long as no reliable protocol is developed for the transformation of hemp, the investigation for the

cellulose synthases involved in bast fibre development will be based on gene expression profile or patterns of protein abundance, assessed by proteomics (Chabi *et al.*, 2017) or by immunohistochemistry (Cai *et al.*, 2011). Functional studies in heterologous systems (thale cress, tobacco) may also be performed.

Many aspects of the biology and chemistry of hemp bast fibres are still unknown. Numerous studies described the chemical composition of flax G-layer, highlighting its complexity (Gorshkova *et al.*, 1996, 2000; His *et al.*, 2001; Morvan *et al.*, 2003; Huis *et al.*, 2012; Mikshina *et al.*, 2015a-b). The chemical characterisation of hemp bast fibres may result in direct applications in the industry, regarding resistance to decay, mechanical properties or plasticity. The lignin content has direct consequences on the quality of technical fibre (Fernandez-Tendero *et al.*, 2017). In this respect, we have provided the biological context of lignification in the hemp hypocotyl (Chapter 2), but a specific analysis of the lignin and monolignol-derived products present in the bast fibres may provide new insights in this respect. Genetic manipulation may be useful to study the functions of genes related to lignification. First of all, it will be of primary interest to decipher the molecular regulation of the hypolignification of the hemp bast fibres. Targeted transformation may be performed by a large panel of methods. For instance, a protocol has been recently set up to regenerate hemp shoots from cotyledons (Chaohua *et al.*, 2016), providing the opportunity for growing plant with stable transformation. In addition to reverse genetics characterisation of genes such as *KNOX*s, *MYB4* or *KFB*s, forward genetics analysis of mutant lines with altered lignin phenotype (such as highly lignified tissue) is a promising method. It has been used in flax with the same objective (Chantreau *et al.*, 2013, 2014). To this end, a population of mutants with different

phenotypes is screened to determine the most suitable candidate for downstream analysis, such as transcriptomics. This method has the advantage to overcome the *Agrobacterium*-mediated -transformation and regeneration of hemp plants, but is time-consuming and requires several steps of optimisation for the production of the mutant population with harmful chemicals, such as ethane methylsulphonate.

The non-cellulosic polysaccharides of the gelatinous cell walls display a huge diversity of molecules, from pectins to hemicelluloses. For instance, the rhamnogalacturonan matrix enrobing cellulose in the G-layer does not display the same composition as in flax. The G-layer of flax reacts with the LM5 antibody (Roach *et al.*, 2011), but not the hemp bast fibres (Chapter 1). As rhamnogalacturonan I is one of the most complex molecule in nature (Mikshina *et al.*, 2015a), this is not surprising that it shows a different structure between these two species. In the RG-I backbone, 20 to 96 % of the rhamnose residues are substituted with a side chain comprising galactans, arabinans, arabinogalactans or xylogalacturonans (Mikshina *et al.*, 2015a). The molecular composition of RG-I is unique for each species, and sometimes for each tissue, in order to feat with the requirements of the plant (such as retain water or interact with other polysaccharides; Mikshina *et al.*, 2015b). A multitude of methods are available to study the cell wall composition, from simple acid hydrolysis (Liu *et al.*, 2015) to fractionated extraction and complex monosaccharide linkage analysis (Pettolino *et al.*, 2012). The pectic fraction mainly extracted from the middle lamellae of hemp bast fibres have been previously characterised (Vignon and Garcia-Jaldon, 1996), but not the polysaccharide fraction from the G-layer. The elastic properties of RG-I within the gelatinous cell wall has been investigated in flax (Mikshina *et al.*, 2015b), as well as its chemical composition (Makshakova *et al.*,

2017). Defining these mechanical properties and their underlying physicochemical network in hemp bast fibre may open avenues for industrial applications and contribute to the global understanding of gelatinous cell walls.

Quantitative genetic is a powerful tool to design cultivars with desired traits, such as fibre yield, fibre quality or seed production. For instance, it has been shown in hemp that sex determination is linked to quantitative trait loci (QTLs, Faux *et al.*, 2016). Cloning the genes responsible for these QTLs will support the breeding effort towards the creation of new cultivars with enhanced traits. A high-density genetic map of ramie has evidenced five cellulose synthase genes and three *WAT1*-related genes which are specifically expressed in the bast fibres, suggesting their importance for reaching a high bast fibre yield (Liu *et al.*, 2017). Such molecular maps have been produced in hemp, but they are non-saturated (Ranalli, 2004) and should therefore be improved.

This thesis has already evidenced some genetic markers which seem to be important for stem development, and therefore straw yield, such as *WAT1*, *NST1* or *CesA*s. In addition to functional characterisation, mapping these genes against traits of interest will advance the genetic basis for the design of future cultivars.

References

Abel, S., Nguyen, M.D., Chow, W., and Theologis, A. (1995). *ASC4*, a Primary Indoleacetic Acid-responsive Gene Encoding 1-Aminocyclopropane-1-carboxylate Synthase in *Arabidopsis thaliana*: STRUCTURAL CHARACTERIZATION, EXPRESSION IN *ESCHERICHIA COLI*, AND EXPRESSION CHARACTERISTICS IN RESPONSE TO AUXIN. *J. Biol. Chem.* 270**,** 19093-19099.

Aloni, R. (1979). Role of Auxin and Gibberellin in Differentiation of Primary Phloem Fibers. *Plant Physiol.* 63**,** 609-614.

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

Bashline, L., Li, S., and Gu, Y. (2014). The trafficking of the cellulose synthase complex in higher plants. *Ann. Bot.* 114**,** 1059-1067.

Berthet, S., Demont-Caulet, N., Pollet, B., Bidzinski, P., Cézard, L., Le Bris, P., *et al.* (2011). Disruption of LACCASE4 and 17 Results in Tissue-Specific Alterations to Lignification of *Arabidopsis thaliana* Stems. *Plant Cell* 23**,** 1124-1137.

Cai, G., Faleri, C., Del Casino, C., Emons, A.M., and Cresti, M. (2011). Distribution of Callose Synthase, Cellulose Synthase, and Sucrose Synthase in Tobacco Pollen Tube Is Controlled in Dissimilar Ways by Actin Filaments and Microtubules. *Plant Physiol.* 155**,** 1169-1190.

Chabi, M., Goulas, E., Leclercq, C.C., De Waele, I., Rihouey, C., Cenci, U., *et al.* (2017). A Cell Wall Proteome and Targeted Cell Wall Analyses Provide Novel Information on Hemicellulose Metabolism in Flax. *Mol. Cell. Proteomics* 16**,** 1634-1651.

Chakraborty, A., Sarkar, D., Satya, P., Karmakar, P.G., and Singh, N.K. (2015). Pathways associated with lignin biosynthesis in lignomaniac jute fibres. *Mol. Genet. Genomics* 290**,** 1523-1542.

Chantreau, M., Grec, S., Gutierrez, L., Dalmais, M., Pineau, C., Demailly, H., *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., *et al.* (2014). Ectopic Lignification in the Flax lignified bast fiber1 Mutant Stem Is Associated with Tissue-Specific Modifications in Gene Expression and Cell Wall Composition. *Plant Cell* 26**,** 4462-4482.

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.* 9, 1312-1324.

Chaohua, C., Gonggu, Z., Lining, Z., Chunsheng, G., Qing, T., Jianhua, C., *et al.* (2016). A rapid shoot regeneration protocol from the cotyledons of hemp (*Cannabis sativa* L.). *Ind. Crop. Prod.* 83**,** 61-65.

Crozier, A., Kamiya, Y., Bishop, G., and Yokota, T. (2000). "Biosynthesis of Hormones and Elicitor Molecules," in *Biochemistry and Molecular Biology of Plants,* eds. B. Buchanan, W. Gruissem & R. Jones. American Society of Plant Physiologists, 850-929.

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

Dayan, J., Schwarzkopf, M., Avni, A., and Aloni, R. (2010). Enhancing plant growth and fiber production by silencing *GA 2-oxidase*. *Plant Biotechnol. J.* 8**,** 425-435.

Didi, V., Jackson, P., and Hejátko, J. (2015). Hormonal regulation of secondary cell wall formation. *J. Exp. Bot.* 66**,** 5015-5027.

Faux, A.-M., Draye, X., Flamand, M.-C., Occre, A., and Bertin, P. (2016). Identification of QTLs for sex expression in dioecious and monoecious hemp (*Cannabis sativa* L.). *Euphytica* 209**,** 357-376.

Fernandez-Tendero, E., Day, A., Legros, S., Habrant, A., Hawkins, S., and Chabbert, B. (2017). Changes in hemp secondary fiber production related to

technical fiber variability revealed by light microscopy and attenuated total reflectance Fourier transform infrared spectroscopy. *PLoS One* 12**,** e0179794.

Fike, J. (2016). Industrial Hemp: Renewed Opportunities for an Ancient Crop. *Crit. Rev. Plant Sci.* 35**,** 406-424.

Gorshkova, T., Brutch, N., Chabbert, B., Deyholos, M., Hayashi, T., Lev-Yadun, S., *et al.* (2012). Plant Fiber Formation: State of the Art, Recent and Expected Progress, and Open Questions. *Crit. Rev. Plant Sci.* 31**,** 201-228.

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

Gorshkova, T.A., Salnikov, V.V., Pogodina, N.M., Chemikosova, S.B., Yablokova, E.V., Ulanov, A.V., *et al.* (2000). Composition and Distribution of Cell Wall Phenolic Compounds in Flax (Linum usitatissimum L.) Stem Tissues. *Ann. Bot.* 85**,** 477-486.

Gorshkova, T.A., Wyatt, S.E., Salnikov, V.V., Gibeaut, D.M., Ibragimov, M.R., Lozovaya, V.V., *et al.* (1996). Cell-Wall Polysaccharides of Developing Flax Plants. *Plant Physiol.* 110**,** 721-729.

Grantham, N.J., Wurman-Rodrich, J., Terrett, O.M., Lyczakowski, J.J., Stott, K., Iuga, D., *et al.* (2017). An even pattern of xylan substitution is critical for interaction with cellulose in plant cell walls. *Nature Plants* 3**,** 859-865.

Guerriero, G., Hausman, J.-F., Strauss, J., Ertan, H., and Siddiqui, K.S. (2016). Lignocellulosic biomass: Biosynthesis, degradation, and industrial utilization. *Eng. Life Sci.* 16**,** 1-16.

Hay, A., and Tsiantis, M. (2010). KNOX genes: versatile regulators of plant development and diversity. *Development* 137**,** 3153-3165.

Hedden, P., and Sponsel, V. (2015). A Century of Gibberellin Research. *J. Plant Growth Regul.* 34**,** 740-760.

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. *J. Histochem. Cytochem.* 49**,** 1525-1535.

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

Johnson, K.L., Jones, B.J., Bacic, A., and Schultz, C.J. (2003). The Fasciclin-Like Arabinogalactan Proteins of *Arabidopsis*. A Multigene Family of Putative Cell Adhesion Molecules. *Plant Physiol.* 133**,** 1911-1925.

Keegstra, K. (2010). Plant Cell Walls. *Plant Physiol.* 154**,** 483-486.

Ko, J.-H., Han, K.-H., Park, S., and Yang, J. (2004). Plant Body Weight-Induced Secondary Growth in *Arabidopsis* and Its Transcription Phenotype Revealed by Whole-Transcriptome Profiling. *Plant Physiol.* 135**,** 1069-1083.

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

Liu, M., Fernando, D., Daniel, G., Madsen, B., Meyer, A.S., Ale, M.T., *et al.* (2015). Effect of harvest time and field retting duration on the chemical composition, morphology and mechanical properties of hemp fibers. *Indus. Crop. Prod.* 69**,** 29-39.

Liu, C., Zeng, L., Zhu, S., Wu, L., Wang, Y., Tang, S., *et al.* (2017). Draft genome analysis provides insights into the fiber yield, crude protein biosynthesis, and vegetative growth of domesticated ramie (*Boehmeria nivea* L. Gaud). *DNA Research***,** dsx047-dsx047.

Macmillan, C.P., Mansfield, S.D., Stachurski, Z.H., Evans, R., and Southerton, S.G. (2010). Fasciclin-like arabinogalactan proteins: specialization for stem biomechanics and cell wall architecture in *Arabidopsis* and *Eucalyptus*. *Plant J.* 62**,** 689-703.

Makshakova, O.N., Gorshkova, T.A., Mikshina, P.V., Zuev, Y.F., and Perez, S. (2017). Metrics of rhamnogalacturonan I with β-(1→4)-linked galactan side chains and structural basis for its self-aggregation. *Carbohydr. Polym.* 158**,** 93-101.

Mikshina, P.V., Petrova, A.A., and Gorshkova, T.A. (2015a). Functional diversity of rhamnogalacturonans I. *Russ. Chem. B+.* 64**,** 1014-1023.

Mikshina, P.V., Petrova, A.A., Faizullin, D.A., Zuev, Y.F., and Gorshkova, T.A. (2015b). Tissue-Specific Rhamnogalacturonan I Forms the Gel with Hyperelastic Properties. *Biochem. (Mosc.)* 80**,** 915-924.

Mokshina, N., Gorshkov, O., Ibragimova, N., Chernova, T., and Gorshkova, T. (2017). Cellulosic fibres of flax recruit both primary and secondary cell wall cellulose synthases during deposition of thick tertiary cell walls and in the course of graviresponse. *Funct. Plant Biol.* 44**,** 820-831.

Mokshina, N., Gorshkova, T., and Deyholos, M.K. (2014). Chitinase-Like and Cellulose Synthase Gene Expression in Gelatinous-Type Cellulosic Walls of Flax (*Linum usitatissimum* L.) Bast Fibers. *PLoS One* 9**,** e97949.

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.

Nakagawa, K., Yoshinaga, A., and Takabe, K. (2014). Xylan deposition and lignification in the multi-layered cell walls of phloem fibres in *Mallotus japonicus* (Euphorbiaceae). *Tree Physiol.* 34**,** 1018-1029.

Nieminen, K., Blomster, T., Helariutta, Y., and Mähönen, A.P. (2015). Vascular Cambium Development. *The Arabidopsis Book***,** e0177.

Oh, E., Zhu, J.Y., Bai, M.Y., Arenhart, R.A., Sun, Y., and Wang, Z.Y. (2014). Cell elongation is regulated through a central circuit of interacting transcription factors in the *Arabidopsis* hypocotyl. *eLife* 3**,** e03031.

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.

Pauwels, L., Morreel, K., De Witte, E., Lammertyn, F., Van Montagu, M., Boerjan, W., *et al.* (2008). Mapping methyl jasmonate-mediated transcriptional reprogramming of metabolism and cell cycle progression in cultured *Arabidopsis* cells. *Proc. Natl. Acad. Sci. U. S. A.* 105**,** 1380-1385.

Pettolino, F.A., Walsh, C., Fincher, G.B., and Bacic, A. (2012). Determining the polysaccharide composition of plant cell walls. *Nat. Protoc.* 7**,** 1590-1607.

Phyo, P., Wang, T., Kiemle, S.N., O'neill, H., Pingali, S.V., Hong, M., *et al.* (2017). Gradients in wall mechanics and polysaccharides along growing inflorescence stems. *Plant Physiol.*

Pickering, K.L., Efendy, M.G.A., and Le, T.M. (2016). A review of recent developments in natural fibre composites and their mechanical performance. *Compos. Part A Appl. Sci. Manuf.* 83**,** 98-112.

Pierik, R., Djakovic-Petrovic, T., Keuskamp, D.H., De Wit, M., and Voesenek, L.a.C.J. (2009). Auxin and Ethylene Regulate Elongation Responses to Neighbor Proximity Signals Independent of Gibberellin and DELLA Proteins in *Arabidopsis*. *Plant Physiol.* 149**,** 1701-1712.

Ralph, J., Brunow, G., Harris, P.J., Dixon, R.A., Schatz, P.F., and Boerjan, W. (2008). "Lignification: are Lignins Biosynthesized via simple Combinatorial Chemistry or via Proteinaceous Control and Template Replication?" in *Recent Advances in Polyphenol Research*. Wiley-Blackwell, 36-66.

Ranalli, P. (2004). Current status and future scenarios of hemp breeding. *Euphytica* 140**,** 121-131.

Ranocha, P., Denancé, N., Vanholme, R., Freydier, A., Martinez, Y., Hoffmann, L., *et al.* (2010). Walls are thin 1 (WAT1), an *Arabidopsis* homolog

of *Medicago truncatula* NODULIN21, is a tonoplast-localized protein required for secondary wall formation in fibers. *Plant J.* 63**,** 469-483.

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.

Roach, M.J., and Deyholos, M.K. (2008). Microarray Analysis of Developing Flax Hypocotyls Identifies Novel Transcripts Correlated with Specific Stages of Phloem Fibre Differentiation. *Ann. Bot.* 102**,** 317-330.

Roach, M.J., Mokshina, N.Y., Badhan, A., Snegireva, A.V., Hobson, N., Deyholos, M.K., *et al.* (2011). Development of Cellulosic Secondary Walls in Flax Fibers Requires ß-Galactosidase. *Plant Physiol.* 156**,** 1351-1363.

Sehr, E.M., Agusti, J., Lehner, R., Farmer, E.E., Schwarz, M., and Greb, T. (2010). Analysis of secondary growth in the *Arabidopsis* shoot reveals a positive role of jasmonate signalling in cambium formation. *Plant J.* 63**,** 811-822.

Shibaoka, H. (1994). Plant Hormone-Induced Changes in the Orientation of Cortical Microtubules: Alterations in the Cross-linking Between Microtubules and the Plasma Membrane. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 45**,** 527-544.

Terashima, N., and Fukushima, K. (1989). "Biogenesis and Structure of Macromolecular Lignin in the Cell Wall of Tree Xylem as Studied by Microautoradiography," in *Plant Cell Wall Polymers*. 399 ed: American Chemical Society), 160-168.

Ursache, R., Nieminen, K., and Helariutta, Y. (2013). Genetic and hormonal regulation of cambial development. *Physiol. Plant.* 147**,** 36-45.

Van Bakel, H., Stout, J.M., Cote, A.G., Tallon, C.M., Sharpe, A.G., Hughes, T.R., *et al.* (2011). The draft genome and transcriptome of *Cannabis sativa*. *Genome Biol.* 12**,** R102-R102.

Vanholme, B., Vanholme, R., Turumtay, H., Goeminne, G., Cesarino, I., Goubet, F., *et al.* (2014). Accumulation of *N*-Acetylglucosamine Oligomers in the Plant Cell Wall Affects Plant Architecture in a Dose-Dependent and Conditional Manner. *Plant Physiol.* 165**,** 290-308.

Vignon, M.R., and Garcia-Jaldon, C. (1996). Structural features of the pectic polysaccharides isolated from retted hemp bast fibres. *Carbohydr. Res.* 296**,** 249-260.

Wang, H., Jiang, C., Wang, C., Yang, Y., Yang, L., Gao, X., *et al.* (2015). Antisense expression of the fasciclin-like arabinogalactan protein FLA6 gene in *Populus* inhibits expression of its homologous genes and alters stem biomechanics and cell wall composition in transgenic trees. *J. Exp. Bot.* 66**,** 1291-1302.

Wang, H., Jin, Y., Wang, C., Li, B., Jiang, C., Sun, Z., *et al.* (2017). Fasciclin-like arabinogalactan proteins, PtFLAs, play important roles in GA-mediated tension wood formation in *Populus*. *Sci. Rep.* 7**,** 6182.

Zhong, R., and Ye, Z.H. (2015). Secondary Cell Walls: Biosynthesis, Patterned Deposition and Transcriptional Regulation. *Plant Cell Physiol.* 56**,** 195-214.

Appendix

Impact of jasmonic acid application on the hemp hypocotyl

Shoot primary growth results from the activity of the shoot apical meristem and leads to the development of primary tissues in stem, leaf and hypocotyl. At the end of the elongation phase, further development of the stem and hypocotyl occurs by secondary growth. Secondary growth results from the activity of the vascular cambium. The cambium initials divide periclinally to produce xylem (inwards) and phloem (outwards) mother cells, and anticlinally to create new initials to follow the increased circumference of the stem (Nieminen *et al.*, 2015). This process is particularly obvious in perennial species such as trees. A high-spatial-resolution RNA-Seq analysis identified gene networks and modules associated with vascular cambium, secondary phloem and secondary xylem development in aspen (Sundell *et al.*, 2017). Secondary growth is also occurring in some annual herbaceous plants, such as thale cress, and has been studied both in the inflorescence stem (Ko *et al.*, 2004) and in the hypocotyl (Ragni and Hardtke, 2014). In the hypocotyl system, secondary growth occurs quickly, roughly 10 days after sowing, and elongation growth and secondary thickening are temporally uncoupled (Baucher *et al.*, 2007; Roach and Deyholos, 2008; Chapter 1). Regulation of secondary growth is achieved through a complex network of genes and phytohormones (Nieminen *et al.*, 2015). For instance, procambium proliferation and cambium identity are tuned by the dodecapeptide TRACHEARY ELEMENT

DIFFERENTIATION INHIBITORY FACTOR (TDIF) which is secreted from the phloem (Nieminen *et al.*, 2015). TDIF induces the expression of the transcription factor WUSCHEL-RELATED HOMEOBOX4 (WOX4).

Phytohormones are key molecules regulating secondary growth (Strabala and MacMillan, 2013; Ursache *et al.*, 2013). The role of auxin has been long proven. The development of the vascular cambium is severely affected when the shoot tip is decapitated, causing the loss of the main auxin source in the plant. In addition, secondary growth triggered by the perception of a weight signal induces the expression of many genes having auxin responsive elements in their promoter region, possibly indicating a role for auxin in this process (Ko *et al.*, 2004). Auxin has also been shown to promote the cambial cell division by TDIF-WOX4 and other signalling pathways (Nieminen *et al.*, 2015). Mutants impaired in cytokinin biosynthesis have strong phenotypes, with stem dwarfism and greatly reduced diameter (Nieminen *et al.*, 2015). Ethylene too has a positive effect on cambial activity. Increased ethylene biosynthesis leads to increased cambial activity while this activity is reduced in a loss-of-function mutant of a gene positively involved in ethylene signalling. However, no phenotype was observed in two mutants with impaired ethylene signalling, suggesting the possibility of another molecular regulation (Nieminen *et al.*, 2015).

Floral induction and/or growth of the inflorescence stem triggers secondary growth in the hypocotyl by basipetally translocated gibberellins (Strabala and MacMillan, 2013). Finally, jasmonates positively regulate secondary growth through the canonical signalling SCF^{COI1} pathway. Silencing the negative regulator of jasmonate signalling JAZ10 results in an earlier and more pronounced

interfascicular cambium activity in *Arabidopsis*, and exogenous application of jasmonic acid on stem triggers the formation of phloem fibres (Sehr *et al.*, 2010).

Lignification is the last step of vascular cell differentiation. Lignin impregnates the extracellular matrix of phloem fibres, xylem fibres and xylem vessels (Barros *et al.*, 2015). The timing and the subcellular localisation of lignin deposition are tightly controlled by specific transcription factors (Wang and Dixon, 2012) and interactions between phytohormones (Didi *et al.*, 2015; Nieminen *et al.*, 2015) mirroring the stage of plant development. Since lignin provides mechanical strength and resistance to gravity to the stem, lignification is extremely important for the ability of the plant to cope with these factors. Phytohormones, e.g. auxin and jasmonic acid, mediate the developmental signals into gene expression patterns (Ko *et al.*, 2004; Sehr *et al.*, 2010) and up-regulate specific genes of the monolignol pathway (Pauwels *et al.*, 2008; Tamaoki *et al.*, 2011).

In the stem, fibres are either associated with the xylem tracheary elements and xylem parenchyma or with phloem sieve elements and phloem parenchyma, both of primary (procambium) and secondary (cambium) origins. Xylem fibres carry a xylan-type secondary cell wall and are rich in lignin (Gorshkova *et al.*, 2012), therefore providing interesting properties for the construction industry (*e.g.*, production of “hempcrete”). The phloem fibres possess gelatinous-type cell wall, characterised by their extreme length (up to 55 mm, Snegireva *et al.*, 2015), important thickness and high crystalline cellulose content (up to 75-80%; Guerriero *et al.*, 2013). They are used for the manufacturing of commercial textiles and composite materials (Gorshkova *et al.*, 2012). Secondary growth is an important feature of hemp development. Indeed, secondary bast fibres

produced by cambial activity in fully developed internodes are far more abundant than primary bast fibres (Blake *et al.*, 2008). Secondary bast fibres are shorter and more lignified than the primary ones (Snegireva *et al.*, 2015; Fernandez-Tendero *et al.*, 2017), causing heterogeneity in the industrial quality of hemp technical fibres (Liu *et al.*, 2015).

In Chapter 1, we have highlighted some of the factors involved in hemp secondary growth using the hypocotyl system. The bioactive jasmonate content and the expression of some genes involved in secondary growth (*NST1*, *WOX4*, *PXC1*) increased at the onset of cambial activity and secondary tissues formation, leading to the up-regulation of the secondary cell wall-related genes, and further thickening of the fibres (cellulose, lignin and xylan deposition). The present study aims at investigating more thoroughly the function of jasmonates in this respect. Jasmonates stimulate phloem fibre formation (Sehr *et al.*, 2010) and lignification (Pauwels *et al.*, 2008). Hemp is a species naturally producing a considerable amount of phloem fibres, by contrast with *Arabidopsis*. We thus take advantage of this feature to further study the role of jasmonates in secondary phloem fibre differentiation and overall hypocotyl lignification. Hypocotyls aged 15 days (corresponding to the maximum endogenous jasmonates content, Chapter 1) will be sprayed either with jasmonic acid or a mock solution. The resulting differences in terms of secondary growth and lignification patterns will be presented to shed light on the biological mechanism leading to lignification and secondary bast fibres biogenesis in this important fibre crop.

I. Material and methods

a. Experimental Set-Up

Hemp hypocotyls of cultivar Santhica 27 were grown in a mixture of compost/sand (1:1 w/w) in controlled conditions. A pilot study determined the optimal concentration of jasmonic acid (JA). Either the JA (concentration ranging from 0.03 mM to 3mM) or the mock solution were sprayed on plantlets aged 15 days after sowing (1 $mL.plant^{-1}$). The JA solution (Sigma-Aldrich) was prepared in phosphate buffer saline (PBS) from a 500-fold stock in pure ethanol. A mock solution, consisting in PBS with a minute amount of ethanol (0.06% in the sprayed solution) to compensate for its presence in the jasmonic acid solution, was sprayed on control plantlets. The measurements were carried out on five biological replicates, each consisting of 12 plantlets. For the five concentrations, the lignin content and the hypocotyl diameter were assessed. After measuring their diameter, the hypocotyls were sampled three days after the treatment and used to prepare the cell wall residue (CWR) for lignin analysis.

According to the results of the pilot study, the impact of JA on hypocotyls was further investigated using a concentration of 0.1 mM. Hypocotyls were sampled 1, 2, 3 and 5 days after JA or mock application at 15 days (H16, H17, H18 and H20, respectively). The diameter of the hypocotyls was recorded before each sampling. This experiment was performed in biological triplicate (20-25 hypocotyls for each biological replicate).

b. Lignin analysis

Lignin analysis has been performed as described in Chapter 2, in H17, H18 and H20. Briefly, lignin content was determined with the acetyl

bromide method using 5 mg of dried CWR. The CWR was digested with 2.6 mL of 25 % acetyl bromide in glacial acetic acid for 2 hours at 50°C using a Hach LT200 system. At the end of the reaction, the solution was transferred to a 50 mL Falcon tube containing 10 mL of 2 M sodium hydroxide and 12 mL of glacial acetic acid. After rinsing the reaction tube with glacial acetic acid, 1.75 mL of 0.5 M hydroxylammonium chloride was added and the total volume adjusted to 30 mL with glacial acetic acid (Hatfield *et al.*, 1999). The absorbance at 280 nm was read with a spectrophotometer. An extinction coefficient of 22.9 g^{-1} L cm^{-1} was used to calculate the lignin content.

Lignin composition has been assessed by nitrobenzene oxidation (NBO) or pyrolysis (py) coupled to GC-MS. The NBO was performed as described in Chapter 2. 5 mg of CWR were digested with 2 mL of 2 M NaOH and 30 µL nitrobenzene at 165°C for one hour (Hach LT200 system). Pyrolysis was performed on ca. 1 mg of CWR at a temperature of 700°C for 30 sec. The thermal desorption unit ranged from 50°C to 300°C with a thermal ramp of 720°C/min, on splitless mode. The cooled injection system was kept at 300°C, with a split ratio of 50:1. The GC transfer line was kept at 300°C. The GC oven program was as follow: 50°C for 3 min, ramp at 5°C/min to 180°C and 10°C/min to 320°C (for 5 min). Detection was done with a mass spectrometer (MS). Peaks specific for *p*-hydroxyphenyl (H, 2 MS peaks), guaiacyl (G, 11 MS peaks) and syringyl (S, 10 MS peaks) lignin subunits were investigated. The calculations were done on the relative molar area of each peak. The molar areas were summed separately for H, G and S peaks.

c. RNA extraction and RT-qPCR

Targeted gene expression analysis was performed on H16, H17, H18 and H20 by RT-qPCR. Total RNA was extracted as described in

Chapter 1, using the RNeasy Plant Mini Kit (Qiagen), and quality-checked using a NanoDrop 1000 Spectrophotometer (Thermo Scientific) and a 2100 Bioanalyzer (Agilent Life Sciences). All the RNAs displayed a RIN above 9.3. Reverse transcription was carried out with the ProtoScript II Reverse Transcriptase (NEB) following the manufacturer's instructions. Primers were designed with Primer3 and validated for the absence of dimers and secondary structures. qPCR runs were performed in 384 well-plates, on a ViiA7 Real-Time PCR System (Applied Biosystems) with the Takyon SYBR Green low ROX (Eurogentec). A melt curve was realised at the end of each experiment to ensure the specificity of the products. Relative gene expressions were determined with the qBasePLUS software v2.5 (Biogazelle).

d. Library preparation and sequencing

Two time points (H18 and H20) were selected for transcriptomics based on the results of targeted gene expression. mRNAs were isolated from 5 µg of total RNA using the Magnosphere UltraPure mRNA purification kit (Takara), following manufacturer's instruction. After separation, mRNAs were quality-checked using the Pico RNA assay in the Bioanalyzer. The residual rRNA contamination was below 5 % for all the samples. The first-strand cDNA synthesis, reverse transcription and library amplification were performed with the SMARTer Stranded RNA-Seq kit (Clontech), using 8 µl of mRNA (for a quantity ranging between 24 and 42 ng of mRNA) and following manufacturer's instruction. The enrichment step was carried out using 11 cycles of PCR. The profile of the libraries was evaluated using the High Sensitivity DNA Assay in the Bioanalyzer (concentration ranging from 19 to 39 nM). Indexing was performed using the Illumina indexes 1-12. Quantification was performed using the KAPA library quantification kit (KAPA) with the qPCR system. The pooled libraries

(20 pM) were sequenced on an Illumina MiSeq in 6 consecutive runs (MiSeq reagent kit V3, 150 cycles) generating 75 base pairs (bp) paired-end reads.

e. Assembly, mapping and data analysis

Raw sequences reads were uploaded in CLC Genomics Workbench 9.0.1. Sequences were filtered and trimmed as follows: sequence length > 35 bp, sequence quality score < 0.01, no ambiguity in the sequence, trimming using Illumina adaptors, hard trim of 5 bp at the 5' end and 2 bp at the 3' end. The *de novo* assembly, with sequences from Chapters 1 and 4, was performed with an auto wording size of 20, an auto bubble size of 50 and a minimum contig length of 300 bp. The reads were mapped back to the assembly with a mismatch, insertion and deletion cost of 3, a length fraction and similarity > 0.95. The assembly was then annotated using Blast2GO PRO version 3.0 against the *Arabidopsis thaliana* non-redundant database. For each library, the mapping was performed with CLC Genomics Workbench 9.0.1 according to the following criteria: a maximum hit per reads of 3, similarity fraction and length fraction > 0.95, a mismatch, insertion and deletion cost of 3. The expression values were then calculated using the RPKM method. Genes with less than 10 mapped reads, with no specifically mapped reads in at least one of the libraries were removed from the dataset. In order to highlight the differentially expressed genes, an ANOVA with 4 groups (H18 and H20 mock or JA) was performed.

f. Phytohormone analysis

Phytohormone analysis was performed on H15, H16, H17, H18 and H20 (mock or JA), with the methods described in Chapter 1. Briefly, 10-15 mg of lyophilised material were extracted with methanol/formic

acid/water (15:1:4; v/v) with prior addition of 10 pmol [^{2}H]-labelled internal standards. Extracts were evaporated in vacuum concentrator and purified through mixed mode reversed phase – cation exchange SPE column (Oasis-MCX, Waters). The purification procedure leads to split each sample into two fractions: fraction A, containing acidic hormones; and fraction B, containing basic hormones. Each fraction was separately analysed by LC-MS.

II. Results and discussion

a. Pilot study

The effect of JA application at various concentrations on the diameter, lignin content and lignin composition of hypocotyls was investigated. JA has a positive effect on the diameter and the lignin content of the hypocotyls, with the exception of the 1 mM concentration for which the statistical test did not detect significant changes (Figures A-1 and A-2)

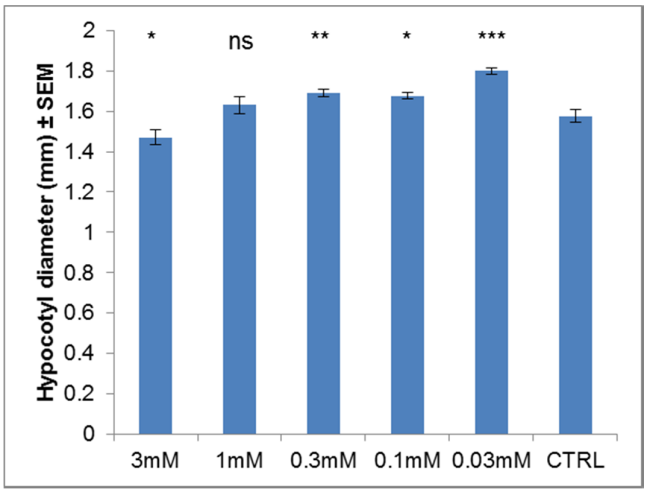


Figure A-1: Impact of JA application on hypocotyl diameter (average ± SEM). A Student's *t*-test is performed for each concentration (*p*-value < 0.05*, 0.01** and 0.001***, ns non-significant).

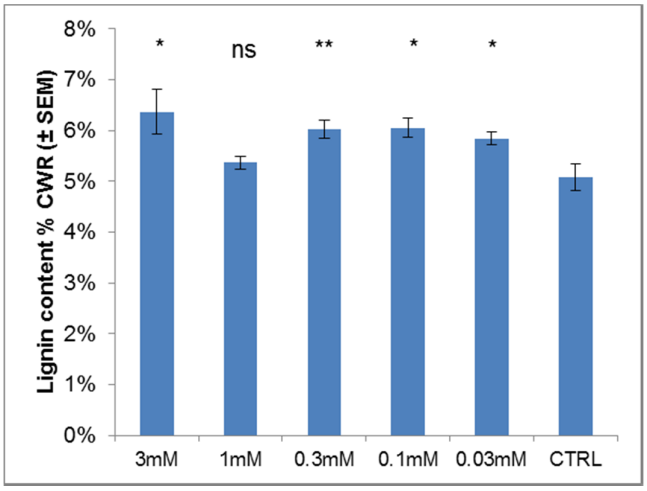


Figure A-2: Impact of JA application on hypocotyl lignin content (average ± SEM). A Student's *t*-test is performed for each concentration (*p*-value < 0.05*, 0.01** and ns non-significant).

These results confirm the positive role of jasmonic acid on secondary growth (Sehr *et al.*, 2010) and monolignol biosynthesis / lignin deposition (Pauwels *et al.*, 2008). To further characterise the impact of JA on lignification, the lignin monomeric composition was investigated by py-GC-MS on the Control and 0.1 mM JA-treated hypocotyls. We observed a slight but significant increase of the S/G ratio in JA-treated hypocotyls (0.59 *vs* 0.53, *t*-test *p*-value = 2.10%), due to a higher molar area of S-related peaks.

b. Phenotype of hypocotyls treated with JA

Hemp plantlets have been sprayed with jasmonic acid at 0.1 mM and sampled after 1, 2, 3 and 5 days. The difference regarding the lignin content in this batch of hypocotyls was minor (Figure A-3), possibly due to the heterogeneity of the plant material.

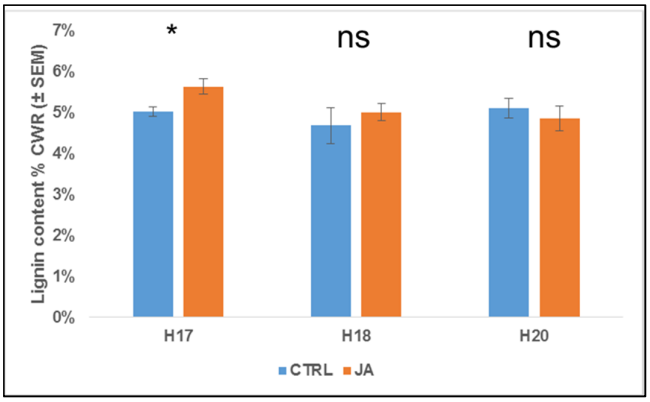


Figure A-3: Impact of JA application (0.1 mM) on hypocotyl lignin content (average ± SEM). A Student's *t*-test is performed for each time-point (* indicates *p*-value < 0.05 and ns non-significant).

No significant differences were observed on the syringyl to guaiacyl ratio (S/G) determined by nitrobenzene oxidation (in H17, H18 and H20, data not shown).

As already observed in the pilot study, the application of JA have increased the diameter of the hypocotyls (Figure A-4).

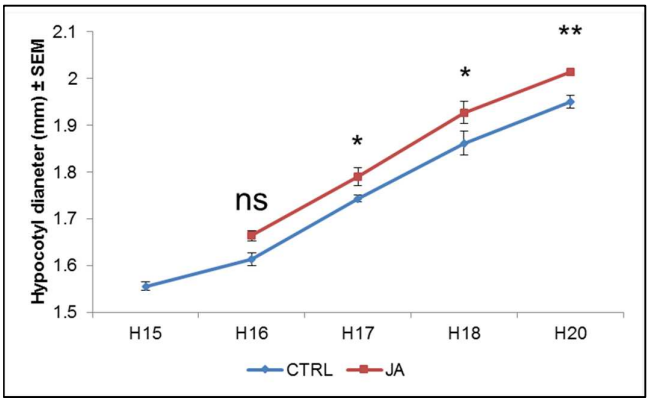


Figure A-4: Impact of JA application (0.1 mM) on hypocotyl diameter (average ± SEM). A Student's *t*-test is performed for each time point (*p*-value < 0.05*, 0.01** and ns non-significant).

In their study, Sehr and colleagues (2010) observed that plants treated with JA have additional phloem fibres. In order to quantitatively characterise this phenomenon, we have counted in an independent experiment the secondary bast fibres in mock- and JA-treated hypocotyls. The average diameter of 32-33 hypocotyls was determined, and a subset of 12 hypocotyls representative of the whole batch was observed by confocal microscopy five days after application (*i.e*., at 20 days). Cross sections of JA-treated plantlets displayed significantly more secondary bast fibres than the mock-treated (505 *vs*. 341, *t*-test *p*-value = 0.93%). Two representative pictures of mock- and JA-treated cross sections are provided Figure A-5.

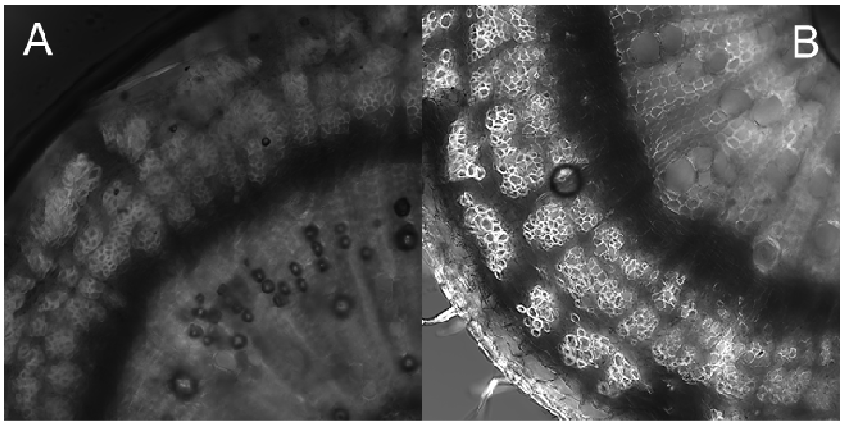


Figure A-5: Representative cross-section of mock- (panel A, *ca*. 330 secondary bast fibres) or JA- (panel B, *ca*. 530 secondary bast fibres) treated hypocotyls.

c. Gene expression analysis

In order to select the 12 samples which will be further investigated by RNA-Seq, a targeted gene expression analysis was performed on 10 genes related to secondary cell wall and lignification on hypocotyls

after JA or mock treatment (H16 to H20). The results are displayed on Figure A-6.

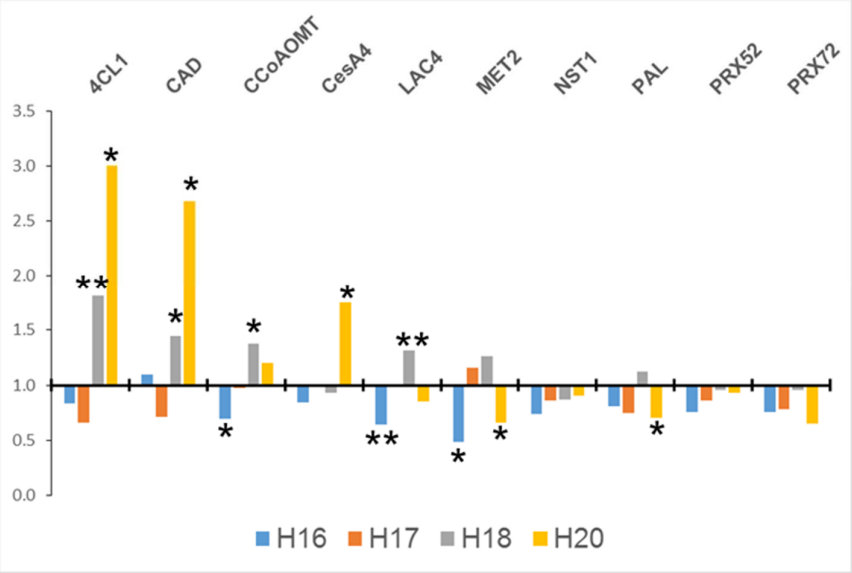


Figure A-6: Gene expression between H16 and H20. The values are fold-change between JA-treated and mock-treated hypocotyls. Values above 1 for genes more expressed in JA-treated hypocotyls (t-test *p*-value < 0.05* and < 0.01**). Abbreviations as in the previous chapters.

Most of the genes of the phenylpropanoids / monolignols pathway (*4CL1*, *CAD* and *CCoAOMT*) were up-regulated by JA in H18 and H20. These two time-points, resulting in 12 samples, have therefore been selected for transcriptomics.

d. Transcriptome overview

Following RT-qPCR analyses, we have investigated by a transcriptomic study the following samples: H18-CONTROL, H18-JA, H20-CONTROL and H20-JA (biological triplicates, resulting in 12 libraries). A pool of the twelve libraries was analysed in six consecutive runs, resulting in an assembly of 37083 contigs (Table A-1).

Table A-1: RNA-Seq assembly statistics

	Length
N75	775 bp
N50	1530 bp
N25	2430 bp
Minimum	240 bp
Maximum	15761 bp
Average	1033 bp
Count	37,083 contigs
Total	38,314,473 bp

The raw reads were processed as described in the Material and Methods section, resulting in the following output (Table A-2).

Table A-2: Statistics of sequence trimming and filtering

	Assembly	
Libraries	Raw sequence reads	Post filtering remaining reads
H18-JA-1	18,157,118	17,108,018
H18-JA-2	25,838,470	23,678,120
H18-JA-3	27,727,368	25,665,142
H18-CTRL-1	21,460,084	20,071,242
H18-CTRL-2	21,040,974	19,260,958
H18-CTRL-3	27,964,642	25,891,622
H20-JA-1	31,124,850	28,929,504
H20-JA-2	19,884,562	18,626,930
H20-JA-3	23,267,070	21,903,520
H20-CTRL-1	44,686,432	41,953,034
H20-CTRL-2	24,116,702	22,746,592
H20-CTRL-3	19,517,040	18,292,346

The clean reads obtained with this experiment were combined with the reads from Chapter 1 and Chapter 4, to obtain a transcriptome assembly of *ca*. 789 M of reads (Table A-3).

Table A-3: Reads used for the mapping

	Total reads
This study	284,127,028
Chapter 1	264,068,902
Chapter 4	241,235,084
TOTAL	789,431,014

The reads of each library were mapped to the obtained reference assembly (Table A-4).

Table A-4: Statistics of the mapping

	Mapping				
Libraries	Total reads	Total reads mapped	Reads mapped with unique hit	Reads mapped with multiple hits	Un-mapped reads
H18-JA-1	17,108,018	14,945,388	14,937,863	7,525	2,162,630
H18-JA-2	23,678,120	20,746,486	20,734,724	11,762	2,931,634
H18-JA-3	25,665,142	22,512,428	22,503,693	8,735	3,152,714
H18-CTRL-1	20,071,242	17,554,375	17,546,687	7,688	2,516,867
H18-CTRL-2	19,260,958	16,902,468	16,893,463	9,005	2,358,490
H18-CTRL-3	25,891,622	22,707,916	22,696,843	11,073	3,183,706
H20-JA-1	28,929,504	25,354,698	25,343,612	11,086	3,574,806
H20-JA-2	18,626,930	16,314,076	16,305,904	8,172	2,312,854
H20-JA-3	21,903,520	19,180,069	19,171,491	8,578	2,723,451
H20-CTRL-1	41,953,034	36,719,155	36,701,414	17,741	5,233,879
H20-CTRL-2	22,746,592	19,958,897	19,950,230	8,767	2,787,595
H20-CTRL-3	18,292,346	16,074,049	16,066,369	7,680	2,218,297

A total of 199 genes were found to be differentially expressed among the four conditions (ANOVA *p*-value < 0.05, min RPKM > 1 and minimum fold change > 2). 51 genes were differentially regulated between H18-CTRL and H18-JA, 38 between H20-CTRL and H20-JA and 55 were not annotated. This is far from the number of genes reported to be JA-sensitive in the literature. The most recent study, performed by RNA-Seq of 15 time-points over a 16-h period in leaves of thale cress, estimates this number to 3611 (Hickman *et al.*, 2017). In their experimental set-up, the 5-week-old plants were treated by dipping the rosette leaves into a mock or methyl jasmonate (0.1 mM) solution. The leaf number 6 (from the base) was harvested at intervals of one-hour and subjected to RNA-Seq. It is therefore likely that the sampling interval of our study is inappropriate, as the transcriptomic response occurs typically within hours. The aim of this experiment was to determine the long-lasting impact of JA on gene expression, which explains the choice of this strategy.

e. Phytohormone analysis

Endogenous phytohormone contents have been measured in H15 and subsequent time-points. With the exception of salicylic acid (SA, Figure A-7), the application of JA did not result in significant changes in terms of phytohormone content.

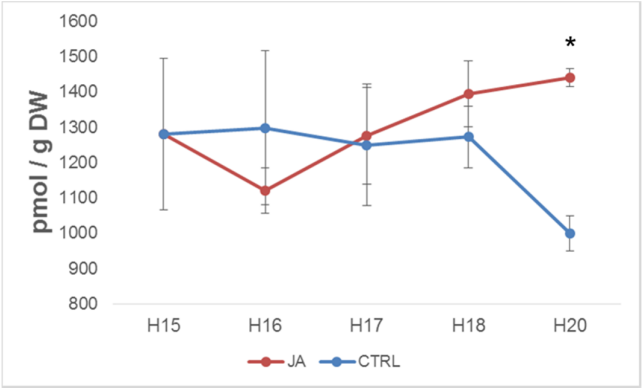


Figure A-7: Salicylic acid content between H15 and H20 (average ± SEM). A Student's *t*-test is performed for each time-point (*p*-value < 0.05*).

JA positively regulates the transcription of several genes of the SA biosynthetic pathway, such as *PAL* (Dempsey *et al.*, 2011). In *Arabidopsis thaliana*, SA strongly antagonizes the JA signalling pathway (Caarls *et al.*, 2015) and represses the expression of key genes of the JA biosynthetic pathway (Leon-Reyes *et al.*, 2010), resulting in the downregulation of a large set of JA-responsive genes (Leon-Reyes *et al.*, 2010).

Conclusion

The aim of this appendix was to further investigate the role of JA during secondary growth and lignification of the hemp hypocotyl. To this end, we have determined with a pilot study that the optimal JA concentration to maximise the effect on these parameters was 0.1 mM. A second experiment has been carried out to highlight the hypocotyl transcriptional reprogramming after JA application. This analysis has been complemented with phytohormone quantification. However, most of these analyses have not provided the expected outcomes: only a few genes were differentially expressed, and no significant change in

the phytohormone profile has been detected. This is likely due to the inadequate sample strategy, which was spanning an interval of several days, whereas the transcriptional reprogramming occurs typically within hours. However, the enhancement of secondary bast fibres biogenesis by JA is a promising element for further investigations and future technical applications. The molecular mechanism leading to this higher cambial activity should be studied with an optimised sampling protocol (shorter interval).

References

Barros, J., Serk, H., Granlund, I., and Pesquet, E. (2015). The cell biology of lignification in higher plants. *Ann. Bot.* 115**,** 1053-1074.

Baucher, M., El Jaziri, M., and Vandeputte, O. (2007). From primary to secondary growth: origin and development of the vascular system. *J. Exp. Bot.* 58**,** 3485-3501.

Blake, A.W., Marcus, S.E., Copeland, J.E., Blackburn, R.S., and Knox, J.P. (2008). In situ analysis of cell wall polymers associated with phloem fibre cells in stems of hemp, *Cannabis sativa* L. *Planta* 228**,** 1-13.

Caarls, L., Pieterse, C.M.J., and Van Wees, S.C.M. (2015). How salicylic acid takes transcriptional control over jasmonic acid signaling. *Front. Plant Sci.* 6.

Dempsey, D.M.A., Vlot, A.C., Wildermuth, M.C., and Klessig, D.F. (2011). Salicylic Acid Biosynthesis and Metabolism. *The Arabidopsis Book* 9**,** e0156.

Didi, V., Jackson, P., and Hejátko, J. (2015). Hormonal regulation of secondary cell wall formation. *J. Exp. Bot.* 66**,** 5015-5027.

Fernandez-Tendero, E., Day, A., Legros, S., Habrant, A., Hawkins, S., and Chabbert, B. (2017). Changes in hemp secondary fiber production related to technical fiber variability revealed by light microscopy and attenuated total reflectance Fourier transform infrared spectroscopy. *PLoS One* 12**,** e0179794.

Gorshkova, T., Brutch, N., Chabbert, B., Deyholos, M., Hayashi, T., Lev-Yadun, S., *et al.* (2012). Plant Fiber Formation: State of the Art, Recent and Expected Progress, and Open Questions. *Crit. Rev. Plant Sci.* 31**,** 201-228.

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.

Hickman, R., Van Verk, M.C., Van Dijken, A.J.H., Mendes, M.P., Vroegop-Vos, I.A., Caarls, L., *et al.* (2017). Architecture and Dynamics of the Jasmonic Acid Gene Regulatory Network. *Plant Cell* 29**,** 2086-2105.

Ko, J.-H., Han, K.-H., Park, S., and Yang, J. (2004). Plant Body Weight-Induced Secondary Growth in *Arabidopsis* and Its Transcription Phenotype Revealed by Whole-Transcriptome Profiling. *Plant Physiol.* 135**,** 1069-1083.

Leon-Reyes, A., Van Der Does, D., De Lange, E.S., Delker, C., Wasternack, C., Van Wees, S.C.M., *et al.* (2010). Salicylate-mediated suppression of jasmonate-responsive gene expression in *Arabidopsis* is targeted downstream of the jasmonate biosynthesis pathway. *Planta* 232**,** 1423-1432.

Liu, M., Fernando, D., Daniel, G., Madsen, B., Meyer, A.S., Ale, M.T., *et al.* (2015). Effect of harvest time and field retting duration on the chemical composition, morphology and mechanical properties of hemp fibers. *Ind. Crop. Prod.* 69**,** 29-39.

Nieminen, K., Blomster, T., Helariutta, Y., and Mähönen, A.P. (2015). Vascular Cambium Development. *The Arabidopsis Book***,** e0177.

Pauwels, L., Morreel, K., De Witte, E., Lammertyn, F., Van Montagu, M., Boerjan, W., *et al.* (2008). Mapping methyl jasmonate-mediated transcriptional reprogramming of metabolism and cell cycle progression in cultured *Arabidopsis* cells. *Proc. Natl. Acad. Sci. U. S. A.* 105**,** 1380-1385.

Ragni, L., and Hardtke, C.S. (2014). Small but thick enough-the Arabidopsis hypocotyl as a model to study secondary growth. *Physiol. Plant.* 151**,** 164-171.

Roach, M.J., and Deyholos, M.K. (2008). Microarray Analysis of Developing Flax Hypocotyls Identifies Novel Transcripts Correlated with Specific Stages of Phloem Fibre Differentiation. *Ann. Bot.* 102**,** 317-330.

Sehr, E.M., Agusti, J., Lehner, R., Farmer, E.E., Schwarz, M., and Greb, T. (2010). Analysis of secondary growth in the *Arabidopsis* shoot reveals a positive role of jasmonate signalling in cambium formation. *Plant J.* 63**,** 811-822.

Snegireva, A., Chernova, T., Ageeva, M., Lev-Yadun, S., and Gorshkova, T. (2015). Intrusive growth of primary and secondary phloem fibres in hemp stem determines fibre-bundle formation and structure. *AoB Plants* 7.

Strabala, T.J., and Macmillan, C.P. (2013). The Arabidopsis wood model—the case for the inflorescence stem. *Plant Sci.* 210**,** 193-205.

Sundell, D., Street, N.R., Kumar, M., Mellerowicz, E.J., Kucukoglu, M., Johnsson, C., *et al.* (2017). AspWood: High-Spatial-Resolution Transcriptome Profiles Reveal Uncharacterized Modularity of Wood Formation in *Populus tremula*. *Plant Cell* 29**,** 1585-1604.

Tamaoki, D., Karahara, I., Nishiuchi, T., Wakasugi, T., Yamada, K., and Kamisaka, S. (2011). Involvement of auxin dynamics in hypergravity-induced promotion of lignin-related gene expression in *Arabidopsis* inflorescence stems. *J. Exp. Bot.* 62**,** 5463-5469.

Ursache, R., Nieminen, K., and Helariutta, Y. (2013). Genetic and hormonal regulation of cambial development. *Physiol. Plant.* 147**,** 36-45.

Wang, H.Z., and Dixon, R.A. (2012). On-Off Switches for Secondary Cell Wall Biosynthesis. *Mol. Plant* 5**,** 297-303.